Build-up and Resetting of Auditory Stream Segregation in Quiet and in Complex-Tone Backgrounds

Nicholas Russell Haywood

Doctor of Philosophy

ASTON UNIVERSITY

August 2009

Aston University

Build-up and Resetting of Auditory Stream Segregation in Quiet and in Complex-Tone Backgrounds

Thirteen experiments investigated the dynamics of stream segregation. Experiments 1-6b used a similar method, where a same-frequency induction sequence (usually 10 repetitions of an identical pure tone) promoted segregation in a subsequent, briefer test sequence (of alternating low- and high-frequency tones). Experiments 1-2 measured streaming using a direct report of perception and a temporal-discrimination task, respectively. Creating a single deviant by altering the final inducer (e.g., in level or replacement with silence) reduced segregation, often substantially. As the prior inducers remained unaltered, it is proposed that the single change actively reset build-up. The extent of resetting varied gradually with the size of a frequency change, once noticeable (experiments 3a-3b). By manipulating the serial position of a change, experiments 4a-4b demonstrated that resetting only occurred when the final inducer was replaced with silence, as build-up is very rapid during a same-frequency induction sequence. Therefore, the observed resetting cannot be explained by fewer inducers being presented. Experiment 5 showed that resetting caused by a single deviant did not increase when prior inducers were made unpredictable in frequency (four-semitone range). Experiments 6a-6b demonstrated that actual and perceived continuity have a similar effect on subsequent streaming judgements – promoting either integration or segregation, depending on listening context. Experiment 7 found that same-frequency inducers were considerably more effective at promoting segregation than an alternating-frequency inducer, and that a trend for deviant-tone resetting was only apparent for the same-frequency case. Using temporal-order judgments, experiments 8-9 demonstrated the stream segregation of pure-tone-like percepts, evoked by sudden changes in amplitude or interaural time difference for individual components of a complex tone. Active resetting was observed when a deviant was inserted into a sequence of these percepts (Experiment 10). Overall, these experiments offer new insight into the segregation-promoting effect of induction sequences, and the factors which can reset this effect.

Nicholas Russell Haywood
Doctor of Philosophy
August 2009

KEY PHRASES: *Auditory perception, Scene Analysis, Deviant Tone, Attention, Perceived Continuity*

Contents

Dedication

For Kate,

Her endless patience and support kept me going.

Acknowledgments

I have really enjoyed the time that I have spent working on this project, thanks in no small part to the people that I have been lucky enough to work with. I am indebted to my supervisor, Brian Roberts, for his unreserved support. On both a personal and a professional level, it has been a real pleasure to work with Brian. Throughout my research, I have benefitted greatly from his knowledge, guidance, encouragement and critical thinking. This thesis has profited from his many helpful comments and suggestions. The document has also been improved thanks to thoughtful feedback from Peter Bailey and Caroline Witton.

I also owe particular thanks to Steve Holmes for his invaluable help with programming. My initial progress would have been delayed a great deal without his Visual Basic know-how. More generally, I am grateful to all of my lab colleagues for providing a friendly and enjoyable working environment. Similarly, I would like to acknowledge my fellow postgraduate students for their friendship and camaraderie. A special mention should go to those who I worked alongside in office 711.

Many of the studies presented in this thesis have also contributed to a number of poster presentations at various conferences (see the appendix section for full details). Thanks goes to all of those who have offered helpful suggestions and comments on my research.

Outside of academia, I am thankful for all of the encouragement that I have received from my family and friends. I am especially grateful to my parents, who have always been a constant source of support. Finally, I would like to thank Kate, who always stood right by me through all of the difficult times.

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Auditory Scene Analysis and Stream Segregation

In everyday listening environments we are often exposed to multiple acoustic events. Examples of this could include the sounds of a busy street or party, or many musical instruments playing together in an orchestra. Normally, we are able to focus on one event in the acoustic environment and hear it separately from all others - for example, we can listen to an individual speaker at a party, or choose to listen to one specific instrument in an orchestra. Most of the time, we are able to do this effortlessly - although this task can become difficult in highly reverberant environments. Despite this apparent ease, our auditory system must solve a complex problem in order for us to follow accurately the target source. Namely, our ears receive a mixture of all the acoustic events happening around us, so we must actively separate the mixture to recover descriptions of distinct perceptual objects. In order to obtain accurate information, each perceptual object must correspond to the acoustic event from which it originated. This process has been termed *auditory scene analysis* (Bregman, 1990). Bregman (1990) made a useful distinction between two interlinked problems that our auditory system must solve for accurate perceptual representation. Firstly, we must correctly group sounds that occur at the same time (termed *simultaneous grouping*). In many circumstances, simultaneously occurring sounds originate from the same acoustic event (e.g., the harmonic components of a musical chord). However, it is also possible that two separate acoustic events simply overlap in time. Only through a careful analysis of the relationships between simultaneously occurring sounds can our auditory system estimate whether sound elements are related or not. Secondly, we must correctly group sounds over time (referred to as *sequential grouping*). For example, we can only follow the melody played by an instrument if the successive notes are grouped together as a single perceptual object (called a *perceptual stream*). Furthermore, this task must often be accomplished in the presence of co-occurring, irrelevant acoustic events. In this circumstance, our auditory system must group all the relevant sounds into a perceptual stream, but also exclude any irrelevant sounds from the stream. As these sounds may be mixed together over time, factors other than temporal proximity must influence the stream-formation process. The parsing of sequential sounds into separate perceptual streams is commonly referred to as *auditory stream segregation* (Bregman, 1990).

Although the distinction between simultaneous and sequential grouping is often useful conceptually and experimentally, there are clear interactions between the two processes. Indeed, in many cases, both aspects of grouping can occur at once. As with grouping principles in general, both sequential and simultaneous grouping can either complement or challenge each other in the process of forming a correct representation of the auditory scene. One way in which simultaneous and sequential grouping can interact is examined in Chapter 9.

1.1.1 Auditory Stream Segregation

The previously described situations (a busy party or an orchestra) are both good examples of stream segregation in an everyday context. However, in such acoustic environments there would normally be multiple acoustic cues that might influence sequential grouping (e.g., the pitch, timbre and/or location of the sounds). Therefore, stream segregation is often studied using isolated pure tones. For such stimuli, the individual factors that influence grouping are reasonably well understood. This section outlines two of the pioneering studies of stream segregation.

A repeating sequence of pure tones which alternate rapidly on a single dimension may be heard as two separate streams, with each stream corresponding to the tones with these alternate properties (a percept which is referred to as *segregation*). In contrast, the alternating tones may group together and form a single perceptual stream (*integration*). The frequency separation between the alternating tones can strongly influence stream segregation. The first empirical study of this was conducted by Miller & Heise (1950). They presented a repeating LH sequence, (where L and H represent low- and high-frequency tones), and participants adjusted the frequency of the high tone. All tones were grouped as an integrated percept when the frequency separation between the tones was small. In contrast, at larger frequency separations the low and high tones were heard as separate low- and high-pitched streams (i.e., segregation).

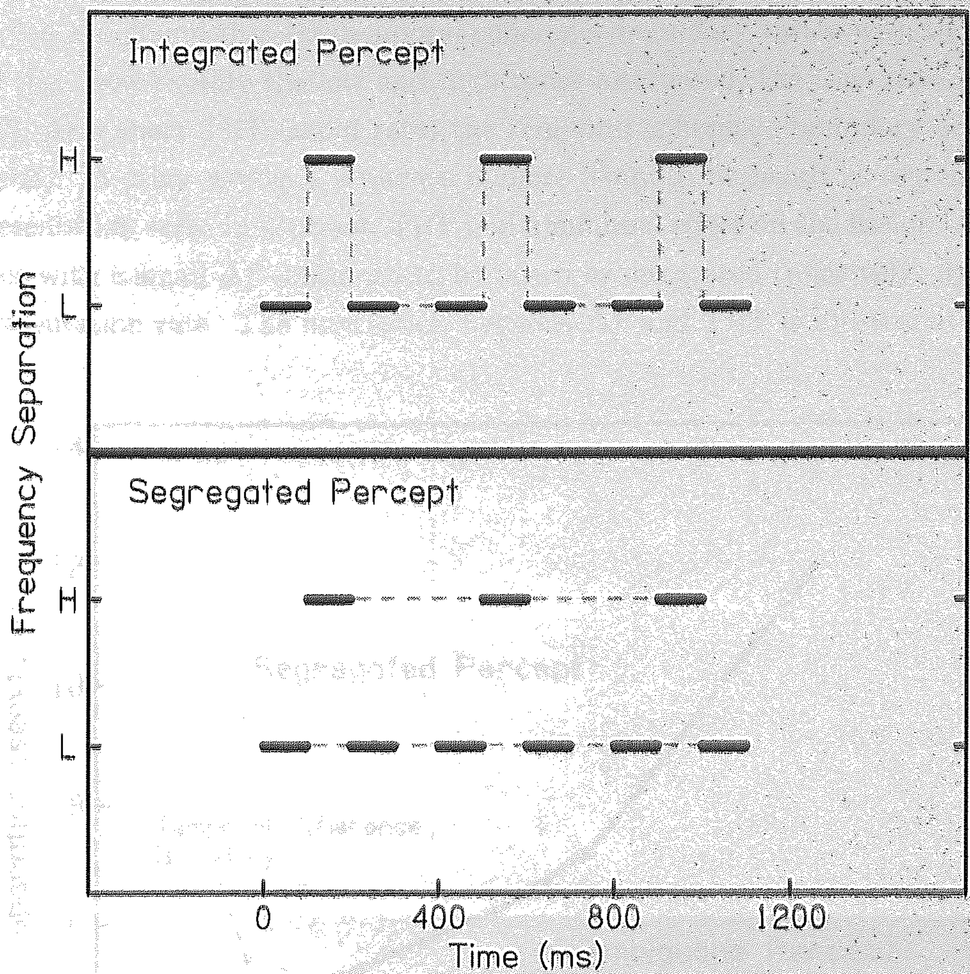

Figure 1.1: **Solid lines represent pure tones of either a low (L) or a high (H) frequency. The tones are arranged in a LHL-LHL-LHL- sequence (originally developed by van Noorden, 1975). Dashed grey lines represent perceptual grouping. In the top panel, the L & H tones are grouped together (integrated), whereas in the bottom panel, two separate streams are perceived, each corresponding to either the L or the H pitched tones.**

The basic finding of Miller & Heise (1950) was studied further in a series of experiments by van Noorden (1975). He presented sequences comprising repeating LHL- tone arrangements (where - represents a silent interval). For such a sequence, an integrated percept is heard as a clear *gallop-like* rhythm, but when the tones segregate the gallop is lost and two separate monotonous streams of isochronous rhythms are perceived (i.e., the tempo of the L stream rhythm is twice as fast as that of the H stream). This clear differentiation in rhythm allows participants to detect easily changes in perceptual grouping. These alternative perceptual representations of a repeating LHL- sequence are illustrated in Figure 1.1. van Noorden (1975) presented repeating LHL- stimuli at various frequency separations (Δf) between the low and high frequency tones. Segregation was always perceived above a certain Δf; this threshold was termed the temporal coherence boundary. Similarly, below a given Δf, integration was always reported (the fission boundary). Between these two boundaries exists an ambiguous region, in which no percept dominates, and the sequence may be heard as either integrated or segregated.

van Noorden also found that the temporal coherence boundary varied with the rate of a sequence (i.e., how rapidly the low and high tones alternated; the tone repetition time - or TRT). At a short TRT (rapid rate) the temporal coherence boundary occurred at a smaller Δf. In other words, a sequence is more likely to be heard as segregated at a faster presentation rate. In contrast, TRT had a minimal effect on the fission boundary - a sequence with a small Δf continued to be heard as integrated (relatively) irrespective of the presentation rate. The interaction between Δf and TRT is illustrated in Figure 1.2.

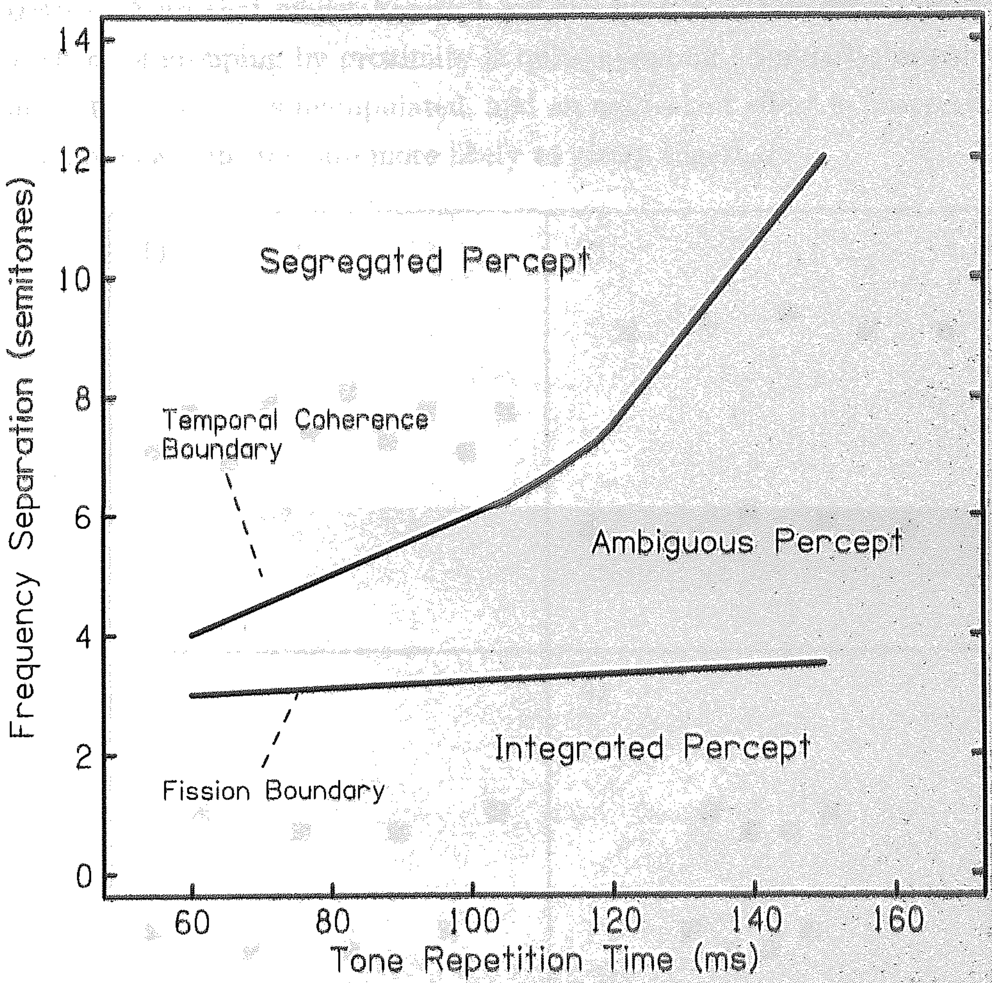

Figure 1.2: **The interaction between frequency separation and presentation rate on the perception of a repeating LHL- sequence. Frequency separation is measured in semitones, and presentation rate is measured in terms of tone repetition time (TRT = the onset-to-onset interval between successive tones). A short TRT corresponds to a fast rate and vice versa. Adapted from van Noorden (1975).**

The findings of Miller & Heise (1950) and van Noorden (1975) have been interpreted by Bregman (1990) in terms of the Gestalt principles of perceptual organisation (e.g., Koffka, 1935). Whilst many of the Gestalt principles are applicable to auditory perception in general, the proximity principle is perhaps most relevant to the streaming of sequential pure-tone stimuli. This principle states that the strength of perceptual grouping increases as elements become closer together. This is represented in Figure 1.3. For the panels 1 and 2, the frequency separation between tones is adjusted. In panel 1, the low- and high-frequency elements both occupy a similar frequency region, and so each successive element is likely to be grouped with its immediate neighbours (i.e., perceptual integration). Indeed, if the frequency separation between these elements fell below the fission boundary, a listener would certainly hear an integrated percept (van Noorden, 1975). However, when a larger frequency separation is introduced in panel 2, two distinct perceptual groups emerge - only the tones which are closer in frequency group together. Note that whilst Figure 1.3 is designed to represent auditory stimuli, the visual effect of grouping by proximity is quite apparent. Similarly, in panels 3 and 4, the time between tones is manipulated, and an equivalent effect is observed. Similar tones which are closer in time are more likely to group together.

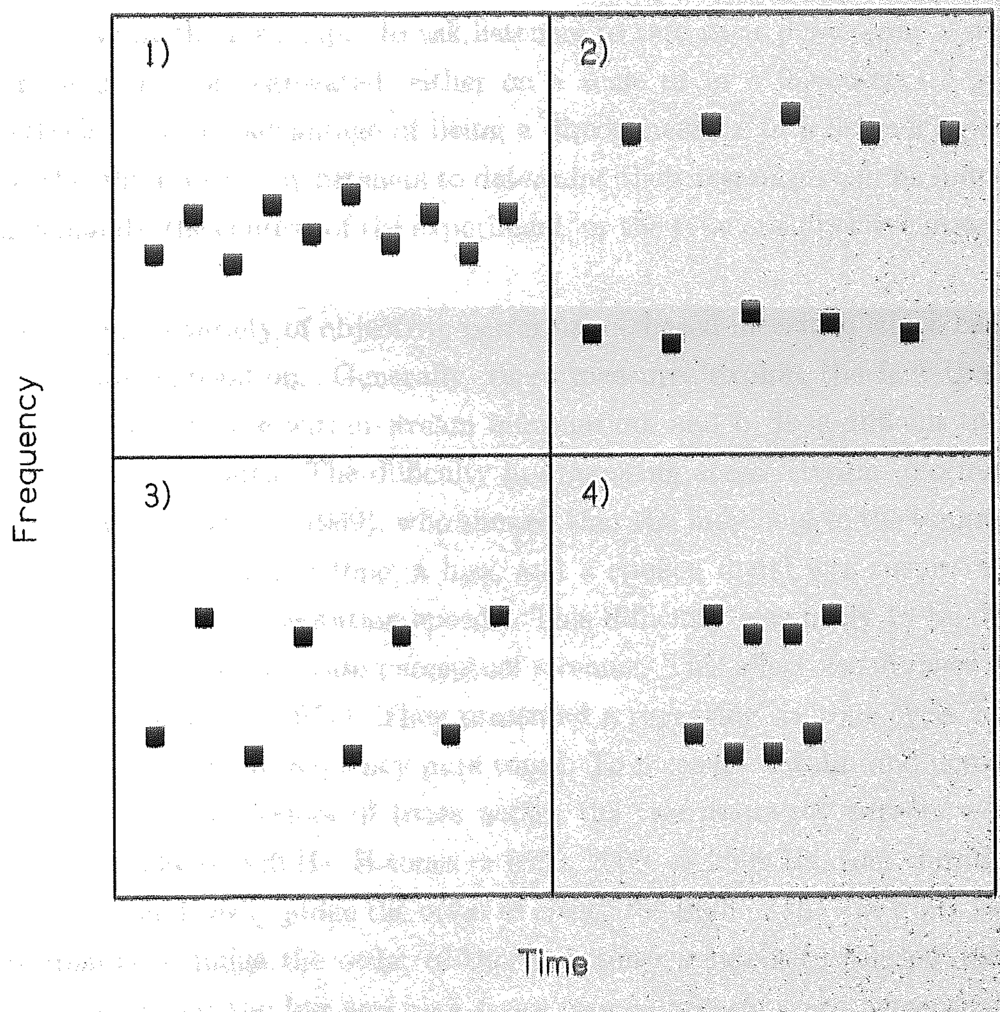

Figure 1.3: **Illustration of grouping by proximity. Panels 1) and 2), the perception of two distinct groups (i.e., segregation) increases as the frequency separation between tones is increased. Panels 3) and 4), stream segregation becomes more likely as the tones are presented at a faster rate. Adapted from Bregman (1990).**

The grouping principle of proximity also relates to the ecological purpose of auditory perception (Bregman, 1990). If we consider a naturally occurring sound source, a very large change in frequency is quite unlikely to occur rapidly. A more likely explanation would be that two separate acoustic events had occurred in succession. In contrast, a smaller or more gradual change in frequency would be likely to occur in an on-going natural sound source. This reasoning is in good accord with the interaction between frequency separation and tone-repetition-time observed by van Noorden (1975). A tone sequence comprising large frequency changes is only heard as integrated if there is a sufficient time-interval between successive tones (Bregman, 1990).

1.1.2 The Measurement of Stream Segregation

Stream segregation can be measured in a variety of ways. The previously described studies by Miller & Heise (1950), and van Noorden (1975), both relied on listeners reporting their perception of a tone sequence (a subjective measure). Perhaps the most straightforward subjective method is simply to ask listeners to rate their perception of a sequence as either integrated or segregated, either on a scale or in a forced-choice procedure. Such methods have the advantage of being a direct measure of a listener's perception. However, the criteria used by listeners to determine their responses can be influenced by the task demands, the context of the experiment, or the type of subjective measure used.

There are also a variety of objective, performance-based measures which can be used to index stream segregation. Generally, these measures exploit the fact that a) it is relatively easy to compare within-stream information, and b) it is difficult to compare information across streams. The difficulty in organising across-stream information was illustrated by Warren et al. (1969), who showed that the judgment of the temporal order of unrelated items (a buzz, a tone, a hiss, and a spoken digit) was surprisingly poor, even at relatively slow presentation speeds. This difficulty was likely to be due to the different sounds forming separate perceptual streams. This effect was further examined by Bregman & Campbell (1971). They presented a repeating six-tone cycle comprising three low-, and three high-frequency pure tones. To measure within- and across-stream grouping, the frequency values of tones *within* the two frequency subsets were varied (L-tones = 350, 430, or 550 Hz, H-tones = 1600, 2000, or 2500 Hz, tone duration = 100 ms). Listeners could easily judge the order of either the high or the low tones separately, but were unable to judge the order of the high tones relative to the low tones. The authors proposed that the low and high tones formed two separate perceptual streams - and that whilst within-stream temporal judgments were easy, across-stream temporal judgments were much more difficult. This general principle also affects listeners sensitivity to temporal relationships more generally (i.e., across-stream timing differences are difficult to detect), and this has been successfully exploited to measure the effects of various tone and sequence parameters on streaming (e.g., Vliegen et al., 1999; Cusack

& Roberts, 2000; Roberts et al., 2002, 2008). The use of temporal discrimination as a measure of stream segregation is discussed in more detail in Chapter 3.

Conversely, some measures have exploited the fact that stream segregation can *benefit* performance in certain tasks. More specifically, streaming can aid performance when task-irrelevant information is heard as segregated from the task-relevant information. For example, Dowling (1973) showed that two interleaved melodies could only be recognized if they occupied separate frequency ranges. In this circumstance, listeners could attend to one of the melodies, whilst the other was heard in the background (both melodies could not be attended simultaneously). When the two melodies occupied the same range, a meaningless jumble was heard. This effect of frequency differences on performance is in good accord with the known effect of frequency separation on stream segregation (van Noorden, 1975).

1.1.3 Peripheral Channelling is not a Pre-Requisite for Stream Segregation

So far, we have seen that the frequency separation between pure tones is a strong cue for stream segregation. However, perceptual properties besides pitch can also promote streaming. For example, a range of studies using complex tones have shown that both pitch and timbre differences can affect stream segregation (for example, Wessel, 1979; Singh, 1987). This point was well illustrated by Singh and Bregman (1997), who measured the difference in fundamental frequency ($F0$) required for the segregation of two alternating harmonic complex tones. They observed that a smaller $F0$ difference was required for segregation when the complex tones also differed in amplitude envelope (rise and fall times of 5 and 95 ms, or of 95 and 5 ms) than for when they did not. This demonstrates that differences besides those in peripheral channel excitation can enhance segregation, but not that these cues are sufficient to promote streaming on their own.

More recent research has addressed this issue and demonstrated that stream segregation can occur in the absence of differences in peripheral excitation. For example, in a melody recognition task, Vliegen and Oxenham (1999) demonstrated streaming in sequences of complex tones which comprised only unresolved harmonics with a common passband. For this arrangement, the notes of the melody and the interleaved distractors were differentiated only by changes in the $F0$ of the tones. In a similar task, Cusack & Roberts (2000) also observed streaming when spectral cues were (largely) removed - sequences comprised interleaved pure tones and narrow-band noises (within a restricted frequency range). Melodic notes were defined by changes in centre frequency. These findings are supported by the fact that stream segregation evoked solely by timbral differences will also *impair* performance in a temporal discrimination task (Vliegen et al., 1999; Cusack & Roberts, 2000; Roberts et al., 2002).

In a review of this literature, Moore & Gockel (2002) concluded that stream segregation is related to the perceptual differences between sounds, and that any salient difference may lead to streaming - irrespective of peripheral channelling cues. This evidence is contrary to earlier theories of stream segregation which proposed that streaming is determined exclusively by patterns of peripheral excitation (Hartmann & Johnson, 1991; Beauvois & Meddis, 1996, 1997). Despite this, it should be noted that spectral differences remain a very strong cue for stream segregation.

1.2 The Build-up of Stream Segregation

When listening to a repeating sequence of unchanging low and high frequency pure tones, the tendency to hear stream segregation increases over time. van Noorden (1975) commented on this effect, but the first direct study of the build-up of segregation over time was conducted by Bregman (1978). He varied the length of a "tone package" and measured perceived segregation (packages comprised either of 12, 24 or 48 repetitions of an HL stimulus). For each trial, one type of package was repeated indefinitely, with each individual package separated by four seconds of silence. In a final condition, the HL stimuli were repeated continuously with no interleaving silences. Perceived stream segregation was measured by a process of titration[1]. Participants increased the rate of the sequence, from an initial tone-repetition-time of 600 ms, until segregation was heard. This rate was taken as a measure of segregation - an increased sequence rate is known to promote streaming, and so hearing segregation at a slower rate is an indication that another factor is also promoting segregation. It was found that as the package length increased, segregation was heard at slower sequence rates. This is evidence that as the number of tones in a package increases, so too does the tendency for streaming to occur.

Subsequent studies have confirmed Bregman's (1978) initial finding by directly measuring the rate at which build-up occurs during a tone sequence. Anstis & Saida (1985) presented relatively long (30 s and 60 s), unchanging and uninterrupted sequences of alternating low and high frequency pure tones. In one experiment, participants were asked to report continuously their perception by holding down one of two keys to indicate either integration or segregation. For all sequences tested, integration was initially heard, but the tendency to report segregation built up over the course of the sequence,

[1] When two factors promote stream segregation, the relative contribution of one factor can be measured by the amount of the second factor required for streaming to occur.

This general finding was true for all the stimuli measured, but the tonal properties of the sequence were shown to affect both the rate and the overall extent of build-up. In two separate experiments, either the frequency separation between tones or the presentation rate was varied. For sequences more likely to be heard as segregated (i.e., those with a large Δf or a rapid TRT) initial build-up was more rapid and the overall probability of reporting segregation was higher (see Figure 1.4). Build-up is typically characterized by an initially rapid increase in segregation (i.e., within the first 5 - 10 s of a sequence), and a much more gradual increase over the remaining portion of the sequence. Owing to the rate of increase of streaming in the first few seconds of a sequence, build-up can be observed in a relatively short sequence. Anstis & Saida (1985) also demonstrated build-up by asking listeners to adjust continuously the rate of a long sequence in order to track the threshold between integration and segregation. This threshold was initially at a fast rate, but over the course of the sequence, the rate had to be increasingly slowed in order for the same threshold to be apparent. As with Bregman's (1978) earlier study, a slower rate indicated an increased difficulty for hearing integration. Therefore, Anstis & Saida (1985) found consistent evidence that the tendency to hear segregation builds up over the course of a sequence.

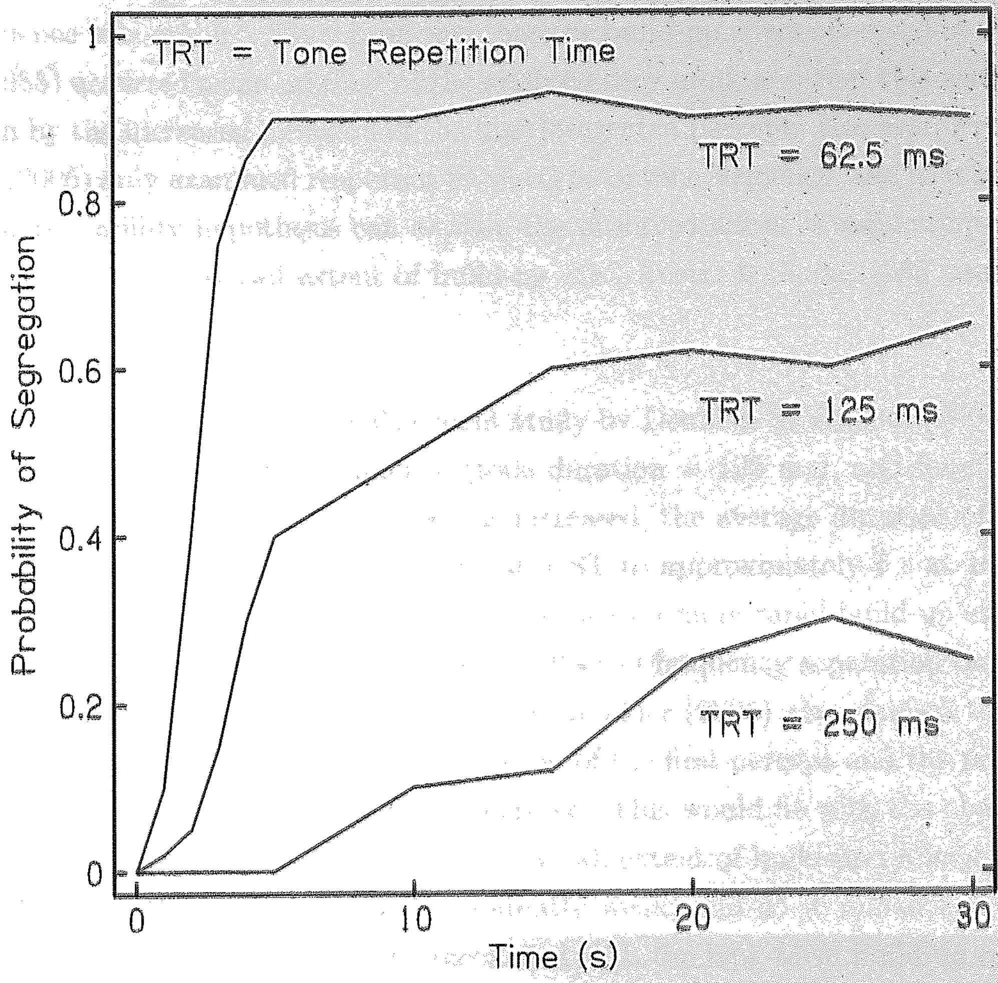

Figure 1.4: **Listeners continuously rated their perception of a 30 s repeating LH sequence (L tones = 800 Hz, H tones = 1200 Hz). Tone presentation rate was varied across trials (tone repetition time = the onset-to-onset interval between successive tones). The probability of segregation was calculated from the data of 5 listeners who each completed 5 repetitions of each condition. Adapted from Anstis & Saida (1985).**

In their study, Anstis & Saida (1985) analysed response data by averaging across trials. Some more recent studies have used a similar direct report of perception, but have analysed the duration of the individual intervals of each integrated or segregated percept (Pressnitzer & Hupé, 2006; Denham & Winkler, 2006). By taking this measure and calculating the average duration of each successive percept, these authors have shown that in long, four minute, tone sequences segregation is never exclusively heard (i.e., the build-up of streaming does not continue to a point were a listener only hears segregation). Pressnitzer & Hupé (2006) presented an HLH- sequence (H = 587 Hz, L = 440 Hz, tone duration = 120 ms), and showed that the rate of switching was constant over the length of the sequence. In other words, there was no long-term trend in the average durations of each percept (integrated or segregated) - on average, each lasted approximately 9 s. The exception to this was that the initial integrated percept lasted substantially longer than all subsequent percepts (around 17 s). The pattern of results observed led Pressnitzer & Hupé (2006) to conclude that perception of an unchanging tone sequence is bi-stable[2]. They proposed that the build-up effect observed by Anstis & Saida (1985) occurred as an artefact of the analysis they used, and that this was primarily driven by the increased duration of the first integrated percept. However, Pressnitzer & Hupé (2006) only examined responses to one type of tone sequence, and so it is unclear how their bi-stability hypothesis can explain the observed effect of sequence properties on both the rate and overall extent of build-up (i.e., Anstis & Saida, 1985 - see Figure 1.4).

This issue was addressed in a subsequent study by Denham & Winkler (2006). They presented a four minute LHL- sequence (tone duration = 125 ms), and found that as the frequency separation between tones was increased, the average duration of the first percept declined (from approximately 20 s at 4 ST to approximately 6 s at 10 ST). A shorter initial integrated percept would correlate with a more rapid build-up effect, and so this finding is consistent with the observed effect of frequency separation on the rate of build-up (Anstis & Saida, 1985). Denham & Winkler (2006) also claimed to observe a positive correlation between the (log) duration of the first percept and the proportion of time that the sequence was heard as integrated. This would fit with the observed effect of increased frequency separation on the overall extent of build-up (Anstis & Saida, 1985). However, this correlation was apparently weak, and so it remains to be seen how the bi-stability hypothesis can be reconciled with the long-term build-up of stream segregation. Furthermore, it is also unclear as to how the bi-stability hypothesis can explain the results observed from other measures of build-up. For example, Anstis & Saida (1985) observed that the presentation rate required for segregation decreased over the course of a tone sequence. If the perception of a tone sequence is bi-stable, one would not expect this long-term change to be observed. Despite these issues, the evidence seems to suggest that the build-up of segregation is at least partly governed by an extended integrated percept at the onset of a sequence.

Currently, the physiological mechanisms responsible for the build-up of stream segregation are not fully understood. Most explanations of build-up have addressed the functioning of peripheral mechanisms (Beauvois & Meddis, 1996; McCabe & Denham, 1997). More specifically, these authors have argued that build-up is influenced by the

[2] The key characteristics of perceptual bi-stability, as defined by Leopold and Logothetis (1999), are; 1) exclusivity (perceptual representations are mutually exclusive - only one representation may be perceived at a given time), 2) inevitability (switches in perception will always occur) and 3) randomness (the successive durations of percepts are uncorrelated). Perceptual bi-stability has primarily been demonstrated in a variety of ambiguous visual stimuli. (e.g., a Necker cube - Necker, 1832, or binocular rivalry - Helmholtz, 1925).

peripheral channelling of alternating tones (i.e., grouping is based on differences in frequency region of excitation). The models of Beauvois & Meddis (1996) and McCabe & Denham (1997) both successfully replicated earlier studies of build-up (e.g., Anstis & Saida, 1985). Despite this, these models must be at least partially incomplete as there is evidence that differences in peripheral excitation are not a pre-requisite for streaming to occur (Vliegen et al., 1999; Cusack & Roberts, 2000). Indeed, a build-up of segregation has been observed in a sequence when sounds comprising unresolved harmonics differed in phase cues, but not in excitation patterns evoked (Roberts et al., 2002). Anstis & Saida (1985) proposed a theory of build-up which was based on adaptation of frequency-motion detectors. However, there is now good evidence that stream segregation can be induced by a preceding sequence of same-frequency tones (see the next section). Based on this, the notion of frequency-motion detectors cannot fully explain the build-up of stream segregation.

Despite the uncertainty regarding the physiological basis of build-up, Bregman (1978, 1990) has proposed a functional explanation for the observed effect. He proposed that at the onset of a sound sequence, the default assumption is that all acoustic elements have arisen from the same acoustic event. However, as the two subsets of sounds have consistently different properties, the stream-formation mechanism becomes increasingly likely to segregate the sounds on the basis of their continued difference from one and other. This evidence-accumulation process is relatively slow, so that changes in grouping are reasonably conservative (in order to prevent excessive fluctuations in the perceptual organization of a sound sequence). This proposed mechanism is in good accord with the evidence that an initial integrated percept lasts substantially longer than all subsequent percepts (Pressnitzer & Hupé, 2006; Denham & Winkler, 2006).

1.2.1 The Segregation-Promoting Effect of a Same-Frequency Induction Sequence

Rogers & Bregman (1993) explored the factors that can promote the build-up of stream segregation. They presented stimuli that were divided into two sections; a relatively long induction sequence and a subsequent, briefer test sequence (there was no break between these two sequences). Participants indicated their perception of the test sequence (which comprised 3 HLH- triplets), and the effect of various induction sequences on streaming was measured. They found that an induction sequence which comprised only H- repetitions was highly effective at promoting segregation in the test sequence (i.e., the induction sequence differed from the test sequence only in that all of the L tones were replaced with a silence of an equivalent duration). From this study, it would appear that the extent of perceived segregation following this single-frequency induction sequence is somewhat similar to that following an equivalent alternating-frequency sequence (in terms of the number of H-tone onsets). This finding indicates that frequency alterations are not required for the tendency to hear stream segregation to build up. However, it should be noted that no study has directly compared build-up following same-frequency and alternating-frequency induction sequences, and so it is currently unclear exactly how alike these two processes are (this issue is explored further in Chapter 8).

Rogers & Bregman (1993) manipulated the tone density and number of tone onsets in a same-frequency induction sequence (tone density = 20% of sequence - and one onset per 400 ms, 40% and 2 onsets, 60% and 3 onsets, or 80% and 4 onsets). They found that a same-frequency induction sequence was most effective at promoting segregation when the H tone density and number of onsets was matched to that in the test sequence (the 60% density case). As induction tone density/number of onsets decreased from this, so too did reported segregation. In contrast, an increase in induction tone density did not increase perceived segregation beyond that when the tone density was matched to that of the test sequence. In a final condition, a single continuous induction tone (i.e., 100% tone density and one onset) did not promote *any* segregation in the test sequence (i.e., perception was statistically identical to that when the test sequence was heard in the absence of any prior build-up). Based on this evidence, Rogers & Bregman (1993) concluded that the total duration of tones does not predict the segregation-promoting effect, but rather that most induction occurred when there was a rough matching in properties between the induction and test sequence tones.

In a second experiment, the tone density and number of tone onsets in the induction sequence was fixed, but the onset times and durations of the individual tones were varied. Regular and temporally unpredictable induction sequences both promoted the same amount of stream segregation. In a third experiment, Rogers & Bregman (1993) presented the test sequence to the right ear only. An HLH- type induction sequence was effective at promoting segregation when it was also presented to the right ear. However, much less build-up was observed when the induction sequence was presented to both ears. This result was attributed to the change in intensity (left ear) and perceived lateralization at the induction/test sequence boundary. Indeed, there is a body of evidence which indicates that abrupt changes between the induction and test sequence can *reset* the build-up of segregation. The concept of resetting is discussed in depth in the next section. In a related condition, segregation was reduced when the standard right-ear induction sequence was accompanied in the left ear by either white noise, or an alternative tone sequence[3]. The evidence that irrelevant sounds presented to one ear can affect streaming

[3] The alternative tone sequence presented to the left ear was an HLH- sequence of dif-

judgments in the other ear suggests that the mechanism responsible for the build-up of streaming is influenced by central factors (i.e., build-up is not solely governed by peripheral cues). Again, a likely explanation for this finding is that the abrupt change in sequence properties at the induction/test sequence boundary had a resetting effect on the build-up of segregation.

1.3 The Decay and Resetting of Stream Segregation

So far, we have seen that the tendency to hear segregation builds up over the course of a sequence. There is also clear evidence that, once stream segregation has built up, certain changes to a sequence can cause grouping to revert back to an integrated percept. There are two general ways in which this can occur; relatively gradually over a silent interval, or suddenly after an abrupt change in the tonal properties of an on-going sequence. These two processes are considered separately below.

1.3.1 The Decay of Build-up during Silence

The decay of build-up was first demonstrated by Bregman (1978). By varying the length of the silent intervals between successive tone packages (silences between 0 to 4 s, tone sequences of 12 HL repetitions), he demonstrated that listeners heard segregation at a faster presentation rate when the intervening silent intervals were made longer. As a faster rate promotes segregation, this is evidence that longer silent intervals reduced streaming in the tone sequences. The rate of this decay was rapid over the first 1.5 seconds of silence, but still apparent at the longest (4 s) silent interval tested. Bregman (1978) also demonstrated that this decay of build-up occurred at the same rate whether the intervening intervals were silent or filled with white noise. A similar study by Beauvois & Meddis (1997) also demonstrated the gradual decay of build-up during a silent interval. A 10-s induction sequence of a repeating low frequency tone was presented to promote segregation in a subsequent test sequence (the test sequence was a brief, 1.44 s, LH sequence). With no interleaving silence, the induction sequence strongly promoted segregation in the test sequence. By measuring perception when a range of silent intervals were inserted between the induction and test sequence (0 to 8 s, in 0.5 s steps), it was demonstrated that perceived segregation in the test sequence decreased as the duration of the silent interval increased. Specifically, the decay of stream segregation in silence was best fitted by an exponential function with a time constant of approximately 3.8 seconds. However, a more detailed analysis of the data revealed a different time-course of decay for musically trained listeners and for non-musicians. For musically

ferent frequencies to, and temporally offset from the right-ear sequence. This alternative sequence was designed to "disguise" the rhythm of the right-ear sequence.

trained listeners, they found a gradual, linear decay of build-up across the range of silent intervals tested. Decay was faster for the other listeners. Overall, Beauvois & Meddis (1997) concluded that the build-up and decay of stream segregation appear to occur at a similar rate.

A somewhat more rapid decay of build-up was observed by Cusack et al. (2004), who inserted a brief silence into an on-going LHL- sequence. They found an almost full decay of stream segregation after the shortest silent interval that they tested (1 s). Furthermore, as decay appeared almost complete after 1 s, the extent of decay was broadly consistent across the range of silent intervals that they measured (from 1 to 10 s). These authors offered two explanations for their findings. Firstly, following a silence, the continuation of a sequence may be heard as a new stream (for which no prior build-up has occurred). Secondly, the silent interval may draw attention to the subsequent sounds, and a change in attentional focus may have a resetting effect. There is some support for this second argument from related literature, which is discussed in section 1.4.

1.3.2 The Sudden Resetting of Build-up

Several studies have shown that a sudden change in an on-going sequence can reset build-up. This was first observed by Anstis & Saida (1985), who applied a sudden alteration to the tonal properties of a repeating LH sequence. Either the ear of presentation was changed, or an equal shift in frequency was applied to both the L- and H-tones. A standard sequence was heard for four seconds (the adapting stimulus), and then the tonal properties were altered for one second (the test stimulus). This five second arrangement was repeated for 90 s, with no break in sequence. Participants adjusted the rate of the *test stimulus*, until it was heard at the threshold between integration and segregation. The average rate was calculated from the last 30 s of each trial. When the test stimulus was no different from the adapting stimulus, a slow rate was required to track this threshold. A slow rate indicates difficulty in hearing integration, therefore this result shows that segregation had built up in the unchanged test stimuli. However, when the test sequence was presented in a different ear or frequency region from the adapting sequence, the same threshold was tracked at a faster rate (indicating that less build-up of stream segregation had occurred for the test stimuli). Resetting was only observed when the frequency change was outside the range from -1 to +3 semitones (ST) from the centre frequency of the adapting stimuli. Whilst this range is roughly equal to the bandwidth of the auditory filter in this frequency region (Glasberg & Moore, 1990), the reason for the +1 semitone offset from the adapting stimuli is unclear.

In a series of experiments, Rogers & Bregman (1998) demonstrated that sudden changes in level or perceived location can also have a resetting effect. Build-up was promoted in a 4.8-s repeating HLH- induction sequence, and then measured in a subsequent test sequence (a 1.2-s repeating HLH- sequence). When there was no change in tonal properties between the two sequences, build-up occurring in the induction sequence resulted in listeners hearing the test sequence as more segregated. However, when the induction sequence was perceived at a different location to the test sequence, less segregation was perceived in the test sequence (perceived location was altered through a change in inter-aural time or level difference over headphones, or through a change in location in a loudspeaker array). Also, a similar resetting effect was observed if the induction sequence was 12 dB less intense than the test sequence (however, when the induction sequence was 12 dB louder than the test sequence, practically no resetting occurred). Roberts et al. (2008) have demonstrated resetting by abrupt changes in level, frequency, or lateralization in an on-going tone sequence when streaming was measured using a temporal discrimination task. Their findings are broadly consistent with the studies which measured resetting through a direct report of perception (Anstis & Saida, 1985; Rogers & Bregman, 1993, 1998). A discussion of temporal discrimination as a measure of stream segregation, and a more detailed evaluation of Roberts et al.'s (2008) study is provided in Chapter 3.

Rogers & Bregman (1998) offered two plausible explanations for the resetting effect of a sudden change. Firstly, build-up occurring in a tone sequence may be specific to tones of those distinct properties - therefore build-up may *fail to transfer* between tones with different characteristics. This explanation is compatible with Bregman's (1978, 1990) functional explanation of build-up. If our auditory system compiles evidence that two distinct acoustic events are occurring, this process should be fairly specific to the sequence which is currently heard. However, Rogers & Bregman (1998) showed that an increase in level had a resetting effect, but that a decrease in level did not. This suggests that the direction of the change can influence resetting (rather than the perceptual difference between the two sequences *per se*). From this evidence, Rogers & Bregman (1998) proposed that resetting may only occur when a new sound source is perceived, as an increase in level could well indicate a new event, but a decrease could not. In other words, a newly perceived event may *actively reset* build-up. Rogers & Bregman (1998) suggested that the perception of a new event could cause the auditory system to re-analyse the acoustic scene, resetting the evidence-accumulation process (Bregman, 1978, 1990). The active resetting effect of a sudden change could also be related to attentional factors, although this link has not been clearly established. The role of attention in streaming is discussed further in the next section.

Whilst these two explanations of resetting (failure to transfer vs. active resetting) may not necessarily be mutually exclusive, to date no study has clearly differentiated between either hypothesis. Rogers & Bregman (1998) partially addressed the issue by measuring the resetting effect of gradual changes. They found that when changes in level (loudness) or lateralization (perceived location) were applied gradually over the course of the induction sequence, only marginal resetting occurred. This contrasts with when the same changes were applied in one single transition, for which a much larger resetting effect was observed. This could be interpreted as evidence that only sudden change promotes resetting (i.e., active resetting). However, for the gradual-change conditions only the first induction tone was set at the maximum difference from the test sequence tones, and all subsequent induction tones became increasingly similar to those in the test sequence. In contrast to this, in the sudden-change conditions all induction tones were set at the maximum difference from the test sequence tones. Owing to this increased similarity, reduced resetting in the gradual-change conditions could simply reflect an increased transferral of build-up between induction and test tones (a point acknowledged by the authors).

1.4 The Role of Attention in Stream Segregation

In his influential book, Bregman (1990) proposed that stream segregation is affected by both bottom-up and top-down processes. He argued that perception of tone sequences that fall below the fission boundary or above the temporal coherence boundary is fully determined by the stimulus properties and is attention-independent. In contrast, the perception of stimuli in the ambiguous region between these boundaries is known to be affected by listening set (i.e., top-down, schema-based factors). These factors are assumed to be effortful, as they use limited attentional resources. It has proved difficult to measure streaming in unattended sequences, as behavioural research is dependent on listeners making responses to indicate their perception. To counter this problem, some behavioural research has directed attention away from a sound sequence and measured streaming indirectly.

1.4.1 Behavioural Studies

The presence of an irrelevant sound sequence can impair performance in difficult non-auditory tasks (this is typically studied using a visual memory task). This impairment is known as the irrelevant sound effect, or the ISE (see Jones, 1995, for a review). It is assumed that the distracting sounds are unattended, as attention is focused towards the memory task. Importantly, the ISE is only observed when the properties of the unattended sounds alternate; an unchanging repetition of the same sound does not impair performance (Tremblay & Jones, 1998). Jones et al. (1999) measured the ISE of a rapidly alternating tone sequence on performance in a letter-recall task. When the frequency separation between the tones was small, performance in the task was impaired. This indicates that the sounds were grouped together, and heard as an alternating-pitch percept. In contrast, at a larger Δf performance was unaffected by the presence of the tone sequence, suggesting that the low and high frequency tones were organized into separate streams of unchanging pitches. This finding suggests that the frequency separation between the unattended sounds influenced their grouping, in the same manner as in an attended sequence (cf. van Noorden, 1975). Furthermore, an effect consistent with this account has been shown when the presentation rate of the tone sequence is varied. For a fixed Δf, the ISE was observed at a slow presentation rate, but not at a faster rate for which stream segregation would be expected to occur (Macken et al., 2003). Whilst these studies suggest that attention is not required for stream segregation; it is difficult to rule out the possibility that participants were paying some attention to the background sounds.

Carlyon et al. (2001) examined the relationship between attention and sequential organization in an attention-switching task. Listeners were presented with a long (21 s) tone sequence to the left ear, and higher-frequency noise bursts to the right ear. The noise bursts began at the sequence onset, but were only presented during the first half of the tone sequence. Participants were required either to ignore the noise bursts and continuously indicate their perception of the tone sequence or to discriminate between the amplitude envelopes of the noise bursts, and then switch to making streaming judgments once the noise bursts ended. When listeners continuously rated the tone sequence, build-up occurred at a rate consistent with previous studies (e.g., Anstis & Saida, 1985). In contrast, when listeners switched from attending the noise bursts to making streaming judgments, rated segregation was greatly reduced compared with the corresponding time in the "ignore" condition. This finding suggests that build-up did not occur during the unattended portion of the tone sequence. Consistent results with Carlyon et al.'s (2001) study were also found when listeners switched attention from a visual or cognitive task to making streaming judgments (Carlyon et al., 2003). Whilst these results may indicate that attention is required for build-up to occur, an alternative explanation suggests that a switch in attention may have a resetting effect. In other words, build-up may have occurred in the unattended portion of the sequence, but the switch in attention may have caused its loss owing to substantial resetting. This explanation was supported in a study by Cusack et al. (2004), in which build up was found to be largely reset when participants briefly switched attention away from an on-going tone sequence. In a related experiment, Cusack et al. (2004) also partially replicated the conditions previously tested by Carlyon et al. (2001). Contrary to the initial finding, Cusack et al. (2004)

found evidence of build-up during the unattended portion of the tone sequence. When listeners switched from a noise-burst discrimination task to reporting streaming in an on-going tone sequence, the probability of listeners hearing segregation was significantly increased in comparison with their initial perception of segregation at the physical onset of the tone sequence (at least for tone frequency separations of either 8 or 10 ST). This finding contrasts with Carlyon et al.'s (2001) argument that build-up does not occur in an unattended tone sequence, although the reason for this discrepancy is unclear.

1.4.2 Electrophysiological Studies

Several electrophysiological studies have also examined the relationship between attention and stream segregation. These experiments have used the mismatch negativity (MMN) component of event-related brain potentials as a measure of streaming. The MMN is elicited when an element of a sound stimulus differs from an established pattern. For this reason, the MMN component is thought to reflect a deviance detection process which is based on the memory of the regularities occurring in an acoustic stimulus (Näätänen & Winkler, 1999). The MMN component typically occurs 150 ms after the onset of a deviant sound, and it is primarily observed in the auditory cortex. Infrequent changes in frequency, tone duration, spatial location or intensity can all elicit an MMN component. Other violations of a repeating tone pattern can also generate an MMN component (see Näätänen et al., 2001, for a review). This component has proved a useful tool in the study of unattended streaming because MMN elicitation does not require the participant to attend or respond overtly to the sounds. Therefore, the MMN component can be used to index the organization of a task-irrelevant sound sequence when attention is directed to a separate task.

The MMN component has been used to test whether streaming builds up in an unattended sound sequence. Sussman et al. (2007) presented a repeating HLL tone sequence at either a small (1 ST) or large (8 ST) frequency separation. To avoid any across-trial build-up effect, the inter-trial H-tone frequency was randomized within a 1397 - 3136 Hz range (the L tone frequency was set relative to this). The level of each L-tone was randomized (from 65 to 92 dB SPL), whilst the H-tones were fixed at 74 dB SPL. In two separate conditions, either the 4^{th} or 10^{th} H-tone of the sequence was 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 intensity-deviant H-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 changed H tone would be heard as a deviation from the otherwise regular intensity H-tone-only stream (and so should elicit the MMN component). In summary, the MMN component would only be expected when the sequence was organized into separate low- and high-pitched streams. To ensure listeners did not attend the tone sequence, the primary task was to detect small intensity changes applied to an on-going band-pass filtered white noise (100 - 1200 Hz, which was presented to a different speaker in a free-field environment). For the large frequency separation, the MMN component was not elicited when the intensity deviation was applied to the 4th H-tone, but it was when the deviation was applied to the 10th tone. No MMN component was observed for the smaller frequency separation. A second experiment demonstrated that a deviant 4th tone could elicit a MMN component when the H-tones were presented alone. These findings imply that streaming built up over the course of the unattended sequence. This evidence is inconsistent with Carlyon et al.'s (2001, 2003) suggestion that build-up only occurs in an attended sequence. It could be that build-up did occur in the unattended portion of Carlyon et al.'s stimuli, but that the switch in attention had a resetting effect (as proposed by Cusack et al., 2004, see section 1.4.4).

Generally, MMN research has indicated that attention is not required for sequential organization (Sussman et al., 1999; Ritter et al., 2000; Winkler et al. 2003a). However, many of these studies only controlled for attention by instructing subjects to either read a book or watch a silent film. For both of these controls, it is difficult to be confident that subjects did not at least intermittently attend the sound stimuli. A more robust control of attention was employed by Winkler et al. (2003b), who measured MMN elicitation when subjects either read a book, or performed a difficult n-back task, which required participants to remember the position of previously presented visual stimuli (either 1- or 3-back task, see Watter et al., 2001). The properties of the task-irrelevant sound sequence were similar to that used by Sussman et al. (2007), except that the deviant H-tone was varied in duration rather than intensity (the low-frequency tones were unpredictable in duration, and also in level). The deviant H tones only elicited an MMN component when there was a large frequency separation between the L- and H-tones. The MMN component was consistently observed when listeners read a book or performed in either of the n-back tasks. The finding that MMN elicitation did not vary with the attentional demands of the primary task supports the argument that streaming is attention-independent. This conclusion is also consistent with the studies which have indirectly measured streaming through the ISE (Jones et al., 1999; Macken et al., 2003).

So far, we have seen evidence that auditory grouping does occur even when attention is focused towards a visual or a cognitive task. A contrasting effect has been shown when listeners attended one sound stream, but ignored two other sets of sounds. Sussman et al.

(2005) presented listeners with a repeating LMH tone sequence (M = medium-frequency tone). Within this sequence, the L and M tone subsets were both presented as repeating, ascending frequency triplets, and these triplets were likely to be heard as segregated (the frequency variations within the L and M subsets were 311-349 Hz, and 880-998 Hz respectively). Occasionally, the order of the tones within a subset was reversed so that a descending frequency triplet was presented. The H tones were of a fixed frequency (2489 Hz), but were subject to irregular frequency deviations (2637 Hz). When the entire sequence was unattended (i.e., participants watched a silent film), changes to the L, M, or H subsets all elicited an MMN component. This indicated that all subsets formed separate streams. In contrast, when listeners were asked to report the H tone deviants, only those deviants elicited a MMN component - no MMN was observed after the L or M subset pattern deviations. The authors concluded that when the H subset of sounds was attended, the additional subsets were not organized into streams. A similar finding was reported in a behavioural experiment by Brochard et al. (1999). For this study, participants detected a temporal irregularity in a same-frequency tone sequence. Performance was impaired when this sequence was accompanied by a subset of different frequency tones. Importantly, increasing the number of task-irrelevant tone subsets had no further detriment on performance. These authors concluded that the task-irrelevant tone subsets were not organized into perceptual streams.

1.4.3 Conclusions

Although no study has proved conclusive, the evidence suggests that stream segregation does occur even when all auditory stimuli are unattended (Jones et al., 1999; Macken et al., 2003; Sussman et al., 1999; Ritter et al., 2000; Winkler et al. 2003a; Winkler et al. 2003b). There is also some evidence that when attention is paid to one subset of sounds, any additional, unattended sound subsets are not organized into perceptual streams (Sussman et al., 2005; Brochard et al., 1999). Carlyon et al. (2001, 2003) showed that when attention was switched from a separate auditory task to a streaming task, there was no evidence of build-up occurring in the unattended portion of the tone sequence. As Sussman et al. (2007) demonstrated that build-up may occur in a fully unattended tone sequence, it seems that a switch in attention towards a sequence may trigger a substantial resetting effect (a similar conclusion was reached by Cusack et al., 2004, see the next section).

1.4.4 The Hierarchical Decomposition Model

Cusack et al. (2004) proposed a model to explain the interactions between attention and stream segregation. Perceptual grouping has previously been considered as hierarchical (see Bregman, 1990; Darwin & Carlyon, 1995), but Cusack et al. (2004) also factored the role of attention into this interpretation. The model is best illustrated by a real world example - imagine simultaneously hearing a piece of music, a conversation and road traffic. Cusack et al. (2004) argued that initial (and perhaps automatic) grouping allows a listener to focus attention on a single event, for example the music. The focus of attention allows for the attended source to be "elaborated" further - so that source-specific perceptual streams become available for segregation (for example the melody played by the guitar). As previously, further elaboration becomes possible once this source is attended to (the high guitar notes), and so on.

This hierarchical decomposition model argues that only attended streams can be further elaborated. Furthermore, if attention is withdrawn from a source, these elaborated streams become no longer perceptually salient. This argument is in good accord with many of the observed effects of attentional switches on streaming (Carlyon et al., 2001, 2003; Cusack et al., 2004). However, it remains unclear as to how this model can be reconciled with the electrophysiological evidence that unattended tone sequences are organized onto perceptual streams (for example, Sussman et al., 2005, 2007).

1.5 Questions for Research

It is widely accepted that the tendency for stream segregation builds up over the course of a sequence (Bregman 1978; Anstis & Saida, 1985), but that a sudden change in sequence properties has a resetting effect (Anstis & Saida, 1985; Rogers & Bregman 1993, 1998; Roberts et al., 2008). Whilst resetting has been demonstrated in a variety of studies, the exact mechanisms responsible for this effect are currently unknown. Previous research has suggested that there are two potential explanations for resetting; 1) that build-up fails to transfer between tones sequences of different properties, or 2) that an abrupt change has an active resetting effect (Rogers & Bregman, 1998). Relating to this, there is also some evidence that a switch in attention may trigger resetting (Cusack et al., 2004). Several of the experiments presented in this thesis aim to differentiate between these alternative explanations of resetting, and to provide a greater understanding of this effect.

The build-up and resetting of stream segregation has been demonstrated in a variety of different tone arrangements. For example, the build-up of segregation has usually been studied in an alternating-frequency sequence (Bregman, 1978; Anstis & Saida, 1985), but a repetition of a same-frequency tone (e.g., a repeating L- arrangement) can also strongly promote segregation in a subsequent LHL- type sequence (Beauvois & Meddis, 1997; Roberts et al., 2008). Consistent findings were reported by Rogers &

Bregman (1993), who used a repeating H- inducer to promote streaming in a subsequent HLH- test sequence. Furthermore, abrupt changes to both these types of sequence have been shown to trigger resetting (Anstis & Saida, 1985; Rogers & Bregman, 1993; 1998; Roberts et al., 2008). Despite these apparent similarities, it is unclear exactly how alike the segregation-promoting effect of a same-frequency tone sequence is to the build-up induced by an alternating-frequency sequence. The principal aim of this thesis is to explore further the factors promoting the build-up (and resetting) of stream segregation arising from the presence of an induction sequence. This includes a comparison of the effects various types of inducer, and the use of subjective and objective measures of stream segregation.

Streaming has most often been studied for sequences of isolated pure tones. As briefly discussed earlier, normal listening environments often contain a mixture of both sequential and simultaneous grouping cues. In order to obtain an accurate perceptual representation of the auditory scene, these cues can interact in a variety of ways (Bregman, 1990). When listening to an on-going complex tone, an abrupt change in the properties of a single frequency component may cause it to stand out from the others. The changed component is heard as a pure-tone-like percept, distinct from the on-going complex tone. Previous studies have shown that increasing the amplitude or altering the perceived location of a component are both effective cues for segregation (Kubovy & Daniel, 1983; Bregman et al., 1994a; Kubovy et al., 1974). Although these tone-like percepts are highly salient, their sequential organization has not previously been investigated. A further aim of this thesis is to determine whether a series of these pure-tone-like percepts can be organized into distinct perceptual streams, and if so whether the effects of build-up and resetting can be demonstrated in such a sequence. These topics are discussed in more detail in Chapter 9.

1.6 Summary of Experiments

The first set of experiments was designed to differentiate between two alternative explanations of the resetting effect of an abrupt change on the build-up of stream segregation - either that build-up fails to transfer between tone sequences of different characteristics, or that a sudden change in itself has an active resetting effect (Rogers & Bregman, 1998). Experiment 1 measured build-up using a direct report of perception, similar to that used by Rogers & Bregman (1993, 1998). The experiment promoted build-up using a same-frequency induction sequence (10 L- cycles), and measured stream segregation in a briefer alternating frequency test sequence (3 LHL- repetitions). Altering only the final induction tone (in frequency, level, duration or replacement with silence) to create a single deviant tone reduced reported segregation in the test sequence. This effect was substantial, and broadly comparable for all of the different changes tested. As the nine prior induction tones were unaltered, the observed reduction in segregation is taken as evidence that the single, abrupt change actively reset build-up (an alternative explanation of this effect was ruled out in Experiment 4a). Using a comparable sequence structure, Experiment 2 aimed to demonstrate active resetting in a temporal-discrimination task (similar to that used by Roberts et al., 2008). When the final induction tone was altered in level or replaced with silence, a resetting effect was observed (although this effect was only very slight for a deviant tone based on level). Both experiments demonstrated a consistent resetting effect, but resetting was more evident in the subjective task of Experiment 1.

Using the same subjective procedure as Experiment 1, Experiments 3a - 3b demonstrated that the extent of resetting increased only slightly with an increase in the size of a frequency change applied to the final inducer (so long as the change was noticeable). As very small frequency changes were effective at triggering discernible resetting, it was concluded that resetting is not directly influenced by the likelihood of a new source being heard, but rather that any noticeable change can have an active resetting effect. By varying the serial position of a deviant tone, Experiment 4a demonstrated that resetting only occurred when the final inducer was replaced with silence (there was no evidence of resetting when same change was applied to either the 4^{th} or the 7^{th} induction tone). This indicates that the resetting effect of a single deviant tone cannot be explained simply by fewer inducers contributing to the cumulative build-up during the induction sequence. This finding was explained in Experiment 4b, which showed that build-up occurs rapidly during a same-frequency induction sequence. Based on this evidence, it was proposed that a same-frequency induction sequence may promote stream segregation by capturing one subset of test-sequence tones into an on-going, pre-established stream (see also Rogers & Bregman, 1993).

Experiment 5 attempted to study the relationship between sequence predictability and resetting by randomizing the frequencies of the first nine induction tones (within a ±2 ST range, last inducer fixed at -2 ST). Reported segregation following the constrained-random induction was reduced by a similar extent to that for when only the final inducer was lowered in frequency (-2 ST). A near-full resetting effect was observed following an induction sequence in which all tones were set to -2 ST from the standard 1 kHz L-tone. These findings were taken as an indication that mean stimulation frequency may be a key determinate of the segregation-promoting effect of an induction sequence. Experiments 6a-6b studied the effect of perceived continuity on the segregation-promoting effect of an induction sequence (perceived continuity refers to the perception of a sound as continuing through a louder sound, despite the original sound being physically absent;

see Chapter 7). It was found that actual and perceived continuity both have a similar effect on test sequence streaming judgements - both could promote either integration or segregation, depending on listening context.

Using a tracking procedure in which listeners continuously reported their perception of a 20 s test sequence, Experiment 7 compared the segregation-promoting effect of a same-frequency induction sequence with the build-up that occurs during an alternating frequency sequence. It was found that same-frequency inducers promoted considerably more segregation than did the alternating-frequency case. Indeed, at test-sequence frequency separations of 6 and 9 ST, most segregation was heard immediately after the same-frequency induction sequence, and the extent of stream segregation declined during the first seconds of the test sequence. Furthermore, deviant-tone resetting was only clearly apparent when a change (in tone duration) was applied to the final tone of the same-frequency induction sequence. There was no clear evidence of resetting when the same change was applied to a single tone of an on-going, alternating-frequency sequence. The results suggest that different mechanisms may be responsible for the increase in stream segregation following both types of sequence.

The final experiments (8 - 10) investigated the sequential grouping of pure-tone-like percepts, created by sudden changes in amplitude (Kubovy & Daniel, 1983; Bregman et al., 1994a), or interaural time difference (Kubovy et al., 1974) for individual components of a complex tone. By replicating the results of a temporal-discrimination measure of the streaming of pure-tone sequences (Bregman & Rudnicky, 1975), experiments 8 and 9 demonstrated that these tone percepts can be organized into distinct perceptual streams. Using the same method, Experiment 10 demonstrated that active resetting could be demonstrated when a deviant tone was inserted into a sequence of these tone percepts (evoked using amplitude increments).

Chapter 2

General Method for Experiments using 2AFC Subjective Measures

All of the experiments presented in this thesis used behavioural measures to examine the dynamics of stream segregation. Several of these experiments used the same general method (experiments 1, and 3a - 6b), and the tone sequences presented in these experiments shared the same common properties. Therefore, for brevity, this section will describe all of these common elements. In addition, the same two control conditions were used in all these experiments and these are also described in this section. The various specific conditions that distinguish the different experiments are described in their appropriate chapters.

2.1 General Structure of the Stimuli

The structure of the stimuli used in experiments 1, and 3a - 6b, was similar to that used by Rogers & Bregman (1993). On each trial, participants reported their perception of a test sequence, which for the majority of trials was preceded by a longer induction sequence. The induction and test sequences combined to form a single, continuous sequence (i.e., without any intermittence). Also, the border between the two sequences was not indicated to the participants. The properties of the induction sequence were experimentally varied, and the effect on a constant set of test sequences were measured (each of the induction sequences was tested across a uniform range of frequency separations for the test-sequence tones). Therefore, any difference in the perception of the set of test sequences can be directly attributed to the effect of the preceding induction sequence. Given that perceptual grouping was only measured in the test sequence, it was possible to test a variety of induction sequence arrangements. For example, one could measure the effect of a same-frequency induction sequence (i.e., repetitions of an L- stimulus) on streaming in the subsequent, alternating-frequency test sequence (i.e., LHL- triplets).

2.2 Experimental Design

2.2.1 Properties of the Test Sequence

The test sequence comprised three LHL- triplets (L = low-frequency tone, H = high-frequency tone, and - = silent interval). This type of arrangement was first used by van Noorden (1975), and is widely used to study stream segregation. For such a sequence, there is a clear difference in rhythm depending on the perceptual grouping of the tones. When all tones form a single, integrated perceptual stream, a clear gallop-like rhythm is heard. In contrast, when the low and high frequency tones segregate, two separate streams are heard, each with a monotonous rhythm. For the current experiments, participants were required to report their perception of the final LHL- triplet of the test sequence as either integrated or segregated, in a two-alternative forced-choice (2AFC) procedure.

Although the test sequence was used to measure the effect of various induction sequences, streaming will also be influenced by the tonal properties within the test sequence itself. For example, the frequency separation between tones is known strongly to influence grouping (van Noorden, 1975). Above a certain frequency separation, stream segregation is always perceived (the temporal coherence boundary, or TCB), and similarly, below a certain frequency separation, integration is always heard (the fission boundary). In between these two boundaries there is an ambiguous frequency region, in which either percept may be perceived. In order to compare effectively the effects of different induction sequences on streaming, the properties of the test sequence must primarily fall within this ambiguous region. To ensure this, the test sequence tones were presented at a variety of frequency separations. Pilot studies confirmed that the frequency range tested fell predominantly within the ambiguous region. The extreme frequency separations were typically heard as strongly integrated or segregated; this ensured that participants were regularly presented with clear examples of these percepts. This range of frequency separations in the test sequence also provided a measure of whether participants were conforming to well-established response patterns (i.e., reporting more segregation with increasing frequency separation).

The frequency separation between tones in the test sequence was measured in semitones (ST), a frequency ratio of approximately 1:1.06. Low frequency tones were fixed at 1 kHz, whilst the frequency of the high tones was varied. The smallest frequency separation tested was 4 ST (high tone = 1260 Hz), which increased in two-semitone steps (6 ST / 1414 Hz, 8 ST / 1587 Hz, 10 ST / 1782 Hz, 12 ST / 2000 Hz) to the maximum separation of 14 ST (2245 Hz). Tone duration was set to 100 ms, including 10 ms raised cosine ramps at the tone onset and offset. Within each LHL- triplet, there was no inter-tone silent interval. The silence at the end of each triplet lasted 100 ms (i.e., equivalent to the duration of one tone). Therefore, the overall duration of the test sequence was 1.2 s (3 LHL- triplets, with each triplet lasting 0.4 s). All tones were presented diotically and at 70 dB SPL.

2.2.2 Properties of the Induction Sequences

Two control induction conditions were presented in each experiment. The standard induction sequence comprised 10 L- repetitions. The tones shared identical properties to the L tones in the test sequence, and the inter-tone silent interval was set to 100 ms (i.e., the total duration of the standard induction sequence was 2 s). Therefore, this induction sequence matched the L tones of the test sequence in both tempo and frequency. A schematic illustration of the standard induction sequence paired with a test sequence is shown in Figure 2.1. A similar same-frequency induction sequence was used by Rogers & Bregman (1993) - the only difference was that these authors used H-only induction tones, and measured perception in an HLH- test sequence. They found that this type of induction arrangement promoted a strong perception of segregation in the subsequent test sequence.

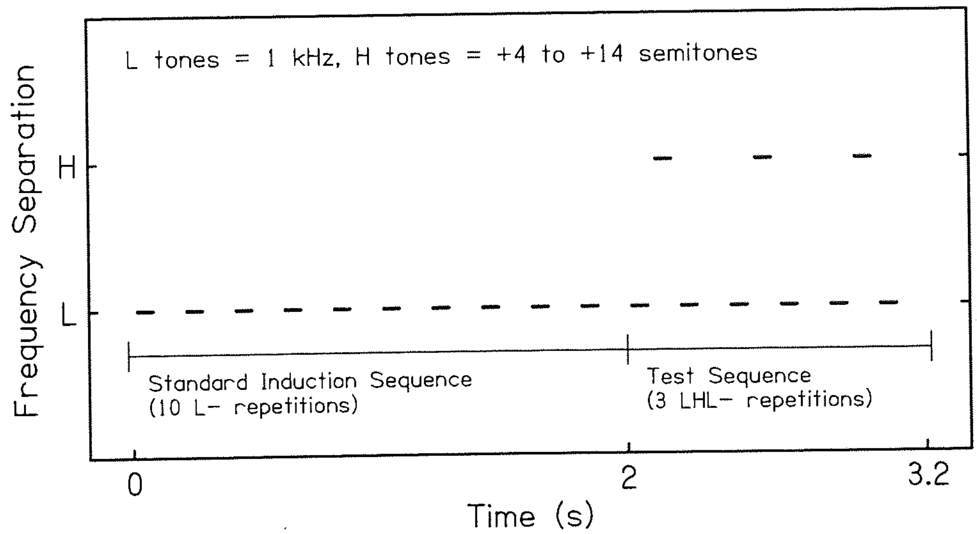

Figure 2.1: **Illustration of a standard induction sequence paired with a test sequence.**

For the second control condition, the test sequence was not preceded by any induction sequence. This was referred to as the no-induction condition. This condition was included to measure streaming within the test sequence alone. Rogers & Bregman (1993, 1998) used a similar condition, in which the test sequence was preceded by a continuous burst of broadband white noise. As continuous white noise is known not to promote streaming (Bregman, 1978), Rogers & Bregman's (1993, 1998) condition is somewhat comparable to the current no-induction arrangement. Whilst neither induction arrangement would be expected to promote test-sequence streaming, the current no-induction condition resulted in the immediate onset of the test sequence at the beginning of the trial. This contrasts with the standard induction condition and the various experimental conditions in which the test-sequence onset occurred typically 2 seconds after the trial onset. Due to this, for the no-induction condition trials the onset of the test sequence would occur earlier than anticipated - especially as the presentation order of trials was randomized within blocks (see the general procedure). This issue is further considered in the context of the results from Experiment 1 (Section 3.2.3).

Despite this issue, when taken together, the two control conditions should demonstrate the effect of a strong segregation-promoting induction sequence, and perception of the test sequence in the absence of any prior build-up, respectively. The experimental induction sequences were all variations of the standard induction sequence (see the summary of experiments in Chapter 1 for more details). Generally, most of the experimental conditions were expected to reduce stream segregation relative to the standard induction case (i.e., by resetting the build-up of segregation). Therefore the extent of any resetting effect could be directly measured by the extent of return to the results for the no-induction condition.

2.3 General Procedure

On each trial, a single combination of an induction sequence and a test sequence was presented. Listeners were required to report their perception of the final LHL- triplet of the test sequence. Participants were asked to avoid trying to listen specifically for either integration or segregation, but rather simply to report which of the two percepts was most dominant (i.e., neutral listening instructions). For an ambiguous tone sequence, van Noorden (1975) noted that it was relatively easy to hear segregation when instructed, but that it was more difficult to hear integration on request. Based on this observation, Rogers & Bregman (1993, 1998) asked subjects to attempt to hear integration when rating their perception of a test sequence. Whilst listening set can strongly affect grouping, pilot studies for this thesis confirmed that listeners made appropriate responses to changes in frequency separation without being primed to listen for a specific percept. Participants pressed either "1" or "2" on a computer keyboard to indicate a perception of either integration or segregation, and then confirmed their selection by pressing "↵" (to reduce the chance of an erroneous response being made). After this confirmation, there was a three second pause before the next trial began (to allow any build-up of streaming to decay before the onset of the next stimuli). This was shorter than the 5 -

6 s silent intervals used by Rogers & Bregman (1993, 1998). However, their experiments used considerably longer induction sequences (4.8 s), which would be likely to promote a stronger tendency to hear stream segregation. Furthermore, there is evidence to indicate that much of any build-up decays in the first second or two of silence (Bregman, 1978; Cusack et al., 2004). Therefore, a shorter silent interval was chosen to reduce the overall duration of the experiment.

For each experiment, the trials were organized into blocks. A trial block comprised a combination of each induction sequence with each of the six frequency separations for the test sequence. As the number of induction sequences tested varied across experiments, so too did the number of trials in each trial block. Typically, experiments tested either five or six induction conditions, which resulted in either 30 or 36 trials per trial block (i.e., 5 or 6 induction conditions $\times$ frequency separations in the test sequence). The order of trial presentation within a block was fully randomized, and the randomization was performed anew for each repetition of a trial block.

After a verbal explanation of streaming, participants were presented with an example of clearly integrated and segregated LHL- test sequences. Once familiar with these stimuli, participants then completed a training session which comprised two trial blocks. This training was given to familiarize participants with responding to the stimuli - there were no criteria for acceptance into the main experiment[1]. The main experiment comprised 20 trial blocks (i.e., 20 repetitions of each possible stimulus). As some experiments contained different numbers of induction conditions, the overall duration of each experiment varied. Typically, a participant took between 2 and $2\frac{1}{2}$ hours to complete an experiment, which was usually split into two separate sessions. Eight participants contributed to the final data set for each experiment.

2.4 Apparatus

All stimuli were synthesized with 16-bit resolution using MITSYN (Henke, 1997) and stored on disk. The stimuli were played back via a Turtle Beach Santa Cruz sound card at 20 kHz sampling rate. The stimuli were presented over Sennheiser HD480-13-II earphones; the overall output level of the sound card was set using the on-board analog attenuator (for coarse adjustment) and digital multiplication (for fine adjustment). The

[1] However, data from two participants in Experiment 1 and one participant in Experiment 5 were rejected. After analysis of the full data set, neither of these participants displayed the typical response pattern of reporting increased segregation with increasing frequency separation in the test sequence. As this was only observed in two listeners, no formal criteria for acceptance were set.

setup was calibrated using a sound-level meter (Brüel & Kjaer, type 2209) coupled to the earphones by an artificial ear (type 4153).

The experiments were run using a specially created program which was made using the Visual Basic programming language (Visual Studio, 2003, version 6.0), incorporating Direct X commands. Participants completed the experiment in a quiet listening environment - either in a double-walled sound-attenuating chamber (Industrial Acoustics 1201A), or in a single-walled chamber (Industrial Acoustics 401A) housed within a quiet room.

Chapter 3

Subjective and Objective Measures of the Resetting Effect of a Single Deviant Tone

3.1 Introduction

The two experiments presented in this chapter examined whether a single deviant tone in a sequence can reset the build-up of stream segregation. Previous studies have demonstrated that resetting can occur when the acoustic properties of an on-going sequence are suddenly altered (Anstis & Saida, 1985; Rogers & Bregman, 1993, 1998; Roberts et al., 2008). Each of these studies used slightly different sequence arrangements and measures of streaming, but all of them promoted build-up using some form of induction sequence and then measured streaming in a subsequent, briefer test sequence. Resetting has been observed when the test sequence differs from the induction sequence in frequency or ear of presentation (Anstis & Saida, 1985), and in perceived location or level (Rogers & Bregman, 1998). Similar resetting effects have also been demonstrated in a temporal discrimination task (Roberts et al., 2008 - see Experiment 2 in this chapter).

The reason for this resetting effect is not currently known. Build-up is thought to reflect a gradual accumulation of evidence in favour of two separate acoustic sources being present (Bregman, 1978, 1990). If this evidence-accumulation process is specific to the sequence being heard, any increased tendency to hear segregation should not transfer to a subsequent sequence comprising tones with different characteristics. In other words, build-up may fail to transfer between tone sequences with different properties. Whilst this explanation seems plausible, Rogers & Bregman (1998) observed that a sudden increase in tone level (+12 dB) during a sequence had a substantial resetting effect, but that a similar decrease in level (-12 dB) did not. This finding indicates that reduced segregation cannot solely be explained by a difference in sequence characteristics, but that the properties of the transition itself must also influence resetting. If we consider a natural acoustic environment, certain sudden changes are likely to signify the onset of a new acoustic event (such as a sudden increase in loudness). Any indication of a new event may trigger a re-analysis of the auditory scene, which could result in a resetting of the evidence-accumulation process. This hypothesis is capable of explaining why resetting is influenced by the direction of a change, as a decrease in loudness is unlikely to indicate a new acoustic event (Rogers & Bregman, 1998).

Relating to this theory, Rogers & Bregman (1998) compared resetting following either an abrupt or a gradual change (i.e. the same change was applied over the course of several tones, rather than in a single transition). They found that only sudden changes resulted in substantial resetting. This finding could indicate that resetting only occurs when a new acoustic event was anticipated (as a gradual change is more likely to be perceived as a change in an on-going sound source). However, a direct comparison of the sudden and gradual change conditions was confounded as the mean properties of the two induction sequences differed. If an otherwise identical change is applied gradually rather than in a single transition, the mean tonal properties of the two sequences will inevitably differ. In Rogers & Bregman's (1998) study, when compared with the sudden change case, the mean tonal properties of the gradual change conditions were more similar to those of the test sequence. Owing to this unavoidable problem, increased segregation in the gradual change conditions may have been due to an increased transferral of build-up (a point acknowledged by the authors). This issue is particularly difficult to address, as to do so would require knowledge of the relative contribution of each induction tone to the overall segregation-promoting effect of the induction sequence. As pointed out by Roberts et al. (2008), there is no such model to weight the contribution of each individual inducer, and so any attempt to do so would be mainly speculative. For this reason, evidence from gradual change conditions cannot differentiate between the two explanations of resetting (failure to transfer vs. active resetting).

More recent research has indicated that attentional factors may influence resetting. Evidence from mismatch negativity (MMN) studies suggests that unattended tone sequences are organized into perceptual streams (Sussman et al., 1999; Ritter et al., 2000; Winkler et al. 2003a, 2003b). Furthermore, Sussman et al. (2007) provided evidence that the build-up of stream segregation may occur in an unattended tone sequence. In contrast, Carlyon et al. (2001, 2003) found that when attention was switched from a separate task towards an on-going tone sequence, there was no evidence that build-up occurred in the unattended portion of the sequence. Similarly, Cusack et al. (2004) demonstrated that streaming was greatly reduced after a brief switch in attention away from an on-going tone sequence. These results seem to suggest that build-up may occur in an unattended sequence, but that a switch in attention to the sequence has a substantial resetting effect on this build-up.

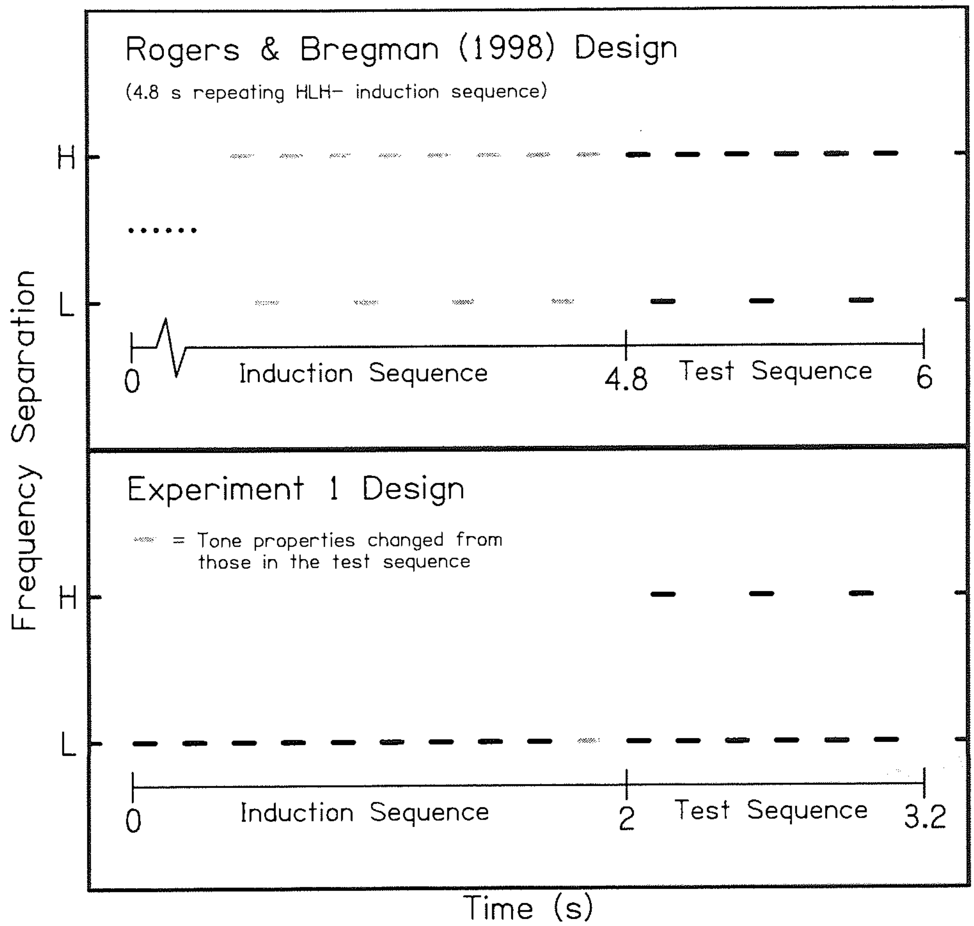

Figure 3.1: **Comparison of Experiment 1 with Rogers & Bregman's (1998) experimental design. Both experiments measured perception in a 3 × LHL- / HLH- test sequence. Rogers & Bregman (1998) promoted build-up using a 4.8 s HLH- induction sequence, whereas Experiment 1 used a 2 s L-only induction sequence. Rogers & Bregman (1998) demonstrated resetting by altering the properties of the entire induction sequence from those of the test sequence. Experiment 1 aimed to demonstrate resetting when only a single induction tone was changed.**

For the current experiments, a same-frequency induction sequence was used to promote stream segregation in a test sequence (see Chapter 2). In the critical condition, resetting was measured when the properties of the last induction tone were altered, but all prior induction tones were identical to the L tones of the test sequence. As the majority of the induction tones matched the test sequence tones, any transferral of build-up should be relatively unaffected by the presence of a single changed tone. Therefore, any reduction in streaming following a deviant tone can be attributed with confidence to the change having an active resetting effect. Figure 3.1 shows an illustration of how the design of Experiment 1 contrasts with Rogers & Bregman's (1998) earlier study of resetting

3.2 Experiment 1

3.2.1 Method

Participants

Eight listeners took part in Experiment 1, including the author. Note that the author participated in experiments 1 and 2, but not in any of the subsequent experiments. Participants were drawn from a university student population, and all reported normal hearing. Two of them were replaced because their results did not show clearly the well established dependence on LH frequency separation for their streaming judgments (see, e.g., van Noorden, 1975).

Stimuli and Conditions

Each trial comprised an induction sequence and a subsequent test sequence (the test sequence comprised three LHL- triplets). As described in the previous chapter, a standard induction condition was designed to promote segregation in the test sequence. This standard induction sequence comprised 10 L- repetitions, where the L tones were matched to those of the test sequence in terms of frequency and tempo (L tones = 1 kHz, 70 dB SPL, tone duration and inter-tone silences = 100 ms). To test resetting, a range of alterations were applied to the last tone of the standard induction sequence (all prior tones were unaltered, see Figure 3.1). This deviant tone was changed either in frequency, level, or duration. With the exception of the specific alteration, all other deviant tone properties were identical to those of a standard L tone. In a final condition, the last induction tone was replaced with an equivalent-duration silent interval. The specific details of the deviant-tone conditions were:

1. *Deviant lowered in frequency - or - Deviant(-3 ST)*

   Anstis & Saida (1985) demonstrated that a sudden shift in the centre frequency of an LH sequence could substantially reset build-up. Resetting was only observed following frequency changes greater than -1 to +3 ST from the centre frequency of the tone sequence (although it is offset from 0 ST, this range is broadly similar to the bandwidth of the auditory filter in this frequency region; Glasberg & Moore, 1990). For the current condition, the last induction tone was lowered in frequency (from 1 kHz) by 3 ST (i.e. the tone was presented at 841 Hz).

2. *Deviant increased in level - or - Deviant(+12 dB)*

   Rogers and Bregman (1998) found that a sudden increase in level between the induction sequence and the test sequence (+12 dB) substantially reset build-up.

For the current condition, the same change was applied only to the deviant tone (standard induction tone = 70 dB SPL, deviant induction tone = 82 dB SPL).

3. *Deviant extended in duration - or - Deviant(+50 ms)*

   For this condition, the duration of the last induction tone was increased from 100 ms to 150 ms. The following inter-tone silence was reduced to 50 ms, to preserve the rhythm of the tone onsets. It has previously been demonstrated that an H-tone-only induction sequence is most effective at promoting segregation when it roughly matches the tone density of the corresponding high-frequency test-sequence tones (Rogers & Bregman, 1993). Using a temporal-discrimination task, Roberts et al. (2008) demonstrated that an induction sequence comprising tones of a longer duration than those of the test sequence was largely ineffective at promoting stream segregation (where induction tones = 150 ms, test-sequence tones = 50 ms). The current condition examined whether a single change in tone duration would have an active resetting effect.

4. *Deviant replaced with silence - or - Deviant(silent)*

   For the final deviant condition, the last induction tone was simply replaced by an equivalent-duration silent interval (i.e. 100 ms). The onset times of the neighbouring tones were unchanged. This resulted in a 300-ms silent interval between the 9th induction tone and the 1st test-sequence tone. Some research has suggested that it takes several seconds for the majority of build-up to decay in a silent interval (Bregman, 1978; Beauvois & Meddis, 1997). In contrast to these findings, Cusack et al. (2004) found that a range of silent intervals (between 1 - 10 s) all substantially reset build-up. The reason for this discrepancy is unclear. Notwithstanding this issue, the silent interval presented in the current condition was briefer than any previously tested, and so any evidence of resetting would extend Cusack et al.'s (2004) previous observation.

Procedure

For each trial, listeners rated the last LHL- triplet of the test sequence as either integrated or segregated (in a 2AFC procedure). Trials were organized into blocks. A trial block comprised a combination of each induction condition with every tone frequency separation for the test sequence (6 × 6). This resulted in 36 trials per block, and in each block the trials were presented in a new randomized order. Each participant completed 20 trial blocks (or 720 trials). Chapter 2 provides a more detailed description of the experimental procedure.

3.2.2 Results

For each listener, responses to each condition were averaged separately, giving a proportional measure of reported stream segregation. Responses across all participants were then averaged to give an overall indication of streaming for each separate condition (i.e. the percentage heard as segregated across listeners). These mean values are shown in Figure 3.2. The mean percentage of trials heard as segregated for frequency separations of 4, 6, 8, 10, 12, and 14 ST were 3.8%, 20.8%, 41.4%, 67.3%, 77.6%, and 92.1%, respectively. The mean percentage of trials heard as segregated following the standard, no-, deviant(-3 ST), (+12 dB), (+50 ms), and (silent) induction conditions were 60.4%, 39.0%, 52.0%, 53.1%, 51.7%, and 46.8%, respectively. Data were analysed using a two-way, repeated-measures analysis of variance (ANOVA). The two independent variables were the six frequency separations for the test sequence (0, 4, 6, 8, 10, 12, or 14 ST), and the six induction conditions (standard induction, no induction, or 4 × deviant inductions). The partial eta-squared value (η_p^2) was used as a measure of effect size. The ANOVA confirmed significant main effects of frequency separation [$F(3, 35) = 44.141$, $p < 0.001$, $\eta_p^2 = 0.863$], and of induction condition [$F(5, 35) = 6.616$, $p < 0.001$, $\eta_p^2 = 0.486$]. The interaction between these two variables was also significant [$F(25, 175) = 2.485$, $p < 0.001$, $\eta_p^2 = 0.262$]. This interaction was probably driven primarily by the floor and ceiling effects at the two extreme frequency separations.

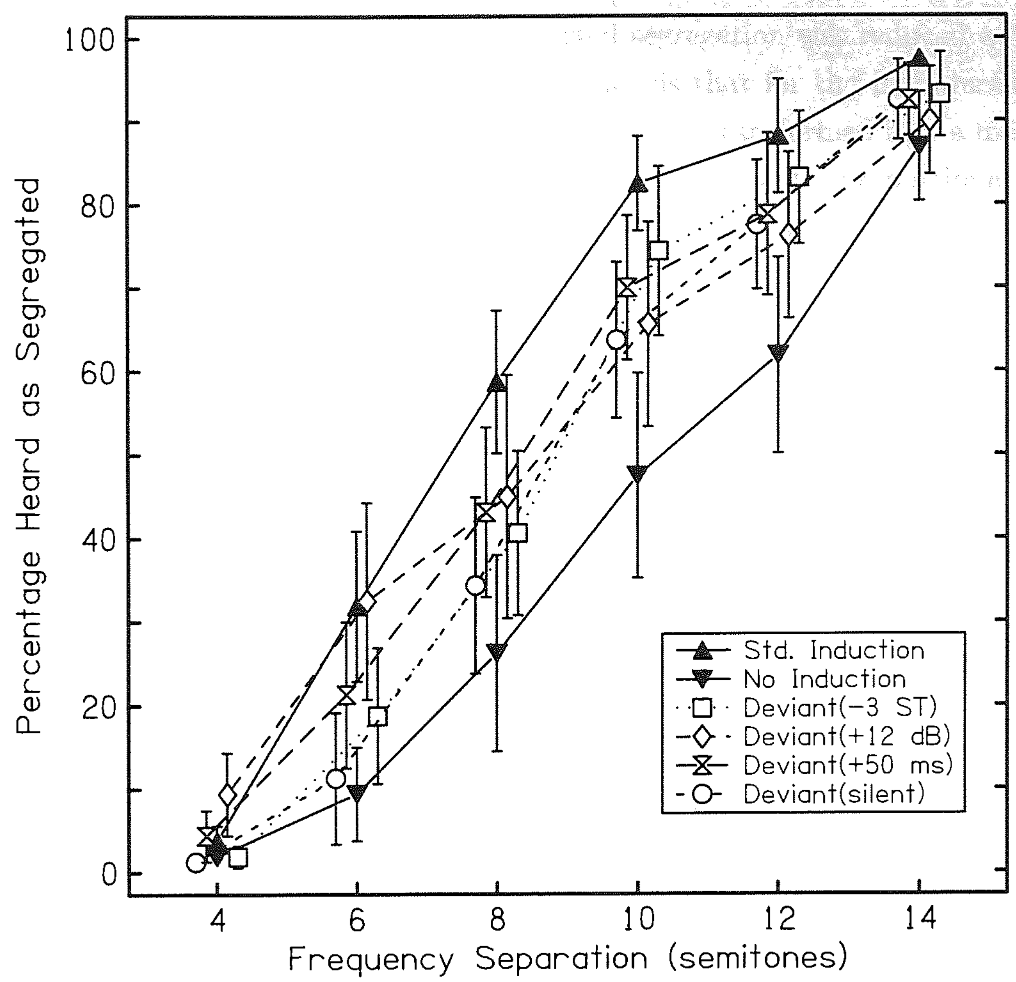

Figure 3.2: **Results from Experiment 1. The effects of induction condition, and frequency separation in the test sequence on reported segregation. Each point represents the percentage of trials reported as segregated (averaged from 8 listeners). The insert identifies the different induction conditions tested (deviant refers to the change applied to the last tone in a standard induction sequence). Error bars represent ±1 inter-subject standard error. As for all subsequent graphs of experimental results, the data are slightly offset (along the abscissa) to allow the reader to distinguish clearly between separate data points and error bars.**

The standard induction condition promoted the greatest amount of stream segregation, and the least was heard in the no-induction condition. However, a careful inspection of Figure 3.2 reveals that the functions relating to reported segregation and frequency separation differed slightly between these two control conditions. For the standard induction condition there was a negative acceleration of reported segregation with increasing frequency separation, whereas a positive acceleration was observed for the no-induction condition. This trend is relatively slight in Experiment 1, but it is also apparent in the subsequent experiments which tested the same conditions using the same method. Therefore, this observation will be further considered in the discussion of this current experiment.

For all of the deviant-tone conditions, reported segregation was reduced away from that for the standard induction condition and towards that for the no-induction case. This shift is represented in Table 3.1, where the data are transformed into a measure of the extent of resetting. The figures quoted in Table 3.1 represent a proportional shift on a scale defined by the standard induction condition representing no change (0.00), and the no-induction condition representing a maximum change (1.00). When subsequently discussed, this measure is referred to as the *extent of resetting.*

For any given test-sequence frequency separation, it is important to note that the relative extent of resetting values will be influenced by the magnitude of the difference between the standard and the no-induction conditions. For a small difference between these control conditions, any two experimental conditions which differ only slightly in absolute terms will promote largely different extent of resetting values. This was compensated for in two ways. Firstly, as clear floor and ceiling effects were observed at the two extreme frequency separations (4 and 14 ST), these data were excluded from the extent of resetting analysis. Secondly, a weighting system was applied when averaging the extent of resetting across the remaining frequency separations. The weight applied to any given frequency separation is calculated as the difference between reported segregation[1] in the standard and the no-induction conditions as a percentage of the *total* difference between these two conditions when reported segregation is summed across *all* applicable frequency separations (i.e., 6, 8, 10, and 12 ST). Therefore, the weight applied to an individual frequency separation can be expressed as:

$$weight_x = 100\left(\frac{\bar{std}_x - \bar{no}_x}{\left(\sum_{6\leq x\leq 12} \bar{std}_{(x)}\right) - \left(\sum_{6\leq x\leq 12} \bar{no}_{(x)}\right)}\right)$$

Where: x = test sequence frequency separation (in semitones), $\bar{std}$ = mean reported segregation in the standard induction condition, and $\bar{no}$ = mean reported segregation in the no-induction condition.

Therefore, the weight applied to the 6 ST test-sequence frequency separation for the current results was calculated as:

$$weight_{(6\,ST)} = 100\left(\frac{31.88 - 9.38}{261.3 - 145}\right) = 100\left(\frac{23.50}{116.30}\right) = 19.35\%$$

[1] Reported segregation refers to the percentage heard as segregated (see the first paragraph of the current results section, and Figure 3.2).

Frequency Separation	Deviant Condition: -3 ST	+12 dB	+50 ms	Silent	*Mean*
6 ST (19.35%)	0.58	-0.03	0.47	0.92	***0.49***
8 ST (27.96%)	0.56	0.42	0.48	0.75	***0.55***
10 ST (30.11%)	0.23	0.48	0.36	0.54	***0.40***
12 ST (22.58%)	0.19	0.45	0.36	0.41	***0.35***
Weighted Mean	***0.38***	***0.36***	***0.41***	***0.64***	***0.45***

Extent of resetting = standard induction (***0.00***) → no induction (***1.00***)

Table 3.1: **Results from Experiment 1. Reported segregation in the deviant-tone conditions was compared with that in the standard induction case. Values are expressed as a proportional shift away from the standard induction condition towards the no-induction condition. This serves as a metric for the *extent of resetting*. In order to provide a stable measure of resetting, the 4 and 14 ST frequency separation conditions were excluded from this analysis, and the across-frequency separation means were weighted (see the main text). Note that *only* the across-frequency means are weighted; each individual value represents the extent of resetting prior to the application of any weight.**

Two-tailed pairwise comparisons were performed on the original (i.e., untransformed) data set using the restricted least-significant-difference (LSD) test (Snedecor & Cochran, 1967). As floor and ceiling effects were observed for the 4 and 14 ST frequency separation conditions, these data were excluded from the pairwise analyses. From this reduced data set, each pairwise comparison is reported with a percentage measure of the difference in reported segregation between the two conditions. There was a significant difference in reported segregation between the standard and the no-induction conditions [difference = 21.4%, $t(7) = 6.019$, $p < 0.001$]. Compared with the standard induction condition, stream segregation was significantly lower in three of the deviant conditions [-3 ST: 8.4%, $t(7) = 2.756$, $p < 0.05$], [+50 ms: 8.8%, $t(7) = 3.755$, $p < 0.01$], and [silent: 13.7%, $t(7) = 6.972$, $p < 0.001$]). There was no significant difference between the standard induction condition and the deviant(+12 dB) condition [7.3%, $t(7) = 1.510$, $p > 0.05$]]. When compared with each other, there were no significant differences between any of the deviant conditions ($p > 0.05$ in all cases).

In summary, the results show that reported segregation was highest following the standard induction sequence. Reported segregation was significantly reduced when the last induction tone was changed in either frequency or duration, or when it was replaced with an equivalent-duration silent interval (a non-significant reduction in segregation was observed when the deviant tone was increased in level). The size of the reduction observed across the set of deviant conditions (i.e., the extent of resetting) was a loss of build-up evoked by the induction sequence in the range one-to-two thirds.

3.2.3 Discussion

For Experiment 1, reported stream segregation rose as the frequency separation between the tones in the test sequence was increased. This finding is in good accord with the previous literature (e.g. Miller & Heise, 1950; van Noorden, 1975). For all of the induction conditions, the standard induction sequence promoted the greatest amount of reported segregation, whilst the least segregation was observed in the no-induction condition. This is also largely consistent with earlier studies (Rogers & Bregman, 1993), and is an indication that the tendency to report segregation built up over the course of the standard induction sequence. As noted in the results section, the functions relating to reported segregation and test-sequence frequency separation differed following the standard and the no-induction conditions. For the standard induction condition, there was a negative acceleration of reported segregation as a function of increasing test-sequence frequency separation, whereas a positive acceleration was observed for the no-induction condition (a similar pattern of results was also present in the subsequent subjective experiments which tested the same conditions).

This trend was not observed by Rogers & Bregman (1993), who tested similar control conditions. A plausible explanation for this difference may be that the 'no-induction-of-segregation' condition employed by Rogers & Bregman (1993) was a continuous burst of white noise which was equivalent in duration to the other induction conditions (white noise is known not to promote stream segregation; Bregman, 1978). For the current no-induction condition, the test sequence was presented in the complete absence of any prior induction sequence, and so the test sequence began immediately at the start of the trial (see Chapter 2). As the order of trials was randomized within blocks, listeners would not have been able to anticipate the earlier onset of the test sequence in the no-induction conditions. This unexpected change in the timing of the test-sequence onset may have influenced participants responses in some way. Despite this issue, there remains a clear difference in reported segregation following the two control conditions, and this pattern of responses is largely consistent with previous studies. Also, the mean values obtained from the extent of resetting measure were weighted to compensate for the different functions of the control conditions (see the results section for a more detailed description of this analysis). Therefore, the currently employed control conditions were considered suitable for use in the subsequent experiments..

When compared with the standard induction condition, reported segregation was reduced for all the deviant-tone conditions (although this decrease was not significant for the deviant(+12 dB) case). Overall, reduced segregation in these conditions is evidence that a sudden change in tonal properties resets the build-up of stream segregation. Furthermore, the extent of resetting was substantial and broadly similar for all of the deviant-tone conditions tested, although it is perhaps noteworthy that the greatest resetting effect was observed when the final inducer was replaced with silence (albeit that this trend was non-significant). This issue aside, the current results generally suggest that any salient, abrupt change to a single tone has a substantial and broadly comparable resetting effect. However, it is difficult to compare the results directly across the different deviant-tone conditions, as all of the changes to the last inducer tone were applied to different acoustic dimensions. Therefore, a systematic examination of whether resetting is affected by the size of a change is conducted in Chapter 4.

Previous studies have demonstrated that resetting occurs when all induction sequence tones differ from those of the test sequence (i.e. in frequency, ear of presentation, perceived location, or loudness, Anstis & Saida, 1985; Rogers & Bregman, 1993, 1998). For the current experiment, similar resetting effects were observed when the first nine induction tones were identical to the L tones of the test sequence, but the properties of the last (10^{th}) induction tone only were altered. This finding suggests that the abrupt change associated with a single deviant tone had an active resetting effect. As Rogers & Bregman (1998) used a relatively similar method to the current experiment, it is possible to compare the extent of resetting found in both these studies[2]. From their reported means, one can calculate the extent of resetting using the same scale to that for the current experiment (where no change from the standard induction sequence = 0.00, and a complete shift to the no-induction case = 1.00, see Table 3.1). Rogers & Bregman (1998) found that when the entire induction sequence was presented at +12 dB from the test sequence, the extent of resetting caused by this change was 0.64. For the current experiment, as shown in Table 3.1, the same change applied to the last induction tone

[2] It should be noted that Rogers & Bregman (1998) used a different method from the current study to measure reported stream segregation. Listeners rated the test sequence as either integrated or segregated in a basic adaptive staircase procedure. Each integrated response by the listener increased the frequency separation between the tones by 1 ST, and each segregated response reduced the frequency separation by the same amount. The final frequency separation at the termination of the staircase was taken as a measure of streaming. Despite this difference, these authors used similar control conditions to the current experiment, and so it is possible to calculate the extent of a resetting effect in the same way as for the current experiment (see results section). Note that because of the adaptive procedure Rogers & Bregman (1998) employed, the across-frequency-separation weighting system used in the current analysis is not applicable to their results.

resulted in an resetting index of 0.36 (which was also non-significant). Furthermore, Rogers & Bregman (1998) also found that sudden changes in perceived lateralization (from extreme right to extreme left) had resetting effects. Substantial resetting was observed when the induction sequence differed from the test sequence in either interaural level difference (or ILD, resetting = 0.45), or in interaural time difference (or ITD, resetting = 0.47). Although neither of these cues was manipulated in the current experiment, the values reported by Rogers & Bregman (1998) closely resemble the mean resetting effect of the various deviant-tone conditions (0.45, on average, see Table 3.1) which were tested here. This suggests that a single change in one tone at the end of the induction sequence can account for a large part of the resetting which was observed in Rogers & Bregman's (1998) study. One exception to this is their finding of a full extent of resetting (≈ 1.00) when the location of a sequence was changed using a loudspeaker arrangement. Anstis & Saida (1985) also found seemingly large resetting effects when a sequence was changed in frequency or ear of presentation. However, these authors did not provide a measure of perception in the absence of any build-up, and so one cannot calculate the proportional extent of resetting observed by Anstis & Saida (1985). Despite this, it would appear that the majority of the resetting effect observed in these previous studies can be replicated when only a single tone is changed. The implication of this finding is considered in the general discussion of the current chapter.

The current experiment also showed that replacing the last induction tone with silence had a substantial resetting effect. This change, in effect, resulted in a 300-ms silent interval between the 9^{th} inducer and the 1^{st} test-sequence tone. As the resetting effect of a similar silent interval was also measured in Experiment 2, the current finding is evaluated more fully in the general discussion at the end of this chapter.

In conclusion, Experiment 1 demonstrated that a variety of changes to a single deviant tone all substantially reset the build-up of reported segregation. There is strong evidence from the literature that streaming can also affect the ability of listeners to perform a range of pattern recognition tasks. Therefore, the subjective resetting effect observed in Experiment 1 should be replicable using a performance measure. This hypothesis was tested in Experiment 2, where streaming was measured in a temporal discrimination task.

3.3 Temporal Discrimination as a Measure of Stream Segregation

The stream segregation of sounds can reduce a listener's sensitivity to their temporal relationships. Generally, it is difficult to report the sequential order of sounds which are organized into separate streams (Warren et al., 1969; Bregman & Campbell, 1971). As discussed in Chapter 1, Bregman & Campbell (1971) demonstrated that listeners were easily able to report the sequential order of *within-stream* tones, but not the order of *across-stream* tones. Related to this, van Noorden (1975) measured the effect of streaming on the detection of delayed L tones in an HLH-HLH-HLH sequence (i.e. L tones were delayed from the mid-point between successive H tones - see Figure 3.3). The smallest detectable delay was measured in two conditions. The three delayed L tones were either included at the end of a long HLH- sequence, or they were part of an isolated 3 × HLH- sequence. Following a long sequence, the frequency separation between the tones strongly affected the detection threshold for the delayed tones - larger frequency separations resulted in an increased threshold. In contrast, when the delayed HLH- triplets were presented in isolation, performance was relatively unaffected by the frequency separation between the tones. These observations are consistent with the evidence that streaming builds up over time, and that the detection of temporal relationships is impaired when tones are heard as segregated. Both Neff et al. (1982) and Vliegen et al. (1999) have also reported similar findings.

It should be noted, however, that the results from these types of temporal discrimination task are not directly analogous to those from experiments where subjects directly report their perception. van Noorden (1975) demonstrated that, at certain frequency separations, tone sequences were always rated as either integrated or segregated (i.e. tone frequency separations below the fission boundary or above the temporal coherence boundary). However, these distinct boundaries are not apparent in temporal discrimination tasks. Instead, performance gradually and smoothly worsens with an increasing tone frequency separation. There are no sudden changes in performance which would signify the presence of either the fission or temporal coherence boundary. Vliegen et al. (1999) speculated on several possible explanations for this difference. Firstly, they proposed that there may only be an indirect relationship between temporal discrimination and stream segregation (i.e. both are affected by the frequency separation between successive tones). However, van Noorden (1975) demonstrated that temporal discrimination was impaired when the delayed tones were preceded by a long tone sequence. If temporal discrimination was solely governed by tone frequency separation, the presence of a prior tone sequence would not affect performance thresholds. Therefore, van Noorden's (1975) finding suggests that organizational factors do affect temporal discrimination. Vliegen et al.'s (1999) second explanation was that, when listeners are trying to hear integration, there may not be an absolute boundary at which integration ceases and segregation occurs. Rather, the likelihood of perceiving integration progressively decreases with an increasing frequency difference between tones. These issues suggest the need for some caution when comparing the results of subjective and objective measures of stream segregation.

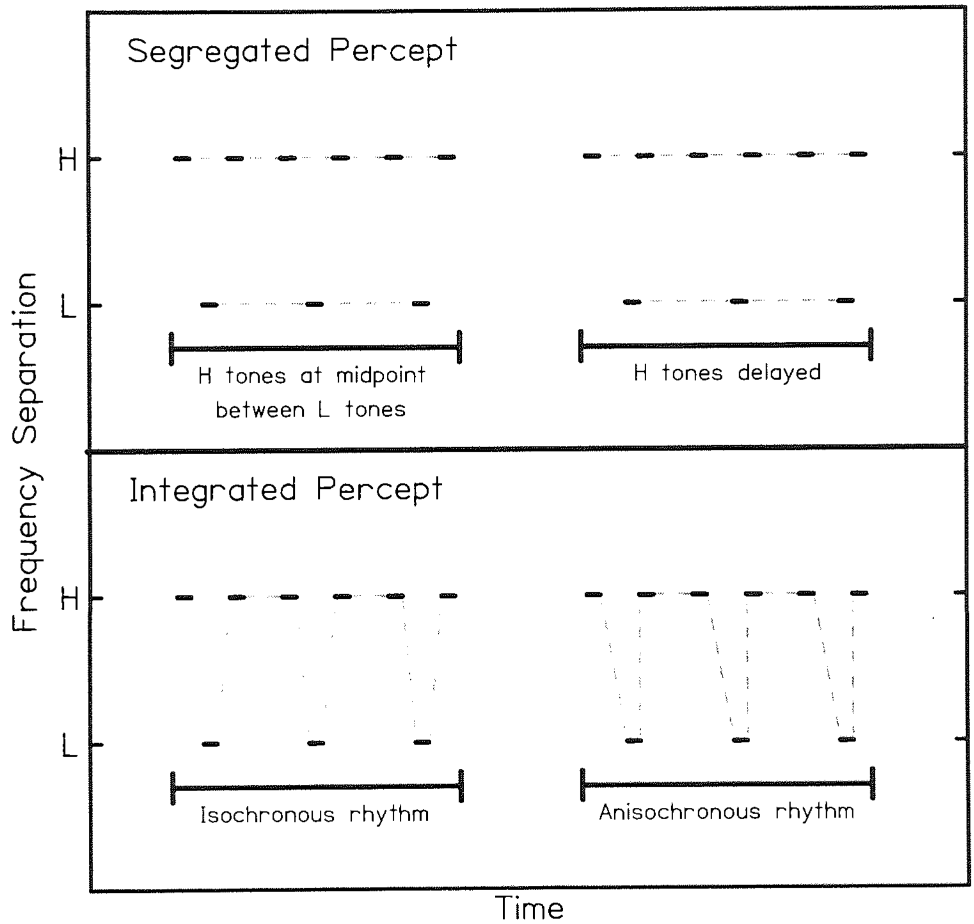

Figure 3.3: **The effect of stream segregation on the detection of delayed L tones.** Segregated Percept: **When the sequence is heard as two streams, the change in rhythm is very difficult to detect.** Integrated Percept: **When tones are heard as a single stream, the rhythmic difference between the isochronous and the anisochronous (delayed) sequence is highly salient.**

As discussed, van Noorden (1975) demonstrated that the build-up of stream segregation can impair the discrimination of the relative timings of tones in an HLH- sequence. Therefore, the factors known to trigger a resetting of build-up (Anstis & Saida, 1985; Rogers & Bregman 1993, 1998) should *improve* performance in a similar temporal discrimination task. This hypothesis was tested in a series of experiments by Roberts et al. (2008). They measured build-up and resetting using an induction/test sequence design in which listeners performed a temporal discrimination task on the test sequence (which comprised an LHLHLH tone arrangement). Listeners were presented with two repetitions of an identical stimulus, except that the H tones were delayed in one of the test sequences but not the other (cf. Figure 3.3). Participants were required to identify which interval contained the delayed H tones, and listeners' sensitivity to this delay was measured in an adaptive procedure (for which the duration of the delay was varied). The threshold was taken as an indication of streaming (i.e. an ability to detect a small delay was an indication that a sequence was easily heard as integrated). Similar to Rogers & Bregman's (1993) studies, a repeating L- induction sequence was used to promote segregation in the test sequence (see also Experiment 1). Roberts et al. (2008) found that, following this induction sequence, the temporal-discrimination threshold increased rapidly with the frequency separation between the test sequence tones. In contrast, when the test sequence was heard in isolation (with no prior induction sequence), temporal discrimination thresholds rose more gradually with increasing tone frequency separation. This is consistent with van Noorden's (1975) earlier finding, and an indication that the repeating L- induction sequence was effective at promoting segregation in the test sequence.

To measure resetting, Roberts et al. (2008) varied the tone properties of the induction sequence from those of the test sequence. As for the earlier studies by Rogers & Bregman (1993, 1998), these changes were applied to all the tones in the induction sequence. The changes tested were: number of tone onsets (several short tones or one extended tone), tone/silence ratio (standard or extended-duration tones), frequency, level, and interaural time difference. All of these changes reduced thresholds towards those for the case when no induction sequence was presented. The resetting effects of these changes were broadly consistent with previous studies of reported perception (frequency - Anstis & Saida, 1985; an increase in level, or a change in perceived location - Rogers & Bregman, 1998). However, the extent of resetting observed was much greater in the temporal discrimination task than for these earlier subjective experiments. Roberts et al. (2008) found an almost full resetting effect following changes in frequency, ITD, or an increase in level, whereas Rogers & Bregman (1998) only observed an (approximately) 50% reduction in rated segregation following similar changes in level or perceived location. Also, an abrupt decrease in level produced around 50% resetting in the temporal discrimination task, but the same change had no resetting effect in a subjective task (Rogers & Bregman, 1998). Similarly, increasing the duration of the induction tones resulted in near-full resetting in the temporal discrimination experiment, but had very little effect on reported stream segregation (Rogers & Bregman, 1993).

In conclusion, the resetting effect demonstrated in a temporal discrimination task appears to be stronger, and to be triggered by a greater variety of changes than for tasks where listeners report their perception of stream segregation. The reason for this discrepancy is unclear, but it could potentially be explained by the different types of induction sequences used in these experiments. Roberts et al. (2008) used a relatively short (2 s) repeating L- induction sequence, whereas Rogers & Bregman (1998) used a longer (4.8 s) HLH- induction sequence. It is likely that the longer induction sequence would promote more build-up, and so it is conceivable that more segregation transferred from the induction sequence to the test sequence (as suggested by Roberts et al., 2008). The contrast in results may also indicate a difference in the segregation-promoting effect of L-only and HLH- type induction sequences (see Chapter 8). Nonetheless, resetting demonstrated in the temporal discrimination task cannot be attributed to a disruption of timing judgments caused by a distracting effect of a sudden change in tonal properties, because the sudden changes resulted in improved performance.

3.4 Experiment 2

Overall, Roberts et al. (2008) concluded that a sudden change in sequence properties triggers resetting and that this can improve subsequent temporal discrimination. Experiment 1 used a subjective measure and demonstrated that a single deviant tone can reset the build-up of segregation. Based on Roberts et al.'s (2008) findings, one would also expect this resetting effect to be apparent when performance is measured in a temporal discrimination task. This hypothesis was tested in Experiment 2.

3.4.1 Method

Participants

Eight listeners successfully completed the training procedure and took part in Experiment 2. All reported normal hearing. Listeners were trained and screened prior to the main experiment (see procedure; two failed this training). Seven of the eight participants who completed the experiment had also taken part in Experiment 1 (including the author).

Stimuli and Conditions (Test Sequence)

Tone properties were generally similar to those used in the subjective experiments (see Chapter 2), except that the standard tone duration was reduced from 100 ms to 60 ms (including 10-ms raised cosine ramps at onset and offset). Standard tone onset-to-onset times were preserved by introducing a compensating 40-ms silence between consecutive L and H tones. Tones were presented diotically and at 70 dB SPL. The test sequence comprised 3 LH repetitions plus a final L tone added at the end of the sequence. The frequency of the L tones was fixed at 1 kHz, and the H tones were set to be either 0, 4, 8, or 12 semitones higher (1000, 1260, 1587, or 2000 Hz, respectively). As in the subjective experiments, the onset-to-onset time of the L tones was 200 ms (resulting in a 140-ms inter-tone interval). For the regular sequence, H tones were inserted at the temporal mid-point between successive L tones (i.e. 40 ms inter-tone silence between successive L & H tones). For the target sequence, the H tones were delayed from this mid-point (see procedure). H tones were delayed rather than advanced, as van Noorden (1975) demonstrated that listeners are more sensitive to that change. The reason for the additional L tone at the end of the test sequence was to ensure that the overall duration of the test sequence was constant irrespective of whether or not the H tones were delayed (i.e. all test sequences lasted 660 ms).

This tone presentation rate and the range of frequency separations used are broadly equivalent to those used by Roberts et al. (2008). These authors found that these parameters were suitable for a similar temporal discrimination task. The test sequence properties used here are also similar to those used in Experiment 1.

Stimuli and Conditions (Induction Sequences)

The effects of four different induction sequences on temporal discrimination in the test sequence were measured. These induction conditions were broadly similar to four of the conditions used in Experiment 1. Two conditions measured performance after the presence or absence of a standard segregation-promoting induction sequence. The other two induction conditions measured the resetting effect of a single deviant tone inserted at the end of the induction sequence. Whilst all of the deviant tones tested in Experiment 1 had a broadly similar resetting effect, there was some variance between conditions. Most noticeably, the +12 dB deviant resulted in the least resetting measured (a 0.36 extent of resetting – Table 3.1), whilst replacing the final inducer with silence resulted in the most (0.64). These two conditions were replicated in Experiment 2, to explore whether a similar trend in the results would be observed using a temporal-discrimination measure.

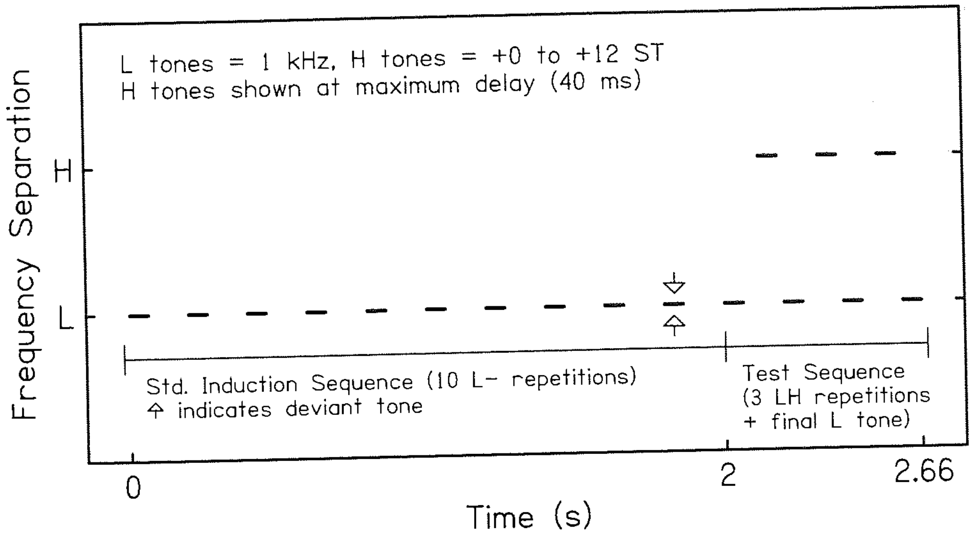

Figure 3.4: **Illustration of a standard induction sequence paired with a target (irregular) test sequence in Experiment 2. The arrows indicate the induction tone which was altered for the deviant-tone conditions.**

1. *Standard induction condition*

   The standard induction sequence comprised 10 L- repetitions. These L tones were identical to those in the test sequence (as described above). As in the test sequence, the L tones were separated by an inter-tone silence of 140 ms (each L-cycle = 200 ms, total induction sequence length = 2 s). Therefore, the induction tones matched the L tones of the test sequence in both frequency and tempo. This induction sequence was expected to promote stream segregation in the test sequence. An illustration of the standard induction sequence paired with a test sequence is shown in Figure 3.4.

2. *No-induction condition*

   For this condition, the test sequence was presented alone, with no preceding induction sequence. This was used as a measure of performance in the absence of any prior build-up effect.

3. *Deviant increased in level - or - Deviant(+12 dB)*

   For this condition, the last tone of the standard induction sequence was altered. The level of this tone was increased from 70 dB to 82 dB SPL. Any resetting effect associated with this change should improve performance relative to that for standard induction condition.

4. *Deviant replaced with silence - or - Deviant(silent)*

   For this condition, the last tone of the standard induction sequence was replaced with an equivalent-duration silent interval. Any resetting effect associated with this change should improve performance relative to that for standard induction condition. Given that this change had the greatest resetting effect in Experiment 1, the improvement in performance may be even greater than that for the Deviant(+12 dB) condition.

Procedure

The apparatus used was identical to that which is described in Chapter 2. On each trial, participants were presented with two stimuli. One interval contained a regular test sequence, and the other contained a target test sequence in which the H tones were delayed. The size of the delay imposed on the H tones was varied adaptively. The order of the stimulus presentation within each trial was randomized. Participants were required to detect which interval contained the stimulus with delayed H tones, in a 2I-2AFC procedure. Visual feedback was given after each trial (indicating either a correct or an incorrect response). Runs were initiated, and trial responses were made, via a computer keyboard. Within each trial, a 1-s silence was inserted between the two stimuli, and successive trials were separated by a minimum of a 3-s silence (i.e. 3 s from a response being made). These silent intervals were chosen to ensure that the majority of build-up would decay between successive stimuli (Bregman, 1978; Cusack et al., 2004), but also to keep the overall duration of the experiment as short as possible (cf. Roberts et al., 2008).

The smallest detectable H-tone delay was estimated using a 3-up 1-down adaptive staircase procedure (Levitt, 1971). This provides a measure of the 79.4% correct point on the psychometric function. The maximum delay imposed on the H tones was 40 ms (any larger delay would have resulted in the H tones overlapping with the L tones, see Figure 3.4). From this maximum, the delay was adjusted on a logarithmic scale - the initial step size was 1.414, but after 2 turn points this was reduced to 1.189. The step size was then held constant for four further turn points (i.e. runs were ended after a total of six turn points). The geometric mean of the delay at the last 4 turn points was taken as the threshold for temporal discrimination. If a participant made six incorrect responses at the maximum (40 ms) delay, the run was terminated and 40 ms was taken as the threshold estimate (this occurred for 12.6% of runs). These criteria for the adaptive procedure are similar to those used previously by Roberts et al. (2002, 2008).

Before the main experiment, participants attended a combined training and screening session. First, participants were presented with examples of regular and of clearly delayed sequences. Once listeners were confident of the rhythmic differences between the two sequences, the session began. The training started with three runs of a stimulus which was expected to result in good temporal discrimination performance (no induction sequence, frequency separation in the test sequence = 0 ST). In order to proceed to the main experiment, listeners were required to detect the delay at a mean threshold of 20 ms or less (two participants failed to meet this criterion). Successful listeners then completed a training session comprising eight runs which were selected to reflect the variety of conditions used in the main experiment. The stimuli for this training session were presented in a fixed order, indicated in terms of induction sequence condition and test-sequence frequency separation, which was: standard induction (0 ST), no-induction (4 ST), standard induction (4 ST), deviant(silent) (4 ST), deviant(+12 dB) (8 ST), no-induction (12 ST), deviant(silent) (8 ST), and standard induction (12 ST).

The main experiment was made up of three blocks of trials. Each block comprised a single run of each induction sequence combined with each test-sequence frequency separation (4 $\times$ 4). This resulted in 16 runs for each trial block (and 48 runs for the whole experiment). The order of presentation within each trial block was randomized anew. The geometric mean of the three estimates was taken as the threshold for each condition. However, if the standard deviation of these three log values was greater than 0.2 (approximately 27.3% of the data), the participant completed an extra run, and the outlying threshold estimate was excluded from the data set. Any additional runs were completed at the end of the main experiment, and were presented in their original order. The total duration of the experiment was approximately 6 hours, and participants typically completed the experiment in three $1\frac{1}{2}$ - 2 hours sessions.

Apparatus

The apparatus was identical to that which is described in Chapter 2.

3.4.2 Results

The results from Experiment 2 are displayed in Figure 3.5. Each point represents the mean temporal discrimination threshold for a given condition (each average was calculated from the data for eight listeners). The graph shows that performance worsened as the frequency separation between the tones in the test sequence increased. The geometric mean thresholds for frequency separations of 0, 4, 8, and 12 ST were 8.8 ms, 15.0 ms, 22.3 ms, and 26.5 ms, respectively. Performance was also affected by the type of induction condition - the geometric mean thresholds for the standard, no-, deviant(+12 dB), and deviant(silent) conditions were 18.9 ms, 14.1 ms, 17.7 ms, and 16.6 ms, respectively. The results were analysed using repeated-measures ANOVA. The independent variables were induction condition (standard induction, no induction, and $2 \times$ deviant-tone induction conditions) and frequency separation (0, 4, 8, and 12 ST). The main effects of frequency separation [$F(3, 21) = 50.045$, $p < 0.001$, $\eta_p^2 = 0.877$] and induction condition [$F(3, 21) = 9.155$, $p < 0.001$, $\eta_p^2 = 0.567$] were both significant. The two-way interaction between these variables was also significant [$F(9, 63) = 3.067$, $p < 0.005$, $\eta_p^2 = 0.305$]. The significant interaction term primarily reflects the finding that thresholds were similar for all induction conditions only in the absence of an HL frequency difference. This is as expected, because differences in the build-up of the tendency for stream segregation cannot be revealed in practice unless there is a difference in frequency between the H and L tones. For the no-induction condition, there was a gradual increase in temporal discrimination thresholds with increasing test-sequence frequency separation. Thresholds rose more rapidly in the other three conditions. It is likely that this difference also would have contributed to the significant interaction observed.

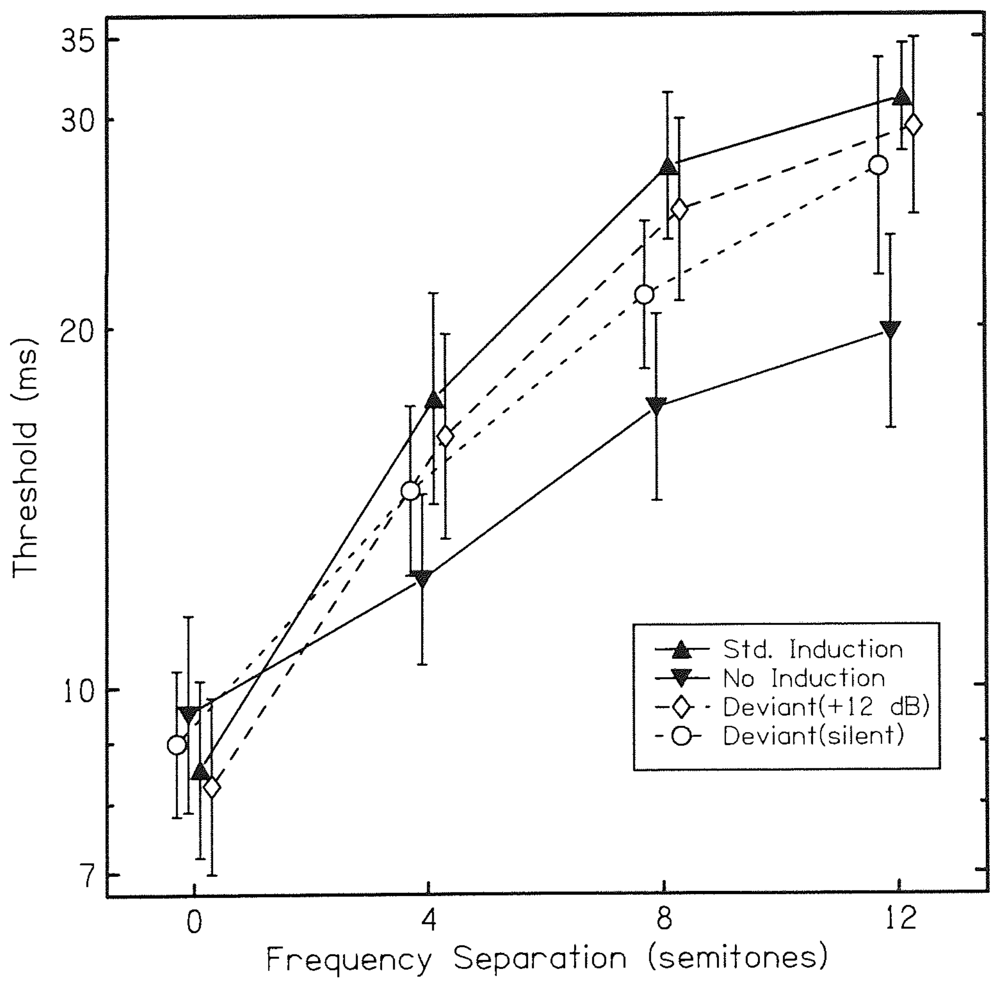

Figure 3.5: Results from Experiment 2. The effects of induction condition and frequency separation in the test sequence on temporal discrimination. Each point represents the average threshold for the detection of delayed H tones for eight listeners. The largest possible delay was 40 ms. The insert identifies the different induction sequences tested (deviant refers to the change applied to the last tone in an otherwise standard induction sequence). Error bars represent ± 1 inter-subject standard error.

From Figure 3.5, it is apparent that the highest temporal discrimination thresholds were recorded for the standard induction condition. In contrast, performance was best in the no-induction condition. A two-tailed pairwise comparison (a restricted LSD test, Snedecor & Cochran, 1967) on the previously described data set confirmed that the predicted difference between the standard and no-induction conditions was significant [difference in mean threshold = 4.8 ms, $t(7) = 4.257$, $p < 0.005$]. Furthermore, when compared with the standard induction condition, performance was somewhat improved for both the deviant(+12 dB) and deviant(silent) conditions. Pairwise comparisons indicated that performance for the deviant(silent) condition was significantly better than for the standard induction condition [2.3 ms, $t(7) = 3.902$, $p < 0.01$], but performance in the deviant(+12 dB) condition was not [1.2 ms, $t(7) = 1.348$, $p > 0.05$].

Frequency Separation	Deviant Condition		
	+12 dB	Silent	*Geometric Mean*
4ST (*27.49%*)	0.21	0.51	***0.33***
8 ST (*36.75%*)	0.18	0.53	***0.31***
12 ST (*35.75%*)	0.12	0.29	***0.19***
Weighted Mean	***0.17***	***0.44***	***0.27***

Extent of resetting = standard induction (***0.00***) → no induction (***1.00***)

Table 3.2: **Results from Experiment 2. Average thresholds for the two deviant-tone conditions were compared to those for the standard induction case. Values are expressed as a proportional shift away from the (log) standard induction threshold towards the (log) no-induction threshold (the *extent of resetting*). As results for the 0 ST condition were non-differentiating, these data were excluded from this analysis. Also, the across-frequency-separation means were weighted in a manner consistent with the results of Experiment 1 (see main text). As for Table 3.1, *only* the across-frequency means are weighted. Each individual value represents the unweighted extent of resetting.**

Finally, the data were used to calculate an extent of resetting for the deviant-tone conditions (a measure analogous to that used in Experiment 1, see Table 3.1). This was achieved by calculating the proportional return in threshold from the standard (0.00) towards the no-inducer case (1.00) in geometric space. In order to compensate for differences between the standard and the no-induction conditions across test-sequence frequency separations, a weighting system was applied to these means. These weights were determined using the same procedure as that employed in Experiment 1, except that the values were calculated in geometric space. These values for each deviant-tone condition at each frequency separation are displayed in Table 3.2. Also note that data from the 0-ST frequency separations are excluded from this table.

3.4.3 Discussion

Experiment 2 found that, as the HL frequency separation was increased, temporal discrimination thresholds rose most rapidly when the test sequence was preceded by the standard induction sequence. The most gradual rise in thresholds was observed for the no-induction condition. As higher thresholds are associated with increased stream segregation, this indicates that streaming built up over the course of the standard induction sequence. This finding is in good accord with Roberts et al.'s (2008) previous temporal discrimination study. The two deviant-tone conditions tested in Experiment 2 were chosen from those which demonstrated the greatest and least extent of resetting in the subjective task used in Experiment 1. These conditions were when the deviant tone was replaced with silence, or was increased in level - for Experiment 1, the extent of resetting for these cases was 0.64 and 0.36, respectively (see Table 3.1). For the current study, thresholds were significantly lower for the deviant(silent) condition when compared with the standard induction condition. This indicates that partial resetting occurred when the last induction tone was replaced with an equivalent-duration silent interval. This observed resetting effect of a brief silence is consistent with that shown in Experiment 1. Only a modest (and non-significant) trend towards resetting was found when the final induction tone was increased in level by 12 dB. This was also reflected in the proportional extent of resetting from the standard to the no-inducer case caused by each deviant case when averaged over the non-zero test-sequence frequency separations (see Table 3.2). These values for the silent and +12 dB deviants were 0.44 and 0.17, respectively. Although these values suggest a smaller tendency for resetting than was found in the subjective task used in Experiment 1, it is notable that the relative extents of resetting are broadly comparable for these deviant types across both experiments. Indeed, in both experiments the effect of the deviant(+12 dB) fails to reach significance. Regardless of the apparent differences in the overall extent of resetting for equivalent changes in the two experiments, reduced thresholds for the deviant-tone conditions cannot be attributed to a disruption of timing judgments, as performance in these conditions was *improved* relative to that for the standard induction case.

Roberts et al. (2008) measured resetting in a temporal discrimination study that was similar to the current experiment. These authors demonstrated that substantial resetting occurred when an induction sequence differed from a test sequence either in frequency, level, lateralization, or temporal characteristics. The current experiment found that this resetting effect could be partially replicated when only the final induction tone was increased in level or replaced with silence. However, the resetting effect of a deviant tone was relatively modest in this experiment. In contrast, Roberts et al. (2008) found a much larger resetting effect when the entire induction sequence was changed from the test sequence (near 100% resetting). Indeed, these authors demonstrated that the effect of resetting on temporal discrimination performance was much greater than for that when listeners directly reported their perception (Rogers & Bregman, 1993, 1998). As previously described, the opposite effect was observed in the current experiments. The reason for the discrepancy between Experiment 2 and Roberts et al.'s (2008) previous finding is unclear. Nonetheless, the findings of Experiment 2 are broadly consistent with Experiment 1, in that both of these studies demonstrated that a single deviant tone can cause the build-up of stream segregation to be partially reset.

3.5 General Discussion

Previous studies have demonstrated that the build-up of stream segregation can be reset when the properties of an on-going sequence are suddenly altered (Anstis & Saida, 1985; Rogers & Bregman, 1993, 1998; Roberts et al., 2008). Rogers & Bregman (1998) offered two explanations for resetting; that build-up may fail to transfer between tone sequences of different characteristics, or that an abrupt change in sequence properties may actively reset build-up. Experiments 1 and 2 have both demonstrated that a single deviant tone can partially reset the build-up of segregation. For the deviant-tone conditions, the first nine induction tones were identical to the L tones of the test sequence, and only the properties of the last induction tone were altered. As the majority of the induction tones matched the test sequence tones, any transferral of build-up should be relatively unaffected by the presence of the single deviant tone. Therefore, any resetting observed for these conditions can be taken as an indication that the abrupt change actively resets build-up. Substantial resetting was most clearly shown in the subjective task of Experiment 1. The temporal-discrimination measure used in Experiment 2 was broadly consistent with the results of Experiment 1, but the observed resetting effect was rather smaller in the former case than in the latter. These results are somewhat inconsistent with previous research which has shown that abrupt changes at the inducer/test boundary can strongly affect temporal discrimination performance (Roberts et al., 2008). Although the reason for this difference is unclear, it would appear that resetting arising from a single deviant tone can be demonstrated more effectively using a subjective task. This method is also much more time-efficient than the adaptive procedure that was used in Experiment 2. For these reasons, the experiments presented in chapters 4 - 7 all used the same general method as for Experiment 1.

Whilst the evidence from experiments 1 and 2 appears to indicate that a sudden change has an active resetting effect, there is a possible alternative explanation for these results. Rogers & Bregman (1998) suggested that build-up may not transfer between tone sequences of different characteristics. If so, a deviant tone may not contribute to the cumulative build-up occurring within an induction sequence. Consequently, reduced segregation for the deviant-tone conditions may simply indicate that less build-up occurred (in comparison to the standard induction case). This seems unlikely given that only a single tone was altered, but stream segregation was greatly reduced (at least for Experiment 1). One would not expect such a large reduction if the altered tone merely did not contribute to build-up. However, this explanation cannot be fully dismissed, as a review of the literature suggests that the rate of build-up within a repeating L- type induction sequence has never been directly measured. Therefore, these issues are explored further in Chapter 5. For the current discussion, it is assumed that the observed effect was due to the deviant tone actively resetting build-up caused by the earlier tones in the induction sequence.

The active resetting effect of an abrupt change has several implications for Bregman's (1978, 1990) functional explanation of build-up. His theory proposed that build-up represents a gradual accumulation of evidence in favour of two separate sound sources being present (so as to avoid excessive fluctuations in perceptual organization). Rogers & Bregman (1998) suggested that certain abrupt changes are likely to indicate the onset of a new sound source, and that the perception of a new acoustic event may reset the evidence-accumulation process. If this theory is correct, resetting should only occur when a changed tone is perceived to originate from a separate sound source. A change which is heard as a continuation of an existing stream should not have a resetting effect. However, for some of the conditions tested in Experiment 1, the deviant tone would be unlikely to segregate from the on-going induction-tone stream (e.g. when a tone was changed only in duration). Despite this, substantial resetting was still observed for these conditions. Similarly, a change in the ITD of an on-going sequence can have a substantial resetting effect (Rogers & Bregman, 1993; Roberts et al., 2008), but ITD differences are not themselves a strong cue for stream segregation (Boehnke & Phillips, 2005). Based on this evidence, it would seem that any salient change has a resetting effect - irrespective of whether the changed tone is heard as a separate acoustic event. This hypothesis is further considered in Chapter 4.

Both experiments 1 and 2 also measured the resetting effect of replacing the final induction tone with silence. Bregman (1978) provided evidence that build-up gradually decays during a silent interval after the termination of a tone sequence. He manipulated the silent interval between four-tone packages (package-silence-package-silence etc...), and asked listeners to increase the rate of tone presentation until segregation was heard.

Generally, longer silent intervals caused listeners to hear stream segregation at a slower rate (indicating a stronger perception of integration). More specifically, the results indicated that the decay of segregation was relatively rapid over the first 1.5 seconds of silence, but that the effect was still apparent after 4 s. Beauvois & Meddis (1997) also reported a broadly similar observation. These authors used a similar sequence structure to that used in Experiment 1, but inserted a silent interval of varying length between the induction sequence and the test sequence. For musically experienced listeners, they found a gradual, and somewhat linear decay of build-up over the range of intervals tested (0-8 s). A lesser, but more rapid, decay was observed for non-musical listeners. For both sets of participants, very little decay was observed at the shortest silence they measured (500 ms).

In contrast to these studies, the current experiments found that reported segregation was greatly reduced when the final induction tone was replaced with an equivalent-duration silent interval (i.e. a 300-ms silence between the 9^{th} induction tone and the first test-sequence tone). Whilst it is difficult to compare directly between studies, the extent of resetting currently observed was clearly much greater than the decay of stream segregation previously associated with a similar duration of silent interval (Bregman, 1978; Beauvois & Meddis, 1997). This suggests that whilst decay may have partially contributed to the current pattern of results, this effect would have been minimal, and so cannot explain fully the observed reduction in reported segregation. Therefore, it is concluded that the 300-ms silent interval must have also had an active resetting effect, and this accounted for the majority, if not all, of the decrease in segregation. The difference between the currently observed results and those of previous studies may be because the current design allowed for a more sensitive measure of resetting. For example, Beauvois & Meddis (1997) used a very long (10 s) induction sequence, which may have had a stronger segregation-promoting effect. Therefore, a longer silent interval may have been required for the decay/resetting of stream segregation to occur.

Notwithstanding this, the current findings are consistent with a study by Cusack et al. (2004). These authors showed that when a range of silent intervals (1 s - 10 s) was inserted into an on-going LHL- sequence, all intervals had a similar (and substantial) resetting effect. This also relates to the studies of Carlyon et al. (2001, 2003), who demonstrated that switching attention from a separate task to an on-going tone sequence resulted in no evidence of any prior build-up before the switch in attention. These authors explained this finding in terms of the hierarchical decomposition model. Briefly, this model proposes that in a multi-source listening environment, there may be some form of automatic stream segregation, but only the attended source is subject to a detailed perceptual representation (see also Brochard et al., 1999). Based on this, Cusack et al. (2004) offered two explanations for the resetting effect of a brief silence. Firstly, they proposed that the tones occurring after a silence may be heard as a new acoustic event, for which no prior build-up has occurred. Alternatively, they suggested that resetting may occur when a listener's attentional focus is changed. Following a silent interval, a listener may pay increased attention to the subsequent onset of the sequence, and this could result in an increased tendency to hear integration. For the current stimuli, deviant-tone resetting may be due to the changed tone temporarily drawing attention away from the on-going sequence. Note that this explanation of resetting is equally applicable to the variety of deviant tones currently tested, and not just to the silent-tone case.

Chapter 4

The Effect of the Size of the Change to the Deviant Tone on Resetting

4.1 Introduction

Experiments 1 and 2 demonstrated that a single, abrupt change to one tone in an on-going sequence can have a partial, and often substantial, resetting effect. In the subjective task used in Experiment 1, participants reported significantly less stream segregation in the test sequence (3 LHL- triplets) when the final tone of the induction sequence (10 L- repetitions) was either lowered in frequency (- 3 ST, standard tone = 1 kHz) or extended in duration (+50 ms, standard tone = 100 ms). A similar resetting effect was observed when the final inducer was replaced with an equivalent-duration silent interval. The resetting effect of a silent interval was also confirmed in the temporal-discrimination task of Experiment 2. There is also evidence from both experiments of a trend towards resetting produced by an increase in tonal level (+12 dB, standard tone = 70 dB SPL). The wide range of changes tested in Experiment 1 all resulted in a broadly similar amount of resetting. This finding could be taken as evidence that *any* salient change, or perhaps even any noticeable change, to an on-going sequence can have an equivalent resetting effect (at least for a subjective task). Whilst this conjecture appears plausible, it is important to note that all of the changes tested were applied to different tonal dimensions (frequency, level, duration, or replacement with silence). The current experiments were designed to examine systematically the relationship between the size of a change applied to a deviant tone and the associated resetting effect. This was tested by manipulating the size of the change across a single dimension (tone frequency).

Bregman (1978, 1990) proposed that build-up represents a gradual accumulation of evidence in favour of two separate sound sources being present. He suggested that at the onset of an alternating tone sequence (e.g., a LH arrangement), the auditory system assumes that both tone subsets have originated from the same acoustic event, and so integration is perceived. Gradually over time, evidence is gathered which indicates that the two subsets of tones are of consistently different properties. From this evidence, the stream-formation mechanism becomes increasingly likely to segregate the sounds on the basis of their continued difference from each other. The relative slowness of this process ensures that perception does not rapidly fluctuate between alternative perceptual organizations. Rogers & Bregman (1998) attempted to expand this theory also to encompass the resetting of build-up. These authors demonstrated that an induction sequence which differed from a subsequent test sequence in either perceived location or in level was relatively ineffective at promoting stream segregation in the test sequence. Rogers & Bregman (1998) offered two explanations for these findings; (1) that build-up does not transfer between tone sequences of different characteristics, and (2) that an abrupt change may indicate the onset of a new acoustic event. They proposed that the perception of a new event may trigger a re-analysis of the auditory scene, which in turn may reset any build-up which had previously occurred.

As only a *single* deviant tone can have a large resetting effect (see experiments 1 & 2), resetting cannot solely be explained as a failure of build-up to transfer between tone sequences of different characteristics. Therefore, the current experiments were designed to explore Rogers & Bregman's (1998) second hypothesis in more depth. If resetting only occurs when a listener perceives the presence of a new sound source, one would expect a relationship between the size and/or the direction of a change and the associated resetting effect. More specifically, if a certain change is *not* heard as signifying the onset of a new event, then it should not have a resetting effect. Rogers & Bregman (1998) reported some evidence that was consistent with this concept. Specifically, an increase in level (+6 dB from 65 dB SPL) at the induction/test sequence boundary had a large resetting effect, whilst a similar decrease in level (-6 dB) had practically no effect on build-up. They concluded that this was because a sudden increase in loudness can indicate the presence of a new sound source, whereas a sudden decrease cannot. Using a temporal-discrimination task, Roberts et al. (2008) tested similar conditions to those examined by Rogers & Bregman (1998). They found that although a decrease in level at the induction/test boundary did have a noticeable resetting effect (as measured by changes in temporal discrimination thresholds), nevertheless it was rather less than for an equivalent increase in level. In summary, there is evidence suggesting that the direction of a level change can influence the extent of resetting - changes which may signify a new acoustic event have a much larger resetting effect than changes which are less likely to indicate a new acoustic source.

If resetting only occurs when a new source is perceived, one would also expect some relationship between the size of a change and the associated resetting effect. If we consider naturally occurring sounds, a sudden, but small change along any given tonal dimension would be unlikely to indicate a new source, whereas a larger abrupt change would (Bregman, 1990, see also Chapter 1). Therefore, one would expect a greater resetting effect as the size of a change is increased. Anstis & Saida (1985) presented evidence which is largely consistent with this argument. These authors applied abrupt frequency changes to an on-going tone sequence. Streaming was induced in a 4-s alternating HL sequence (frequency separation = 2 ST, centre frequency of sequence = 1 kHz, see the next paragraph for a more complete description of these stimuli) and the centre frequency of a subsequent 1-s test sequence was manipulated (range = ± 0 to 12 ST). This 5-s arrangement was repeated throughout a 90-s sequence. For the induction sequence, the tone onset-to-onset time was fixed at 125 ms. Participants adjusted the rate of the test sequence so that the perceptual border between integration and segregation was heard. The mean rate over the final 30 s of the sequence was taken as a measure of the threshold between integration and segregation. A slow rate was required when the centre frequency of the test sequences was shifted within a range of -1 to +3 ST from the induction sequence. As slow rate indicates a difficulty in hearing integration, this finding was taken as evidence that build-up successfully transferred between the induction and test sequences in these conditions. In contrast, the larger frequency changes resulted in listeners hearing the same threshold at a faster rate which did not greatly vary as the size of the frequency change was further increased (i.e., the test sequences were more easily heard as integrated). This suggests that the changes in centre frequency outside of a -1 to +3 ST range all reset the build-up of stream segregation.

Anstis & Saida (1985) observed a finely tuned range in which no resetting occurred (changes within -1 to +3 ST from 1 kHz), and this range was considerably larger than the smallest detectable difference in this frequency region (approximately 2-3 Hz, Moore, 2003; Sek & Moore, 1995). Note however that Anstis & Saida (1985) created alternating-frequency tone sequences by applying a square-wave frequency modulation (FM) to a continuous carrier tone. Based on this, the lack of observed resetting following small changes in (carrier) frequency may have been influenced in some way by the adaptation of specialized frequency-modulation sensitive channels (Kay & Matthews, 1972; Green & Kay, 1973; Gardner & Wilson, 1979). Anstis & Saida's (1985) induction sequences comprised a 1 kHz carrier, frequency modulated at a rate of 4 Hz (i.e., a 125-ms onset-to-onset time between low- and high-frequency tone components). Subjects manually adjusted the test sequence modulation rate - to between 3 and 5 Hz on average (and the carrier frequency was varied across conditions). Green & Kay (1973) also used a 1 kHz carrier frequency and similar modulation rates, and they reported evidence of adaptation to frequency modulation. Specifically, after exposure to a FM tone (12 s), thresholds for detecting subsequent FM at the same carrier and modulation frequencies were elevated in comparison to control conditions (see also Regan & Tansley, 1979). However, it is important to note that a wide range of subsequent studies have clearly demonstrated build-up and resetting in sequences of discrete pure tones (Bregman, 1978; Rogers & Bregman 1993, 1998: Carlyon et al., 2001; Roberts et al., 2008). Therefore, it would seem unlikely that these effects can be explained simply in terms of adaptation to FM. Despite this, adaptation to FM may have had some influence on Anstis & Saida's (1985) findings regarding the resetting effects of abrupt changes in carrier frequency.

It is interesting to note that Anstis & Saida (1985) found that frequency changes which failed to trigger resetting broadly fell within the bandwidth of the auditory filter at that frequency region (where 1 $ERB_N \approx 1$ ST, Glasberg & Moore, 1990) - except that the boundary for resetting was offset by +1 ST. This offset can potentially be explained by the properties of the induction sequence. As the induction sequence modulation rate was fixed, the final 'tone' before the test sequence was *always* an H-tone, and the test sequence always began with an L-tone. This would result in a smaller frequency change between the final induction tone and the first test tone when the centre frequency of the test sequence was *increased* than when it was decreased by the same amount. One might speculate that this subtle difference may have influenced the continuity between the induction and test sequence, and so may have also influenced the extent of resetting. When evaluating Anstis & Saida's (1985) study, it is also important to note that the presentation rate could vary between the induction and test sequence (as the rate of the test sequence was adjusted by the listener). There is evidence that a change in the number of tonal onsets per unit time may affect the dynamics of build-up (Rogers & Bregman, 1993), and that a sudden change in tonal duration can have a resetting effect (Roberts et al., 2008; Experiment 1). Therefore, it is difficult to assess the relative contributions of the change in centre frequency *and* the change in presentation rate on the resetting effect observed by Anstis & Saida (1985).

The current experiments were designed to explore further the relationship between the size of a change to a single tone and the associated resetting effect. Specifically, only the last induction tone was lowered in frequency, and streaming was measured in a subsequent test sequence. As the frequency of the deviant tone was only lowered in comparison to the standard L-tones, the current experiments did not address the effect of the direction of a frequency change. This decision was made to avoid a potential confounding effect. Namely, a frequency increase applied to the deviant tone would often result in an increased similarity between the deviant tone and the test-sequence H-tones. As there appears to be a complex relationship between the direction of a frequency change and resetting (see Anstis & Saida, 1985), this issue was considered to be beyond the scope of the current experiments.

4.2 Experiment 3a

4.2.1 Method

The general method and procedure for Experiment 3a is described fully in Chapter 2. On each trial, participants heard an induction sequence which was immediately followed by a test sequence (test sequence = 3 $\times$ LHL- triplets, L-tone = 1 kHz, H-tone = +4 to +14 ST, tone duration = 100 ms, inter-triplet silence = 100 ms). Participants were required to indicate their perception of the final LHL- triplet as either integrated or segregated in a 2AFC procedure.

Participants

Eight listeners took part. Four of these participants had previously taken part in Experiment 1, and all reported normal hearing.

Induction Conditions

Both the standard induction condition and the no-induction condition were used in this experiment (these conditions are described in detail in Chapter 2). The standard induction sequence comprised 10 L- repetitions (L tone = 1 kHz, tone duration = 100 ms, inter-tone silence = 100 ms). All five experimental conditions were a modification of this standard induction sequence. For these conditions, the final induction tone was lowered to one of a range of different frequencies (see Figure 4.1). Except for this change in frequency, the deviant tone was otherwise identical to a standard L-tone. The deviant tone was set either to 0.5, 1, 3, 6, or 12 ST lower than the standard 1-kHz tone (972, 944, 841, 707, and 500 Hz, respectively). This range of frequency changes was chosen as it was similar to that previously tested by Anstis & Saida (1985). Note that the bandwidth of the auditory filter centred on 1 kHz is approximately 133 Hz (Glasberg & Moore, 1990). Therefore, for the current experiment, the frequency changes of -0.5 and -1 ST both fell within 1 ERB_N of the standard 1-kHz tone, but the others did not.

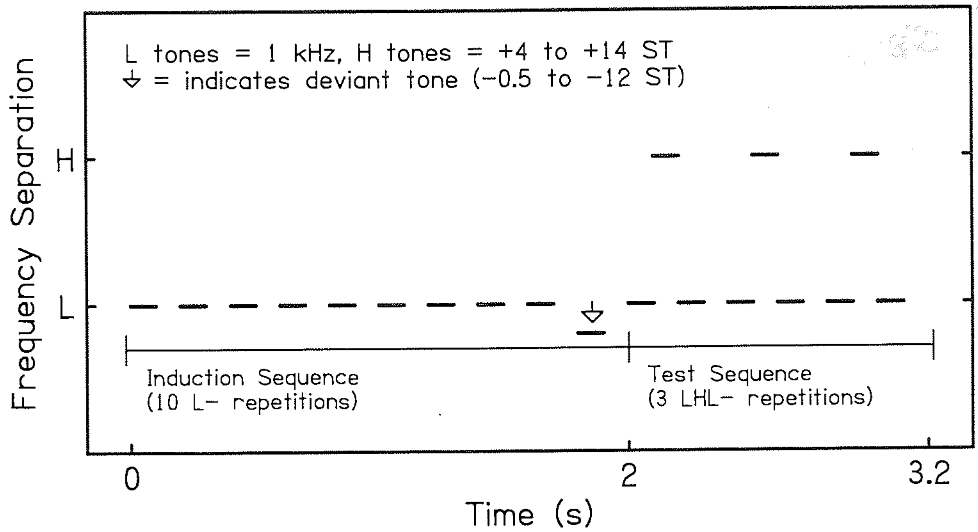

Figure 4.1: **Illustration of the induction conditions tested in experiment 3a. A standard induction sequence is paired with a subsequent test sequence. The arrow denotes the induction tone that was lowered in frequency.**

Procedure and Apparatus

The procedure and apparatus used for the current experiment were fully described in Chapter 2. In summary, trials were organized into blocks; each block comprised 42 trials - a combination of each of the seven induction conditions with each of the six test-sequence frequency separations. For the main experiment, each participant was presented with 20 trial blocks in total (i.e., 840 trials).

4.2.2 Results

The results were averaged to give a single measure of the percentage of trials which were heard as segregated for each combination of induction condition and frequency separation. Responses from all participants were then averaged to give an overall percentage of trials which were heard as segregated. These data are presented in Figure 4.2. The mean percentages reported as segregated for frequency separations of 4, 6, 8, 10, 12, and 14 ST were 7.1%, 22.9%, 71.4%, 80.7%, 85.0%, and 96.4%, respectively. The means for the standard, no-, deviant(-0.5 ST), (-1 ST), (-3 ST), (-6 ST) and (-12 ST) conditions were 60.8%, 30.9%, 52.9%, 53.0%, 52.7%, 50.8%, and 48.5% respectively. Results were analysed using a two-way, repeated-measures ANOVA. The two independent variables were the six frequency separations for the test sequence (4, 6, 8, 10, 12, or 14 ST), and the seven induction conditions (standard induction, no induction, or 5 × deviant-tone induction). The ANOVA confirmed significant main effects of frequency separation [$F(5, 35) = 76.003$, $p < 0.001$, $\eta_p^2 = 0.916$], and of induction condition [$F(6, 42) = 22.047$, $p < 0.001$, $\eta_p^2 = 0.759$]. The interaction between these two variables was also significant [$F(30, 210) = 5.395$, $p < 0.001$, $\eta_p^2 = 0.435$], probably due mainly to floor and ceiling effects at the two extreme frequency separations.

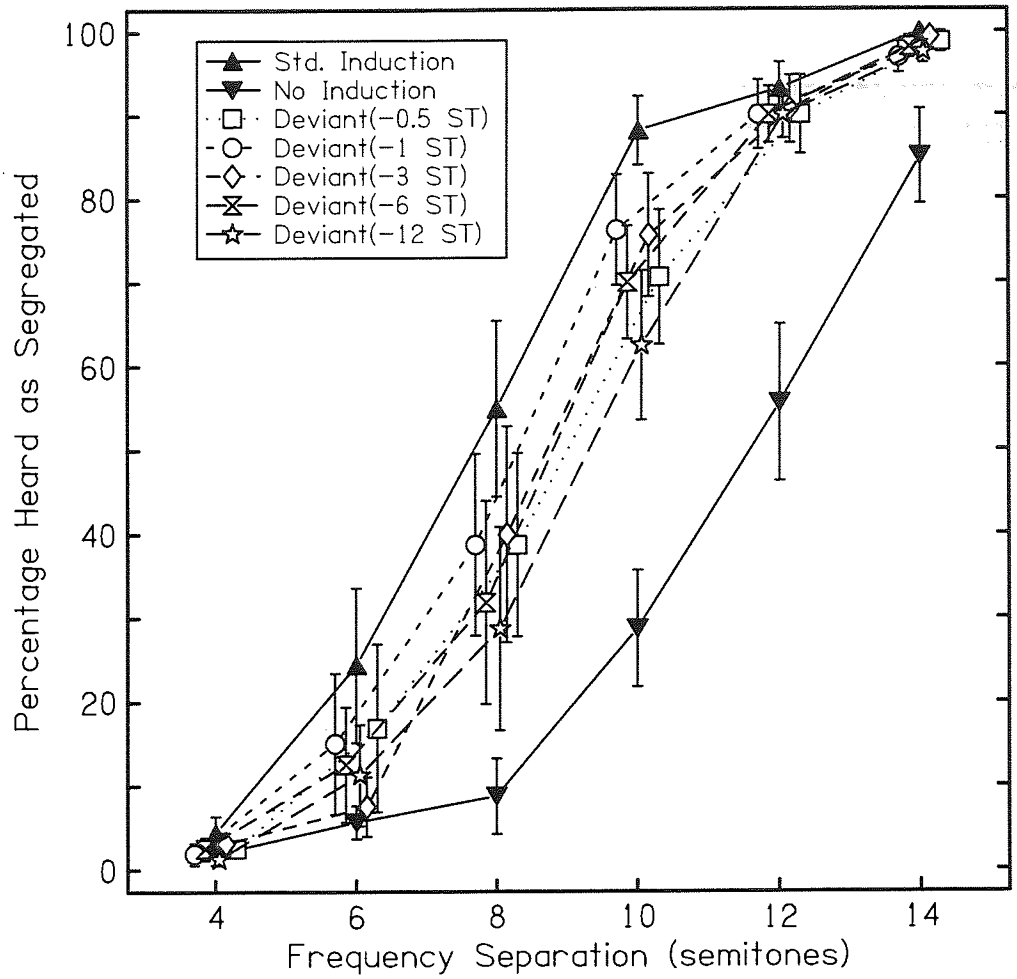

Figure 4.2: **Results from Experiment 3a. The effects of induction condition, and frequency separation in the test sequence on reported segregation. Each point represents the percentage of trials reported as segregated (averaged across 8 listeners). The insert identifies the different induction conditions tested (deviant refers to the change applied to the last tone in an otherwise standard induction sequence). Error bars represent ± 1 inter-subject standard error.**

Overall, the standard induction condition promoted the greatest amount of stream segregation, whilst the least was heard in the no-induction condition (see Figure 4.2). As for experiments 1 and 2, the shift away from the standard induction condition was calculated as an extent of resetting. This extent of resetting was calculated on a proportional scale defined by the standard induction condition representing no change (0.00), and the no-induction condition representing a maximum change (1.00). These values are displayed in Table 4.1. In a manner consistent with Experiment 1, the extent of resetting is not calculated for test sequence frequency separations of 4 and 14 ST, and the mean values are weighted to compensate for the across-frequency-separation differences in the magnitude of the difference between the standard and the no-induction conditions. Conditions in which there was a smaller difference between reported segregation in the standard and in the no-induction conditions made a reduced contribution to the across-frequency-separation mean (for further details see the results section of Experiment 1;

Chapter 3).

Frequency Separation	Deviant Condition					
	-0.5 ST	-1 ST	-3 ST	-6 ST	-12 ST	*Mean*
6 ST *(11.58%)*	0.40	0.50	0.90	0.63	0.70	***0.63***
8 ST *(28.57%)*	0.35	0.35	0.32	0.50	0.57	***0.42***
10 ST *(36.68%)*	0.30	0.20	0.21	0.31	0.43	***0.29***
12 ST *(23.17%)*	0.08	0.08	0.07	0.08	0.08	***0.08***
Weighted Mean	***0.27***	***0.25***	***0.29***	***0.35***	***0.42***	***0.32***

Extent of resetting = standard induction (***0.00***) → no induction (***1.00***)

Table 4.1: **Results from Experiment 3a. Reported segregation in the deviant-tone conditions was compared to that in the standard induction case. Values are expressed as a proportional shift away from the standard induction condition towards the no-induction condition (the *extent of resetting*). As floor and ceiling effects was observed for the 4 and 14 ST conditions, these data were excluded from this analysis. Also, the across-frequency separation means were weighted to reflect the magnitude of the difference between the standard and the no-induction conditions at each test-sequence frequency separation. As in tables 3.1 and 3.2 individual values are displayed as unweighted, as only the across-test-sequence-frequency means are weighted.**

An inspection of Table 4.1 reveals that there was some interaction between the magnitude of deviant-tone resetting and the test-sequence frequency separation. For all deviant tone conditions, the extent of resetting decreased as the test-sequence frequency separation increased - from a large resetting effect at 6 ST (a 0.63 extent of resetting on average) to much less resetting in the 12 ST conditions (extent of resetting = 0.08). For the no-induction condition, there was a strong positive acceleration of reported segregation with increasing test sequence frequency separation - in clear contrast to the negative acceleration observed in all other conditions (see Figure 4.2). This pattern of results was considered in the discussion of Experiment 1, and is likely to have contributed to the currently observed variance in the extent of resetting as a function of test-sequence frequency separation.

Two-tailed pairwise comparisons were performed on the original (i.e., untransformed) using the restricted LSD test (Snedecor & Cochran, 1967). Owing to floor and ceiling effects, the data from the 4 and 14 ST frequency separations were excluded from these pairwise comparisons. There was a significant difference between the standard and the no-induction conditions [difference = 29.9%, $t(7) = 7.561$, $p < 0.001$]. When compared with the standard induction condition, reported segregation was significantly lower in all of the deviant-tone conditions ($p < 0.01$ in all cases). When compared with each other, there were no significant differences between most of the deviant-tone conditions ($p > 0.05$ in most cases). The one exception was the significantly greater segregation observed for the deviant(-12 ST) than for the deviant(-1 ST) condition [4.5%, $t(7) = 2.492$, $p < 0.05$].

In summary, the results show that the standard induction sequence was highly effective at promoting stream segregation in the test sequence, and that the extent of segregation was significantly reduced when the last induction tone was lowered in frequency. These findings are in good accord with those of Experiment 1, as both experiments demonstrated that a frequency change can have a substantial resetting effect (both experiments tested a -3 ST change – see discussion). Despite the fact that only two of the deviant-tone conditions produced changes in reported segregation that were significantly different from each other, there was a clear trend for the extent of resetting to rise as the size of the frequency change imposed on the deviant tone was increased, particularly beyond -3 ST (see Figure 4.2, and also Table 4.1).

4.2.3 Discussion

The results of Experiment 3a are consistent with the known effect of frequency separation on the stream segregation of pure-tone sequences (Miller & Heise, 1950; van Noorden, 1975). The finding that the standard induction sequence promoted the greatest amount of streaming, whilst the least segregation was reported in the no-induction condition is also consistent with earlier studies (Rogers & Bregman, 1993; Experiment 1), and is evidence that the tendency to hear segregation built up over the course of the standard induction sequence. When compared with the standard induction condition, reported segregation was significantly reduced for all the deviant-tone conditions. This active resetting effect of a single deviant tone is consistent with that found in Experiment 1. Indeed, both experiments 1 and 3a tested the same deviant-tone condition, for which the final inducer was lowered in frequency by 3 ST. Although this change had a clear and significant resetting effect in both of these experiments, the resulting proportional extent of resetting was somewhat larger in Experiment 1 (0.38) than in Experiment 3a (0.29 - see tables 3.1 and 4.1, respectively).

Experiment 3a was designed to measure whether resetting varied with the size of the frequency change applied to the deviant tone (changes from -0.5 ST to -12 ST). Although only two of the deviant-tone conditions produced significantly different reported segregation from each other, there was nevertheless a clear trend for a gradual increase in the extent of resetting as the size of the frequency change was increased. The extent of resetting rose from 0.27 for a -0.5-ST frequency change to 0.42 for the -12-ST case (see Table 4.1). The only exception to this trend was the -1 ST condition (extent of resetting = 0.25), in which reported segregation was marginally reduced in comparison to the smaller change in the -0.5 ST case (0.27). These findings are considered further in the general discussion of this chapter.

Irresepective of this issue, these current findings clearly contrast with those reported by Anstis & Saida (1985). These authors observed that substantial resetting only occurred when the centre frequency of an on-going HL sequence was shifted by a value greater than -1 to +3 ST (HL frequency separation = 2 ST, initial centre frequency = 1 kHz). Changes within this frequency range did not have any resetting effect, and all larger changes had a broadly similar resetting effect (range tested = 0 to +/- 12 ST). Any frequency change that fell within the same auditory filter bandwidth as the initial sequence did not have a resetting effect (as 1 $ERB_N \approx 1$ ST, Glasberg & Moore, 1990). The current findings differ from those of Anstis & Saida's (1985), as frequency changes that fell unequivocally within the same auditory filter bandwidth as the on-going sequence *did* have a significant resetting effect (i.e., changes of -0.5 and -1 ST). The reasons for these differences from Anstis & Saida's (1985) results are considered in the general discussion for this chapter. It is interesting to note that even the smallest change tested (-0.5 ST) produced a reasonably large extent of resetting (0.27). This could be evidence that *any* noticeable frequency change can reset build-up, even if the change is only very slight. Experiment 3b was designed to test this idea further by examining whether even very small frequency changes can have a discernible resetting effect.

4.3 Experiment 3b

4.3.1 Method

Experiment 3b used the same subjective procedure as for Experiment 3a. On each trial, listeners were required to indicate their perception of the final LHL- triplet of a test sequence as either integrated or segregated (test sequence = 3 LHL- triplets). For the majority of trials, the test sequence was preceded by an induction sequence, the properties of which were varied systematically. The full details of this method are described in Chapter 2

Participants

Eight listeners took part, four of whom had previously taken part in Experiment 3a. All reported normal hearing.

Stimuli and Conditions

Both the standard and no-induction conditions were used in this experiment (see Chapter 2). The standard induction sequence comprised 10 L- repetitions (L tone = 1 kHz, tone duration = 100 ms, inter-tone silence = 100 ms). Each of the four experimental induction conditions was a modification of the standard induction sequence. As with Experiment 3a, the final induction tone was lowered in frequency for the experimental conditions, but otherwise identical to a standard L-tone (see Figure 4.1 for an illustration of a similar condition). Relatively small frequency changes were tested, and so the values are expressed in Hz rather than in semitones. Specifically, the last tone was set either to 997.5, 995, 990, or 980 Hz. Note that all of these frequency changes are smaller than those previously tested in Experiment 3a, where the smallest change used (-0.5 ST) corresponded to a frequency of 972 Hz. Therefore, all of the changes tested in the current experiment well fell within 1 ERB_N of the standard 1 kHz tone (Glasberg & Moore, 1990).

Procedure and Apparatus

The procedure and apparatus used for the current experiment were fully described in Chapter 2. Each trial block comprised a combination of each of the six induction conditions with each of the six frequency separations for the test sequence (i.e., 36 trials). For the main experiment, each participant was presented with 20 trial blocks (i.e., 720 trials in total).

4.3.2 Results

Responses for each combination of induction condition and test-sequence frequency separation were averaged to give an overall indication of streaming for each separate condition (i.e., the percentage heard as segregated). These mean values are shown in Figure 4.3. The mean percentages reported as segregated for frequency separations of 4, 6, 8, 10, 12, and 14 ST were 10.2%, 19.5%, 46.8%, 79.9%, 89.2%, and 97.3%, respectively. The means for the standard, no-, deviant(997.5 Hz), (995 Hz), (990 Hz) and (980 Hz) conditions were 65.0%, 30.9%, 64.3%, 64.2%, 60.7%, and 57.7% respectively. The data were analysed using a two-way, repeated-measures ANOVA. The two independent variables were the six frequency separations for the test sequence (0, 4, 6, 8, 10, 12, or 14 ST), and the six induction conditions (standard induction, no induction, or 4 × deviant-tone induction). This analysis confirmed a significant main effect of frequency separation [$F(5, 35) = 46.152$, $p < 0.001$, $\eta_p^2 = 0.868$], and of induction condition [$F(5, 35) = 15.816$, $p < 0.001$, $\eta_p^2 = 0.693$]. Once again, the interaction between these two variables was also significant [$F(25, 175) = 8.534$, $p < 0.001$, $\eta_p^2 = 0.549$].

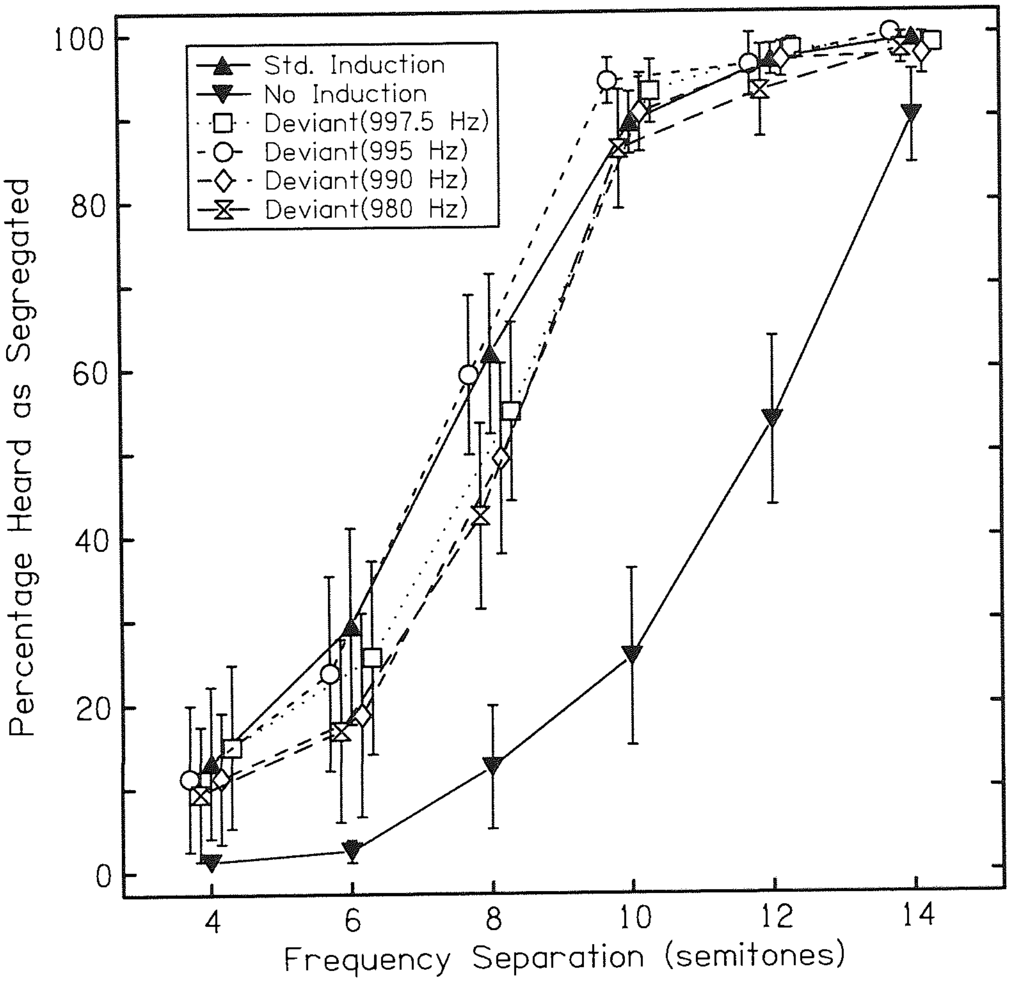

Figure 4.3: **Results from Experiment 3b. The effects of induction condition, and frequency separation in the test sequence on reported segregation. Each point represents the percentage of trials reported as segregated (averaged across 8 listeners). The insert identifies the different induction conditions tested (deviant refers to the change applied to the last tone in an otherwise standard induction sequence). Error bars represent ±1 inter-subject standard error.**

On average, the standard induction condition promoted the greatest amount of stream segregation, whilst the least was heard in the no-induction condition (see Figure 4.3). As for the earlier experiments, the resetting effect of the deviant-tone conditions was calculated as a proportional extent of resetting from the standard induction condition towards the no-induction case (i.e., where standard induction represents 0.00 and the no-induction condition represents 1.00 on the scale). These values are displayed in Table 4.2. For this table, the extent of resetting is not calculated for test sequence frequency separations of 4 and 14 ST. Also, the mean values are weighted to reflect the magnitude of the difference between the standard and the no-induction conditions at the various test-sequence frequency separations (see Experiment 1).

Frequency Separation	Deviant Condition: 997.5 Hz	995 Hz	990 Hz	980 Hz	*Mean*
6 ST *(14.68%)*	0.14	0.21	0.40	0.47	***0.30***
8 ST *(26.96%)*	0.14	0.05	0.25	0.39	***0.21***
10 ST *(34.81%)*	-0.06	-0.08	-0.02	0.05	***-0.03***
12 ST *(23.55%)*	-0.03	0.02	0.00	0.09	***0.02***
Weighted Mean	***0.03***	***0.02***	***0.12***	***0.21***	***0.10***

Extent of resetting = standard induction (***0.00***) → no induction (***1.00***)

Table 4.2: **Results from Experiment 3b. Reported segregation in the deviant conditions are expressed as a proportional shift away from the standard induction condition towards the no-induction condition (the *extent of resetting*). See the main text, or the caption for Table 4.1 for more details on this measure. Note that a (small) negative value represents a case where there is a shift *away* from the no-induction sequence (i.e., reported segregation was greater than for the standard induction condition). Note that *only* the across-test-sequence means are weighted.**

Two-tailed pairwise comparisons were performed on the original (i.e., untransformed) data set using the restricted LSD test (Snedecor & Cochran, 1967). As before, the data from the 4 and 14 ST frequency separations were excluded from these pairwise comparisons. There was a significant difference between the standard and the no-induction conditions [difference in reported segregation = 31.4%, $t(7) = 2.773$, $p < 0.001$]. Compared with the standard induction condition, reported stream segregation was significantly lower only in the deviant(980 Hz) condition [7.3%, $t(7) = 2.773$, $p < 0.05$], and in the deviant(990 Hz) condition [4.3%, $t(7) = 5.286$, $p < 0.05$]. For the smaller changes in frequency imposed on the deviant tone, the reduction in segregation compared with the standard induction condition was not significant [995 Hz: 0.8%, $t(7) = 0.552$, $p > 0.05$]; [997.5 Hz: 0.7%, $t(7) = 0.973$, $p > 0.05$].

4.3.3 Discussion

The general pattern of the results from Experiment 3b for increasing HL frequency separation was again consistent with previous studies (e.g., Miller & Heise, 1950; van Noorden, 1975). As for the previous subjective experiments, the standard induction sequence promoted a strong perception of stream segregation in comparison with the no-induction case. This finding is taken as clear evidence that the tendency to report segregation built up over the course of the standard induction sequence.

For the deviant-tone conditions, only changes of 10 or 20 Hz (i.e., deviant tone = 990 or 980 Hz) had a significant resetting effect, and this effect was quite small (in the range 0.12 - 0.21; see Table 4.2). The smaller changes in frequency (deviant tone = 995 or 997.5 Hz) did not produce a significant resetting effect. It should be noted that this may be due to listeners not detecting these small frequency changes. Several studies have measured frequency discrimination thresholds for pure tones; for a tone presented at a frequency of 1 kHz (at 60 - 70 dB SPL), the smallest noticeable difference is approximately 2 to 3 Hz (Moore, 2003; see also Sek & Moore, 1995). Considering that for the current experiment, participants were not required actively to detect the deviant tone, it seems very likely that the smallest frequency changes used (2.5 and 5 Hz) may simply not have been noticed by the listeners. If so, the current finding suggests that *any noticeable frequency change applied to the final induction tone can partially reset the build-up of stream segregation.* This idea is considered further below.

4.4 General Discussion

Experiments 3a and 3b have both provided evidence consistent with the notion that a noticeable change in pitch on the final induction tone can significantly, and in some cases substantially, reset the build-up of stream segregation. Experiment 3b measured the resetting effect of small frequency changes, and found that changes of only 10 or 20 Hz from 1 kHz (i.e., 1% or 2%) had a modest, but significant resetting effect. Smaller frequency changes (2.5, or 5 Hz) did not trigger any discernible resetting - this was probably due to these changes being undetected by the listener (Moore, 2003; Sek & Moore, 1995). For Experiment 3a, a range of larger frequency changes (range; -0.5 to -12 ST from 1 kHz) all had a significant and often large resetting effect. The majority of these frequency changes were not significantly different from each other, although there was a trend for the largest frequency changes (-6 and -12 ST) to promote the greatest extent of resetting (see Table 4.1). Indeed, if we examine the combined results of experiments 3a and 3b, there is evidence for a general increase in the extent of resetting as the frequency of the deviant tone was progressively lowered. These results are reproduced in Figure 4.4, where each data point represents the mean proportional resetting effect of a deviant tone for each separate size of frequency change. Figure 4.4 shows that for frequency changes above 10 Hz (Experiment 3b), the extent of resetting generally increased as a function of the (log) frequency change applied to the deviant tone.

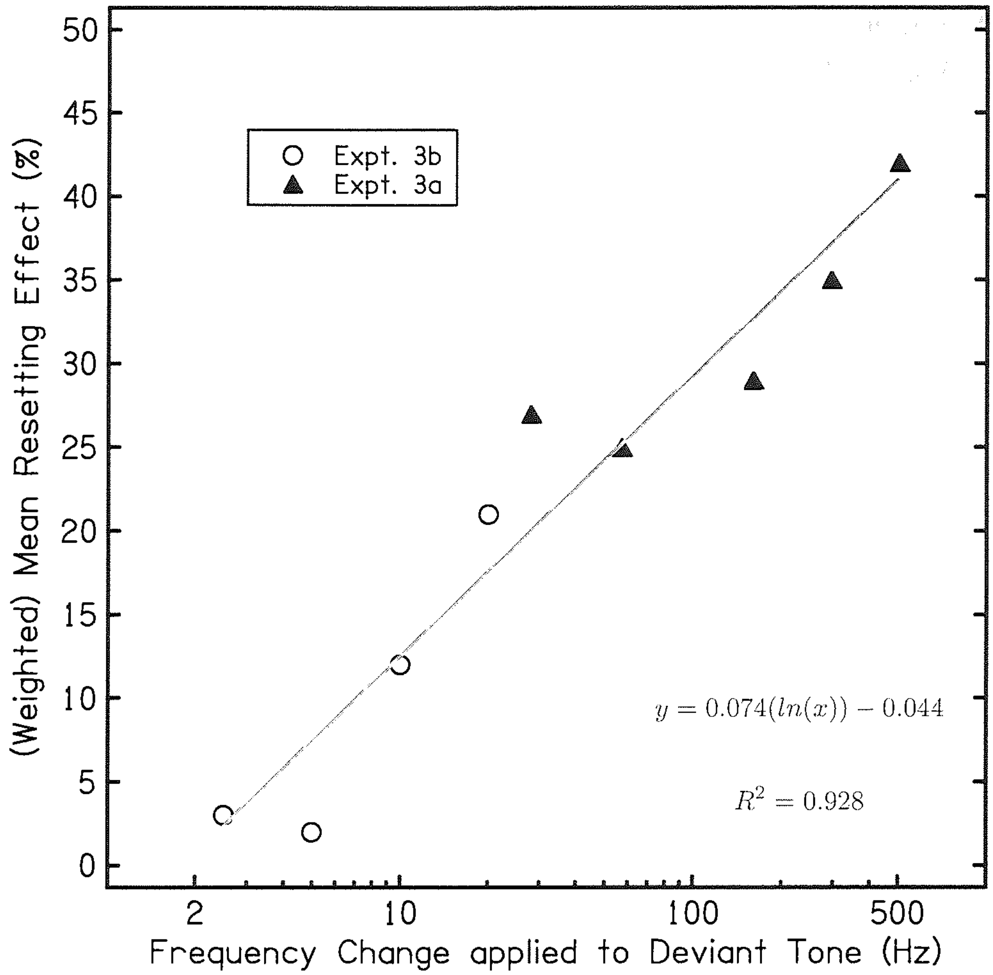

Figure 4.4: **Summary of the results from Experiment 3a and 3b (see insert). The mean resetting effect of a deviant tone (the extent of resetting) is shown as a function of the size of the frequency change for the deviant tone. Data are reproduced from Tables 4.1 & 4.2. Each data point represents the weighted extent of resetting across test-sequence frequency separations of 6, 8, 10 & 12 ST (data from 4 & 14 ST are excluded, see text). A linear regression line is fitted to illustrate the relationship between the (log) frequency change and the extent of resetting.**

The results of experiments 3a and 3b contrast with those of Anstis & Saida (1985) in two ways. Firstly, for the current studies, frequency changes which fell *within* the same auditory filter bandwidth as the on-going induction sequence could partly reset the build-up of stream segregation (standard L-tone = 1 kHz, 1 $ERB_N \approx 1$ ST, Glasberg & Moore, 1990). Secondly, the extent of resetting mainly rose in a linear fashion as the size of the frequency change was increased on a log scale (see Figure 4.4). In comparison, Anstis & Saida (1985) found that changes in centre frequency which fell within roughly the same filter bandwidth as the initial sequence did not appear to trigger any resetting, and larger changes all resulted in a broadly similar (and large) resetting effect.

There are three possible explanations for these differences in results. First, in the current studies, only a single deviant L-tone was changed in frequency whereas Anstis

& Saida (1985) shifted the frequency of the entire test sequence away from that of the induction sequence. Given that none of the induction tones matched the frequencies of the test-sequence tones, this could mean that Anstis & Saida (1985) observed a more complete resetting effect (see also Chapter 6). However, because these authors did not measure test-sequence streaming in the absence of any prior build-up, it is difficult to calculate the actual extent of the (seemingly large) resetting effect which they observed. A second reason why the two sets of results differ could be that Anstis & Saida (1985) measured streaming by asking listeners to adjust the presentation rate of the test sequence, whilst the rate of the induction sequence was kept constant. The resulting change in rhythm at the induction/test sequence boundary may have in itself have affected subsequent perception, as a change in temporal properties such as tone duration can have a substantial resetting effect (Roberts et al., 2008; see also Experiment 1). Finally, the current studies promoted test-sequence segregation by presenting a same-frequency induction sequence (i.e., 10 L- repetitions), whereas the induction sequences used by Anstis & Saida (1985) comprised an alternating-frequency tone sequence (i.e., a 4-s repeating HL arrangement). Although both of these types of induction sequence are effective at promoting subsequent stream segregation (cf. Rogers & Bregman, 1993, 1998), it remains unclear exactly how alike these two processes are. This issue is considered further in chapters 5 and 8.

The current experiments were primarily undertaken to test Rogers & Bregman's (1998) hypothesis that resetting occurs only when a new sound source is perceived. If this theory were correct, one would expect only relatively large frequency changes on the deviant tone to have an appreciable resetting effect. This is because small frequency changes can commonly occur in a natural sound source, whereas abrupt large changes are more likely to originate from a separate acoustic event (Bregman, 1990). Whilst this general trend was observed in experiments 3a and 3b - the largest frequency changes did appear to have an increased resetting effect - the current results are not entirely consistent with Rogers & Bregman's (1998) hypothesis. This is because the increase in resetting with the size of the frequency change was relatively gradual, and very small (but noticeable) frequency changes did have some resetting effect. If resetting only occurs when a new source is detected, one would not expect this pattern of results.

From the current experiments, it is quite apparent that the resetting process must be independent of the process of stream formation. van Noorden (1975) demonstrated that, below a given tone frequency separation, an on-going LHL- sequence is always perceived as integrated (the fission boundary, see Figure 1.2). Generally, sequences where the tones are separated by less than 3 ST should exclusively be heard as integrated (although the fission boundary does vary slightly with the presentation rate of the sequence). However, for the current experiments, frequency changes which fell far below the fission boundary have been shown to have an appreciable, and sometimes substantial, resetting effect. If resetting only occurs when there is sufficient evidence to form a new stream, then any frequency change below the fission boundary should not have a resetting effect. Instead, it would appear that resetting is not directly influenced by the likelihood of the deviant tone being heard as segregated from the induction tone stream.

Roberts et al. (2008) also noted that the resetting effect of a sudden change in a particular tonal property appears to be relatively unrelated to the ability of that property to establish stream segregation in the first place. For example, abrupt changes in perceived location arising from ITD cues at the induction/test sequence boundary can have a large resetting effect (Rogers and Bregman, 1998, Roberts et al, 2008). However, there is evidence indicating that ITD cues are a relatively weak cue for stream segregation. For example, Boehnke and Phillips (2005) measured the stream segregation of noise bursts in an on-going left-right-left-right sequence. These authors used two measures of streaming - a subjective report of perception (the proportion of time for which the sequence was heard as segregated, cf. Anstis & Saida, 1985), and a measure of temporal discrimination (cf. Roberts et al., 2002). The noise bursts were lateralized by presentation to different ears, or through ILD (15 dB), or ITD (0.5 ms) cues. In both tasks, differences in ear-of-presentation were an effective cue for segregation, as were ILD cues. In comparison, ITD cues failed to promote stream segregation - performance/reported segregation in these conditions was largely the same as that when the noise bursts were presented diotically.

In conclusion, the current experiments have demonstrated that any noticeable frequency change which is applied to the final induction tone can partially reset the build-up of stream segregation. The size of this frequency change does have a slight influence on the extent of resetting - there is some evidence non-significant trend for larger frequency changes to have a greater resetting effect. This evidence is partially consistent with Rogers & Bregman's (1998) hypothesis that resetting may be triggered when a new acoustic event is perceived. However, because very small changes in frequency (10 or 20 Hz) result in partial resetting, it would seem unlikely that these changes would be perceived as a separate (new) event. Instead, it would appear that any noticeable change to a sequence can trigger some resetting.

Chapter 5

The Effect of the Serial Position of a Deviant Tone on Resetting

5.1 Introduction

So far, the experiments presented in this thesis have demonstrated that a salient deviant tone can substantially reset the build-up of stream segregation. This effect has been shown in experiments 1 - 3b, where build-up was promoted in a repeating, same-frequency induction sequence and then measured in a subsequent test sequence. For Experiment 1, stream segregation generally was reduced substantially when the final induction tone was noticeably altered (in frequency, duration, or level, or when the final tone was replaced with a silence). For these conditions, the mean extent of resetting was 0.45 (see Table 3.1). So far, this finding has been interpreted as evidence that a salient, abrupt change can actively reset the build-up of stream segregation. However, there are two alternative, but related, explanations for the observed decrease in segregation following a deviant tone that merit consideration. The build-up of stream segregation is relatively gradual (Bregman, 1978, Anstis & Saida, 1985), and Rogers & Bregman (1998) suggested that this process may be quite specific to the tone sequence which is currently heard. If so, build-up may fail to transfer between sequences that comprise different tonal characteristics (e.g., in frequency, level, or perceived location, see Rogers & Bregman, 1998, Roberts et al., 2008). This hypothesis is of relevance to the previous studies of this thesis for two reasons. First, if the deviant tone does not contribute to the build-up occurring during the induction sequence, then the loss (real or effective) of the 10^{th} induction tone increases the time interval between the end of the induction sequence and the first L-tone of the test sequence. Any increase in the interval between the induction and test sequence should also result in some decay of build-up (Bregman, 1978; Beauvois & Meddis, 1997). However, the decay of segregation typically occurs over several seconds, and the loss of the final inducer would have only resulted in a 300-

ms interval between the 9^{th}inducer and the first test tone. Therefore, the large extent of resetting observed following a single deviant tone in experiments 1 - 3b cannot be fully explained in terms of a gradual decay of segregation (see the general discussion of Chapter 3).

The second explanation relates to this issue, but is somewhat more difficult to dismiss. If build-up does not transfer between tone sequences of different characteristics, a single deviant tone would be unlikely to contribute to the cumulative build-up caused by the other tones during an induction sequence. Therefore, reduced segregation for the deviant-tone conditions of experiments 1 - 3b may simply reflect the fact that less build-up occurred in these induction sequences (as opposed to increased decay of build-up or to the deviant tone having an active resetting effect).

Experiments 1 and 2 provide a clear example of this argument, as both included a condition in which the final induction tone was replaced with an equivalent-duration silent interval. This actually reduced the number of induction tones from ten to nine. As the number of tones in the sequence was decreased, one might predict that less build-up would occur. By extension, any single deviant tone would effectively reduce the number of identical induction tones from ten to nine. However, as build-up is most rapid during the *initial* portion of a tone sequence (Bregman, 1978, Anstis & Saida, 1985), one would not expect such a large decrease in segregation simply if the *final* inducer in a set of ten did not contribute to build-up. Despite this, the hypothesis cannot be fully dismissed, as these previous studies have only measured the rate of build-up occurring in a repeating LH type sequence (Bregman, 1978, Anstis & Saida, 1985). In contrast, experiments 1 - 3b all measured build-up and resetting in a repeating, same-frequency induction sequence (i.e., 10 L- repetitions). Whilst such a sequence can have a strong segregation-promoting effect, the rate of build-up has never been directly studied for this type of configuration[1], and it may not necessarily be equivalent to that for a sequence of tones which alternates in frequency. The two experiments presented in this chapter were designed to address these issues.

1

Rogers & Bregman (1993) manipulated the tone rate (the number of tone onsets) in a same-frequency induction sequence of constant overall duration. These authors found an H-tone induction sequence which matched the tone rate of the test-sequence H-tones was highly effective at promoting test-sequence segregation. Slower induction tone rates were less effective at promoting streaming, and faster tone rates were no more effective than the matched rate. Whilst this study may offer some insight into the rate of build-up in a same-frequency induction sequence, for most induction conditions there was also a noticeable change in H-tone density (and therefore rhythm) at the induction/test sequence boundary. This abrupt change may have also affected streaming by having a resetting effect (Roberts et el., 2008; see also Jones et al., 1981).

Experiment 4a presented induction sequences in which a single tone was replaced with a silent interval of equivalent-duration (which is referred to as a "silent tone" for convenience). Experiments 1 & 2 demonstrated that when this change was applied to the last induction tone, stream segregation was substantially reduced (see Chapter 3). For Experiment 4a, the serial position of a silent tone within an induction sequence was varied. If the decrease in test-sequence segregation was simply due to the presence of fewer tones, a similar decrease should be apparent irrespective of the position of the silent tone. Alternatively, if the silent tone reduced stream segregation by actively resetting build-up, the serial position of the silent tone should strongly affect how the test sequence is perceived. More specifically, one would expect resetting to be most evident when the silent tone occurs near to the onset of the test sequence. A silent tone that occurs earlier in the induction sequence is less likely to have such an apparent resetting effect, as build-up should re-occur owing to the subsequent induction tones.

5.2 Experiment 4a

5.2.1 Method

The method used in Experiment 4a was identical to that described in Chapter 2. In summary, on each trial listeners were presented with a combination of an induction sequence and a subsequent test sequence (test sequence = 3 LHL- triplets, L tone = 1 kHz, H tone = +4 to +14 ST, tone duration = 100 ms, inter-triplet silence = 100 ms). In a 2AFC procedure, listeners were required to indicate their perception of the final LHL- triplet as either integrated or segregated.

Participants

Eight listeners took part in this experiment, all of whom reported normal hearing. Four of these participants had also taken part in at least one of the previous subjective experiments.

Induction Conditions

Both the standard induction condition and the no-induction condition were used in this experiment (these are described in detail in Chapter 2). The standard induction sequence comprised 10 L- repetitions (L tone = 1 kHz, tone duration = 100 ms, inter-tone silence = 100 ms). All three experimental conditions were a modification of this induction sequence. For these conditions, a single tone was replaced with a silent interval of equivalent duration. This so-called "silent tone" could replace either the 4^{th}, 7^{th} or 10^{th} (i.e., the last) tone of the induction sequence (see Figure 5.1). For each of these conditions, the properties of all the other induction tones were identical to those of the standard induction case - therefore a silent tone resulted in a 300-ms silence between the two neighbouring L tones. Note that the condition in which the 10^{th} tone was replaced with silence was identical to a condition tested in Experiment 1 (and it was also very similar to a condition tested in the temporal discrimination task used in Experiment 2).

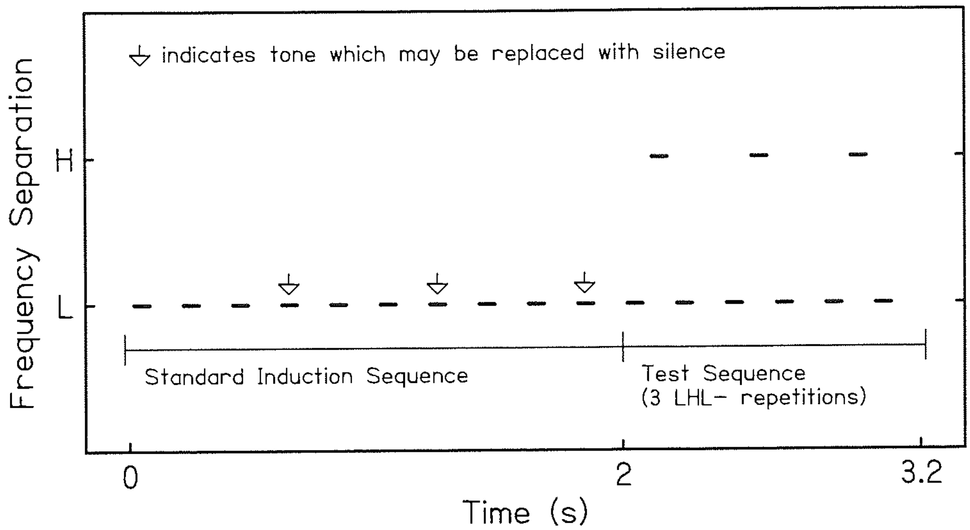

Figure 5.1: **Illustration of the induction conditions tested in Experiment 4a. A standard induction sequence is paired with a subsequent test sequence. The three arrows indicate the set of induction tones from which one tone was selected and replaced with a silent interval of equivalent duration.**

Procedure and Apparatus

The procedure and apparatus used were identical to that described in Chapter 2. The stimuli were organized into trial blocks. Each block comprised a combination of each of the five induction conditions with each of the six frequency separations in the test sequence (i.e., 30 trials per block). For the main experiment, each participant was presented with 20 trial blocks in total (i.e., 600 trials).

5.2.2 Results

For each listener, responses to each condition at each separate frequency separation were averaged to yield a measure of reported segregation. Responses across all participants were then averaged to give an overall indication of streaming for each combination of induction sequence and frequency separation (i.e., the percentage heard as segregated across listeners). These mean values are shown in Figure 5.2. The mean percentage reported as segregated for frequency separations of 4, 6, 8, 10, 12, and 14 ST were 1.3%, 10.4%, 30.3%, 57.3%, 63.9%, and 88.5%, respectively. The means for the standard, no-, 4^{th}-silent, 7^{th}-silent, and 10^{th}-silent induction conditions were 51.0%, 23.0%, 51.1%, 50.2%, and 34.2%, respectively. Data were analysed using a two-way, repeated-measures

ANOVA. The two independent variables were the six frequency separations for the test sequence (4, 6, 8, 10, 12, or 14 ST), and the five induction conditions (standard induction, no induction, or 3 × deviant induction). The ANOVA confirmed a significant main effect of both frequency separation [$F(5, 35) = 49.071$, $p < 0.001$, $\eta_p^2 = 0.875$] and induction condition [$F(4, 28) = 37.401$, $p < 0.001$, $\eta_p^2 = 0.842$]. There was also a significant interaction between these two variables [$F(20, 140) = 9.753$, $p < 0.001$, $\eta_p^2 = 0.582$]. This interaction is likely to be due in part to floor and ceiling effects at the 4 ST and the 14 ST frequency separations (respectively), but also due to the overall positive acceleration of reported segregation with increasing frequency separation observed for the no-induction and 10th-silent conditions (in comparison to the more negative acceleration observed for the other induction conditions).

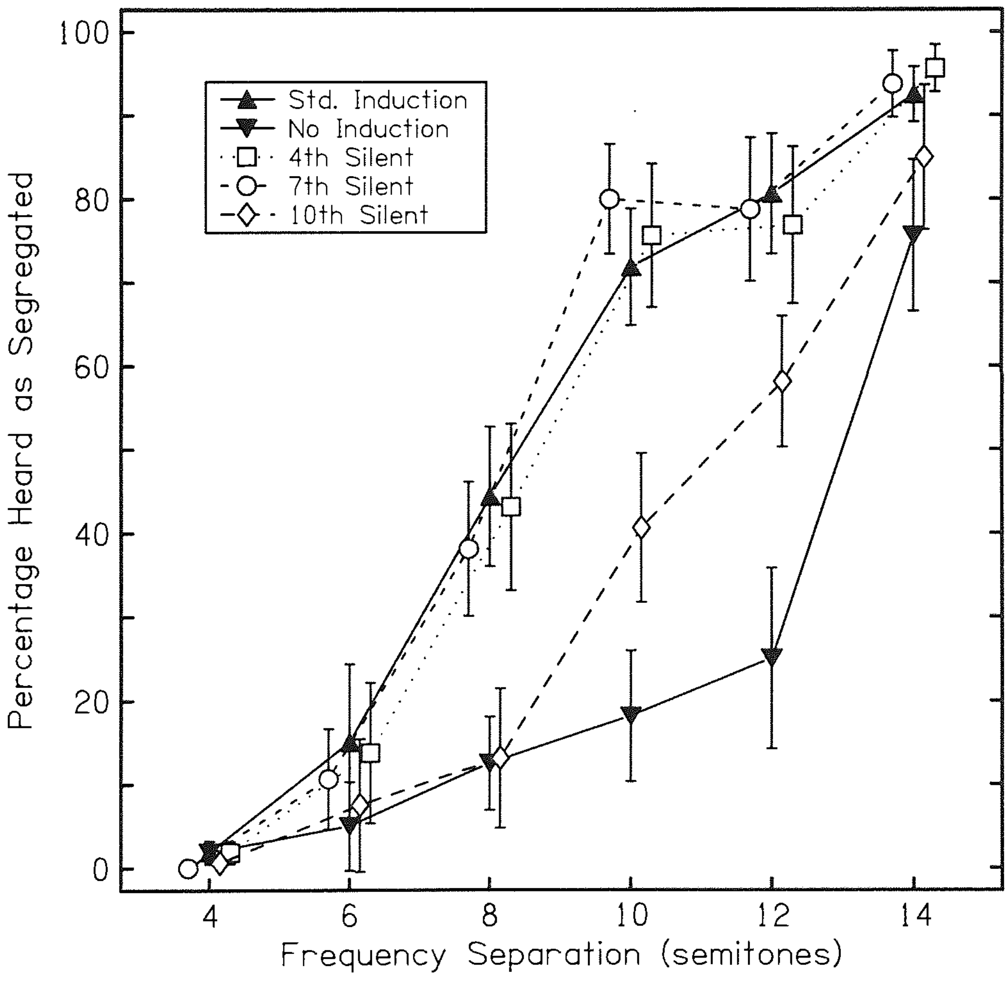

Figure 5.2: **Results from Experiment 4a. The effects of induction condition and frequency separation in the test sequence on reported segregation. Each point represents the percentage of trials which were reported as segregated (averaged across 8 listeners). The insert identifies the different induction conditions tested. The error bars display ±1 inter-subject standard error.**

As shown in Figure 5.2, the standard induction, the 4th-silent, and the 7th-silent conditions all promoted a strong tendency for listeners to report stream segregation. For the 10^{th}-silent condition, reported segregation was decreased relative to the standard induction case towards that for the no-induction case. As for the earlier subjective experiments, the shift away from the standard induction condition associated with the experimental conditions was calculated as an extent of resetting. Data from the 4 ST and 14 ST frequency separation conditions were excluded from this analysis, and a weighting system was applied to the calculation of the across-test-sequence-frequency mean values (see Chapter 3). This weighting system was devised so that frequency separations at which there was a smaller difference between reported segregation in the standard and in the no-induction conditions made a reduced contribution to the mean. This procedure is described in detail in the results section of Experiment 1. These values are displayed in Table 5.1. From both Table 5.1 and Figure 5.2, it is clear that resetting was only apparent when the 10^{th} (i.e., final) induction tone was replaced with silence.

Frequency Separation	Deviant Condition: 4th silent	7th silent	10th silent	***Mean***
6 ST *(6.61%)*	0.13	0.44	0.75	***0.44***
8 ST *(21.07%)*	0.04	0.20	0.98	***0.41***
10 ST *(35.54%)*	-0.07	-0.15	0.58	***0.12***
12 ST *(36.78%)*	0.07	0.03	0.40	***0.17***
Weighted Mean	***0.02***	***0.03***	***0.61***	***0.22***

Extent of resetting = standard induction (***0.00***) → no induction (***1.00***)

Table 5.1: Results from Experiment 4a. Reported segregation in the deviant conditions was compared with that in the standard induction case. Values are expressed as a proportional shift away from the standard induction condition towards the no induction case (the *extent of resetting*). In order to provide a stable estimate of resetting, the data from the 4 ST and 14 ST test-sequence frequency separation conditions are excluded from this analysis. As described in the main text, a weighting system was applied to the calculation of across-frequency means, but each individual value represents the extent of resetting prior to the application of any weight. A negative value indicates a shift *away* from the no-induction case (i.e., reported segregation was greater than for the standard induction condition).

Two-tailed pairwise comparisons (restricted LSD test, Snedecor & Cochran, 1967) were conducted on the original, untransformed data set to compare reported stream segregation for each of the silent conditions with that for the standard induction case. These comparisons are also reported with the relevant percentage difference in reported segregation. 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 significant difference in reported segregation between the standard and the no-induction conditions [difference in reported segregation = 28.0%, $t(7) = 9.170$, $p < 0.001$]. The tests confirmed that there was no significant difference between the standard induction and the 4th-silent conditions [0.1%, $t(7) = 0.186$, $p > 0.05$], or between the standard induction and the 7th-silent conditions [0.8%, $t(7) = 0.340$, $p > 0.05$]. However, there was a significant difference between the standard induction and the 10th-silent conditions [16.8%, $t(7) = 4.989$, $p < 0.001$]. Indeed, reported segregation for the 10th-silent condition was significantly different from that for all the other induction conditions tested ($p < 0.005$ in all cases).

In summary, these results show that the standard induction sequence was effective at promoting stream segregation. Furthermore, an equivalent amount of segregation was heard following the 4th-silent and 7th-silent induction conditions. Reported segregation was reduced only when the 10th (the last) induction tone was replaced with a silence. In this case, the extent of resetting was approximately two thirds.

5.2.3 Discussion

Listeners' responses were consistent with the known effect of frequency separation on stream segregation for sequences of pure tones (e.g., van Noorden, 1975). As in all of the previous experiments, the standard induction sequence had a strong segregation-promoting effect, and the least streaming was observed for the no-induction condition. When the final induction tone was replaced with a silence of equivalent duration, reported segregation was significantly reduced when compared with the standard induction case. This finding is consistent with the results of experiments 1 and 2, where a broadly equivalent resetting effect was observed when the final induction tone was replaced by silence. Importantly, when either the 4th or the 7th induction tone was replaced with silence, reported segregation did not differ from the standard induction case.

This finding demonstrates that the observable resetting effect of a silent tone is highly dependent on its position within an induction sequence. Each of the three experimental conditions comprised nine induction tones, whilst the standard induction sequence comprised ten tones. If reduced stream segregation in the nine-tone experimental conditions was merely a consequence of reduced overall build-up, we would expect a similar reduction for all of these conditions. Instead, reduced segregation was only observed when the last induction tone was replaced with a silence. This indicates that the observed reduction in segregation is not due merely to the presence of fewer induction tones. As the decrease in segregation following the silent tone is too large to be explained as a gradual decay of segregation (Bregman, 1978; Beauvois & Meddis, 1997 – see Chapter 3), it is concluded that the silent tone at the inducer/test boundary must act to reset build-up. Furthermore, experiments 1 – 3b also measured resetting for cases where the final inducer was altered on some acoustic dimension (as opposed to being replaced by silence). The current findings can also be generalized to these conditions - the reduced stream segregation cannot be attributed simply to the deviant tone failing to contribute to build-up. Instead, the current results support the hypothesis that a noticeable deviant tone must actively reset the build-up of stream segregation.

For Experiment 4a, there was *no* evidence of resetting when either the 4^{th} or the 7^{th} induction tone was replaced with silence. One possible explanation is that some resetting takes place in those cases but is obscured by build-up re-occurring during the subsequent induction tones. If so, this recovery in build-up must occur very rapidly during the same-frequency induction sequence (indeed, the recovery of build-up following resetting may be more rapid than the original build-up itself). One would not expect such a rapid rate of build-up in an alternating-frequency tone sequence (e.g., a repeating LH sequence, Bregman, 1978, Anstis & Saida, 1985). This suggests either that the rate of build-up for an L-tone-only sequence is more rapid than that for an alternating-frequency sequence, or that the recovery of build-up after a single perturbation is more rapid than original build-up. The rate of build-up in a repeating, same-frequency induction sequence has never been directly measured, and so this issue was investigated in Experiment 4b.

5.3 Experiment 4b

5.3.1 Method

As for Experiment 4a, the method used in Experiment 4b was identical to that described in Chapter 2. In summary, on each trial listeners judged the final LHL- triplet of a test sequence as either integrated or segregated (test sequence = 3 LHL- triplets, L tone = 1 kHz, H tone = +4 to +14 ST, tone duration = 100 ms, inter-triplet silence = 100 ms). For the majority of trials, the test sequence was preceded by a segregation-promoting induction sequence, the properties of which were varied as described below.

Participants

Eight listeners took part in Experiment 4b, all of whom reported normal hearing. None of the participants had taken part in any of the previous experiments presented in this thesis.

Stimuli and Conditions

This experiment used both of the control conditions that were described in Chapter 2 (the standard induction and no-induction conditions). The three experimental induction sequences differed from the standard case only in that each contained fewer L- tones. Specifically, these induction sequences contained either 6, 3, or 1 L- tone(s). As each L-tone cycle lasted 200 ms, the overall duration of these tone sequences differed (standard induction = 2000 ms, 6 inducers = 1200 ms, 3 inducers = 600 ms, 1 inducer = 200 ms). However, to keep the overall length of each induction sequence constant at 2000 ms, the initial portion of each experimental induction condition was filled with continuous band-pass filtered noise. The noise continued until 100 ms before the onset of the first L tone (i.e., for the 6, 3, and 1 inducer cases, the duration of the noise was 700, 1300, and 1700 ms, respectively). An illustration of the induction conditions is provided in Figure 5.3.

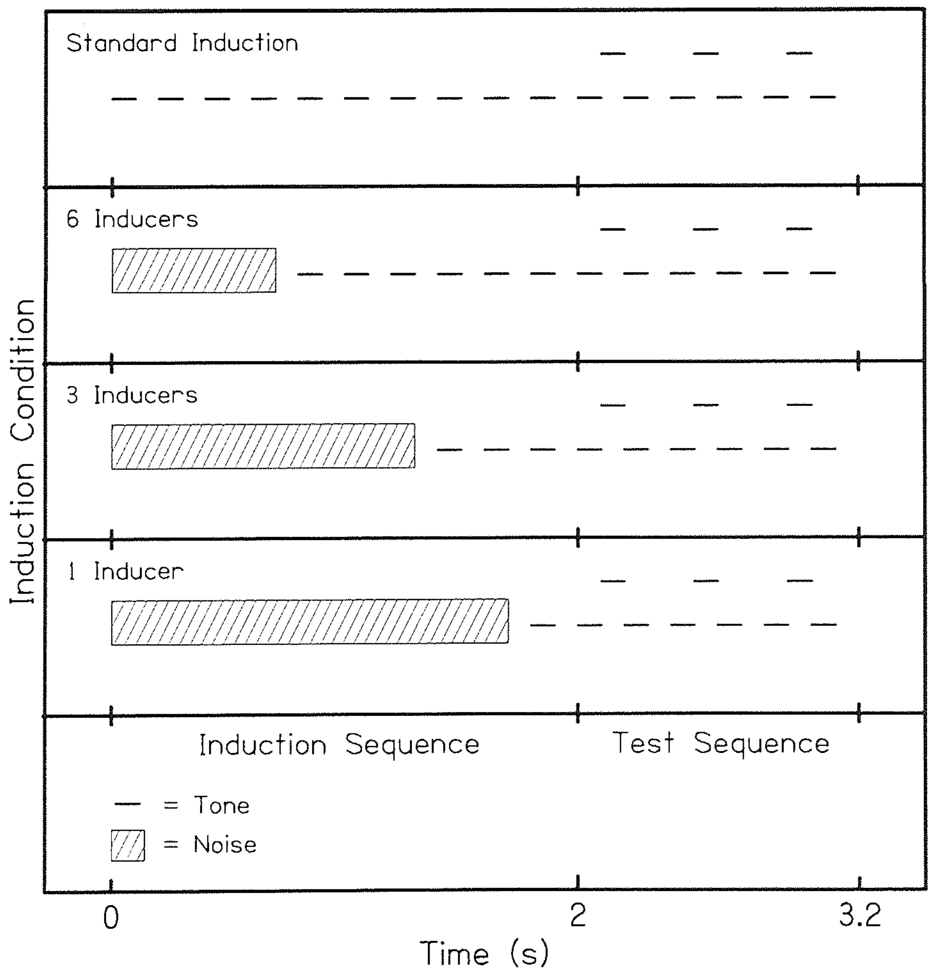

Figure 5.3: Illustration of the induction conditions tested in Experiment 4b. Each panel displays a different induction sequence paired with a subsequent test sequence. The solid lines represent pure tones, and the shaded boxes represent band-pass filtered noise.

The noise used was filtered with a bandwidth of four semitones, which was centred on 1 kHz (range = 891 - 1122 Hz, spectral roll-off = 80 dB per octave). Band-pass filtered noise was chosen in preference to broadband noise in order to reduce the loudness for the same spectrum level. The noise band was digitally created by combining equal-amplitude sinusoids with random starting phases distributed at 2-Hz intervals across the specified bandwidth. For this arrangement, the 2 Hz spacing of sinosoids would result in a 2 Hz periodicity of the noise. Therefore, the number of noise cycles would have increased as the overall duration of the noise-burst was extended. However, informal listening suggested that this repetition of noise cycles was not discernible perceptually. 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.

A brief pilot study indicated that the band-pass filtered white noise chosen to precede the tone sequences did not appear to affect subsequent streaming judgments. Broadband (white) noise is known not to induce subsequent stream segregation (Bregman, 1978; Rogers & Bregman, 1993, 1998), and the same appeared to be true for the narrower noise band used here (which sounded more "noise-like" than "tonal"). This observation is also supported by the fact that any long, continuous stimulus with only a single onset (even a pure tone) is known to be ineffective at promoting subsequent stream segregation (Rogers & Bregman, 1993; Roberts et al., 2008).

The current experimental induction conditions were chosen as they related to those of the previous experiment. For Experiment 4a, two separate conditions tested the effect of replacing either the 4^{th} or the 7^{th} inducer of a ten-tone sequence with a silent interval of equivalent duration. This resulted in either six or three induction tones being heard after the silent tone.

Procedure and Apparatus

The procedure and apparatus were identical to that described in Chapter 2. Each trial comprised a combination of an induction sequence and a subsequent test sequence. These trials were organized into trial blocks, each block comprising a combination of the five induction conditions and the six frequency separations for the test sequence. For each listener, the main experiment comprised 20 trial blocks (i.e., 600 trials in total).

5.3.2 Results

The results were averaged to give a single measure of the percentage of trials which were heard as segregated for each combination of induction condition and frequency separation. These mean values are shown in Figure 5.4. The mean percentage reported as segregated for frequency separations of 4, 6, 8, 10, 12, and 14 ST were 2.0%, 27.6%, 61.5%, 80.5%, 89.2%, and 96.6%, respectively. The means for the standard induction, no-induction, 6 inducer, 3 inducer, and 1 inducer conditions were 64.2%, 45.2%, 67.2%, 65.3%, and 56.0%, respectively. These data were then analysed using a two-way, repeated-measures ANOVA. The two independent variables were the six frequency separations for the test sequence, and the five induction conditions (standard induction, no-induction, or 3 $\times$ experimental induction conditions). The ANOVA confirmed a significant main effect of both frequency separation $[F(5,35) = 56.707,\ p < 0.001,\ \eta_p^2 = 0.890]$ and induction condition $[F(4,28) = 7.690,\ p < 0.001,\ \eta_p^2 = 0.523]$. There was also a significant interaction between these two variables $[F(20,140) = 3.587,\ p < 0.001,\ \eta_p^2 = 0.339]$. As with the previous subjective experiments, this appears to be due primarily to floor and ceiling effects at the two extreme frequency separations.

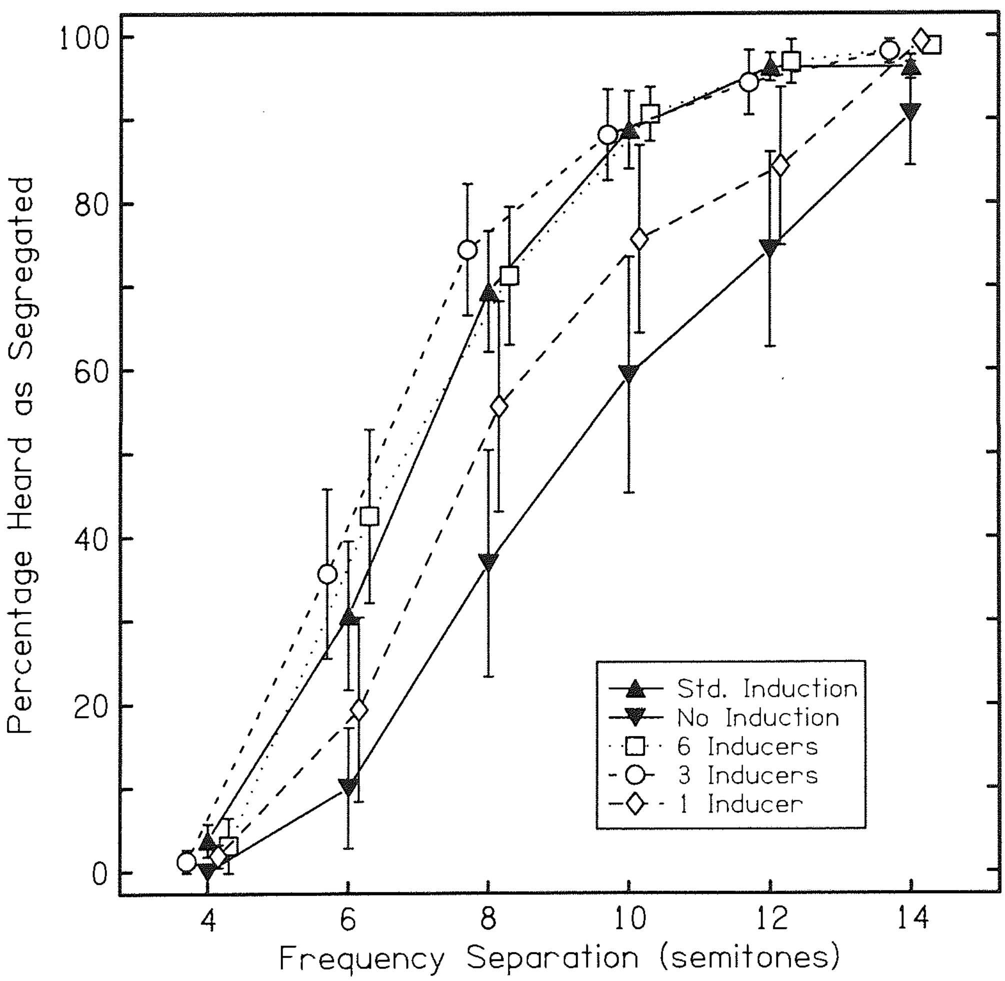

Figure 5.4: **Results from Experiment 4b. The effects of induction condition and frequency separation in the test sequence on reported segregation. Each point represents the percentage of trials reported as segregated (averaged across 8 listeners). The insert identifies the different induction conditions tested. Error bars represent ± 1 inter-subject standard error.**

As can be seen in Figure 5.4, the standard induction, 6-inducer, and 3-inducer conditions all promoted a similar and strong tendency for listeners to report stream segregation. Least segregation was reported in the no-induction condition. For the 1-inducer condition, reported segregation was roughly midway between that for the standard and no-induction cases. For Table 5.2, the data from the experimental conditions are transformed into a measure of build-up that is relative to both the standard and no-induction cases. This relative measure is analogous to the "extent of resetting" calculated for tables 3.1 - 5.1. The key difference is that the figures quoted in Table 5.2 represent a proportional shift on a scale defined by the no-induction condition representing no prior build-up (0.00), and the standard induction condition representing a large segregation-promoting effect (1.00). Therefore, this scale is the inverse of the extent of resetting used in the previous tables. This change was made, as the concept of resetting is not applicable to the currently tested conditions.

As for the extent of resetting measure, the relative build-up measure will also be affected by the magnitude of the difference in reported segregation between the standard and the no-induction conditions. Namely, if only a small difference existed between these two control conditions, any two experimental conditions which differed only marginally in absolute terms would produce largely different relative build-up values. In order to compensate for this issue and provide a stable measure of relative build-up, the data from the 4 ST and 14 ST frequency separations are not reported and a weighting system was applied when averaging relative build-up across the remaining test-sequence frequency separations. The weight applied to each frequency separation is calculated in exactly the same way as for the extent of resetting analysis, and so the reader is referred to the results section of Experiment 1 for a complete description of this procedure.

		Induction Condition			
		6 inducers	3 inducers	1 inducer	*Mean*
Frequency Separation	6 ST (*19.76%*)	1.57	1.24	0.45	***1.09***
	8 ST (*31.14%*)	1.06	1.15	0.58	***0.93***
	10 ST (*28.14%*)	1.06	0.98	0.55	***0.86***
	12 ST (*20.96%*)	1.03	0.91	0.46	***0.80***
	Weighted Mean	***1.15***	***1.07***	***0.52***	***0.91***

Relative build-up = no induction (***0.00***) → standard induction (***1.00***)

Table 5.2: **Results from Experiment 4b. Reported segregation was compared with that in the no-induction case (0.00). Values are expressed as a proportional increase in segregation, towards the standard induction case (1.00). A weighting system was applied to the calculation of across-frequency means, but each individual value represents the unweighted relative build-up. Note that a value greater than 1.00 indicates that reported segregation was *greater* than that for the standard induction condition.**

Two-tailed pairwise comparisons (restricted LSD test, Snedecor & Cochran, 1967) were used on the original, untransformed data set to compare the experimental conditions to the standard induction condition. There was a significant difference between the standard and the no-induction conditions [difference in reported segregation = 19.0%, $t(7) = 2.735$, $p < 0.05$]. There was no significant difference between the standard induction condition and the 3-inducer condition [1.1%, t(7) = 0.515, p > 0.05], but the small difference between the standard induction and the 6-inducers condition was significant [3.0%, $t(7) = 2.573$, $p < 0.05$]. Despite the larger difference in means, corresponding to a resetting effect of about a half, there was no significant difference between the standard induction and 1-inducer conditions [8.2%, $t(7) = 1.542$, $p > 0.05$]. This result is probably due to the relatively large variance across listeners observed for the 1-inducer condition. Finally, all four of the L-tone induction conditions promoted significantly more stream segregation than the no-induction case ($p < 0.05$ in all cases).

5.3.3 Discussion

The data from Experiment 4b were again in agreement with the known effect of pure-tone frequency separation on stream segregation (e.g., van Noorden, 1975). The standard induction sequence (ten L- repetitions) had a strong segregation-promoting effect, and this was consistent with previous studies (experiments 1 - 4a, Rogers & Bregman, 1993, 1998). The induction sequences that comprised either six or three L- repetitions were also effective at promoting segregation. Indeed, the percentage of trials heard as segregated following the 3- and 6-inducer conditions was numerically a little higher than for the standard induction case, and this difference was significant in the latter case. Furthermore, even the induction sequence containing only a single L- cycle also increased the perception of stream segregation with respect to the no-induction condition - albeit to a lesser extent than did all of the other induction conditions (about half).

Given that the six-tone induction sequence promoted statistically more stream segregation than did the standard, ten-tone case, one might speculate that the preceding band-pass filtered noise had some segregation-promoting effect in the experimental induction conditions. Whilst this explanation could potentially explain the current findings, it should be noted that an extended tone with only one onset is largely, if not entirely ineffective at promoting test sequence segregation (Rogers & Bregman, 1993; Roberts et al., 2008). Considering that the noise in the current experiments also had only a single onset, and was also of substantially different timbral properties to the tones, one would not predict that the noise would be effective at promoting segregation in the test sequence. Further study would be needed to determine whether a continuous noise had some form of marginal segregation-promoting effect.

These findings of Experiment 4b provide some insight into the results from the previous experiment. For Experiment 4a, no resetting effect was observed when either the 4^{th} or the 7^{th} tone of a ten-tone induction sequence was replaced with an equivalent-duration silent interval. The current findings suggest that if these changes did substantially reset build-up, the segregation-promoting effect of the subsequent induction tones would have entirely obscured this effect. It was also speculated that the recovery of build-up following resetting may be more rapid than the original build-up itself. Whilst this may be the case, the initial rate of build-up in a same-frequency induction sequence is clearly very rapid, and so can account for the results of Experiment 4a.

It is interesting to note that induction sequences which contained ten, six, or three L-cycles all promoted a very similar degree of reported segregation. This strongly contrasts with previous results on the rate of build-up obtained for on-going, alternating-frequency tone sequences (e.g., a repeating LH type sequence). For such stimuli, build-up is typically rapid and progressive over the first 5 - 10 seconds after the onset of the sequence. After this initial rapid increase, the rate of build-up then becomes more gradual, but still remains apparent over the entire course of a 60-second sequence (Anstis & Saida, 1985, see Figure 1.4). In contrast, it appears that the cumulative segregation-promoting effect of a same-frequency induction sequence has very different dynamic properties. Reported stream segregation did not vary greatly following induction sequences of either three, six, or ten L- cycles. This could indicate that the segregation-promoting effect of a same-frequency induction sequence reaches a maximum level after only three rapid L-repetitions, and that any subsequent L- tones cannot promote any further increase in segregation. To investigate this suggestion further, one would need to measure stream segregation after induction sequences that comprised more than ten L- repetitions.

5.4 General Discussion

The experiments presented in this chapter have provided some insight into the dynamics of build-up and resetting during a same-frequency induction sequence (i.e., a sequence comprising repetitions of an L- tone). For Experiment 4a, it was shown that resetting only occurred when the final induction tone of a ten-tone sequence was replaced with silence - there was no evidence of resetting when the same change was applied to either the 4^{th} or 7^{th} induction tone. Experiment 4b demonstrated that build-up is very rapid in a same-frequency induction sequence, and that induction sequences of 3, 6, or 10 L- repetitions all promoted a broadly equivalent amount of reported stream segregation. Taken together, these experiments demonstrate that the segregation-promoting effect of a same-frequency induction sequence is strongly, perhaps even entirely, governed by the properties of the final three induction tones prior to the onset of the test sequence.

Rogers & Bregman (1993) proposed that during a same-frequency induction sequence, identical tones are heard (e.g., L tones only), and it is logical to assume that a listener will group these tones into an on-going perceptual stream. During the test sequence, a second subset of tones is added (i.e., the H tones). Given that the listener is already grouping the L tones as a distinct auditory object, the new subset of tones may be excluded from this on-going stream (i.e., stream segregation is heard). In other words, at the onset of the test sequence, an L-tone-only stream has already been established, and this may increase the likelihood of the H tones being heard as a separate stream. This proposal is illustrated in Figure 5.5 (see top two panels). Here, this hypothesis is elaborated upon, as the current data lead to three predictions regarding the segregation-promoting effect of a same-frequency induction sequence. Briefly, these predictions are:

1. In order for a same-frequency induction sequence to have a maximum segregation-promoting effect, a clear perceptual stream must be formed from these induction tones before the onset of the test sequence.

2. Once formed, this perceptual stream must be heard to be continuous with one of the two subsets of tones in test sequence. Any change that disrupts this smooth

continuation of the on-going stream into the test sequence should also reduce the segregation-promoting effect of the induction sequence.

3. So long as the first two criteria are met, any additional induction tones should have relatively little effect on streaming in the test sequence.

These three predictions are discussed more fully in the next few paragraphs.

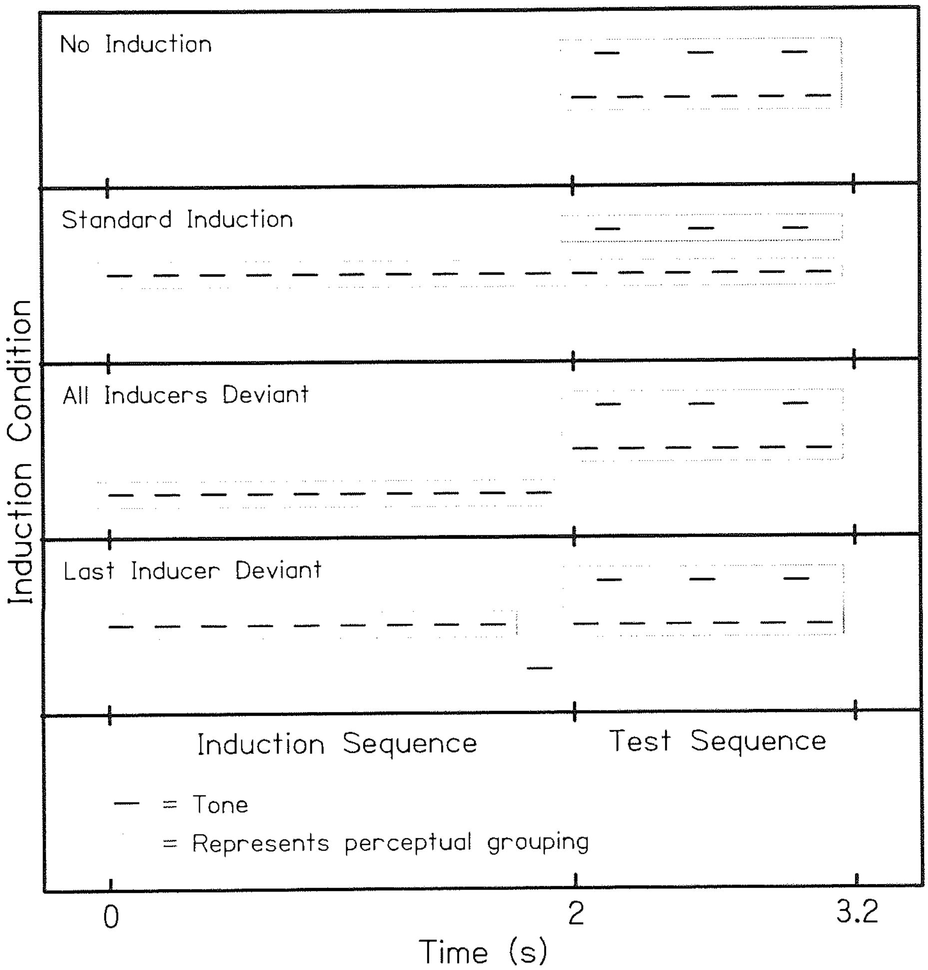

Figure 5.5: The effect of a same-frequency induction sequence on stream segregation. (Panel 1) When heard in isolation, a listener is not primed to segregate the test-sequence tones. (Panel 2) A clear perceptual stream is formed during the induction sequence, which corresponds to the L tones. This stream continues into the test sequence, and so increases the likelihood of the H tones being heard as segregated. (Panel 3) A perceptual stream is formed during the induction sequence, but this stream does not continue into the test sequence (owing to the change in tonal properties). As a result, the induction sequence has little segregation-promoting effect (see Roberts et al., 2008). (Panel 4) The final induction tone is noticeably altered, and this change disrupts the on-going stream. As a result, the segregation-promoting effect of the induction sequence is reduced (see experiments 1 - 4a).

1) At the onset of a test sequence, a listener will only be primed to hear stream segregation if one subset of tones is already grouped into an on-going stream. Experiment

4b demonstrated that an induction sequence comprising 3 L- cycles had a strong, and seemingly maximal, segregation-promoting effect. In comparison, an induction sequence comprising only a single L tone promoted noticeably less segregation (albeit that this difference was not significant). This may indicate that either two or three L-tone repetitions may be required before a perceptual stream is fully formed.

This hypothesis also relates to evidence obtained from electrophysiological studies. The mismatch negativity (MMN) ERP component can be elicited by any noticeable change in a repetitive auditory stimulus. This MMN component is thought to reflect a deviance detection process which is based on the memory of the regularities occurring in an acoustic stimulus (see Näätänen et al., 2001, for a review). Crucially, 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. For Experiment 4b, a three-tone induction sequence was highly effective at promoting segregation in the subsequent test sequence. From the MMN evidence, we know that after this induction sequence, a regular L-tone pattern should have been established in memory. Importantly, the MMN component also correlates strongly with perceptual organization[2]. Therefore, the MMN evidence is consistent with the current hypothesis that three repetitions of an identical sound are required before a listener forms a clear perceptual stream.

2) In order to promote segregation, the perceptual stream that is formed during the same-frequency induction sequence must continue smoothly into one subset of tones in the test sequence. For example, in the current experiments, the properties of the standard induction tones matched precisely the L tones of the test sequence (in terms of tone frequency, duration, and rhythm). Therefore, the L-tone stream formed during the standard induction sequence should seamlessly continue into the test sequence (see the second panel of Figure 5.5). This could explain why the standard induction sequence is highly effective at promoting segregation. In contrast, when the induction tones do not match all the properties of a subset of tones in the test sequence, the perceptual continuity between induction and test sequence should be reduced (cf. Roberts et al., 2008; see the third panel of Figure 5.5 for a related illustration). Similarly, test-sequence segregation is also reduced when only the final induction tone is noticeably altered (i.e.,

[2] A wide range of studies have demonstrated that MMN elicitation strongly correlates with the sequential organization of a tone sequence (e.g., streaming in a repeating LH type sequence - see Sussman et al., 1999; Winkler et al.,2003b; Sussman et al., 2007). These complex studies are difficult to describe concisely, so the reader is referred to Chapter 1 where a more complete explanation is provided.

experiments 1 - 4a - see the bottom panel of Figure 5.5). This break in continuity may have a resetting effect, as the on-going L stream is disrupted immediately prior to the onset of the test sequence. Indeed, Cusack et al. (2004) made a similar suggestion - that resetting may be due to the disruption of an on-going stream.

3) Once an on-going stream has been established in an induction sequence of constant tonal properties, any additional induction tones should not increase segregation of the test sequence any further. For the current induction sequences, there are no competing perceptual organizations for the identical L tones - they will always be integrated into a single stream. Therefore, once this stream is established, there is no reason for the strength of this percept to change over time. Consistent evidence was observed in Experiment 4b, where induction sequences of 3 - 10 tones all promoted a highly similar degree of stream segregation.

A study by Bregman & Rudnicky (1975) supports the current hypothesis that a same-frequency induction sequence promotes stream segregation by capturing one subset of test-sequence tones into an on-going stream. These authors presented two tone pairs, and asked listeners to report whether these pairs were the same or different. Each pair comprised two "target" tones, A and B, presented sequentially either as AB or BA. These pure tones differed in frequency (2200 Hz or 2400 Hz), but were otherwise identical (tone duration = 65 ms, no inter-tone silence). This task was relatively easy, but it became more difficult when the second tone pair was flanked by two "distractor" tones (i.e., the sequence was heard as DABD; D = distractor tone, D frequency = 1460 Hz). Bregman and Rudnicky (1975) suggested that this difficulty may indicate that that the distractor tones grouped with the target tones, and so reduced the salience of the direction of the frequency change. Furthermore, performance could be improved by the addition of a sequence of "captor" tones (i.e., CCCCDABDCCCC). This improvement occurred only when the captors matched the distractors in terms of frequency and tempo; when the captors were of a lower frequency (590 Hz), they did not aid performance. Bregman & Rudnicky (1975) attributed this effect to the matching captor tones pulling the interfering distractor tones into a separate auditory stream from the target tones. Bregman & Rudincky's (1975) results and their proposed explanation both fit well with the current argument - that the same-frequency induction tones promote segregation by capturing the L tones of the test sequence into a pre-established, on-going stream. The findings of Bregman & Rudnicky (1975) are considered in more detail in Chapter 9.

In conclusion, the current experiments have offered further insight into why a same-frequency induction sequence can promote stream segregation a subsequent test sequence. It has been proposed that during such an induction sequence, a perceptual stream is formed which may continue into the test sequence, and so capture one subset of tones into the on-going stream. This may result in an increased tendency to hear the other tone subset as segregated from the established stream. This hypothesis is developed further in Chapter 8: Experiment 7 directly compares the build-up in the tendency for stream segregation brought about by a same-frequency induction sequence and by an alternating-frequency tone sequence.

Chapter 6

Build-up and Resetting for Induction Sequences with Constrained-Random Frequency Changes

6.1 Introduction

In the discussion of Chapter 5, a hypothesis was presented to explain why a same-frequency induction sequence (i.e., 10 L- repetitions) can be so effective at promoting segregation in a subsequent test sequence (i.e., 3 LHL- triplets). Briefly, this hypothesis stated that during the induction sequence, a listener will form a single perceptual stream corresponding to the L-tone inducers. This L-tone-only stream is heard to continue into the test sequence, and so the H tones are likely to be excluded from the already established stream. In other words, only the L-tones of the test sequence are captured into the on-going induction-tone stream, and so the likelihood of hearing stream segregation is increased (see also Rogers & Bregman, 1993).

A related argument was proposed by Bregman & Rudnicky (1975), who asked listeners to report the presentation order of two target tones. When these target tones were positioned sequentially between two distractor tones of a lower frequency, the task became very difficult, but performance improved when the distractors were flanked by a sequence of additional, same-frequency captor tones whose frequency matched that of the distractors. Bregman & Rudnicky (1975) argued that the captor tones may have aided performance by pulling the distractor tones into a pre-established stream, allowing the target tones to be perceived separately.

A single abrupt change to the last inducer can partially reset the segregation-promoting effect of a same-frequency induction sequence (see experiments 1 - 4a). For Experiment

4a, it was demonstrated that observable resetting only occurred when the final induction tone was altered (replaced with silence) - the same change applied to an earlier induction tone had no apparent resetting effect. These results were explained in Experiment 4b, which demonstrated that build-up can rapidly (re)occur in a same-frequency induction sequence. Indeed, only a single L-inducer can have a substantial segregation-promoting effect. Taken together, these findings suggest that the properties of the final induction tone have a considerable influence on subsequent streaming of the test sequence. Despite this it should be noted that, in comparison to longer tone sequences, a single deviant tone is only *partially* effective at resetting build-up, and that a single induction tone is only *partially* effective at promoting test-sequence segregation. Therefore, the segregation-promoting effect of a same-frequency induction sequence is not determined entirely by the properties of the final induction tone alone. The current experiment was designed to examine how the properties of the earlier induction tones also affect test-sequence streaming.

One aim of the current experiment was to compare directly the resetting effect of a single deviant tone with the segregation-promoting effect of an induction sequence where, on one acoustic dimension, all the tones differ from those of either subset in the test sequence (see Roberts et al., 2008). Based on previous research, one would predict that a single frequency-deviant tone would have only a partial resetting effect on the build-up of stream segregation (experiments 1, 3a, 3b), whereas an induction sequence in which all tones are uniformly lower in frequency than the L tones would largely fail to promote any test-sequence streaming (Roberts et al., 2008). It should be noted that no previous study has directly compared these two types of induction sequence.

Several authors have argued that the resetting effect of an abrupt change may be linked to attentional factors (Cusack et al., 2004; Denham & Winkler, 2006; Roberts et al., 2008). For an alternating LH, or LHL- type tone sequence, the initial default percept is that of integration, and the tendency to hear stream segregation gradually builds up over time (Bregman, 1978; Anstis & Saida, 1985). According to Cusack et al.'s (2004) hierarchical decomposition model, this default perception will be heard when attention is switched to a new, or to an *on-going* tone sequence. Therefore, a single deviant tone may trigger resetting by briefly drawing attention away from the on-going sequence, and thus (at least partly) restore a listener's default perception of integration. If Cusack et al.'s (2004) theory were correct, one would predict that deviant-tone resetting is not directly influenced by the absolute properties of the deviant tone *per se*, but rather that resetting will be primarily determined by the ability of the deviant tone to attract attention. In other words, a changed tone that does not attract attention should not have a resetting effect. Whilst it is difficult to manipulate the salience of a deviant tone without also manipulating its acoustic properties, there is some evidence to support this argument. Rogers & Bregman (1993) demonstrated that a same-frequency induction sequence that was temporally unpredictable (in tone duration and inter-tone silences) was as effective at promoting test-sequence segregation as a regular, predictable sequence (where both induction sequences comprised an identical number of tone onsets, and the same overall tone density)[1]. Despite this evidence, Experiment 1 demonstrated that a single temporally deviant tone can have a substantial resetting effect (where a standard tone = 100 ms, deviant tone = 150 ms). Therefore, it would appear that a temporally unpredictable induction sequence may be effective at promoting stream segregation, but that an unexpected change to the temporal properties of a single tone in an otherwise regular sequence can have a substantial resetting effect. This is consistent with the suggestion that resetting only occurs when attentional focus is changed. If we consider an unpredictable induction sequence, a listener cannot anticipate the exact properties of each incoming tone. Therefore, although the properties of the final induction tone may differ from all the preceding tones, because the sequence is unpredictable, the final tone cannot be heard as a deviation from an established pattern. In contrast, when only the final tone is altered in an otherwise predictable sequence, the change is highly salient because it clearly violates an established pattern. Therefore, a single deviant tone should be likely to attract attention and, according to Cusack et al.'s (2004) theory, should also have a resetting effect.

The role of predictability in sequential grouping was considered by Jones in 1976 (see also Jones & Boltz, 1989). Jones (1976) proposed a theory which stated that streaming is governed by rule-based predictability. The basis of this theory was that auditory stimuli are perceptually represented across different dimensions (e.g., pitch, location, or time) in a multidimensional space. Jones proposed that, for a sequence of sounds, the auditory system can detect certain types of regularities across this space. Through an attentional process, these regularities are then used as rules to predict subsequent incoming sounds. If a sound falls within the predicted area of multidimensional space, it is heard as integrated into the on-going stream, whereas the sound is heard as segregated

[1] In a pilot study, which was not developed into a full experiment, Roberts et al. (2008) measured the effect of a temporally unpredictable induction sequence on performance in a temporal discrimination task (see Chapter 3). For this induction condition, the onset time of each tone was randomly advanced or delayed by a certain amount, but each tone was separated by a minimum of 10 ms. This unpredictable sequence also shared the same number of tones, tone duration, and tone density as the regular induction sequence. These authors observed that the unpredictable induction sequence had no appreciable segregation-promoting effect, a finding in clear contrast to that of Rogers & Bregman (1993). Roberts et al. (2008) speculated that this difference may represent a resetting effect which is only apparent in the context of a temporal-discrimination task.

if it falls outside the predicted area. In Jones's (1976) theory, time was regarded as a crucial dimension. Specifically, she argued that attention can only shift between different sounds at a limited speed. Therefore, a change across any given dimension must be proportional to a change on the time dimension. For example, at a rapid presentation rate, attention could only follow successive tones with small differences in frequency (cf. Bregman's, 1990, argument about primitive grouping and the notion of persistence in acoustic events). This argument is consistent with the known effect of frequency separation and presentation rate on stream segregation (van Noorden, 1975, see Figure 1.2).

Jones et al. (1981) provided some evidence that rhythmic predictability may influence sequential grouping. These authors used an experimental design similar to Bregman & Rudnicky's (1975) temporal discrimination task. They showed that the captor tones became less effective at grouping with the distractors when the two types of tones differed in rhythm. However, there are a range of findings which are difficult to explain using Jones's (1976) theory. Jones (1976) predicted that any increased knowledge of a sequence should aid a listener's ability to hear integration. However as previously discussed, Rogers & Bregman (1993) demonstrated that temporal predictability did not affect the segregation-promoting effect of a same-frequency induction sequence (where streaming was measured subjectively in a subsequent test sequence). In a related study, French-St. George & Bregman (1989) presented 30-s sequences of alternating high and low frequency subsets of tones (i.e., standard pattern = L1-H1-L2-H2-L3-H3-L4-H4, where 1 - 4 represent slight frequency variations). When compared to this standard pattern, neither the predictability of the frequency variations (1, 2, 3, 4 versus random - e.g., 4,1,3,2) nor of the timing of the tones (isochronous versus non-isochronous) affected the streaming of the tones into two frequency ranges (L & H). However, the authors admitted that the listeners may have been unable to predict the standard pattern well enough to provide an adequate contrast to the unpredictable conditions of the experiment. Finally, it is well established that alternating LH tone sequences are initially heard as integrated and that the tendency to hear stream segregation increases over time (Bregman, 1978; Anstis & Saida, 1985). Bregman (1990) noted that this type of sequence would become increasingly predictable over time - and so, according to Jones's theory (1976), should become *less* likely to segregate. As the opposite pattern is observed, Bregman (1990) concluded that streaming may reflect grouping by similarity (e.g., in properties such as event density or frequency) rather than an effect of predictive factors.

The purpose of the current experiment was to compare the segregation-promoting effect of three different types of induction sequence. Each of these sequences was a modification of the standard induction sequence (10 L- repetitions, see Chapter 2). In one condition, only the final induction tone was lowered in frequency (-2 ST). Similar conditions were tested in Experiment 3a, and it was found that a single frequency-deviant tone at the inducer/test boundary can have a large resetting effect. In a second condition, all of the induction tones were lowered in frequency by 2 ST (cf. Bregman & Rudnicky, 1975; Roberts et al., 2008). One important difference is that these previous studies both used a much larger decrease in frequency – either approximately 16 ST (Bregman & Rudnicky, 1975), or 24 ST (i.e., two octaves; Roberts et al., 2008). In a final condition, the first nine induction tones were each drawn randomly from a restricted frequency range ($\pm$2 ST with respect to the L tones), whilst the final tone was fixed at -2 ST. French-St. George & Bregman (1989) demonstrated that frequency predictability did not affect the grouping of two tonal sub-sets (where frequency variations were applied to the L and H tone subsets). For each of these two frequency subsets, the tones were varied within an (approximately) 3 ST range. Based on this finding, the unpredictable nature of the current constrained-random induction sequence (frequency range = 4 ST) should have little or no effect on the grouping of its constituent tones. However, as the sequence is largely unpredictable, a listener should be unable to anticipate the frequency of each incoming tone, and so the final induction tone (-2 ST) should not be especially salient. If Cusack et al. (2004) were correct in assuming that resetting occurs when attention is drawn to a sequence, the frequency changes in the unpredictable sequence should not have any resetting effect.

The (constrained) randomization of inducers across a $\pm$2 ST range also affected the spread of peripheral excitation during the induction sequence. This is an important consideration, as the resulting spread of frequency is greater than the auditory filter bandwidth associated with this frequency range (as 1 ERB_N for a centre frequency of 1 kHz $\approx \pm$1 ST, Glasberg & Moore, 1990). The resulting weakening of excitation within the 1-kHz filter bandwidth may have some effect on the segregation-promoting effect of the induction sequence. Furthermore, the spread of excitation above the standard L-frequency of 1 kHz also resulted in some increased similarity between the induction tones and the H-tones of the test sequence. These issues are considered in further detail in the general discussion section.

6.2 Experiment 5

6.2.1 Method

Experiment 5 used the same subjective procedure as described in Chapter 2. In summary, on each trial listeners were required to indicate their perception of the final LHL-triplet of a test sequence as either integrated or segregated (test sequence = 3 LHL-triplets, L tone = 1 kHz, H tone = +4 to +14 ST, tone duration = 100 ms, inter-triplet silence = 100 ms). For the majority of trials, the test sequence was preceded by an induction sequence, the properties of which were varied experimentally.

Participants

Eight participants contributed to the final data set of Experiment 5. Four of these listeners had also contributed to at least one of the earlier subjective experiments, but all were unaware of the purpose of the current experiment. Data from one participant was replaced because their results did not show clearly the well established dependence on LH frequency separation for their streaming judgments (see, e.g., van Noorden, 1975). All participants reported normal hearing.

Induction Conditions

Both the standard induction condition and the no-induction condition were tested in Experiment 5. Both of these induction sequences were described in detail in Chapter 2. The standard induction sequence comprised 10 L- repetitions (L tone = 1 kHz, tone duration = 100 ms, inter-tone silence = 100 ms). For the no-induction condition, the test sequence was presented alone, in the absence of any prior induction tones. The three experimental induction conditions were each a modification of the standard induction case (see Figure 6.1). The resetting effect of a single deviant tone was measured in the deviant(-2 ST) induction condition. For this condition, the final induction tone of the standard induction sequence (i.e., the 10^{th} tone) was lowered in frequency by 2 ST, from 1 kHz to 891 Hz. Except for this frequency change, the deviant tone was otherwise identical to a standard L tone. For the second experimental induction condition, *all* of the standard induction tones were lowered by 2 ST to 891 Hz. This arrangement is referred to as the all(-2 ST) condition.

For the final induction condition, the first nine tones were each set to a different constrained-random frequency, whilst the 10^{th} (final) tone was fixed at -2 ST. The frequency values of the first nine tones were partially randomized within a ±2 ST range with a rectangular distribution and centred on 1 kHz. More specifically, each of the first nine tones was randomly assigned one of 13 frequency values between 891 and 1122 Hz. These values were equally spaced in one-third ST intervals (i.e., -2 ST, -1.67 ST, -1.33 ST...0 ST...+1.33 ST, +1.67 ST, +2 ST, with respect to the 1-kHz reference value). To ensure that the induction sequence remained unpredictable in frequency, no two adjacent induction tones were permitted to share the same frequency value. For all three of the experimental induction conditions, the final induction tone was fixed at -2 ST. Therefore, any differences in the segregation-promoting effect of these sequences can be attributed to the effect of the first nine induction tones.

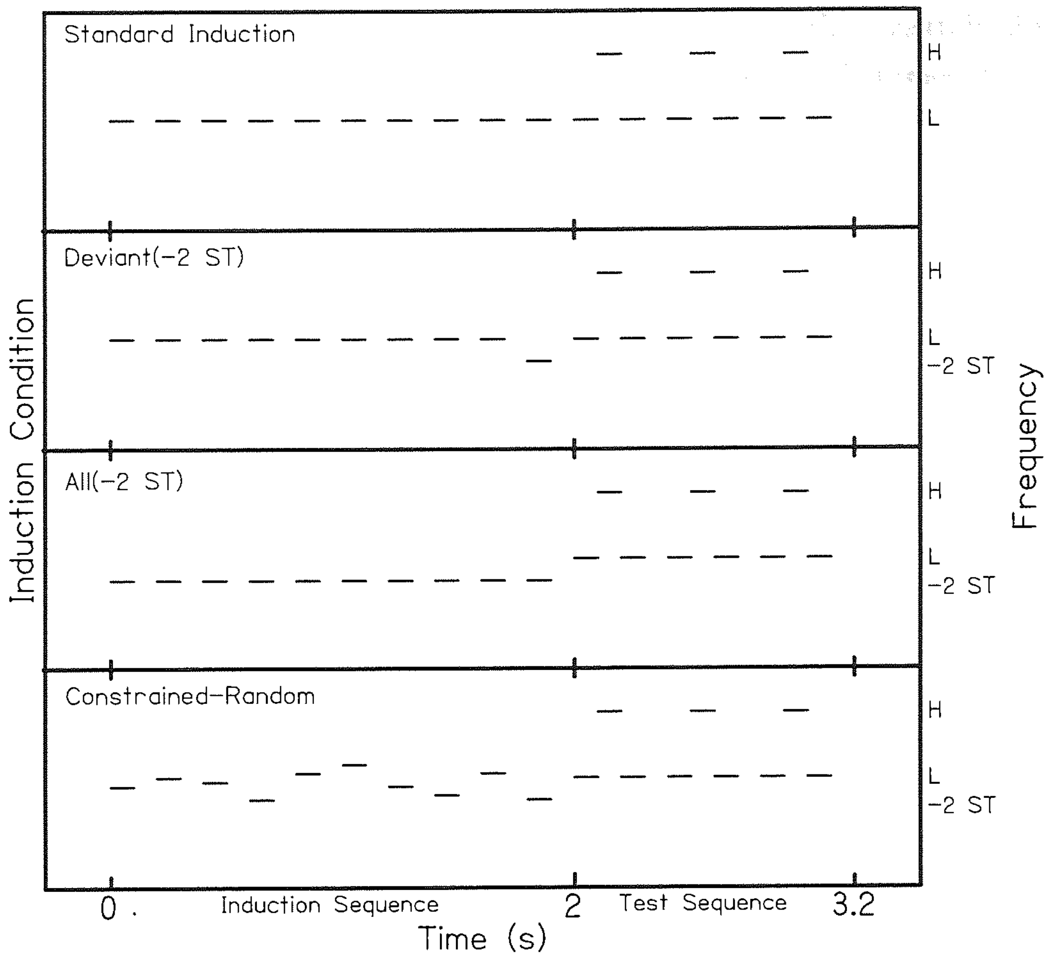

Figure 6.1: **Illustration of the induction conditions tested in Experiment 5. Each panel displays a different induction sequence paired with a subsequent test sequence.**

Procedure and Apparatus

The procedure and apparatus used for the current experiment were fully described in Chapter 2. Each trial block comprised a combination of each of the five induction conditions with each of the six frequency separations for the test sequence (i.e., 30 trials in total). For the main experiment, each participant was presented with 20 trial blocks in total (i.e., 600 trials). One hundred different constrained-random induction sequences were pre-prepared before data collection. During an experiment, one of these sequences was randomly chosen whenever a listener was presented with a trial which required a constrained-random induction sequence.

6.2.2 Results

Results were averaged to give a single measure of the percentage of trials which were heard as segregated for each combination of induction condition and frequency separation. (i.e., the percentage heard as segregated). These mean values are shown in Figure 6.2. The mean percentage reported as segregated for frequency separations of 4, 6, 8, 10, 12, and 14 ST were 9.5%, 29.6%, 60.9%, 79.5%, 86.9%, and 95.7%, respectively. The means for the standard, no-, deviant(-2 ST), all(-2 ST), and the constrained-random induction conditions were 77.8%, 44.1%, 66.5%, 49.2%, and 64.3%, respectively. These data were analysed using a two-way, repeated-measures ANOVA. The two independent variables were the six frequency separations for the test sequence (0, 4, 6, 8, 10, 12, or 14 ST), and the five induction conditions (standard induction, no induction, or three experimental conditions). The ANOVA confirmed significant main effects of frequency separation [$F(5, 35) = 57.124$, $p < 0.001$, $\eta_p^2 = 0.891$], and of induction condition [$F(4, 28) = 16.038$, $p < 0.001$, $\eta_p^2 = 0.696$]. The interaction between these two variables was also significant [$F(20, 140) = 5.117$, $p < 0.001$, $\eta_p^2 = 0.422$]. Once again, this interaction was probably driven primarily by the partial floor and ceiling effects at the two extreme HL frequency separations (see Figure 6.2).

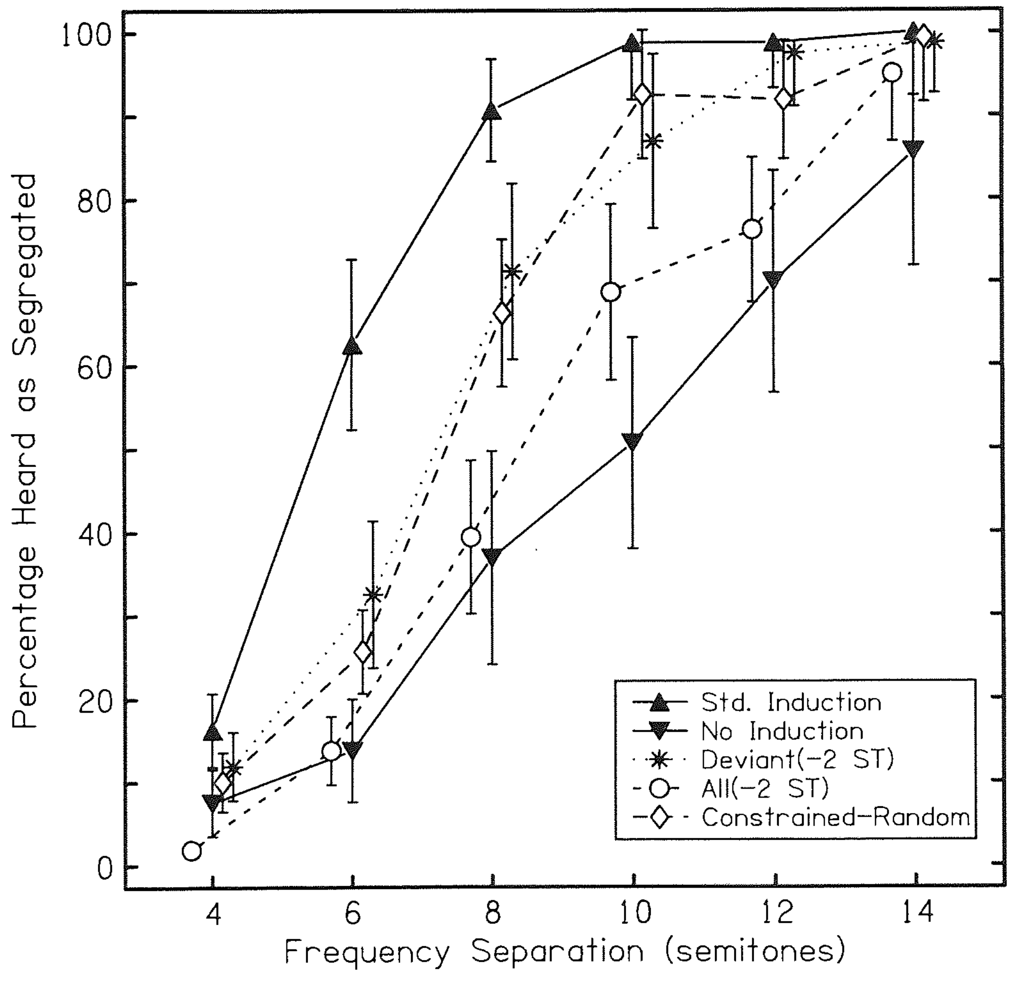

Figure 6.2: Results from Experiment 5. The effects of induction condition and frequency separation in the test sequence on reported segregation. Each point represents the percentage of trials which were reported as segregated (averaged across 8 listeners). The insert identifies the different induction conditions tested. The error bars display ±1 inter-subject standard error.

As with the earlier experiments, the standard induction condition promoted the greatest amount of stream segregation and the least was heard in the no-induction condition. For all of the experimental conditions, segregation was reduced away from the standard induction condition towards that for the no-induction case. This shift in reported segregation is displayed in Table 6.1, where the data are transformed into a measure of the extent of resetting (as described in previous results sections – see Table 6.1).

Frequency Separation		Induction Condition			
		Deviant(-2 ST)	All(-2ST)	Constrained-Random	***Mean***
	6 ST *(27.18%)*	0.62	1.00	0.76	***0.79***
	8 ST *(29.97%)*	0.36	0.95	0.45	***0.59***
	10 ST *(26.83%)*	0.25	0.62	0.13	***0.33***
	12 ST *(16.03%)*	0.04	0.78	0.24	***0.36***
	Weighted Mean	***0.35***	***0.85***	***0.41***	***0.54***

Extent of resetting = standard induction (***0.00***) → no induction (***1.00***)

Table 6.1: **Results from Experiment 5. Reported segregation in the deviant conditions was compared with that in the standard induction case. Values are expressed as a proportional shift away from the standard induction condition towards the no-induction condition (the *extent of resetting*). To provide a stable measure of resetting, data from the 4 ST and 14 ST frequency separation conditions were excluded from this analysis. Also, a weighting system was applied so that frequency separations in which there was a smaller difference between reported segregation in the standard and in the no-induction conditions made a reduced contribution to the across-frequency-separation mean. This procedure is described in detail in the results section of Experiment 1 (Chapter 3). Note that individual values are displayed as unweighted.**

Two-tailed pairwise comparisons were performed on the original (i.e., untransformed) data set using the restricted least-significant-difference (LSD) test (Snedecor & Cochran, 1967). Once again, the non-discriminating 4 and 14 ST conditions were excluded from these pairwise analyses. From this reduced data set, each pairwise comparison is accompanied with a measure of the difference in reported segregation between the two conditions (expressed as a percentage). There was a significant difference in reported segregation between the standard and the no-induction conditions [difference = 33.7%, $t(7) = 5.053$, $p < 0.005$]. Compared with the standard induction condition, stream segregation was significantly lower in all of the deviant conditions [deviant(-2 ST): 15.6%, $t(7) = 5.590$, $p < 0.001$], [all(-2 ST): 38.1%, $t(7) = 6.503$, $p < 0.001$], [constrained-random: 18.6%, $t(7) = 6.642$, $p < 0.001$]. When compared with each other, reported segregation in the all(-2 ST) condition was significantly lower than both the deviant(-2 ST) condition [22.5%, $t(7) = 3.402$, $p < 0.05$], and the constrained-random condition [19.5%, $t(7) = 3.605$, $p < 0.01$]. Both of the latter conditions resulted in a very similar amount of reported stream segregation [3.0%, $t(7) = 0.718$, $p > 0.05$]. Furthermore, reported perception for the all(-2 ST) condition was quite similar to that for the no-induction case, and the difference between them was not significant [6.7%, $t(7) = 1.094$, $p > 0.05$].

6.3 General Discussion

Listeners' responses conformed to the known effect of tone-frequency separation on stream segregation (van Noorden, 1975). When compared to the no-induction case, the standard induction sequence had a strong segregation-promoting effect (consistent with experiments 1 - 4b). When only the final induction tone was lowered in frequency (by 2 ST), reported segregation was significantly reduced away from the standard induction case towards that for the no-induction condition (i.e., a 0.35 extent of resetting – see Table 6.1). This finding is taken as evidence that the deviant tone partially reset the build-up of stream segregation. This result is in accord with experiments 1 and 3a, where similar frequency-deviant tones had somewhat equivalent resetting effects (Experiment 1: -3 ST deviant = 0.38 resetting, Experiment 3a: -1 ST deviant = 0.25 resetting, -3 ST deviant = 0.29 resetting). For the final two conditions of Experiment 5, the last inducer was also fixed at - 2 ST, but the properties of the earlier induction tones were varied. These conditions were tested to provide a measure of how the properties of the first nine induction tones also influence the segregation-promoting effect of an induction sequence.

For the all(-2 ST) condition, every induction tone was lowered in frequency by 2 ST (i.e., all were set to 891 Hz). This induction sequence promoted little streaming in the subsequent test sequence; indeed, reported segregation was not significantly different from that for the no-induction condition. Using a temporal-discrimination task, Roberts et al. (2008) found that an induction sequence which was substantially lowered in frequency with respect to the L tones did not promote streaming in the test-sequence (where the inducers = 250 Hz, test sequence L-tones = 1 kHz). Following Anstis & Saida (1985), Roberts et al. (2008) concluded that build-up is frequency-specific, and so failed to transfer from the induction sequence to the test sequence. It is interesting to note that the currently tested all(-2 ST) induction condition resulted in very little test-sequence segregation, and that this is largely consistent with Roberts et al.'s (2008) study which measured segregation after a much larger frequency change at the induction/test sequence boundary (a two octave change). This is in good accord with the study by Anstis & Saida (1985), which demonstrated that frequency changes to an on-going LH sequence all had a similar, and seemingly large, resetting effect, so long as the changes were greater than -1 to +3 ST from the 1-kHz centre frequency of the tone sequence.

One purpose of the current experiment was to compare directly the all(-2 ST) with the deviant(-2 ST) condition. Whilst a single changed induction tone can have a large resetting effect, it would appear that a single deviant tone will never fully reset build-up (at least for the variety of deviant tones tested in experiments 1 - 3b, including replacement by silence). In contrast, when all of the inducers were lowered by -2 ST, the extent of resetting was very large (0.85 - Table 6.1). Based on this finding it is clear that, whilst the properties of the final induction tone can have a substantial effect on subsequent streaming judgments, test-sequence segregation is not solely determined by the properties of the last inducer tone.

These findings are consistent with the hypothesis presented in the general discussion of Chapter 5. This hypothesis proposed that, for the current tonal arrangement, an induction sequence may promote segregation by capturing the L tones of the test sequence into the on-going induction-tone stream. This capturing effect may increase the likelihood of hearing the second tone subset (i.e., the H tones) as segregated from the established stream (see also Bregman & Rudnicky, 1975; Rogers & Bregman, 1993). The findings of the current experiment are generally compatible with this explanation. If we consider the all(-2 ST) case, all of the induction tones consistently differ from any of the test-sequence tones (in terms of frequency). As there is a change in sequence properties at the induction/test sequence boundary, the induction-tone stream would not seamlessly continue into the test sequence (essentially, build-up fails to transfer; Roberts et al., 2008). In contrast, when only the final tone is lowered in frequency, the induction-tone stream would otherwise directly correspond to the L tones of the test sequence. Whilst a deviant tone may partially disrupt the continuity of the induction-tone stream into the test sequence, the fact remains that the induction-tone stream would otherwise strongly promote segregation. Based on this difference, one would not expect the resetting effect of a single deviant tone to match the general failure of build-up to transfer from an induction sequence comprising different tonal properties (in this case, in terms of frequency).

Cusack et al. (2004) proposed that resetting may be linked to attentional factors, so that when attention is switched to a new, or to an on-going tone sequence, the initial default perception is heard (i.e., integration). Based on this theory, a single deviant tone may trigger resetting by temporarily drawing attention away from the on-going sequence. If this theory were correct, one would predict that deviant-tone resetting is not directly influenced by the absolute properties of the deviant tone *per se*, but rather, that resetting will be primarily determined by the ability of the deviant tone to attract a listener's attention. This was tested in the current experiment in the constrained-random condition. For this condition, the frequencies of the first nine induction tones were randomized (range = ±2 ST, centred on 1 kHz), and the final induction tone was fixed at -2 ST. It was predicted that these induction tones would still form a coherent single stream (see French-St. George & Bregman, 1989) with a mean stimulation frequency of approximately 1 kHz. Importantly, because a listener would be unable to predict the precise frequency of each incoming tone, the final induction tone should not be heard as a deviation from an established pattern. Therefore, the attentional salience of the final induction tone should be reduced (even though the final inducer is identical to that presented in the deviant(-2 ST) condition). If resetting is linked to attentional factors (Cusack et al., 2004), then one would predict that little or no resetting would occur for the constrained-random induction condition. Whilst the constrained-random induction was somewhat effective at promoting streaming, significantly more segregation was reported for the standard induction case. Indeed, streaming judgments following the constrained-random induction closely resembled those for the deviant(-2 ST) condition. There are several factors which could account for this outcome, and these are considered in turn below.

The results show that the constrained-random condition promoted less build-up of the tendency for stream segregation than did the standard induction sequence. This may be because the constrained-random tones were not always heard as a single coherent stream. French-St. George & Bregman (1989) demonstrated that two tone subsets (L & H) segregated into two streams even though frequency changes were applied to each tone within a subset (for which semi-random frequency changes occurred within a 3-ST range). For the current study, the constrained-random frequency variations occurred within a slightly broader range (4 ST), and so may have been heard occasionally as segregated. However, 4 ST was the *maximum* possible difference in frequency between two successive tones, and because of the method of randomization used, such large jumps in frequency would be unlikely to occur often. Therefore, frequency changes corresponding to values above the fission boundary (van Noorden, 1975) would be relatively infrequent. For this reason, one would not expect a strong perception of segregation during any given constrained-random induction sequence.

However, for the constrained-random condition, the frequency difference between the 9^{th} and the 10^{th} induction tones would have been larger, on average, than the mean frequency difference between the first nine inducers. This is because the frequency of the final inducer was fixed at -2 ST, which was the *very lowest* frequency value permitted by the randomization process applied to the initial nine tones (where the randomization range = ±2 ST, centred on 1 kHz). One may speculate that a relatively larger frequency change on average between the 9^{th} and the 10^{th} induction tones may have had some form of resetting effect - especially as this larger change would occur close to the induction/test sequence boundary (see Experiment 4a).

An alternative account for reduced build-up following the constrained-random induction sequence relates to the associated spread of peripheral excitation. The spread of excitation across a relatively large (4 ST) range may have weakened the segregation-promoting effect on the induction sequence. This point is especially relevant, as the spread of excitation exceeded the auditory filter bandwidth (where 1 $ERB_N \approx \pm 1$ ST, Glasberg & Moore, 1990). Indeed, Anstis & Saida (1985) demonstrated that build-up does not transfer between LH tone sequences which differed in centre frequency by more than 1 ERB_N (approximately). If an increased spread of stimulation frequency also acts to reduce the segregation-promoting effect of an induction sequence, this effect would be quite distinct from the observed active resetting effect of a single deviant tone. Therefore, the similarity in reported segregation following both the constrained-random and the deviant-tone induction conditions may be coincidental, and reflect two different processes. Specifically, it may be the case that the final inducer in the constrained-random case did cause less resetting, as it was less salient than in the deviant-tone case, but that this effect has been obscured by a roughly equal reduction in build-up arising from the greater spread of peripheral excitation in the former condition. This potential confound prevents a definitive answer to Cusack et al.'s (2004) hypothesis about resetting through attentional switching.

The spread of peripheral excitation for the constrained-random condition would also have had another important effect – namely resulting in an increased proximity between the induction tones and the test-sequence H-tones (as the inducers could be presented at frequencies above the standard L-frequency of 1 kHz). One might speculate that this may have reduced the likelihood of hearing test-sequence segregation, although a review of the literature suggests that this prediction has not previously been tested. However, any increased proximity would be *relative* to the H-tone frequency, as the inducer tones never extended above +2 ST from 1 kHz, whilst the test sequence tones were varied from +4 to +14 ST from 1 kHz. Based on this, one would predict that if increased inducer/H-tone proximity did reduce test-sequence streaming, then this should be most apparent for the smaller test-sequence frequency separations. As there was clearly no indication of any such interaction (see Figure 6.2), it would appear that any increased proximity between the inducers and the test-sequence H-tones did not affect reported streaming (at least for the current listening context).

Despite the greater spread of peripheral excitation, the constrained-random induction condition remained reasonably effective at promoting test-sequence segregation (in comparison with the no-induction case). Indeed, the induction tones for the all(-2 ST) condition fell within the same frequency range as the constrained-random tones, albeit at the lower edge, but very little build-up was observed for this condition. This finding would appear to indicate that mean stimulation frequency is an important determinant of the segregation-promoting effect of an induction sequence. For the standard induction case, the mean stimulation frequency was exactly 1 kHz, and this sequence was most effective at promoting streaming. For both the constrained-random condition, and the deviant(-2 ST) condition, the mean stimulation frequency of the induction sequence was close to 1 kHz. The issue of resetting aside, both of these sequences induced a comparable extent of streaming in the test sequence. For the all(-2 ST) condition, the mean stimulation frequency was 891 Hz - much lower than for all of the other induction conditions. Based on these findings, one might speculate that the ability of the induction tones to capture the test-sequence L-tones into an on-going stream is influenced by the similarity in mean stimulation frequency between the inducers and the test-sequence L-tones.

Chapter 7

The Effect of Physical and Perceived Continuity on the Build-up of Stream Segregation

7.1 Introduction

Generally, an induction sequence is most effective at promoting test-sequence segregation when the properties of the induction tones match those of a subset of test-sequence tones (in the case of experiments 1 - 5, the standard induction tones were identical to the L tones of the test sequence). If the rate of tonal onsets is not equal for the induction sequence and the corresponding tonal subset of the test sequence, the segregation-promoting effect of the induction sequence may be reduced. This was demonstrated by Rogers & Bregman (1993), who manipulated the tone rate (number of tone onsets) in a same-frequency induction sequence of constant duration. These authors found that induction sequences comprising slower tone rates (i.e., fewer tone onsets) became increasingly less effective at promoting streaming. Faster tone rates were no more effective at promoting stream segregation than a tone rate matched to the H subset of test sequence tones. Rogers & Bregman (1993) also measured streaming after a continuous induction tone (which for the purpose of distinguishing clearly the conditions used here in Experiment 6b will be referred to as an *extended* induction tone). They found that an induction sequence comprising a single extended tone (duration = 4.8 s) that ended immediately before the onset of the first test sequence tone did not promote any test-sequence segregation. Roberts et al. (2008) reported a similar finding using a temporal-discrimination measure of streaming. Given that a single extended tone (i.e., 100% tone density) was ineffective at promoting stream segregation, it is apparent that tone onsets are crucial for build-up to occur, and that an induction sequence promotes optimal segregation when there is an approximate matching of tone onsets and density per unit time.

Experiments 6a and 6b used Rogers & Bregman's (1993) finding that an extended induction tone does not promote streaming to explore the effect of perceived continuity on the build-up of stream segregation. Perceived continuity (or the continuity illusion) refers to the perception of a sound as continuing through a louder masking sound, despite the physical absence of the original sound (Miller & Licklider, 1950; Thurlow, 1957). The continuity illusion has been demonstrated in a wide variety of acoustic stimuli, such as speech (Warren, 1970), or tonal stimuli (e.g., Dannenbring, 1976; Warren, 1982; Ciocca & Bregman, 1987). An example of perceived continuity is illustrated in Figure 7.1. In panel A, there are two identical tones, A(1) and A(2), which are separated by a silence. A listener would hear this stimulus as two discrete, successive tones. In panel B, the properties of the two tones remain the same, but the inter-tone silence is filled with a louder burst of noise (B) which is centred on the same frequency as the tones. In this circumstance, the A(1) tone is heard to continue through the noise and into A(2) - so that the listener perceives a single, continuous A tone (even though the A tone is not physically present throughout the duration of the noise). In such an arrangement, the louder noise-burst masks the offset of A(1), and the onset of A(2). Warren (1982) argued that the continuity illusion reflects a perceptual compensation for masking - that is, the auditory system aids perception by restoring the masked portion of the tone. The role of masking in perceived continuity is considered further later in this introduction.

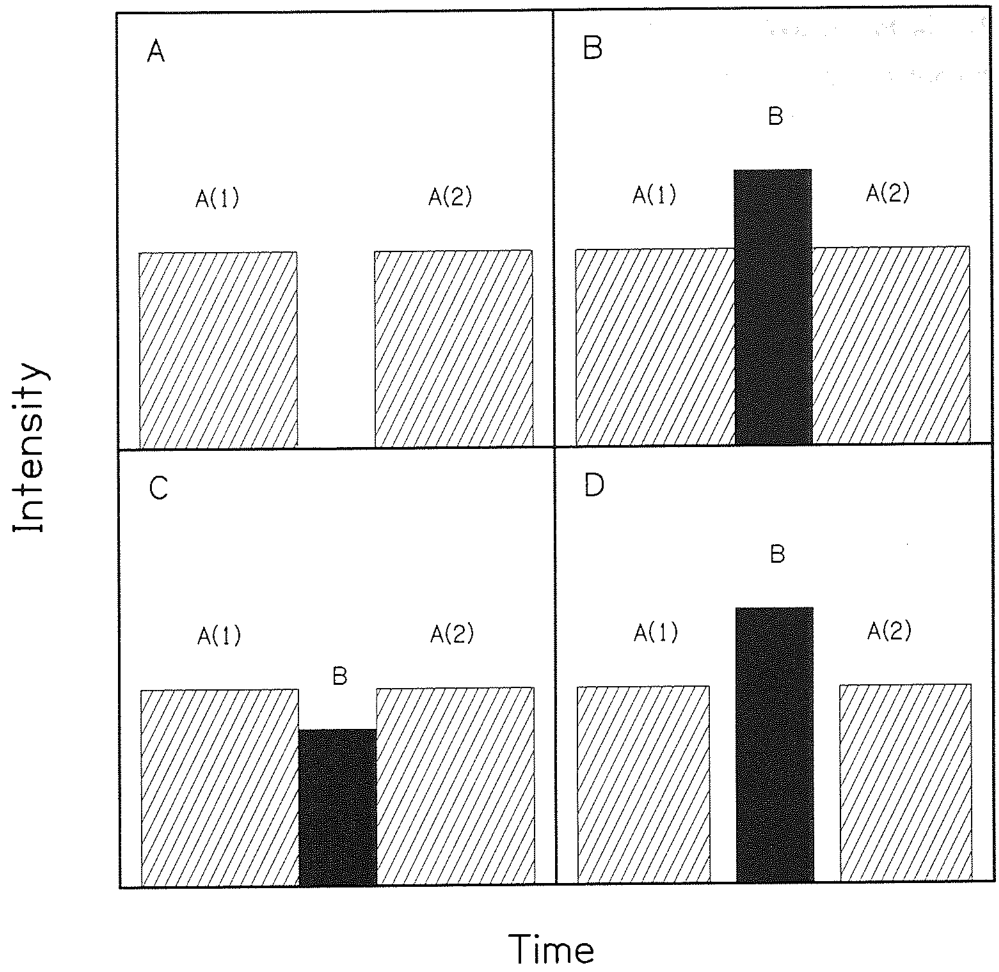

Figure 7.1: **Illustration of perceived continuity (adapted from Bregman, 1990).** Panel (A): **two discrete, identical tones are heard, A(1) and A(2).** Panel (B): **The silence between the tones is filled with a louder noise, B - a listener would perceive the tonal stimulus to continue through the noise, despite being physically absent.** Panel (C): **The noise, B, is quieter than the tones - the perception of continuity is lost.** Panel (D): **The tones and the noise are separated by brief silences - the tones are not heard to continue through the noise.**

Bregman (1990) extended Warren's (1982) interpretation in relation to auditory scene analysis. He suggested that in an ordinary listening environment, it is common for a brief loud sound to interrupt temporarily an on-going quieter sound. In such a circumstance, it would be incorrect to assume that the quieter sound had ceased during the louder sound, especially if the initial sound is perceived to continue after the finish of the louder sound. However, it would not be appropriate to perceive continuity when there is evidence that the initial sound did not continue through the louder sound. From this perspective, Bregman proposed four fundamental rules which govern the perception of continuity:

1. The masking rule. During the louder sound (B), there should be some neural activity that is indistinguishable from that which would occur if the quieter sound (A) were actually present. For example, if the B sound is remote in frequency

from the A sound, a listener will not perceive continuity (Warren et al., 1972). In this circumstance, during the B sound, there is no neural evidence which could be interpreted as representing the continuing A sound. Similarly, if the A sounds are louder than the B sounds, then the perception of continuity is lost. An example of this is illustrated in panel C of Figure 7.1.

2. There should be no evidence that the B sound is simply filling in a silence between A(1) and A(2). Continuity will not be perceived if there is evidence that A(1) actually ceases during the presentation of B. For example, if a brief silence is inserted between the A and B sounds, the A sounds are not perceived as continuous (Warren et al., 1972, see panel D of Figure 7.1).

3. For the perception of continuity, the A(1) and A(2) sounds should be perceived as arising from the same acoustic event. In other words, if the two A sounds were separated by a silence (rather than by the B sound), the A sounds would be heard as integrated into the same perceptual stream (see Ciocca & Bregman, 1987; Bregman et al., 1999).

4. The change from A to B should not be heard as A(1) gradually transforming into B and then back to A(2). The B sound must be heard as an interruption of the A sound, as opposed to a smooth change of a single on-going sound. This is related to the idea that abrupt changes signal new events, whereas gradual ones signal changes in an on-going event.

As previously discussed, Warren (1982) argued that the continuity illusion reflects a perceptual compensation for masking. More recent research has examined the neural basis of this effect. Using extracellular recording, Petkov et al. (2007) measured neural responses in the primary auditory cortex (A1) of awake macaque monkeys (note that the continuity illusion has previously been demonstrated in monkeys using behavioural measures - Miller et al., 2001; Petkov et al., 2003). Petkov et al. (2007) identified three populations of neurons, those which responded to tone onsets, those which responded to offsets, and those which responded to a sustained tone. To summarize briefly their complex study, in two conditions these authors presented two pure tones (tone duration = 172 ms, level = 45 dB SPL, tone frequency = set to the best frequency of the recorded activity), separated by a brief (56 ms) silence or a broadband noise. When the noise was present, the tones should have been perceived as continuous (as the stimuli were identical to those previously studied by Petkov et al., 2003). A single continuous tone was presented in a control condition (tone duration = 400 ms). For the noise condition, the three groups of neurons studied all responded as if the tone were physically present throughout the noise; the offset responders did not encode the offset of the first tone, the sustained responders continued to respond as if the tone continued through the noise, and the onset responders did not encode the onset of the second tone. Petkov et al. (2007) interpreted these findings as evidence to support two of Bregman's (1990) rules of perceived continuity; namely that (1) during the noise, there should be some neural activity which is indistinguishable from that which would have occurred if the tone had actually continued, and (2) there should be no evidence that the noise is simply filling in a silence between the two tones.

The current experiments were designed to examine the effect of perceived continuity during an induction sequence on the perception of a subsequent test sequence. It would appear that no previous study has examined the effect of perceived continuity on the sequential grouping of a subsequent sequence of tones. Given that a perceptually continuous tone appears to be neurally encoded in the same manner as an actually continuous tone (Petkov et al., 2007), one would expect perceived and actual continuity to have a similar effect on *subsequent* streaming judgments (when both are heard in the same listening context). This hypothesis was tested by exploiting Rogers & Bregman's (1993) finding that an extended induction tone is ineffective at promoting test sequence segregation. An induction sequence of six, standard tones (cf. experiments 1-5) would be expected to promote streaming, but if these tones were perceived to form a single extended tone, then this segregation-promoting effect should be lost.

7.2 Experiment 6a

7.2.1 Method

The general method and procedure for Experiment 6a is described in detail in Chapter 2. On each trial, participants heard a combination of an induction sequence and a subsequent test sequence (test sequence = 3 LHL- triplets, L tone = 1 kHz, H tone = +4 to +14 ST, tone duration = 100 ms, inter-triplet silence = 100 ms). Participants indicated their perception of the final LHL- triplet as either integrated or segregated in a 2AFC procedure.

Participants

Eight participants took part in Experiment 6a. Five of these listeners had previously taken part in one or more of the previous subjective experiments. All listeners reported normal hearing.

Induction Conditions

The standard induction sequence and the no-induction sequence were used in this experiment, but the standard induction sequence was slightly modified from that described in Chapter 2. Rather than comprising 10 L- repetitions, the current case comprised *6* L- repetitions. Other than this difference, all other sequence properties were identical to those described in Chapter 2 (L tone = 1 kHz, tone duration = 100 ms, inter-tone silence = 100 ms; see the first panel of Figure 7.2). Therefore, the total duration of the standard induction sequence was 1200 ms. This change was made because the current experiment tested an induction condition comprising a single, extended L tone (described below). Rogers & Bregman (1993) tested a similar condition, and noted that participants found listening to an extended tone a monotonous and unpleasant experience. Therefore, the duration of all induction sequences was reduced to alleviate this experience of monotony during the extended condition. This alteration of the standard induction condition is not problematic, because Experiment 4b demonstrated that a six-tone induction sequence remained highly effective at promoting test-sequence segregation.

The final control condition measured the segregation-promoting effect of a single extended induction tone. This inducer comprised one long L tone which terminated 100 ms before the onset of the test sequence (tone duration = 1100 ms), and is illustrated in the second panel of Figure 7.2. The duration of this extended tone was briefer than the 4.8- and 2.0-s tones tested by Rogers & Bregman (1993) and Roberts et al. (2008), respectively. However, a pilot study confirmed that the current extended tone was similarly ineffective at promoting test-sequence segregation.

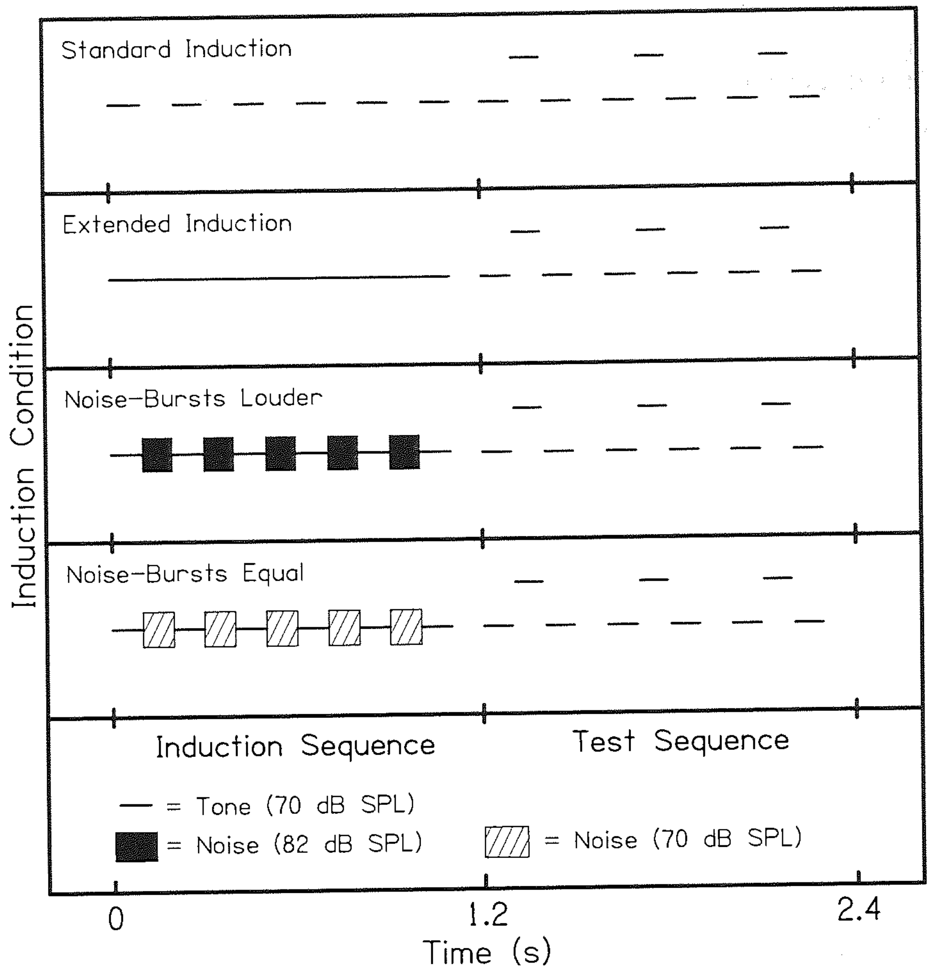

Figure 7.2: **Illustration of the induction conditions tested in Experiment 6a (the no-induction condition is not shown). Each panel displays a different induction sequence paired with a subsequent test sequence. The solid lines represent pure tones, and the shaded boxes represent band-pass filtered white noise. The properties of these stimuli are briefly described in the bottom panel of the figure.**

Two induction conditions were included specifically to measure the effect of perceived continuity on streaming judgments in the test sequence. For both of these conditions, the standard induction sequence was modified so that the silences between induction tones were replaced with band-pass filtered noise. As for the extended-induction condition, the silent gap between the final induction tone and the first test-sequence tone was preserved. If the induction tones were perceived to continue through the noise bursts, a listener should perceive a tone similar to the extended inducer (i.e., 1100-ms in duration). In contrast, if the tones are not perceived as continuous, the induction sequence should be heard as a six discrete tones which are separated by noise bursts. In this circumstance, it was predicted that the tonal portion of the induction sequence would behave like the standard case. The two experimental induction conditions manipulated whether the listener perceived the tones to continue through the noise-bursts or not. This was achieved by altering the overall level of otherwise identical noise-burst stimuli in relation to masked threshold.

The noise used was flat-spectrum and band-pass filtered with a bandwidth of twelve semitones centred on 1 kHz (range = 707 - 1414 Hz, spectral roll-off = 80 dB per octave). The noise was digitally created by combining equal-amplitude sinusoids of random starting phases. These sinusoids were distributed at 2-Hz intervals across the bandwidth of the noise. The noise had 10-ms raised cosine ramps at onset and offset. For the noise-bursts trials, an identical 100 ms noise-burst was repeated five times. As all stimuli were pre-recorded, this 'frozen' noise-burst was unchanged between trials and conditions (except for the 12 dB attenuation). For the noise-bursts-louder condition, the noise was presented at 82 dB SPL. Informal listening confirmed that, for this arrangement, the tones were perceived to continue through the noise. For the noise-bursts-equal condition, the noise was presented at 70 dB SPL (i.e., the same level as for the tones). In this circumstance, perceived continuity was not heard.

In summary, it was predicted that the standard induction sequence would strongly promote stream segregation of the test sequence, but that the extended induction sequence would promote little or none (Rogers & Bregman, 1993; Roberts et al., 2008). When the six induction tones were perceived to continue through the interleaved noise-bursts, the equivalent of a single extended induction tone should be heard. In this case, reported segregation was expected to be similar to that following the actual extended induction sequence (i.e., no build-up). In contrast, when the tones were not perceived to continue through the noise, six discrete induction tones should be heard. Therefore reported segregation for this condition was predicted to resemble that after the standard induction sequence (i.e., strong build-up should occur).

Procedure and Apparatus

The procedure and apparatus used were identical to that described in Chapter 2. Each trial comprised a combination of an induction sequence and a subsequent test sequence. These trials were organized into trial blocks, each block comprising a combination of the five induction conditions and the six frequency separations for the test sequence (i.e., 30 trials). For each listener, the main experiment comprised 20 trial blocks (i.e., 600 trials in total).

7.2.2 Results

The results were averaged to give a single measure of the percentage of trials which were heard as segregated for each combination of induction condition and frequency separation. These mean values are shown in Figure 7.3. The mean percentages reported as segregated for frequency separations of 4, 6, 8, 10, 12, and 14 ST were 6.6%, 18.9%, 33.1%, 54.8%, 72.5%, and 91.3%, respectively. The means for the standard, no-, extended, noise-bursts-louder, and noise-bursts-equal induction conditions were 73.9%, 32.9%, 37.3%, 36.7%, and 50.2%, respectively. These data were analysed using a two-way, repeated-measures ANOVA. The two independent variables were the six frequency separations for the test sequence (4, 6, 8, 10, 12, or 14 ST), and the five induction conditions (standard induction, no induction, or the three experimental induction conditions). The ANOVA confirmed significant main effects of frequency separation [$F(5,35) = 54.439$, $p < 0.001$, $\eta_p^2 = 0.886$], and of induction condition [$F(4,28) = 6.616$, $p < 0.001$, $\eta_p^2 = 0.766$]. Once again, the interaction between these two variables was also significant [$F(20,140) = 5.366$, $p < 0.001$, $\eta_p^2 = 0.434$].

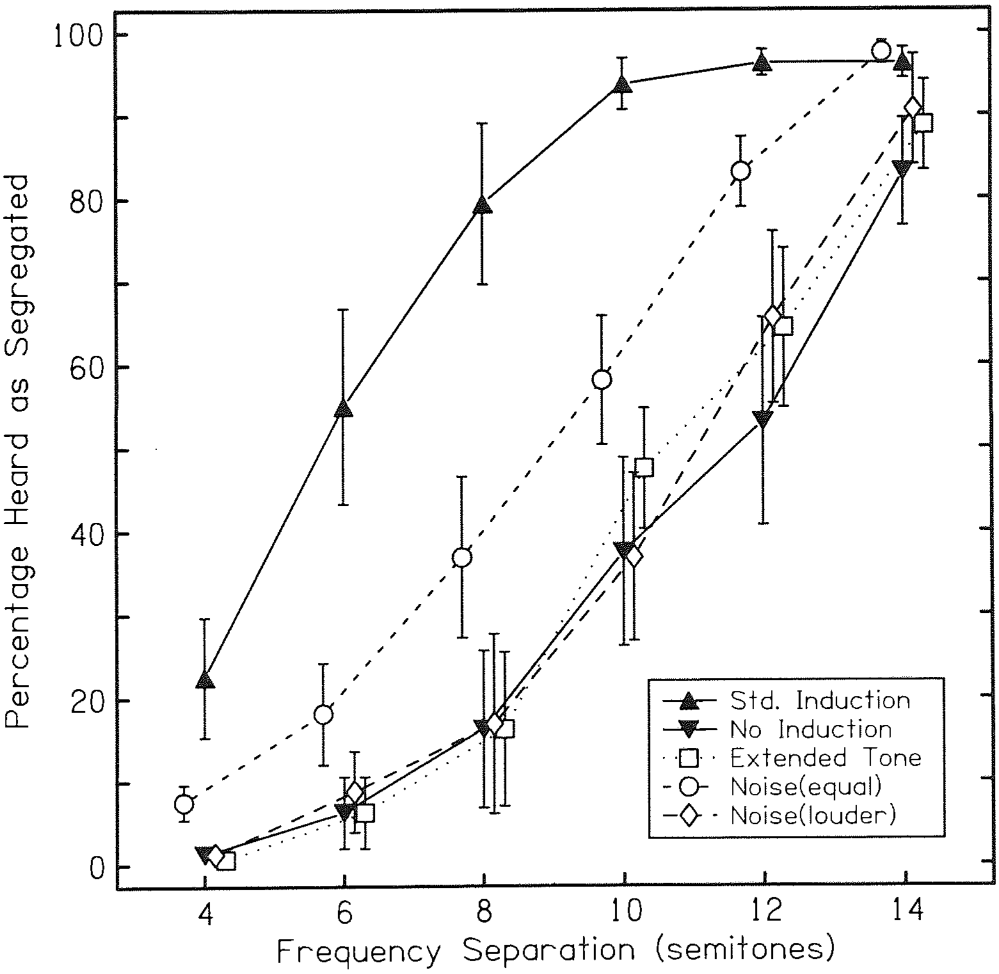

Figure 7.3: **Results from Experiment 6a. The effects of induction condition and frequency separation in the test sequence on reported segregation. Each point represents the percentage of trials which were reported as segregated (averaged across 8 listeners). The insert identifies the different induction conditions tested. The error bars display ±1 inter-subject standard error.**

As previously, the standard and no-induction conditions promoted the greatest and least stream segregation, respectively (see Figure 7.3). For all of the other experimental conditions, reported segregation was reduced in comparison to standard induction case.

For Table 7.1, the data from the experimental conditions are transformed into a measure of build-up which is relative to both the standard and no-induction cases. For this table, the 4- & 14-ST frequency separation cases were not analysed, and a weighting system was applied when calculating the across-frequency mean values. The weights were calculated so that conditions in which there was a smaller difference between reported segregation in the standard and in the no-induction conditions made a reduced contribution to the across-frequency-separation mean. This procedure is consistent with that used for the previous experiments, and is explained in detail in the results section of Experiment 1 (Chapter 3). The figures quoted in Table 7.1 represent a proportional shift on a scale defined by the no-induction condition representing no prior build-up (0.00), and the standard induction condition representing a maximum segregation-promoting effect (1.00).

Frequency Separation	Induction Condition: Extended	Noise-bursts louder	Noise-bursts equal	*Mean*
6 ST *(23.08%)*	0.00	0.05	0.24	***0.10***
8 ST *(29.88%)*	0.00	0.01	0.33	***0.11***
10 ST *(26.63%)*	0.17	-0.01	0.37	***0.18***
12 ST *(20.41%)*	0.26	0.29	0.70	***0.42***
Weighted Mean	***0.10***	***0.07***	***0.40***	***0.19***

Relative build-up = no induction (0.00) → standard induction (1.00)

Table 7.1: Results from Experiment 6a. Reported segregation was compared with that in the no-induction case (0.00). Values are expressed as a proportional increase in segregation, towards the standard induction case (1.00). Note that a (small) negative value indicates that reported segregation was less than that for the no-induction case. To provide a stable measure of relative build-up, only data from the 6, 8, 10 and 12 ST frequency separations are included, and the across-frequency means are weighted (see the main text). Note that individual values are displayed as unweighted.

Two-tailed pairwise comparisons were performed on the original (i.e., untransformed) data set using the restricted LSD test (Snedecor & Cochran, 1967). As with all of the previous subjective experiments, the data for the 4- and 14-ST frequency separation conditions were excluded from the pairwise analyses. From this reduced data set, each pairwise comparison is reported with a percentage measure of the difference in reported segregation between the two conditions. There was a significant difference between the standard and the no-induction condition [difference = 41.0%, $t(7) = 6.867$, $p <$

0.001]. When compared with the standard induction case, stream segregation was significantly lower in all four of the experimental conditions; [extended induction: 36.6%, $t(7) = 7.845$, $p < 0.001$], [noise-bursts-louder: 37.2%, $t(7) = 6.152$, $p < 0.001$], and [noise-bursts-equal: 23.6%, $t(7) = 5.578$, $p < 0.001$]. Indeed, neither the extended induction nor the noise-bursts-louder conditions were significantly different from the no-induction case [extended: 4.3%, $t(7) = 1.814$, $p > 0.05$], [noise-burst louder: 3.8%, $t(7) = 1.572$, $p > 0.05$]. In contrast, significantly more stream segregation was reported for the noise-bursts-equal condition than for the no-induction condition [17.3%, $t(7) = 2.841$, $p < 0.05$], although it should be acknowledged that the difference between the noise-bursts-equal and noise-bursts-louder conditions did not quite reach significance [13.5%, $t(7) = 2.176$, $p > 0.05$].

In summary, the results show that reported stream segregation was highest following the standard induction sequence. Neither the extended inducer nor the noise-bursts-equal inducer promoted any significant segregation (i.e., in comparison to the no-induction case). The noise-bursts-equal condition promoted significantly more stream segregation than the no-induction case, but reported segregation was still reduced compared with the standard case.

7.2.3 Discussion

The results of Experiment 6a are consistent with the known effect of tone-frequency separation on stream segregation (Miller & Heise, 1950; van Noorden, 1975). As for experiments 1 - 5, the standard induction sequence promoted the greatest amount of reported streaming and the no-induction condition promoted the least. As expected, the extended induction sequence failed to promote any appreciable stream segregation in the test sequence (i.e., streaming judgments were similar to those following the no-induction condition). This is consistent with two previous studies (Rogers & Bregman, 1993; Roberts et al., 2008), which both used similar stimuli and also found that an extended tone did not promote test-sequence segregation.

The aim of Experiment 6a was to examine the effect of perceived continuity during the induction sequence on streaming during the test sequence. This was measured in two experimental conditions. In the noise-bursts-louder condition, loud noise bursts were inserted between the six standard induction tones. In this arrangement, listeners should have perceived the tones to continue through the noise, and therefore heard a single tone, similar to the extended inducer. The results confirmed that reported segregation for this condition was indeed very similar to that for the actual extended induction condition. This finding would appear to indicate that both perceived and actual continuity have a broadly equivalent effect on subsequent perception - namely that neither promote stream segregation in the current configuration. A second condition was included to measure the effect of the interleaved noise-bursts in the absence of any perceived continuity (the noise-bursts-equal condition). As the tonal portion of this induction sequence should have been perceived like the standard induction sequence, it was expected that this condition would be effective at promoting streaming in the subsequent test sequence. The results were partially consistent with this argument - reported segregation for the noise-bursts-equal condition was significantly greater than for the no-inducer condition. However, substantially more stream segregation was reported for the standard induction condition than for the noise-bursts-equal condition. This indicates that the presence of the noise bursts affected streaming judgments even in the absence of perceived continuity.

One might speculate on two possible explanations for why streaming was reduced in the noise-bursts-equal condition in comparison with the standard induction case. First, the noise bursts may have partially masked the tone onsets. Rogers & Bregman (1993) suggested that the number of tone onsets was an important factor in determining the build-up occurring during an induction sequence. Therefore, any partial masking of the tone onsets may in itself have reduced the segregation-promoting effect of the induction sequence. Second, during the noise-burst induction sequence, one would hear a regular, alternating pattern (tone-noise-tone...). However, the final induction tone was followed by a silent interval rather than by a noise burst. It may be that this abrupt change from the established pattern in itself had a partial resetting effect. It has been speculated that a shift in attentional focus may trigger a resetting of build-up (Cusack et al., 2004), and it has been demonstrated that a single changed tone can have a resetting effect, including replacement with silence (e.g., Experiment 1). Based on this, it is possible that a single change applied to the inter-tone interval (as opposed to a tone itself) at the inducer/test boundary may also have a similar resetting effect. These two hypotheses were addressed in Experiment 6b, which investigated the effect of perceived continuity in a different listening context.

For the extended inducer, there was a brief silence between the end of the inducer and the first test-sequence tone. A series of pilot studies examined the effect of "filling in" this silence with the on-going induction tone (i.e., by merging the inducer and the first test-sequence tone into a single, *fully continuous* tone). Following this inducer, there was a clear tendency for listeners to report increased test-sequence segregation (in clear contrast to perception after the extended inducer). The reason for this difference is somewhat unclear, but some speculative explanations are proposed in the current discussion (section 7.3.3). Regardless of this issue, the segregation-promoting effect of a fully continuous tone allowed for an exploration of the role of perceived continuity on test-sequence streaming in a new listening context. If the silence following an extended tone was filled by loud noise, a single, fully continuous tone should be heard (as the inducer would be perceived to continue through the noise). In this circumstance, reported perception should be similar to that for the fully continuous condition. In contrast, if the tone is not heard to continue through the noise, it should be heard like an extended inducer of the form used here (i.e., as terminating before the onset of the first test-sequence tone) and behave accordingly.

7.3 Experiment 6b

7.3.1 Method

The method and procedure for Experiment 6b is fully described in Chapter 2, and was also identical to that used in Experiment 6a.

Participants

Eight participants took part in Experiment 6b. Two of these listeners had previously taken part in Experiment 6a. All participants reported normal hearing.

Induction Conditions

Three of the six induction conditions tested in Experiment 6b were identical to their counterparts in Experiment 6a. These were the standard induction condition (six L-repetitions, tone frequency = 1 kHz, tone duration = 100 ms, inter-tone silence = 100 ms, overall induction sequence duration = 1200 ms), the no-induction condition (test sequence presented without any prior induction tones), and the extended induction sequence (one 1100-ms L tone followed by a 100-ms silence). The standard and the extended induction conditions are illustrated in the top two panels of Figure 7.4.

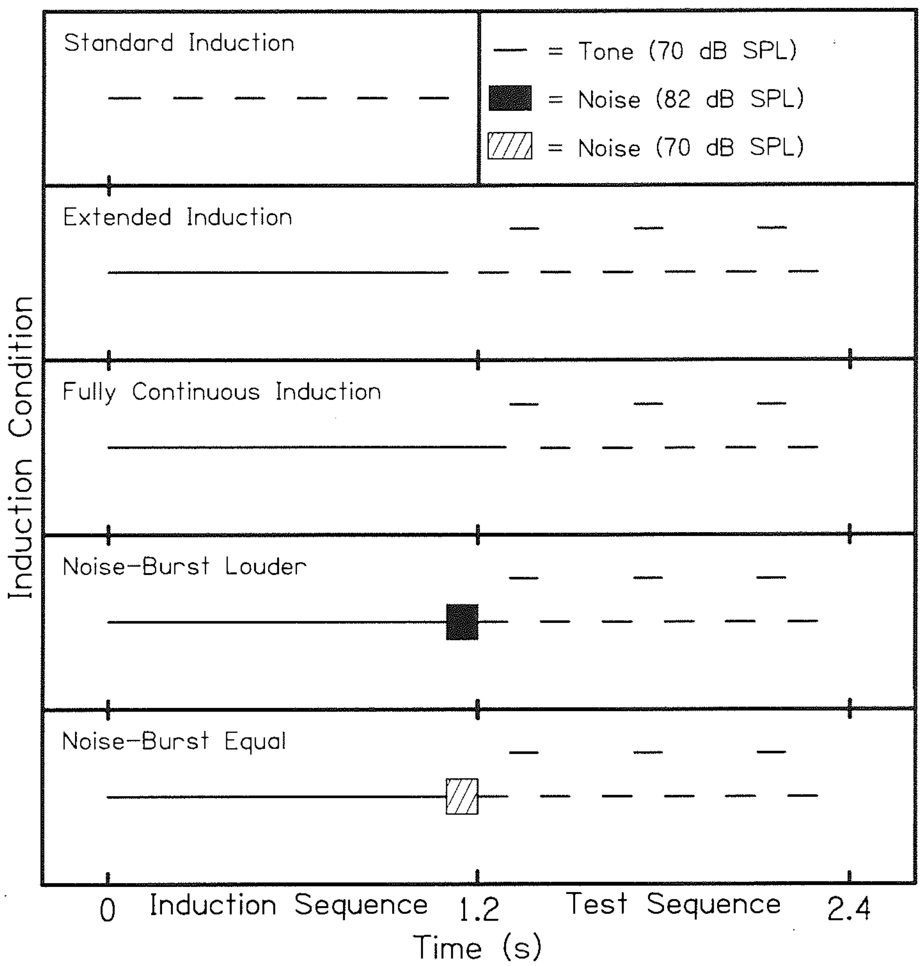

Figure 7.4: **Illustration of the induction conditions tested in Experiment 6b (the no-induction condition is not shown). Each panel displays a different induction sequence paired with a subsequent test sequence (except for the standard induction condition which is displayed without a subsequent test sequence). The solid lines represent pure tones, and the shaded boxes represent band-pass filtered white noise. The properties of these stimuli are briefly described in the top-right panel of the figure.**

A fourth induction condition comprised a single, fully continuous L tone. This was created by extending backwards the first test-sequence L tone by 1200 ms so that there was no intervening silence at the (arbitrary) inducer/test sequence boundary (i.e., tone duration = 1300 ms in total). This novel stimulus arrangement is illustrated in the third panel of Figure 7.4. Note that here the test sequence effectively begins with an H-tone. The final two conditions were designed to measure the effect of perceived continuity on judgments of test-sequence streaming. Both of these conditions were modifications of the extended induction sequence, in which the 100-ms silent interval following the inducer was filled with band-pass filtered noise. The noise-burst-louder condition was designed so that a listener would hear the inducer to continue through the noise-burst, resulting in the induction tone being perceived as fully continuous. In the final condition, the noise burst was presented at the same level as the tones (noise-burst-equal), so that the inducer tone should not be heard to continue through the noise. The last two conditions are illustrated in the bottom two panels of Figure 7.4.

The noise bursts were identical to those used in Experiment 6a. For the noise-burst-louder condition, the steady-state portion of the noise was presented at 82 dB SPL. For the noise-burst-equal condition, the noise was presented at 70 dB SPL. Informal listening confirmed that perceived continuity was heard in the noise-burst-louder condition, but not in the noise-burst-equal condition.

Procedure and Apparatus

The procedure and apparatus used for the current experiment were fully described in Chapter 2. To summarize, trials were organized into blocks comprising 36 trials - a combination of each of the six induction conditions with each of the six test-sequence frequency separations. For the main experiment, each participant was presented with 20 trial blocks in total (i.e., 720 trials).

7.3.2 Results

Responses from all listeners were averaged to give an overall indication of reported streaming for each separate induction condition at each different test-sequence frequency separation (i.e., the percentage heard as segregated). These mean values are shown in Figure 7.5. The mean percentage reported as segregated for frequency separations of 4, 6, 8, 10, 12, and 14 ST, were 4.8%, 23.0%, 54.3%, 75.2%, 88.5%, and 94.3%, respectively. The means for the standard, no-, extended, fully continuous, noise-burst-louder, and noise-burst-equal induction conditions were 71.5%, 45.9%, 49.7%, 64.2%, 57.7%, and 51.1%, respectively. These data were analysed using a two-way, repeated-measures ANOVA. The two independent variables were the six frequency separations for the test sequence (4, 6, 8, 10, 12, or 14 ST), and the six induction conditions (standard induction, no induction, or four experimental conditions). The ANOVA found a significant main effect of frequency separation [$F(5, 35) = 123.670$, $p < 0.001$, $\eta_p^2 = 0.946$] and of induction condition [$F(5, 35) = 12.710$, $p < 0.001$, $\eta_p^2 = 0.645$]. The interaction between these two variables was also significant [$F(25, 175) = 4.642$, $p < 0.001$, $\eta_p^2 = 0.399$]. Once again, this interaction was probably driven primarily by the floor and ceiling effects at the two extreme frequency separations.

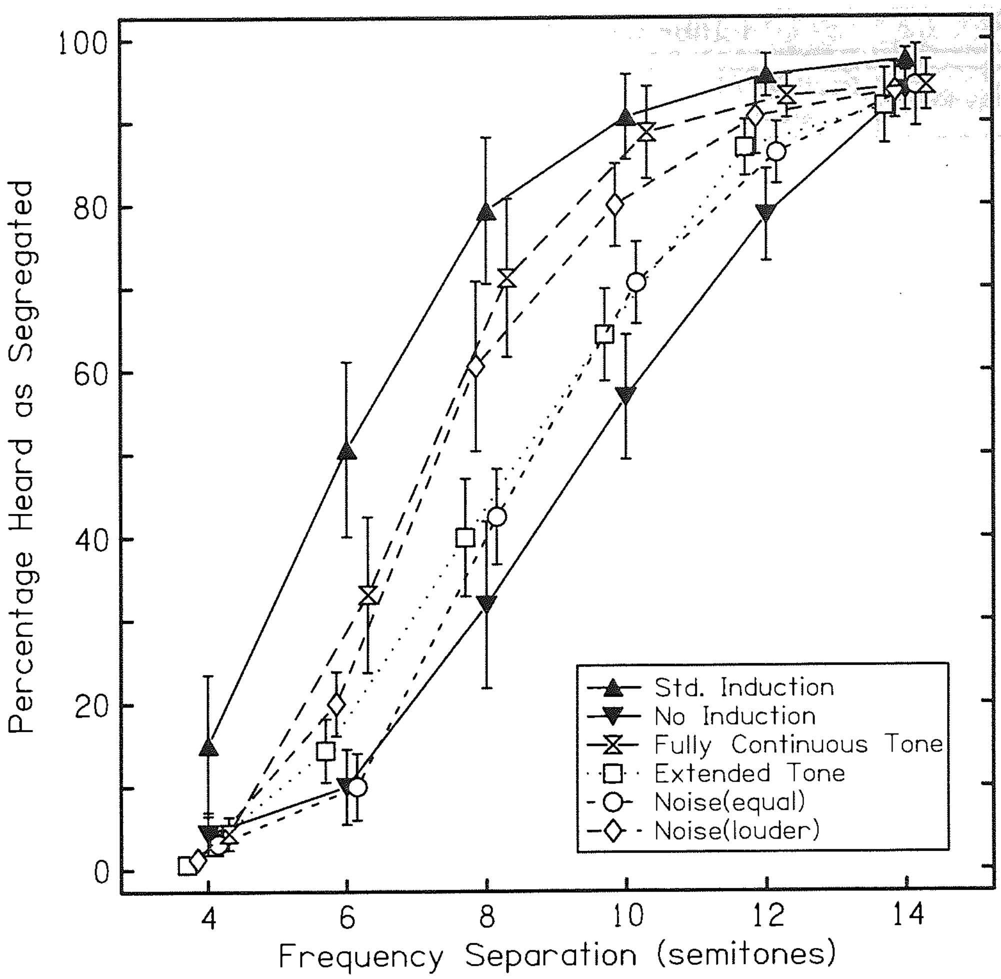

Figure 7.5: **Results from Experiment 6b. The effects of induction condition and frequency separation in the test sequence on reported segregation. Each point represents the percentage of trials which were reported as segregated (averaged across 8 listeners). The insert identifies the different induction conditions tested. The error bars display ±1 inter-subject standard error.**

Consistent with the previous experiments, the standard and the no-induction conditions promoted the greatest and least reported stream segregation, respectively (see Figure 7.4). For each of the four experimental conditions, reported segregation was reduced in comparison with the standard induction case. For Table 7.2, the data from the experimental conditions are transformed into a measure of build-up which is relative to both the standard and no-induction cases (see the results section of Experiment 6a, or Table 7.1).

Frequency Separation	Induction Condition: Fully continuous	Extended	Noise-burst louder	Noise-burst equal	*Mean*
6 ST *(29.28%)*	0.57	0.11	0.25	0.00	***0.23***
8 ST *(34.23%)*	0.83	0.17	0.61	0.22	***0.46***
10 ST *(24.32%)*	0.94	0.22	0.69	0.41	***0.57***
12 ST *(12.16%)*	0.85	0.48	0.70	0.44	***0.62***
Weighted Mean	***0.78***	***0.20***	***0.54***	***0.23***	***0.44***

Relative build-up = no induction (***0.00***) → standard induction (***1.00***)

Table 7.2: **Results from Experiment 6b. Reported segregation was compared with that in the no-induction case (0.00). Values are expressed as a proportional increase in segregation, towards the standard induction case (1.00). To provide a stable measure of relative build-up, data from the 4 and 14 ST frequency separations are excluded from this analysis, and the across-frequency means are weighted (see Chapter 3). Each individual value is displayed as an unweighted measure of relative build-up.**

Two-tailed pairwise comparisons were performed on the original data set (see Figure 7.5) using the restricted LSD test (Snedecor & Cochran, 1967). As floor and ceiling effects were observed for the 4- and 14-ST frequency separation conditions, these data were excluded from the pairwise analyses. When compared with the standard induction condition, reported stream segregation was significantly lower in four of the five other cases ($p < 0.05$). The one exception was the fully continuous condition, for which reported segregation approached that for the standard inducer [difference = 7.5%, $t(7) = 2.280$, $p > 0.05$]. There was a significant difference between the standard and the no-induction conditions [34.7%, $t(7) = 4.454$, $p < 0.005$]. Although reported segregation for the noise-burst-louder condition was significantly different from that for all the other induction conditions, the results most closely resembled those for the fully continuous condition and the difference in means for this comparison was small [8.8%, $t(7) = 4.250$, $p < 0.005$]. Therefore, it is reasonable to conclude that the fully continuous and noise-burst-louder conditions each had a roughly similar and quite substantial segregation-promoting effect. In contrast, far less reported segregation was observed for both the extended condition and the noise-burst-equal condition. When compared with the no-induction case, significantly more segregation was reported for the extended inducer [7.0%, $t(7) = 2.447$, $p < 0.05$] but not for the noise-burst-equal condition [8.0%, $t(7) = 2.121$, $p > 0.05$]. However, when compared with each other, there was no significant difference between the extended and noise-burst-equal conditions [1.0%,

$t(7) = 0.479$, $p > 0.05$]. Therefore, it is concluded that neither the extended nor the noise-burst-equal conditions were effective at promoting test-sequence streaming.

In summary, the fully continuous inducer was almost as effective at promoting stream segregation as the standard case, whereas the extended inducer was considerably less effective. Moreover, this substantial difference was significant [20.2%, $t(7) = 3.398$, $p < 0.05$]). When compared with these two conditions, reported segregation for the noise-burst-louder condition most closely resembled that for the fully continuous case, whilst for the noise-burst-equal condition it most closely resembled that for the extended case (see also the relative build-up measure displayed in Table 7.2).

7.3.3 Discussion

The results of Experiment 6b are consistent with the known effect of tone-frequency separation on stream segregation (van Noorden, 1975). The standard induction sequence promoted the greatest amount of streaming, whilst the least segregation was reported in the no-induction condition. This is also consistent with earlier studies (experiments 1 - 6a; Rogers & Bregman, 1993), and is evidence that the tendency to hear stream segregation built up considerably over the course of the standard induction sequence. The extended induction condition promoted little test-sequence streaming; a trend which is generally congruent with Experiment 6a, and with previous studies (Rogers & Bregman, 1993; Roberts et al. 2008). However, it should be noted that the extended inducer did promote significantly more segregation than the no-induction case. This result contrasts with Experiment 6a, which measured exactly the same conditions but found no significant difference in reported segregation. In contrast to the significant, but relatively marginal segregation-promoting effect of the extended inducer, the fully continuous inducer was significantly more effective at promoting subsequent test-sequence streaming. Indeed, reported segregation following the fully continuous inducer was not statistically different from that following the standard induction case.

Whatever the exact basis for the different effects on stream segregation of the extended and fully continuous inducers (see the next paragraph), Experiment 6b was intended primarily to examine the effect of perceived continuity on test-sequence streaming. In two conditions, the silence at the end of the extended induction sequence was filled with an equivalent duration noise burst. When this noise was presented at a level above that of the tones, so that it was loud enough to mask the tone had it continued, it was predicted that listeners would perceive the extended tone to continue through the noise, and so hear the sequence as similar to the fully continuous condition. This prediction was only partially supported by the results. Although reported segregation following the noise-burst-louder condition was most similar to that following the fully continuous condition, there remained a significant difference between these two conditions. In a final condition, the noise burst was presented at an equal level to the tones, which was insufficient to support perceived continuity. For this condition, the tonal portion of the induction sequence should have been perceived similarly to the extended inducer. The results matched this prediction - the noise-burst-equal condition was neither significantly different from the extended inducer condition, nor from the no-induction condition. Given that significantly more segregation was reported for the noise-burst-louder condition than for the noise-burst-equal condition, from Experiment 6b it is concluded that perceived continuity promoted stream segregation in the current listening context.

It is unclear as to why the fully continuous inducer was so effective at promoting segregation, whilst the extended tone was less effective. Perhaps the most likely explanation relates to the fact that the fully continuous inducer also encompassed the first test-sequence tone, and so the properties of the test sequence were changed in comparison with all of the other conditions. Instead of three discrete LHL- triplets, the test sequence would have been heard as an HL- tone pair followed by only two LHL- triplets. Importantly, when an LHL- tone triplet forms an integrated percept, a salient gallop-like rhythm is heard. Even for a strongly integrated percept, the alteration applied to the first triplet may have reduced the salience of the galloping rhythm (as the test sequence comprised only three triplets in total). If listeners were using the perception of a "gallop" as the criterion to judge integration, the disruption of this percept may have encouraged listeners to report segregation.

There is a second potential explanation for the segregation-promoting effect of the fully continuous inducer. If we consider the extended inducer, there was a 100-ms silence between the offset of the inducer and the first test-sequence tone. It is possible that an extended tone *is* highly effective at promoting stream segregation, but that the subsequent brief silence has a resetting effect. There is some support for this argument, as active resetting can be triggered by a single change to a tone (Experiment 1), and this resetting effect is most evident when a change is applied to the final portion of the induction sequence (Experiment 4a). Furthermore, this resetting argument could be generalized to the previous studies which also found that an extended tone terminating before the onset of the test sequence failed to promote subsequent segregation (Rogers & Bregman, 1993; Roberts et al., 2008). Indeed, the 4.8-s extended tone tested by Rogers & Bregman (1993) was also presented at a lower level than the test sequence tones (extended tone = 59.5 dB SPL, test sequence tones = 62.5 dB SPL). Whilst this was only a modest difference (3 dB), it has previously been established that an abrupt level change can have a substantial resetting effect (Experiment 1; Rogers & Bregman, 1998; Roberts et al., 2008). However, there are several reasons to doubt the validity of this resetting explanation. Any reduction in the number of induction-tone onsets appears to result in less test-sequence segregation (Rogers & Bregman, 1993 - where induction sequences comprised 1 – 48 onsets). As this trend was generally observed across the range of conditions tested, one would not expect an induction sequence comprising only a single onset to have much, if any, segregation-promoting effect. Also, even if the extended inducer was effective at promoting segregation, the subsequent silence would have to produce a near-full resetting effect to account for the fact that reported segregation following the extended tone was largely similar to that for the no-induction case (see also Rogers & Bregman, 1993; Roberts et al, 2008). The results from experiments 1 and 3a suggest that a single change will not have such a complete resetting effect (where the change was applied to the final induction tone and the extent of resetting observed was approximately 0.45 on average in Experiment 1). Owing to these issues, the proposed resetting hypothesis does appear somewhat unlikely.

7.4 General Discussion

The current experiments were designed to investigate the effects of perceived continuity on the sequential organization of subsequent tones. The rules which govern perceived continuity are reasonably well understood (see Bregman, 1990), but to my knowledge, no previous study has examined how perceived continuity affects the perception of a subsequent tone sequence. Previous evidence suggests that a perceptually continuous tone is represented neurally in a similar way to an actually continuous tone (Petkov et al., 2007). Based on this evidence, it was predicted that both perceived and actual continuity should both have a very similar effect on subsequent streaming judgments. For Experiment 6a, listeners were presented with two tone-only induction sequences - a standard induction sequence and an extended induction sequence. To examine the effect of perceived continuity, the inter-tone silences between the standard tones were filled in with louder noise, so that the tone was heard as continuous (i.e., as an extended tone). It was found that neither the physically nor the perceptually extended tone promoted any segregation of the test sequence. This finding is taken as evidence that both perceived and actual continuity have the same effect on streaming judgments. However, when the noise bursts were not loud enough to promote perceived continuity, segregation was not fully restored to that of the standard induction case. This suggested that the mere presence of the noise bursts affected streaming judgments, even in the absence of perceived continuity. It was suggested that this effect may have been due to the partial masking of the tone onsets or, more speculatively, that the presence of the noise bursts between the inducer tones but not between the last inducer and the first test-sequence tone may have affected the dynamics of streaming by triggering resetting at the inducer/test boundary.

For Experiment 6b, it was found that a fully continuous tone was very effective at promoting stream segregation (in contrast to an extended tone which was considerably less effective). This difference allowed for the measurement of the effect of perceived continuity, by replacing the silence between the extended tone and the test sequence with a noise burst of equivalent duration. It was found that the interrupted tone that was perceived to be fully continuous promoted a somewhat similar extent of reported segregation as did the genuine fully continuous tone (although these two conditions were statistically dissimilar, with less segregation heard in the perceived continuity condition). Importantly, when the noise burst was equal in level to the tones (i.e., when no perceived continuity was heard), reported segregation was broadly equivalent to that for the extended induction sequence (and also for the no-induction sequence). This finding can be related back to Experiment 6a. For that study, the noise bursts which were presented at an equal level to the tones affected subsequent streaming judgments (reported segregation was reduced to some extent). It was speculated that this finding may have indicated that the noise bursts affected performance by partially masking the tone onsets. If this were true, one would also expect that the noise-burst-equal condition tested in Experiment 6b would have been dissimilar from the extended induction tone, due to a similar partial masking effect (as the properties of the tones and noise were identical between experiments). However, this was not the case, and so it would appear that the effect of the less intense noise bursts cannot be explained solely by the partial masking of tonal onsets. Therefore, the reduced streaming observed in Experiment 6a when the induction tones were interspersed with equal-level noise must have been due (at least in part) to some other effect of the noise bursts on the dynamics of streaming (e.g., active resetting).

Overall, it is concluded that perceived continuity, rather than the partial masking of tone onsets, is primarily responsible for the observed effects of filling the silent intervals between inducer tones with noise. In summary, it appears that perceived continuity can a similar effect on subsequent streaming judgments as does actual continuity. Specifically, perceived continuity can either promote integration or segregation, depending on the listening context.

Chapter 8

A Comparison of Build-up and Resetting in Same-Frequency and Alternating-Frequency Induction Sequences

8.1 Introduction

The previous experiments presented in this thesis have consistently shown that a same-frequency induction sequence (standard induction sequence = 10 L- repetitions) is highly effective at promoting stream segregation in a subsequent test sequence (3 LHL- triplets), and this pattern of results is in good accord with previous research (Rogers & Bregman, 1993; Beauvois & Meddis, 1997; Roberts et al., 2008). The primary aim of Experiment 7 was to compare directly the segregation-promoting effect of a repeating L-tone induction sequence with the build-up of stream segregation that occurs during an on-going LH, or LHL-, tone sequence (Bregman, 1978; Anstis & Saida, 1985). For the purpose of the current chapter, a repeating L-tone stimulus (where each tone shares identical properties) is referred to as a same-frequency sequence, whilst an LHL- arrangement is referred to as alternating-frequency sequence. Whilst these two types of sequence are each generally effective at inducing streaming in a subsequent test sequence (see Rogers & Bregman, 1993, 1998), their relative segregation-promoting effects may not necessarily be equivalent, and to my knowledge, no previous study has made this comparison.

For an alternating-frequency sequence, the rate of build-up is typically rapid during the first 5 - 10 seconds after the tone-sequence onset, but is more gradual thereafter (Bregman 1978; Anstis & Saida, 1985, see also Figure 1.4). Bregman's (1978, 1990) functional explanation of build-up proposed that, at the onset of an alternating-frequency sequence, the default perception is that both sets of sounds arose from the same acoustic event (i.e., integration). However, the two subsets of sounds have consistently different properties and, on the basis of this continued difference, the stream-formation mechanism becomes increasingly likely to favour segregation. This proposed evidence-accumulation process is relatively gradual, in order to prevent excessive fluctuations in the perceptual organization of the stimuli.

In contrast to an alternating-frequency arrangement, there are no competing perceptual organizations during a same-frequency sequence - instead, a single, integrated stream will always be heard. It has been proposed that a same-frequency induction sequence may promote test-sequence streaming by capturing one subset of tones into the on-going induction stream (Rogers & Bregman, 1993; current Chapter 5; see also Bregman & Rudnicky, 1975). In other words, at the onset of the test sequence an L-tone-only stream has already been formed, and this may cause the H tones to segregate from the established stream. Some evidence in support of this argument was provided by Experiment 4b, for which the number of L- repetitions during an induction sequence was varied. It was found that one L- repetition had a partial, but significant, segregation-promoting effect. Longer sequences of three, six, or ten L- repetitions all resulted in a large and broadly similar extent of reported stream segregation. This extremely rapid form of build-up contrasts strongly with the more gradual rate of build-up observed during an alternating-frequency tone sequence (Bregman, 1978; Anstis & Saida, 1985). The considerably faster build-up observed for the same-frequency induction sequence was taken as evidence that once the stream was established, the tendency to hear segregation did not increase any further when additional induction tones were present (due to the lack of competing perceptual organizations during the induction sequence).

If a same-frequency induction sequence does promote test-sequence streaming by capturing one subset of tones into an on-going stream, reported segregation may not necessarily be equivalent to the build-up occurring during an alternating-frequency sequence. For an alternating-frequency sequence, there are usually two competing perceptual organisations - either integration or segregation. This is generally true, except for a sequence with an extreme frequency separation and/or presentation rate, which may be heard either as solely integrated or segregated (van Noorden, 1975, see Figure 1.2). For most LHL- sequences, segregation is never heard exclusively, even after a very long (four minute) sequence (Pressnitzer & Hupé, 2006, Denham & Winkler, 2006). Therefore, it is conceivable that the capturing effect associated with a same-frequency induction sequence may promote a stronger perception of stream segregation than would normally be observed during such an alternating-frequency sequence. This hypothesis was tested in Experiment 7 in two ways. First, two different conditions presented either a same- or an alternating-frequency induction sequence. As both of these sequences were of equal duration (2 s), and the L-tone density was matched between the two types of inducer; this allowed for a direct comparison of reported segregation after both types of inducer. Second, the segregation-promoting effect of the same-frequency induction sequence was compared with the extent of build-up occurring during a long test sequence. The approach was broadly similar to that of the subjective procedure used in experiments 1 and 3a-6b (see Chapter 2), in that each trial began with a 2-s induction sequence, during which no responses were made, and participants were then required to report their perception of a subsequent test sequence. However, unlike the previous experiments, the test sequence was a 20-s long repeating LHL- sequence. During the test sequence, listeners continuously reported their perception as either integrated or segregated. This tracking procedure was similar to that used by Anstis & Saida (1985). This method allowed for the measurement of long-term changes in perception, and so one would predict that reported segregation would build-up over the course of the test sequence (in addition to the effect of any particular induction sequence). Therefore, this method allowed a comparison of the initial segregation-promoting effect of a same-frequency induction sequence with the build-up occurring during a long, on-going alternating-frequency sequence.

There was also a second aim for Experiment 7 - this was to determine whether deviant-tone resetting could be observed in an on-going LHL- sequence. Experiments 1 - 5 have all consistently demonstrated that an abrupt change applied to the final tone of a same-frequency induction sequence can substantially reset test-sequence segregation. This finding has been interpreted as evidence that a disruption to an on-going stream actively resets build-up, and that this effect may be linked to attentional factors (Cusack et al., 2004, see Chapter 3). It is not known whether a similar, single deviant tone inserted into an on-going LHL- sequence would also have an active resetting effect. For an alternating-frequency sequence, previous research has only demonstrated that an abrupt change in global sequence properties can trigger a resetting of build-up (demonstrated for frequency changes, Anstis & Saida, 1985; and for changes in level or perceived location, Rogers & Bregman, 1998). One might predict that a noticeably changed tone in an on-going LHL- sequence would still be heard as a deviation from an established pattern, and so should also have a substantial active resetting effect.

In summary, Experiment 7 addressed two related issues. First, the experiment compared the segregation-promoting effect of a same-frequency induction sequence with that of an alternating-frequency counterpart. It was speculated that the same-frequency induction sequence may promote a stronger perception of stream segregation than does the alternating-frequency case. Related to this, the experiment was also designed to compare the initial segregation-promoting effect of a same-frequency induction sequence with the dynamics of build-up over a long alternating-frequency test sequence. Second, the experiment also aimed to compare the resetting effect of a single deviant tone in both a same-frequency and an alternating-frequency induction sequence. To my knowledge, the resetting effect of a single deviant tone on the build-up of stream segregation during an on-going alternating-frequency sequence has not been tested before.

8.2 Experiment 7

8.2.1 Method

Participants

Thirteen listeners took part in Experiment 7. All reported normal hearing. One listener was rejected from the final data set, as their responses did not conform to the well established effects of pure-tone frequency separation on streaming judgements or show any signs of build-up (van Noorden, 1975; Anstis & Saida, 1985). Five of the twelve accepted participants had previously contributed to one or more of the other experiments presented in this thesis.

Stimuli (Properties of the Test Sequence)

The test sequence comprised a 20-s presentation of a repeating LHL- pure tone arrangement. As with the previous experiments, the low-frequency tones were fixed at 1 kHz; whilst the frequency of the H-tones was varied between conditions. The frequency separations tested were 3 ST, 6 ST and 9 ST (high tone = 1189, 1414, and 1682 Hz, respectively). Tone duration was fixed at 100 ms, including 10-ms raised cosine ramps at tone onset and offset. For each LHL- triplet, there was no inter-tone silent interval. The silence at the end of each triplet lasted 100 ms, equivalent to the duration of one tone. Therefore, the overall duration of each triplet was 400 ms. Accordingly, the 20-s test sequence comprised 50 LHL- triplets. All tones were presented diotically and at 70 dB SPL.

Stimuli (Properties of the Induction Sequences)

On each trial, the test sequence was preceded by a 2-s induction sequence. In total, five different induction sequences were tested (see Figure 8.1 for an illustration of each induction condition). These conditions are described below:

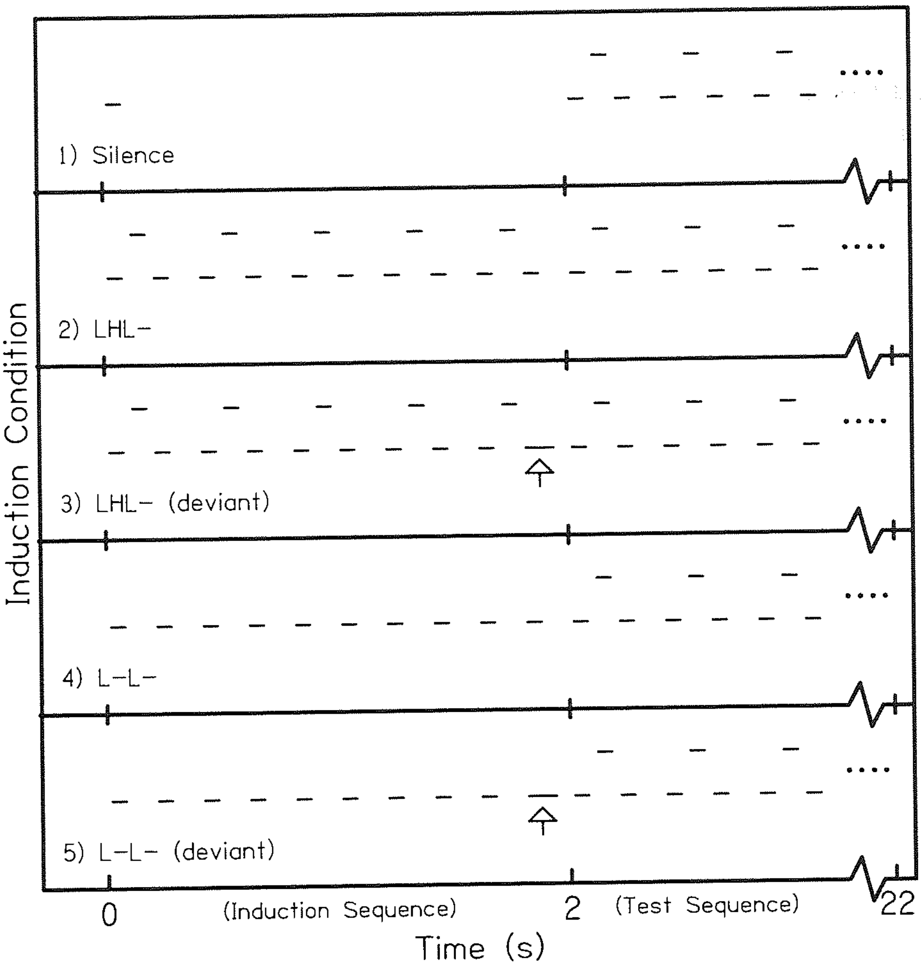

Figure 8.1: **Illustration of the induction conditions tested in Experiment 7. Each induction sequence is paired with a subsequent test sequence (note that the test sequence continued for 20 s). An arrow indicates an induction tone made deviant by extending its duration from 100 ms to 150 ms (the subsequent silent interval was reduced by 50 ms in order to preserve the rhythm of the sequence).**

1. *Silent induction condition.*

   This condition was included as a measure of streaming in the test sequence in the absence of any prior build-up. Only a single L tone was presented at the onset of the stimulus (frequency = 1 kHz, tone duration = 100 ms). Following the tone, the remaining portion of the induction sequence was filled by a 1900-ms silent interval. As listeners were required to make a rapid response once the test sequence began (see procedure), the single L-inducer acted as a "warning tone", included to help participants anticipate the onset of the test sequence. Although a single L-inducer can be partially effective at promoting segregation (Experiment 4b), this effect should decay almost completely during the subsequent silent interval (Bregman, 1978; Cusack et al., 2004), and so this condition was not expected to promote any test-sequence segregation.

2. *LHL- induction condition.*

   Five LHL- triplets were presented in the induction sequence. These triplets were identical to those presented in the test sequence (L-tones = 1 kHz, H-tones = 3, 6, or 9 ST higher, to match the subsequently presented test sequence). Therefore, the inducer/test boundary was defined arbitrarily; the combined induction and test sequences were heard as a continuous stimulus of constant properties (e.g., in tone frequencies and durations). For this induction condition, one would predict that build-up at the (arbitrary) onset of the test sequence would result in reported segregation equivalent to that which occurs two seconds later in the silent induction condition.

3. *LHL-(deviant) induction condition.*

   This condition differed from the LHL- induction condition in only one respect. The final L-tone of the last (fifth) triplet was extended in duration from 100 ms to 150 ms. To preserve the rhythm of the sequence, the subsequent inter-tone silence was reduced from 100 ms to 50 ms. This condition was included to measure the resetting effect of a deviant tone on the build-up of stream segregation in an ongoing LHL- sequence.

4. *L-L- induction condition.*

   This induction sequence was identical to the standard induction condition used in many of the previous subjective experiments (see Chapter 2). The induction sequence comprised 10 L- repetitions. The L-tones shared identical properties with those of their counterparts in the test sequence (frequency = 1 kHz, tone duration = 100 ms), and the inter-tone silent interval was set to 100 ms. Note, therefore, that all properties of the L tones – including tone density and timing – were identical to those of their counterparts in the LHL- induction condition. The previous experiments reported in this thesis have demonstrated that this type of induction arrangement promotes a strong perception of segregation in a subsequent test sequence (experiments 1 - 5).

5. *L-L-(deviant) induction condition.*

   This induction condition differed from the L-L- induction condition only in that the final L-only repetition was altered. As for the LHL-(deviant) condition, 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. Based on evidence from Experiment 1, which tested an identical induction sequence, one would predict that the deviant tone would substantially reset the segregation-promoting effect of the L-L- induction sequence.

General Procedure

On each trial, a single combination of an induction sequence and a test sequence was presented. Each trial was initiated when a participant pressed "↵" on a computer keyboard. At the start of the test sequence, 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. Then, for the remaining duration of the test sequence, participants were asked to press the appropriate key every time their perception of the test sequence changed. The computer program noted the exact time of each key press. Participants received an on-screen visual cue which indicated the start of the test sequence, and prompted them to begin responding. As with the previous subjective experiments, participants were asked to avoid listening actively for either integration or segregation, but rather simply to report which of the two percepts was most dominant (see Chapter 2). 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 trial onset). This inter-trial silence was longer than the 3 s used in the previous subjective experiments, because the current trials presented considerably longer tone sequences. Previous evidence suggests that, even for long sequences, the vast majority of build-up should decay during a 5-s silence (Bregman, 1978; Cusack et al., 2004).

For the main experiment, the trials were organized into blocks. Each trial block comprised a combination of each of the five induction sequences with each of the three frequency separations for the test sequence. This resulted in 15 trials per trial block (i.e., 5 induction conditions $\times$ 3 frequency separations). Within each block, the order of trial presentation was fully randomized. After a verbal explanation of streaming, participants were presented with a clear example of an integrated and of a segregated LHL- sequence. Once familiar with these stimuli, participants then completed a training session comprising a single trial block. This training was given to familiarize participants with the stimuli - there were no criteria for acceptance into the main experiment. The main experiment comprised 10 trial blocks, and this typically took a participant between 2 and $2\frac{1}{2}$ hours to complete. Therefore, the experiment was split in two separate sessions.

Apparatus

The apparatus used was identical to that described in Chapter 2, except that the experiment was run using the Media Control Functions (MCF) stimulus presentation software (Ahad, 2000). This software was used as it supports the precision measurement of the timing of key presses.

8.2.2 Results

Dynamic effects: Time bin analysis

Before any further analysis, response data from each trial were first divided into 1-s time bins (e.g., 0-1 s, 1-2 s . . . 19-20 s). This bin size was chosen as it provided a good balance between temporal resolution and the averaging required to smooth out moment-to-moment random fluctuations in key press timings. The percentage of each time bin heard as segregated was calculated from the timings of the individual key presses. An important consideration for this analysis was that the timing of each first response varied between trials (as participants were required to indicate their initial perception of the test sequence). To compensate for this, for each individual trial a time-bin average was only recorded if the first response of a participant had already occurred in a previous 1-s interval or occurred within the first 500 ms of the current time bin. For the 0-1 s time bin, only 15% of trials met this criterion. In comparison, for the 1-2 s time bin, the criterion was met in 75% of trials. Owing to the limited data available for the first (0-1 s) time bin, it was excluded from all subsequent analysis and graphical representation.

These time-bin data from each separate condition were then averaged for each individual listener (as participants completed 10 repetitions of each condition). For some trials, however, variation in the timing of the first response occasionally caused the failure of the early time bins to meet the criteria for acceptance into the data set (as described above). Owing to this, the across-trial average for each individual time bin was calculated exclusively from the time bins which did meet the acceptance criteria (minimum = 75%). This ensured that there were no missing time-bin values from each participant's averaged results. The data for each condition were then averaged across the tweleve listeners to yield a percentage heard as segregated as a function of time for each of the induction conditions. These data are displayed in Figure 8.2. The time indicated on the abscissa corresponds to the end of the time bin (e.g., 5 s indicates the mean for the 4-5 s time bin). Three panels display the results separately for each of the three frequency separations. For the majority of the induction conditions, participants were most likely to hear integration at the beginning of the test sequence, and the tendency to report stream segregation increased over time. Generally speaking, this result is consistent with the findings of previous research (Bregman, 1978; Anstis & Saida, 1985). However, there are some exceptions to this pattern. Most noticeably, the L-L- induction condition promoted a strong, initial tendency to hear stream segregation. Indeed, for frequency separations of 6 and 9 ST, reported stream segregation actually decayed over the time course of the test sequence. This finding is examined in more detail later.

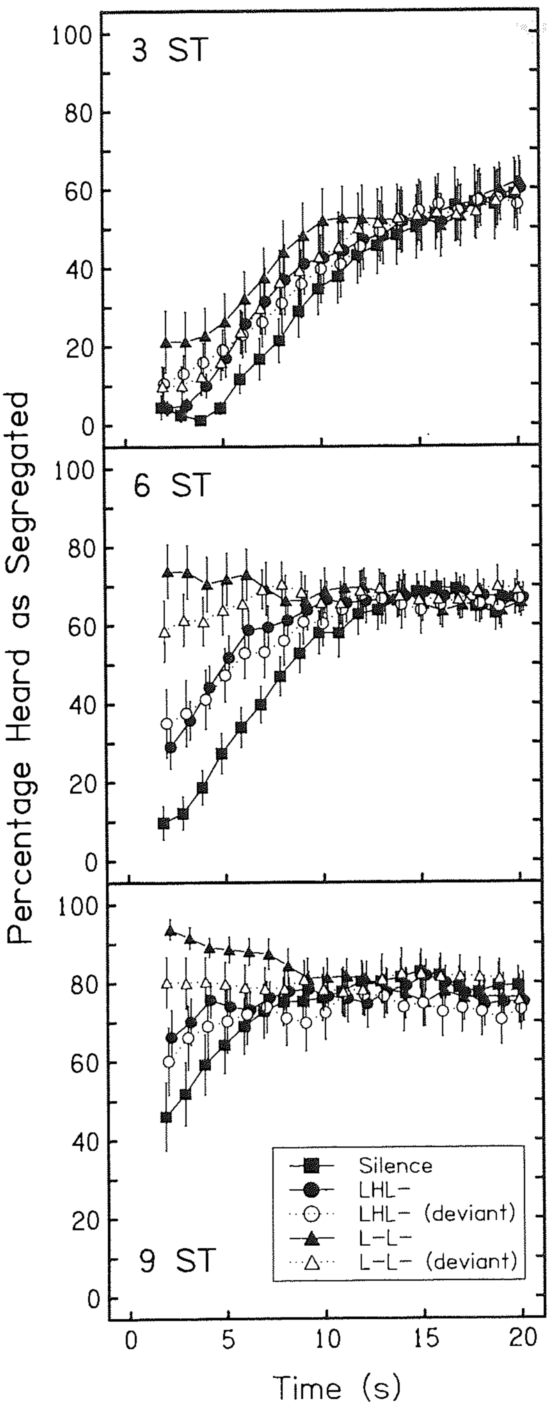

Figure 8.2: Results from Experiment 7 for 12 listeners. Responses for each separate trial are averaged into 1-s time bins, and then averaged across all trials. Results for each frequency separation are displayed in separate panels, and the insert identifies the different induction conditions tested. Data for the first second are excluded owing to the limited number of responses made during this interval (see text for a full explanation). Note that the time indicated on the abscissa corresponds to the end of the corresponding time bin.

An inspection of Figure 8.2 shows that the greatest differences between induction conditions were evident during the first 11 s or so of the test sequence. After this time, responses became relatively homogeneous. This observation was supported by a three-way, repeated-measures ANOVA which analysed responses over the final nine seconds of the test sequence. The three independent variables were frequency separation (3, 6, or 9 ST), induction condition (5 $\times$ conditions), and time interval (time bins from 11-12 s to 19-20 s). There were significant main effects of frequency separation [$F(2, 22) = 6.832$, $p < 0.005$, $\eta_p^2 = 0.383$] and of time interval, [$F(8, 88) = 2.727$, $p < 0.01$, $\eta_p^2 = 0.199$] but there was no significant main effect of induction condition [$F(4, 44) = 0.489$, $p > 0.05$, $\eta_p^2 = 0.043$]. The frequency separation $\times$ time interval interaction was significant [$F(16, 176) = 9.094$, $p < 0.001$, $\eta_p^2 = 0.453$], this is likely to be due to a difference in the rate of build-up between the test sequences of different frequency separations. These findings suggest that build-up continued into the latter portion of the test sequence, and that the extent of reported segregation was influenced by the frequency separation between the low- and high-frequency tones. These observations are in good accord with previous measures of build-up (e.g., Anstis & Saida, 1985). However, this analysis clearly demonstrated that the properties of the induction sequence did not influence responses in the latter portion of the test sequence. Given this, the next analysis will focus exclusively on the responses during the earlier section of the test sequence.

A three-way ANOVA compared responses over the first ten seconds of response data (recall that the 0-1 s time bin was excluded, as explained above). The three independent variables were frequency separation, induction condition, and time interval (time bins from 1-2 s to 10-11 s). The ANOVA confirmed significant main effects of frequency separation [$F(2, 20) = 39.161$, $p < 0.001$, $\eta_p^2 = 0.797$], of induction condition [$F(4, 40) = 13.747$, $p < 0.001$, $\eta_p^2 = 0.579$], and of time interval [$F(9, 90) = 28.335$, $p < 0.001$, $\eta_p^2 = 0.739$]. Each of the two-way interactions was also significant. These were: frequency separation $\times$ induction condition [$F(8, 80) = 2.828$, $p < 0.01$, $\eta_p^2 = 0.220$], frequency separation $\times$ time interval [$F(18, 180) = 7.110$, $p < 0.001$, $\eta_p^2 = 0.416$], and induction condition $\times$ time interval [$F(36, 360) = 8.858$, $p < 0.001$, $\eta_p^2 = 0.470$]. Finally, the three-way interaction term (frequency separation $\times$ induction condition $\times$ time interval) was also significant [$F(72, 720) = 4.349$, $p < 0.001$, $\eta_p^2 = 0.303$]. Furthermore, all three main effects, and all but one of the interactions, remain significant if the silent induction condition is excluded from the analysis. The only exception was the frequency separation $\times$ time interval interaction, where [$F(6, 66) = 1.616$, $p > 0.05$, $\eta_p^2 = 0.128$]. This shows that the significant effects shown in the main analysis were not simply due to the inclusion of the silent induction condition.

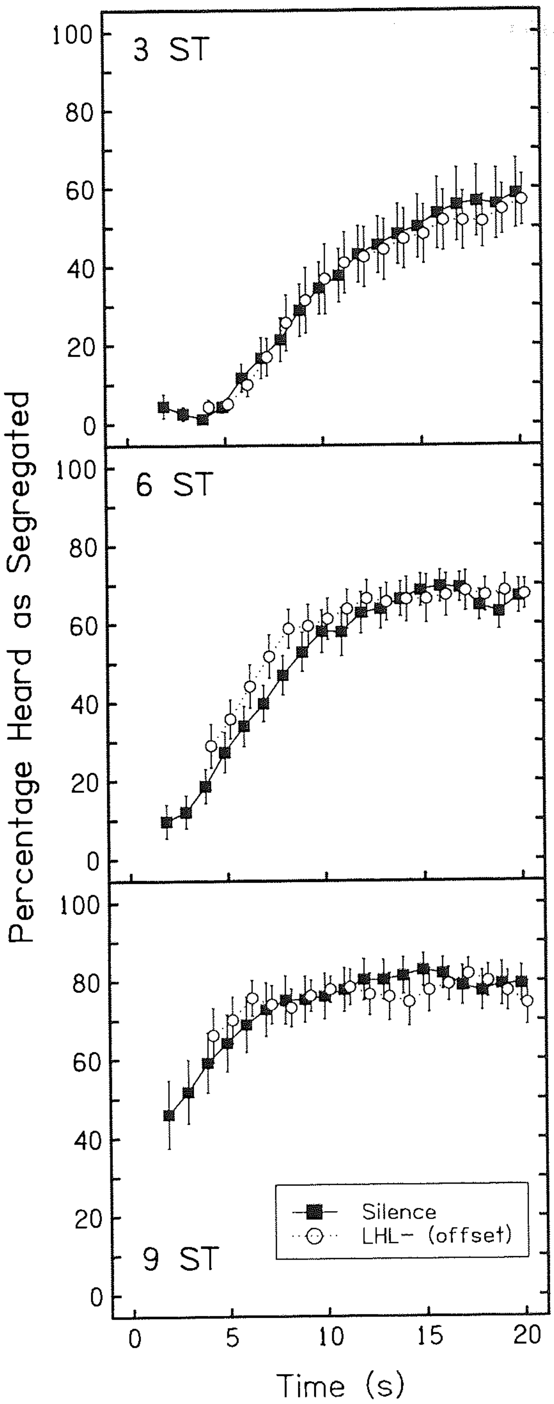

Figure 8.3: Time-aligned results from Experiment 7 for 12 listeners. Responses for each separate trial are averaged into 1-s time bins, and then averaged across all trials. Results for each frequency separation are displayed in separate panels. The insert identifies the two induction conditions displayed. The data presented are the same as for Figure 8.2, except that responses for the LHL- condition are offset by +2 s. This offset was applied to illustrate that build-up during the LHL- induction sequence was largely equivalent to that occurring during the (physically identical) test sequence.

A key aim of the current experiment was to compare the segregation-promoting effect of a same-frequency induction sequence with the build-up occurring in an alternating-frequency sequence. As the 2-s alternating-frequency (LHL-) induction sequence was identical to the subsequent test sequence, one would expect two-seconds-worth of equivalent build-up to have occurred by the onset of the test sequence. Therefore, mean stream segregation in the LHL- condition should be very similar to that occurring two seconds later in the silent inducer condition. To test this, the LHL- data was offset by +2 s and compared with those from the silent-induction case. These data are displayed in Figure 8.3 and were compared using a three-way ANOVA (excluding the first two time bins, for which data was available only for the silent-induction case). The three variables were frequency separation, induction condition (silence or offset LHL-), and time interval (17 conditions: 3 - 4 s bin to 19 - 20 s bin for the silent induction condition and 1 - 2 s bin to 17 - 18 s bin for the LHL induction condition, as defined before the +2 s offset). As expected, the ANOVA confirmed a main effect of frequency separation [$F(2, 22) = 22.518$, $p < 0.001$, $\eta_p^2 = 0.672$], and of time interval [$F(16, 176) = 59.685$, $p < 0.001$, $\eta_p^2 = 0.844$]. Crucially, there was not a significant main effect of induction condition [$F(1, 11) = 1.422$, $p > 0.05$, $\eta_p^2 = 0.114$], and similarly, neither of the two interactions associated with the induction-condition variable were significant (p > 0.05 in both cases).

For the same-frequency induction condition (L-L-), reported stream segregation was much greater than for the alternating-frequency case (at least, for the initial portion of the sequence). From Figure 8.2, it is apparent that reported segregation in the former condition actually decreased over time for frequency separations of 6 or 9 ST. This indicates that the same-frequency induction sequence promoted stream segregation to a greater extent than is normally heard for an on-going alternating-frequency sequence of the same length. Paired-sample t-tests were run to determine for this condition whether the decrease in segregation over time was significant. For each frequency separation, the t-tests compared initial segregation (1-2 s bin) to that after 10 s (10-11 s bin). These tests confirmed that stream segregation increased significantly for the 3 ST case [+31.4%, $t(11) = 3.275$, $p < 0.01$], and that there was no significant change for the 6 ST case [-4.2%, $t(11) = 0.707$, $p > 0.05$]. However, the decrease in stream segregation for the 9 ST case was significant [-12.0%, $t(11) = 2.572$, $p < 0.05$]. In comparison, for the LHL- 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%, $t(11) = 7.424$, $p < 0.001$], but not for the 9 ST case [+9.7%, $t(11) = 1.226$, $p > 0.05$].

Resetting effects: First-response analysis

The second part of the analysis addressed how the various induction sequences affected the initial perception of the test sequence. This was achieved by analysing only the initial response on each trial (i.e., the first key press only, with no time bin analysis or any other such measure). This measure was chosen because the exact timing of the first key press varied from trial to trial. The removal of the time dimension ensured that responses from every trial contributed to the data set. This is a reasonable approach, because the majority of first responses occurred early in the test sequence (i.e., across all listeners and conditions, 75% of responses were made within 2 s after the onset of the test sequence). Note also that the results of the first-response analyses presented here could be reproduced reasonably well by performing the same tests on the 1-2 s time bin data as described in the previous section.

For each separate condition, responses were averaged for each listener to derive a percentage of first responses which were heard as segregated. These values were then averaged across the twelve listeners, and these data are displayed in Figure 8.4. The percentage of trials heard initially as segregated for frequency separations of 3, 6, and 9 ST, were 11.1%, 42.1%, and 68.3%, respectively. The means for the silence, LHL-, L-L-, LHL-(deviant) and the L-L-(deviant) induction conditions were 22.9%, 32.7%, 60.1%, 35.0%, and 51.1%, respectively. These data were analysed using a two-way, repeated-measures ANOVA. The two independent variables were the three frequency separations for the test sequence and the five induction conditions. The ANOVA confirmed significant main effects of frequency separation [$F(2,22) = 82.082,\ p < 0.001,\ \eta_p^2 = 0.882$], and of induction condition [$F(4,44) = 19.421,\ p < 0.001,\ \eta_p^2 = 0.638$]. The interaction between these two variables was also significant [$F(8,88) = 8.862,\ p < 0.001,\ \eta_p^2 = 0.446$]. This interaction may have been driven by a partial floor effect (i.e., primarily integrated responses) for the 3-ST cases.

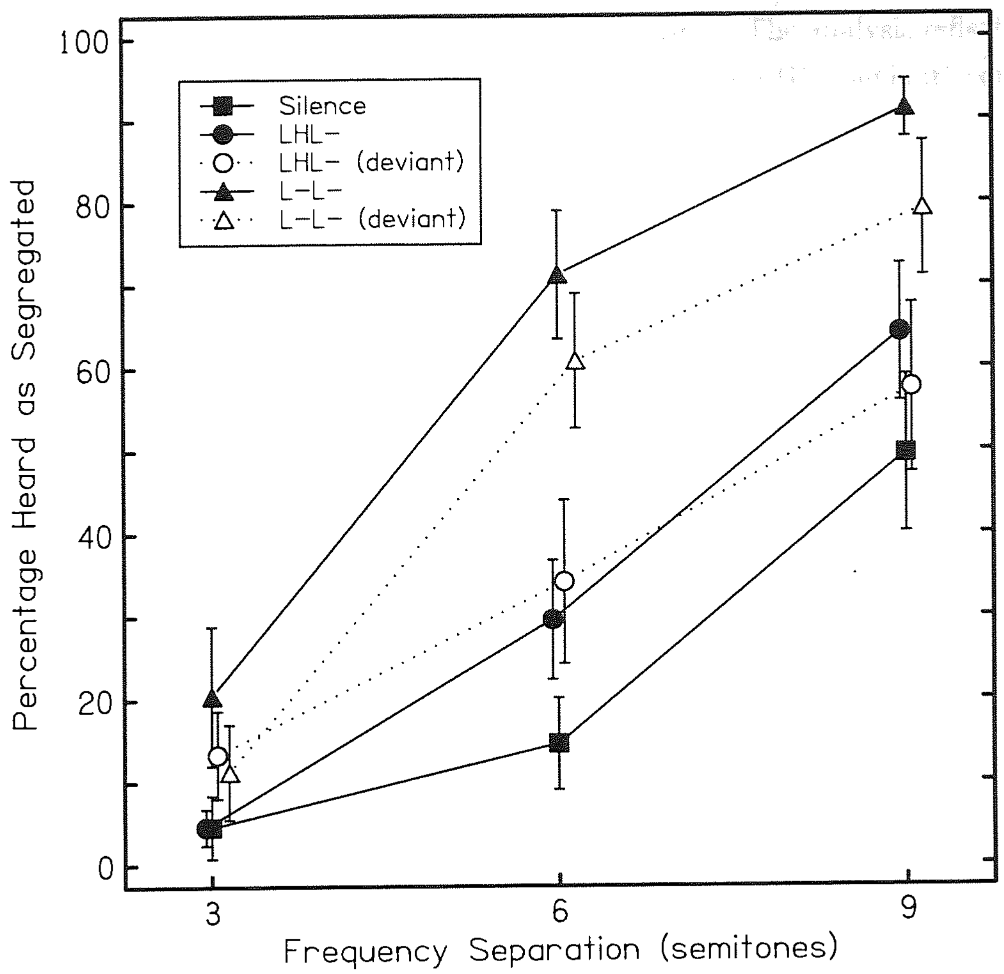

Figure 8.4: **First response data from Experiment 7. The effects of induction condition and frequency separation on initial reported segregation in the test sequence are shown (averaged across 12 listeners). Each point represents the percentage of first responses which were segregated (i.e., means derived from first key-press data only). The insert identifies the different induction conditions tested. The error bars display $\pm$ 1 inter-subject standard error**

Two-tailed pairwise comparisons were performed on this data set using the restricted LSD test (Snedecor & Cochran, 1967). One purpose of the experiment was to examine the resetting effect of a single deviant tone on both the L-L- and the LHL- induction conditions. For the same-frequency induction condition, there was a decrease in stream segregation for the first responses of about 10% when the final inducer was extended in duration, but this difference between the L-L- and L-L-(deviant) conditions only approached significance [$t(11) = 2.038$, $p = 0.068$]. Nonetheless, the general trend of these results remains consistent with the resetting effect of a temporally deviant tone that was demonstrated in Experiment 1.

For the LHL-(deviant) condition, the same duration increase was applied to the final L inducer of the alternating-frequency induction sequence. From Figure 8.4, it is apparent that this deviant tone had only a small and inconsistent effect on initial streaming judgments across the frequency separations tested. The analysis reflected this; there was no significant difference between the LHL- and the LHL-(deviant) conditions [$t(11) = 0.070,\ p > 0.05$].

8.3 General Discussion

The current experiment was designed to compare the dynamics of the build-up and resetting of stream segregation for same-frequency and alternating-frequency induction sequences. Experiment 7 studied these issues through the use of a tracking procedure, which continuously measured perception over the course of a relatively long (20 s) test sequence. On each trial, the test sequence was preceded by an induction sequence, the properties of which were varied experimentally. The silent induction condition was not expected to have any influence on the perception of the test sequence. This prediction was largely supported, as the dynamics of test-sequence build-up were broadly consistent with those observed in previous studies which measured streaming in the absence of an induction sequence (Bregman, 1978; Anstis & Saida, 1985; Carlyon et al., 2001). The current results did, however, differ slightly from those previously observed. Specifically, earlier studies have mainly shown that the overall tendency to report segregation (or the asymptotic percentage of trials heard as segregated) rose as the frequency separation between the tones was increased (Anstis & Saida, 1985; Carlyon et al., 2001; Cusack et al., 2004; Snyder et al., 2008). For example, using similar stimuli to those in Experiment 7, Carlyon et al. (2001) observed that the percentage of trials heard as segregated approached different asymptotes as frequency separation was increased – from approximately 40% for 4 ST, to nearly 100% for 10 ST. For the current study, the same trend was clearly observed, but this difference was not quite as prominent – the percentage heard as segregated increased from around 60 % for 3 ST to 75% - 80% for 9 ST. Nonetheless, the current results are largely consistent with a study by Cusack (2005), who observed only a marginal increase in stream segregation across the range of frequency separations 3 – 7 ST. Furthermore, for the current study, the rate of build-up over the first ten seconds or so increased with frequency separation, and all of the ANOVAs confirmed a significant main effect of frequency separation. These observations are in good accord will all previous research (e.g., Anstis & Saida, 1985). Therefore, it is concluded that the build-up of stream segregation observed following the silent induction condition was broadly consistent with that reported in previous research.

The properties of the alternating-frequency induction sequence (the LHL- condition) were exactly the same as those of the subsequent test sequence. Unsurprisingly, during this induction sequence, build-up occurred at the same rate as during the test sequence. This was confirmed when the data were offset by two seconds to account for the build-up occurring during the (arbitrary) induction sequence (see Figure 8.3). When offset, there was no significant difference between the LHL- condition and the silent induction case. In contrast, the same-frequency induction sequence (L-L- condition) promoted substantially more stream segregation over the first several seconds than did the alternating-frequency case. Indeed, following the same-frequency inducer, the tendency to report segregation actually decreased over the course of the test sequence for frequency separations of 6 and 9 ST cases (although this effect was only significant for the 9-ST case). Therefore, for these frequency separations, the same-frequency induction promoted a greater extent of segregation than would normally occur for an alternating-frequency sequence of corresponding duration.

The build-up of stream segregation during an alternating-frequency sequence is thought to represent a gradual accumulation of evidence in favour of two sources being present (Bregman, 1978, 1990). Despite this, even a very long ambiguous tone sequence (i.e., one which falls between the fission and temporal coherence boundaries; van Noorden, 1975) will never be heard exclusively as segregated (Pressnitzer & Hupé, 2006; Denham & Winkler, 2006). It has been proposed that a same-frequency induction sequence may promote segregation by capturing one subset of test-sequence tones into an on-going stream (Rogers & Bregman, 1993; see also Bregman & Rudnicky, 1975, and Chapter 5). For the current experiment, the stream formed during the L-tone induction sequence may have been heard to continue into the test sequence, and so increased the tendency for the H tones to be excluded from this stream. The current results can be interpreted in the context of this argument. If this argument were correct, one would assume that the capturing effect would be most evident at the onset of the test sequence, when the induction-tone stream would still clearly be perceived. During the test sequence, perception will inevitably begin to switch between integration and segregation (Pressnitzer & Hupé, 2006), and so this capturing effect will not continue indefinitely. Over time, any initial tendency to hear a sequence as segregated will gradually decay. This argument is in good accord with the homogeneity of responses after (approximately) 10 s for all induction conditions.

The clearly different pattern of results following either a same- or an alternating-frequency induction sequence may also suggest a different mechanism is responsible for the segregation-promoting effect of both types of sequence. It has been proposed that for an alternating-frequency sequence, segregation increases because of a build-up of evidence in favour of two streams (Bregman, 1978; Bregman, 1990). In contrast, the proposed capturing effect of a same-frequency sequence may cause an attentional bias to hear stream segregation. During the same-frequency induction sequence, attention must be drawn to the L tones (as these are the only stimuli present), and this attentional bias may continue into the test sequence, so that the additional H-tones are heard as arising from a new source and so are excluded from the pre-established stream. This suggestion relates to the hierarchical decomposition model proposed by Cusack et al. (2004). These authors proposed that in a multi-source listening environment, there may be some automatic segregation, but that attentional focus strongly influences which source is subject to a more complete elaboration of its perceptual representation (see also Brochard et al., 1999).

The suggestion that the initial perception of an alternating-frequency sequence has a strong influence on the dynamics of build-up also relates to Pressnitzer & Hupé's (2006) argument that auditory streaming is a bi-stable percept. These authors measured the durations of successive percepts in a tracking procedure, and demonstrated that the duration of the initial integrated percept lasted substantially longer than all other subsequent percepts. Furthermore, aside from the first percept, these authors found no long-term trend towards reporting stream segregation. Therefore, Pressnitzer & Hupé (2006) proposed that the perceptual representation of an alternating-frequency sequence is bi-stable, with no one percept ever dominating perception. However, this study only measured the perception using a single condition, (LH frequency separation = 5 ST, tone duration = 120 ms). It also remains unclear as to how bi-stable perception can explain the long-term build-up of segregation, as the effect remains apparent over the course of sequences as long as 30 s or even 60 s (Anstis & Saida, 1985). These issues aside, one can still relate the relative importance of a "first-response bias" to the current study. If a first-response bias towards integration influences the build-up of segregation in an alternating-frequency sequence, then a similar bias (except towards segregation) may be responsible for the currently observed decay of stream segregation following the same-frequency induction conditions. This suggestion that the dynamics of streaming are strongly influenced by a first-response bias is supported by the fact that the responses for all induction conditions became relatively homogeneous after approximately 10 s. Unfortunately, this hypothesis was difficult to test using the current results, as the test sequence only lasted for twenty seconds in order to keep down the length of the experiment for the listeners. Indeed, participants typically did not report many switches in perception over the course of a trial (average = 2.9 percepts reported per trial). In addition, the duration of the final percept was truncated by the end of a trial. Hence, an adequate comparison of the duration of the initial percept with subsequent ones was not possible for the current dataset.

The current results can also be related to a study by Snyder et al. (2008). These authors examined the effect of preceding context on streaming judgments. Participants reported their perception of a stimulus once only, at the end of a relatively long (10.8

s) LHL- tone sequence. The L-tone frequency was fixed and the H-tone frequency was varied (+ 0, 4, 7, or 12 ST). It was found that, with increasing frequency separation for the *previous* trial, less streaming was reported in the *current* trial (a contrastive effect). Though weaker, this context effect was still evident for the trial before last. A second experiment demonstrated that this context effect occurred regardless of listening "set" - i.e., whether listeners were instructed to attempt to hear integration or segregation. In a final experiment, Snyder et al. (2008) asked listeners to report continuously their perception of the sequence. The same context effect was apparent throughout the entire duration of the trial - listeners consistently reported more stream segregation during the current trial as the frequency separation for the previous trial was decreased. This context effect was not simply due to response bias, because the perception of segregation during the previous trial (as opposed to the physical frequency separation), did not cause less streaming to be reported during the current trial. Snyder et al. (2008) proposed that the observed context effect may reflect auditory sensory memory (Cowan, 1984) or neural adaptation (Micheyl et al., 2005).

These findings merit consideration in relation to the current same-frequency induction condition, as Snyder et al. (2008) demonstrated a contrastive effect following a same-frequency sequence (i.e., the case where the frequency separation was 0 ST). This could potentially explain why the current L-only induction sequence was so effective at promoting test-sequence segregation. However, there are several reasons suggesting that the current results cannot be explained solely in terms of a contrastive context effect. Snyder et al. (2008) demonstrated a context effect across clearly separated trials (minimum inter-sequence silent interval = 1.44 s), and found that a prior same-frequency sequence increased reported segregation in the current trial by a similar extent for the entire duration of the current sequence. In clear contrast, the segregation-promoting effect of the same-frequency induction sequence in the current study was most apparent at the onset of the test sequence, and this effect diminished over time (most evidently for the 9 ST case). Indeed, the same-frequency induction sequence could drastically alter the dynamics of streaming (evident from the observed decay of segregation). No such effect was observed in Snyder et al.'s (2008) study. Whilst the contrastive effect following a tone sequence comprising a 0-ST frequency separation did lead to an increased tendency to report segregation during the current trial, the *relative* dynamics of build-up over the course of the trial remained unchanged. These differences appear to indicate that the current same-frequency condition influenced perception through grouping with 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. Therefore, it would appear that the current set of results reflect perceptual grouping, as opposed to 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 the results).

The current study also addressed the active resetting effect of an alteration applied to the final induction tone. The deviant tone tested was one extended in duration from 100 ms to 150 ms. Using a related subjective measure, Experiment 1 demonstrated that this change substantially resets the segregation-promoting effect of a same-frequency induction sequence. Exact replicas of these same-frequency induction conditions were tested in Experiment 7 and, although the effect of the deviant tone in a same-frequency induction sequence was not significant, a clear trend towards resetting was observed. Overall, the general pattern of results was in accord with those observed for Experiment 1.

Of particular interest was whether the same deviant tone would also have a resetting effect when inserted into an on-going, alternating-frequency 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 alternating-frequency induction sequence for which no deviant tone was present. The only example of the deviant tone reducing reported segregation away from that in the standard LHL- induction case was observed at the 9 ST frequency separation - and this effect was very marginal and non-significant (see Figure 8.2).

A possible explanation for the lack of resetting could relate to how the LHL- induction sequences were 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 be stronger still in the initial 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 currently tested temporal change). One might predict, in the LHL- context, that a *deviant triplet* would have a much more substantial active 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., reduced in frequency). Even when the 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 was heard as an LHH^{+} - arrangement (where H^{+} = an even higher-frequency tone, so that an ascending percept is heard). Unlike the currently tested temporal change, this deviant tone would substantially alter the perception of the triplet, and so this more salient change may have a greater resetting effect. Importantly, for both of these proposals, only a brief change would be applied to the on-going sequence, and so any reduction in segregation would be evidence in favour of active resetting.

In conclusion, the current experiment clearly demonstrated that the extent of reported stream segregation differed following a same-frequency and an alternating-frequency induction sequence. The increase in reported segregation following the L-only sequence is attributed to the induction tones capturing one subset of test sequence tones into an on-going stream (see also Rogers & Bregman, 1993; Bregman & Rudnicky, 1975). Concerning the second aim of the experiment, the results have not provided unequivocal evidence of resetting for either type of induction sequence. Despite this, the trend observed for same-frequency induction condition was largely consistent with that of previous studies. For the alternating-frequency inducer, there was little or no evidence of resetting, possibly due to the perceptual grouping of the low and high frequency tones.

Chapter 9

Sequential Grouping of Pure-Tone Percepts Evoked by the Segregation of Components from a Complex Tone

9.1 Introduction

The experiments presented in chapters 3 - 8 have all examined the dynamics of stream segregation using pure-tone stimuli. These studies have addressed the various factors that influence the segregation-promoting effect of an induction sequence on streaming judgments in a subsequent test sequence. From these studies, it would appear that a same-frequency induction sequence may promote segregation by capturing one subset of test-sequence tones into an on-going stream, and that this causes the second subset of tones to segregate from the established stream (see also Rogers & Bregman, 1993). In the context of the standard induction condition tested in the previous experiments, the exclusively low-frequency induction tones tended to capture perceptually the subsequent L-tones of the test sequence, causing the H-tones to be excluded from that stream.

As discussed in chapters 5 and 8, this argument also relates to the findings of Bregman & Rudnicky (1975), and their interpretation of those findings. As the current experiments used a similar experimental design to their study, their method and results are briefly re-summarized here. Bregman & Rudnicky (1975) presented two pairs of pure tones, each comprising two target tones, A and B (ordered either AB or BA). These pure tones differed in frequency (2200 Hz or 2400 Hz), but were otherwise identical (tone duration = 65 ms, no inter-tone silence). Listeners reported whether the presentation order was either the same (AB-AB, or BA-BA), or different (AB-BA, or BA-AB). This task was relatively easy when the targets were presented in the absence of any other tones, but it became much more difficult when the second target pair was flanked by two distractor tones (i.e., DABD; D = distractor tone, distractor frequency = 1460 Hz). Performance improved when the targets and distractors were embedded within a sequence of captor tones which matched the frequency and rhythm of the distractors (i.e., CCCCDAB-DCCCC). Presenting the captor sequence at a lower frequency (590 Hz) did not aid performance. The distractor-frequency and lower-frequency captor conditions tested by Bregman & Rudnicky (1975) are illustrated in Figure 9.1. Bregman & Rudnicky (1975) argued that successful performance depended on a listener grouping the target tones separately from the distractors. They proposed that the distractor-frequency captors may have pulled the interfering distractor tones into a separate auditory stream from the target tones. If so, the lower-frequency captors would not have aided performance as they would be unlikely to group with the distractor tones (owing to the lack of frequency proximity).

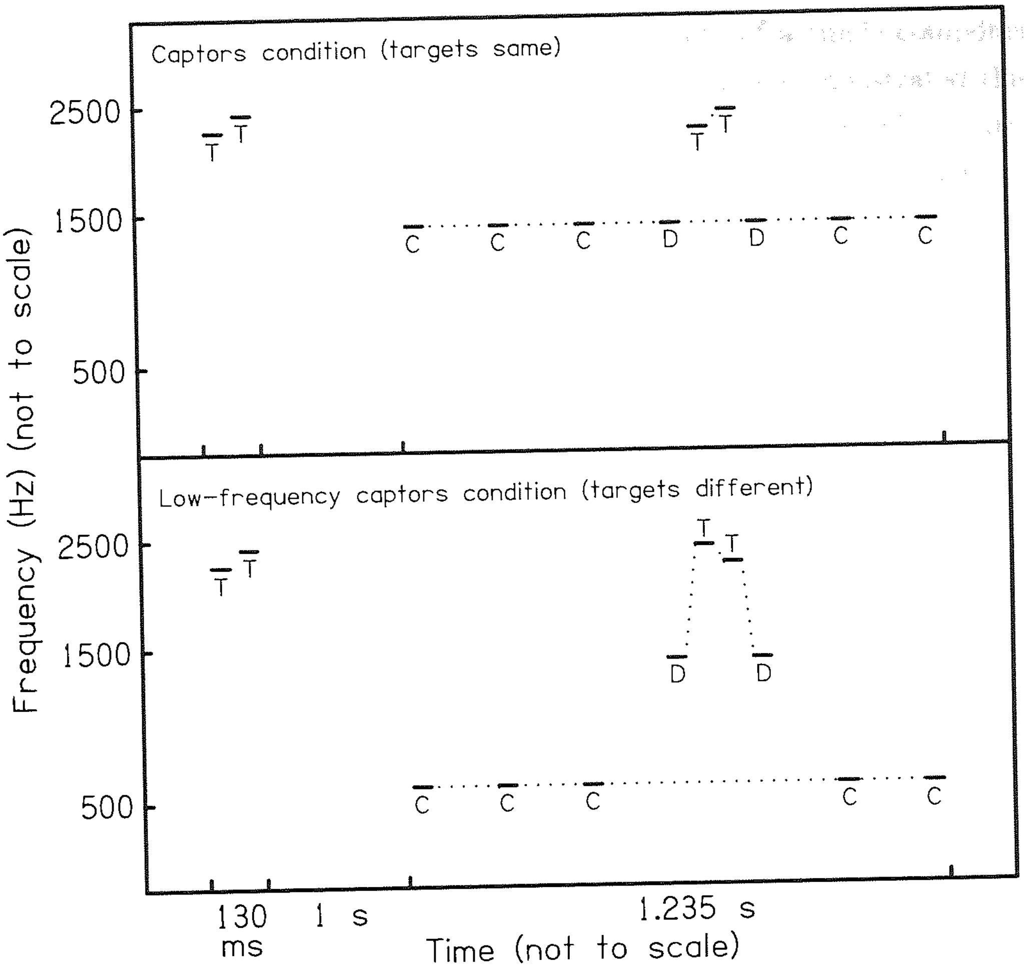

Figure 9.1: **Illustration of two conditions tested by Bregman & Rudnicky (1975). The task was to report whether the two pairs of target tones (T) were the same or different (in terms of the direction of the frequency change).** Captors condition **(top panel); the distractors (D) and captor (C) tones group to form a single perceptual stream, and the targets are heard as segregated from this stream (dotted lines represent perceptual grouping). Performance is good.** Low-frequency captors condition **(bottom panel); the captors no longer group with the distractors, and so the distractors group with the targets. Performance is impaired as the direction of the target frequency change is less salient.**

The majority of research has studied stream segregation using exclusively pure-tone stimuli (e.g., Miller & Heise, 1950; van Noorden, 1975; Bregman & Campbell, 1971), although streaming has also been demonstrated in a wide range of isolated sound stimuli (such as complex tones, van Noorden, 1975). Despite this, there remains a class of acoustic stimuli for which streaming has received little attention. A simultaneously presented set of harmonics is typically heard as a unified percept (a complex tone). However, if an abrupt change is applied to a single harmonic, the changed component may be heard as a pure-tone-like percept, separate (or segregated) from the on-going complex tone. What is not known is whether the sequential organization of these tone-like percepts is similar to that of discrete pure tones. This question is examined by experiments 8 - 10.

Kubovy & Daniel (1983) referred to the "hearing out" of a single component of a complex tone as concurrent pitch segregation. These authors demonstrated this effect for sudden changes in amplitude, ITD, or phase. As Experiment 8 only studied tone-like percepts created from amplitude changes, ITD and phase cues for concurrent pitch segregation are considered later (in section 9.3). To demonstrate amplitude-based segregation, Kubovy & Daniel (1983) presented an eight-component musical chord (duration > 2 s). When a single component was briefly lowered in amplitude, and then restored (decrement duration = 80 ms), the altered component was heard temporarily as a pure-tone-like percept, clearly distinct from the on-going complex tone (see also Kubovy, 1981). This is true despite the amplitude of the altered component never exceeding that of the other (unaltered) components. These authors presented a series of amplitude decrements, each successively applied to a different component of the on-going complex tone. Listeners could easily hear a scale-like sequence of tone-like percepts. As the scale was clearly perceived, this would indicate that the pitch of each tone-like percept corresponded to the frequency of the altered component (see also Bregman et al., 1994a, discussed below). Kubovy & Daniel (1983) also noted that a relatively small amplitude decrement (3 dB) remained effective at generating a tone-like percept. Furthermore, Kubovy (1981) remarked that subjects were only able to hear the tone-like percepts when the components were separated in frequency by more than a critical band.

Bregman et al. (1994a) conducted a more systematic study of the factors influencing the salience of tone-like percepts arising from amplitude increments. These authors presented listeners with a 3.5-s complex tone, which comprised the first five harmonics of a 500-Hz fundamental (presented at 65 dB SPL). After 500 ms, a single component was incremented in amplitude for 1 s, and then, 500 ms later, a second 1-s amplitude increment was applied to a different harmonic. The amplitude increments were applied to the 3^{rd} and 4^{th} harmonics in either ascending or descending order, and listeners were required to report the presentation order of the increments. The experiment examined the effect of onset time and of increment size on the discriminability of presentation order. Listeners were generally good at the task, suggesting that the amplitude increments were heard as segregated from the complex tone (i.e., heard as tone-like percepts). The size of the increment was varied and performance was good at all the levels tested (1 dB - 6 dB), but did improve slightly with increasing increment size. The onset duration of increments was also varied (30 ms - 970 ms) and it was found that faster rise times led to better performance.

These findings also relate to the experiments of Bregman et al. (1994b). These authors studied the effects of the suddenness either of onset or offset on the discrimination of the order of pitches of individual tones in a 1 s, four-tone cluster of overlapping pure tones, where the presentation order was either MLHM or MHLM (where L = 750 Hz, M = 800 Hz, and H = 850 Hz). The onset times of these tones was varied (60, 80, or 100 ms onset asynchrony), but all ended at the same time. All tones also shared the same linear rise and fall times. Generally, it was found that, for the range of amplitude increments tested, faster rise times (range = 10 – 640 ms) resulted in improved discrimination of the order of L- and H-tone onsets (see also Pastore et al., 1982). The only exception to this trend was that performance was slightly worse for the 10-ms than for the 40-ms rise times. This was attributed to the fastest rise time having some disruptive effect, as Bregman et al. (1994b) noted that the trend disappeared over the course of the experiment (as listeners became more familiar with the stimuli). Using similar stimuli, Bregman et al. (1994b) also measured the judgment of offset asynchrony (i.e., when all tones began at the same time). A generally similar trend of results was observed (faster offsets = improved discrimination), but in absolute terms, listeners were much more sensitive to onset order than offset order.

Bregman et al. (1994a, 1994b) argued that these findings indicate that a sudden amplitude change may "reset" the pitch analysis system. In such a circumstance, they proposed that the amplitude-incremented harmonic perceptually dominates the new analysis (in comparison to the unchanged components), and so stands out as a tone-like percept. As more rapid increments created the most salient tone-like percepts, this was taken as evidence that abrupt changes were most effective at resetting the pitch analysis system. Given that these results were observed for amplitude changes from silence (Bregman et al., 1994b), or from an on-going complex-tone (Bregman et al., 1994a), it was argued that any sudden amplitude change is effective at resetting the pitch analysis system. Bregman et al. (1994b) related the proposed resetting of the pitch-analysis system to the resetting effect of an abrupt change on the build-up of stream segregation in a sequence of discrete pure tones (Rogers & Bregman, 1998). Whilst Bregman et al. (1994b) acknowledged that there may be other interpretations for their data, these authors did speculate that the functional significance of these forms of resetting may be similar - namely that the perception of a new event may cause a reanalysis of all acoustic stimuli.

9.2 Experiment 8

Experiment 8 explored whether pure-tone-like percepts created by applying amplitude increments to an on-going complex tone are organized into perceptual streams. Whilst previous studies have demonstrated that listeners are able to report the presentation order of two successive increments (Bregman et al., 1994a), this does not provide a rigorous test of sequential grouping, as there were no competing perceptual organizations (indeed, these authors did not claim to be studying sequential grouping). The current experiment explored sequential organization by replicating the experimental conditions used by Bregman & Rudnicky (1975), but using a sequence of amplitude increments applied to an on-going complex-tone instead of a sequence of discrete pure tones. As streaming is best demonstrated in sequences of relatively brief stimuli, the increments used in Experiment 8 were much briefer (50 ms) than those used by Bregman et al. (1994a) (1 s). The current amplitude increments were designed to be highly salient, as the experiment concerned the perceptual organization of sequences, as opposed to the ability to detect the individual increments.

9.2.1 Method

Participants

Twelve listeners contributed to the main experiment. All reported having normal hearing. None of the listeners had previously taken part in any other auditory-perception experiments.

Stimuli

On each trial, subjects heard a complex tone comprising the first seven, resolved, harmonics of a 300-Hz fundamental (Plomp & Mimpen, 1968; Moore, 2003). The complex tone lasted for 1.2 s in total, including linear onset and offset ramps of 10 ms each. All of the harmonics began in sine phase and each was set to 65 dB SPL. From this complex tone, brief pure-tone-like percepts were created by applying an amplitude increment to an individual harmonic. The specific sequence of amplitude increments varied with each of the conditions tested, but each individual increment was created in the same way. Specifically, a 6-dB amplitude increment was applied to an individual harmonic for 50 ms (including 10-ms linear onset and offset ramps). These amplitude increments were sufficiently large to create clear pure-tone-like percepts.

Conditions

Four different conditions were tested in Experiment 8. For all of these conditions, two target increments occurred in rapid succession on the 5^{th} (1500 Hz) and the 6^{th} (1800 Hz) harmonics. Note that the semitone difference between these targets was approximately 3.2 ST. This difference was larger than the (approximately) 1.5 ST difference used in the previous experiment by Bregman and Rudnicky (1975). The onsets of these increments were 800 ms and 850 ms after the onset of the complex tone (i.e., in immediate succession). These two increments could be sequentially arranged as either ascending or descending in frequency. For the targets-only condition, no other amplitude increments were present. In the other three conditions, the target increments were embedded in a sequence of additional task-irrelevant increments. For the distractors condition, two increments on the 4^{th} harmonic (1200 Hz) occurred immediately before and after the target increments (onset times = 750 ms and 900 ms after the onset of the complex tone). For the matched-captors condition, the target and distractor increments were present, and were accompanied by five additional captor increments. The captor increments occurred on the same harmonic as the distractor increments, and shared the same onset-to-onset time (150 ms). Four captor increments occurred before the distractor and target increments, and one final captor occurred afterward (after 150, 300, 450, 600, and 1050 ms from the onset of the complex tone). Therefore, the captor and distractor increments were heard as an isochronous and monotonous sequence. Using pure-tone stimuli, Jones et al. (1981) demonstrated that captor tones were most effective at capturing the distractors when both tone sets formed an isochronous sequence (in comparison to when there was a difference in rhythm). Finally, the remote-captors condition was identical to the matched-captors condition, except that the captor increments were applied to the 2^{nd} harmonic (600 Hz), whilst the distractor increments remained on the 4^{th} harmonic (1200 Hz). This configuration corresponds to Bregman & Rudnicky's (1975) low-frequency captors condition. A schematic of the remote-captors condition is shown in Figure 9.2.

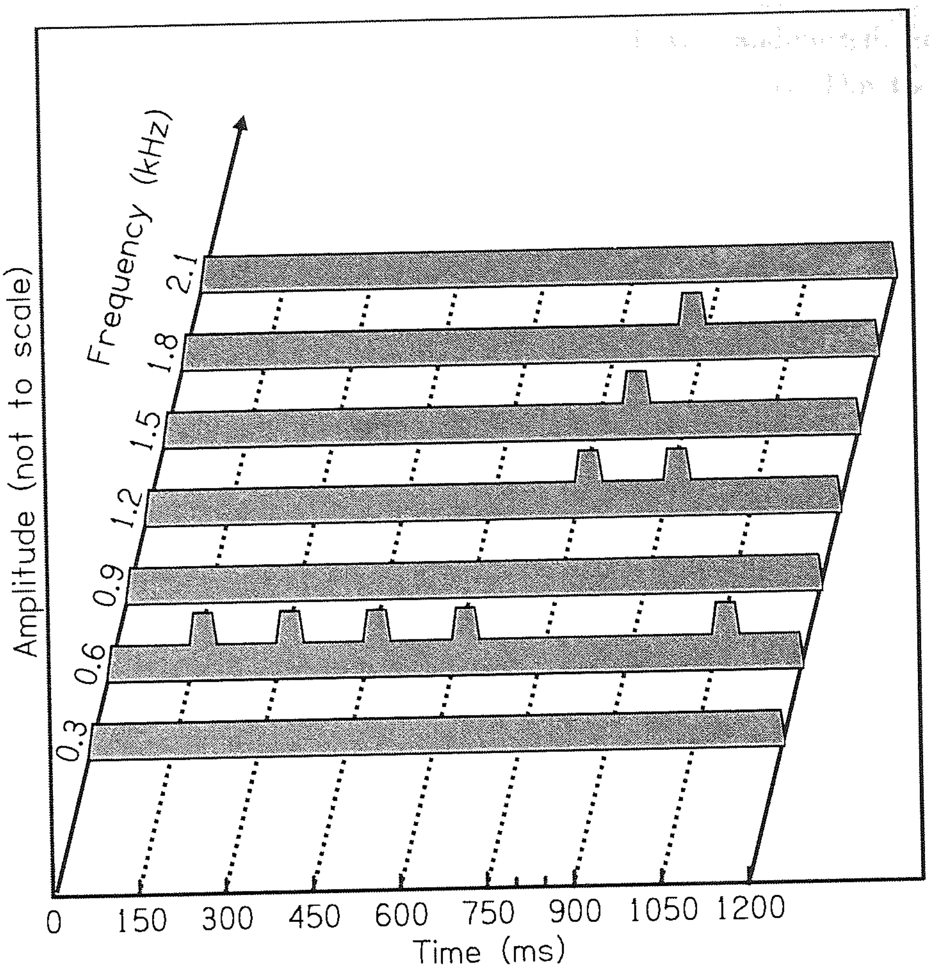

Figure 9.2: **Illustration of the remote-captors condition tested in Experiment 8. Each tone-like percept is created by the application of an amplitude increment to a single component of the harmonic complex tone. The two target increments are on the 5th and 6th harmonics (1500 & 1800 Hz), the two distractors are on the 4th harmonic (1200 Hz), and the five captors are on the 2nd harmonic (600 Hz).**

Procedure

On each trial, a single presentation of one condition was heard. At the end of the stimulus, the subject was asked if the pitch pattern of the target increments had gone up or down. Subjects reported either an 'ascending' or a 'descending' percept (2AFC). There was a 500-ms pause following the response, after which the next trial began automatically. No feedback on performance was given (the same was true for Bregman & Rudnicky's (1975) earlier experiment). Before the main experiment, listeners attended a combined screening and training session. This session comprised 20 repetitions of the targets-only condition, which were divided equally into ascending and descending configurations (the trial order was randomized). Subjects were required to identify correctly >65% of trials to proceed to the main experiment. Out of 14 listeners, two failed this training - leaving 12 participants to contribute to the main experiment. The main experiment was divided into twenty blocks. Each block contained an ascending and descending configuration for each of the four conditions (4 conditions × 2 configurations

= 8 trials per block). The trial order within each block was randomized. For the main experiment, each subject completed 20 blocks (160 trials in total). This took about 30 minutes, and was completed in one session.

It should be noted that the current task differed slightly from that used by Bregman & Rudnicky (1975). For their experiment, listeners were asked whether two sets of target tones were the same or different, but for the current experiment, listeners were required to report the presentation order of a single set of target tones (ascending or descending). A pilot study confirmed good performance in the targets-only condition of the current, simpler design.

Apparatus

The apparatus used was identical to that described in Chapter 2.

9.2.2 Results

For each condition, results from all listeners were averaged to give an overall percentage of correct responses. For the targets-only, distractors, matched-captors, and remote-captors conditions, these means were 78.8%, 67.3%, 77.7%, and 66.5%, respectively (chance level = 50%). These means are displayed with corresponding inter-subject error bars in Figure 9.3. Performance was better for the targets-only and matched-captors conditions than for the distractors and remote-captors conditions. These results were assessed using a one-way, repeated-measures ANOVA. This analysis confirmed a significant main effect of the four conditions tested [$F(3, 33) = 9.721,\ p < 0.001,\ \eta_p^2 = 0.469$].

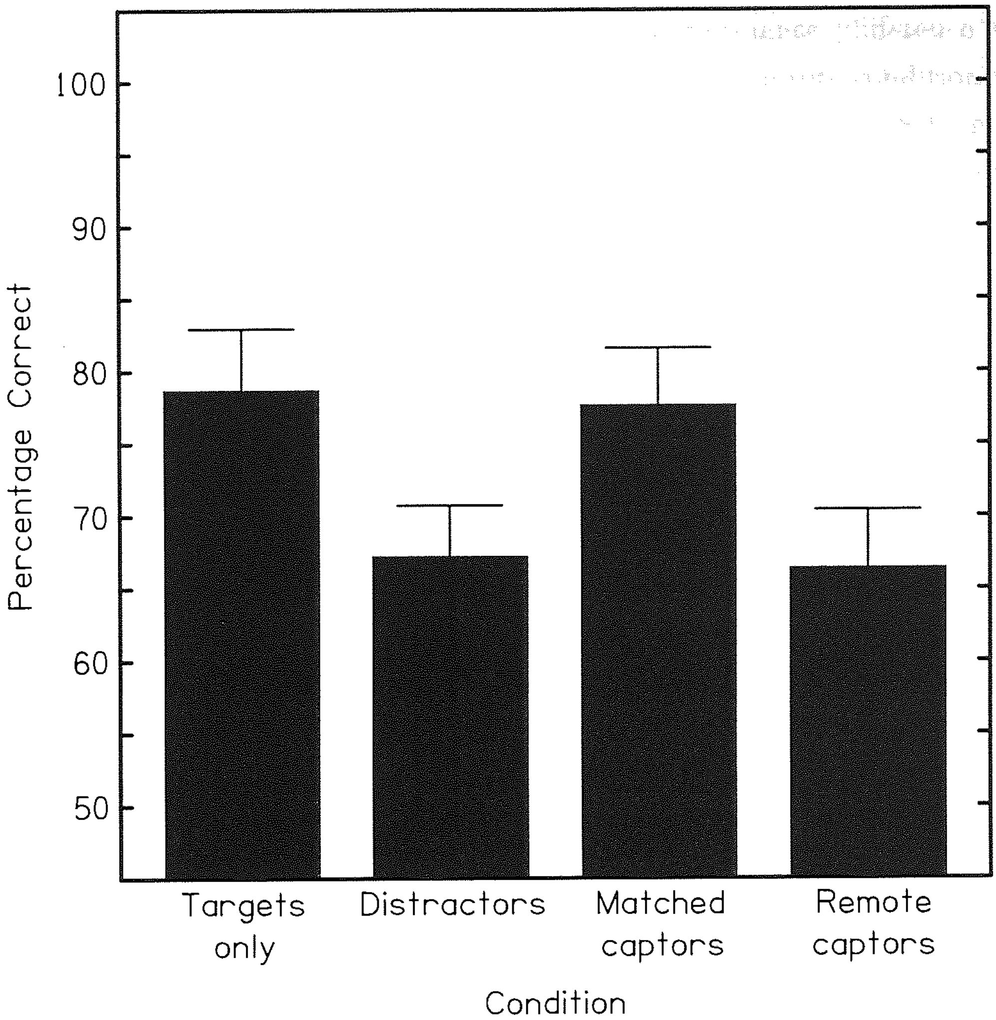

Figure 9.3: **Results from Experiment 8. Performance for each condition is expressed as an overall percentage of correct responses (from 12 participants). The error bars display +1 inter-subject standard error.**

Multiple comparisons were conducted using the Student-Newman-Keuls (S-N-K) test. As this test is appropriate for multiple comparisons of between three to five groups (Howell, 1992), it was considered a suitable analysis for the current experiments (8, 9, and 10). For each comparison the calculated q value was compared to tables of critical q values representing either the 0.05 or the 0.01 significance levels (Glantz, 2005; Sheskin, 2004). These statistics are reported along with a measure of the difference in performance between the two conditions (expressed as a percentage). There was a significant difference in performance between the targets-only and the distractors conditions [difference = 11.5%, $q(k = 3,\ df_{error} = 33) = 5.431,\ p < 0.01$]. This indicates that the presence of distractors significantly impaired performance. Adding captor increments on the same harmonic as the distractors restored performance to a level almost identical to that in the targets-only condition [difference = 1.1%, $q(2, 33) = 0.494,\ p > 0.05$], and there was a significant difference between the matched-captors and distractors conditions [difference = 10.4%, $q(2, 33) = 4.937,\ p < 0.01$]. The presence of captors on a different harmonic (one octave below) than the distractors did not aid performance [difference = 0.8%, $q(2,33) = 0.395$, $p > 0.05$], and performance in the remote-captors condition was significantly below that for the targets-only condition [12.3%, $q(4,33) = 5.826$, $p < 0.01$]. Finally, performance in all conditions was found to be above chance, as all conditions were significantly different from a comparison value of 50% when assessed using one-sample t-tests ($p > 0.05$ in all cases).

9.2.3 Discussion

Experiment 8 examined whether a sequence of pure-tone-like percepts created by applying brief amplitude increments to single components of a complex tone is organized into perceptual streams. It was found that participants were good at discriminating the presentation order of two target amplitude increments presented in isolation. This finding is consistent with an experiment by Bregman et al. (1994a), who demonstrated that listeners were generally good at detecting the presentation order of two successive amplitude increments on different harmonics of a complex tone (increment duration = 1 s). The current study extended this finding by demonstrating that detection of relatively brief amplitude increments (current experiment = 50 ms duration) is also good.

The key finding is that listeners appear to organize sequences of pure-tone-like percepts evoked by sudden amplitude increments in essentially the same way as they organize sequences of discrete pure tones. The presence of the distractor increments significantly impaired performance, but this impairment was lost when the targets and distractors were accompanied by captor increments on the same harmonic as the distractors. This result indicates that the distractors grouped with the captors, and so were heard as part of an on-going isochronous stream, perceptually segregated from the targets. This argument is supported by the fact that captor increments on a lower harmonic than the distractors did not aid participants - performance was similar to that for the distractors condition. This is taken as evidence that the captors were only effective at grouping with the distractors when they were of a similar frequency - in other words, sequential grouping of distractor and captor increments is frequency specific. These results are largely consistent with Bregman & Rudnicky's (1975) study, which found similar results using isolated pure-tone stimuli. This argument could have been further strengthened if the current experiment had also tested performance in a set of pure-tone-only conditions. An attempt to provide this comparison failed, because it proved difficult to replicate Bregman & Rudnicky's (1975) basic findings. A pilot study used very similar stimuli, and an identical method to that tested by Bregman & Rudnicky (1975). Despite this, almost all listeners performed at a near-perfect level for all conditions. The reason for this is unclear, but Bregman & Rudnicky (1975) did note that a number of their participants did also perform at a very high level (>90% correct in all conditions).

Although relative performance was very similar to that observed by Bregman & Rudnicky (1975), listeners generally performed better in the current experiment. Bregman & Rudnicky (1975) rejected a number of participants whose performance was either near-perfect or near-chance. However, for their unfiltered data set, performance for the distractors condition was 63%, whilst performance for the matched-captors condition was 70% (note that they did not test a targets-only condition, owing to their use of a same-different task). For the current experiment, performance was 67.3% and 77.7% correct for the corresponding conditions, respectively. Whilst this is quite surprising, as the current judgments were made against the background of an on-going complex tone, there are several factors which could explain this difference. First, one should note that two listeners were removed from the current experiment due to poor performance in the training session. Second, the frequency difference between the target increments (3.2 ST) was larger than the difference between the two target tones of Bregman & Rudnicky's study (1.5 ST). Third, the apparent improvement could be due to the difference in design between the two experiments - for the current study, listeners were required to detect the order of a single pair of target tones, whereas Bregman & Rudnicky (1975) asked listeners to determine whether two sets of target tones were either the same or different. One might expect better performance in the current design, as listeners were only required to detect the order of a single pair of target tones. Indeed, for the current experiment, performance for the targets-only condition was not perfect (78.8% correct). Therefore, one might predict worse performance if a listener were required to make two of these decisions for each trial, although sensitivity to a *difference* between two sets of target tones may not necessarily require an explicit judgement of the tone order within each pair. Despite these differences, the current results remain largely consistent with those of Bregman & Rudnicky (1975). Therefore, it is concluded that tone-like percepts arising from amplitude increments form distinct perceptual streams in much the same fashion as isolated pure tones.

As noted in the introduction, tone-like percepts can also be created using abrupt changes in ITD (Kubovy & Daniel, 1983). If the explanation offered for the results of the current experiment is correct, then the means by which simultaneous grouping cues are used to create these tone-like percepts should not matter. Specifically, one would predict that the current results could be replicated if the tone-like percepts were created using ITD cues rather than amplitude-increment cues. This hypothesis was tested in Experiment 9.

9.3 Tone-Like Percepts Evoked by ITD and Monaural Phase Cues

Kubovy & Daniel (1983) demonstrated that salient tone-like percepts can be created by applying an ITD change to a single component of an on-going complex tone. These authors presented an eight-component chord (i.e., components equally spaced on a logarithmic scale), where the component frequencies corresponded to the C-major scale. Aside from brief ITD changes, the chord was otherwise presented with no ITD (and so heard to occur straight ahead). Part-way through the on-going tone, an ITD was imposed on a single component by the application of a 45 $^\circ$ phase difference between the two ears (so that the component was heard on the right side of space). This brief ITD resulted in a pure-tone-like percept, heard with a distinctly different lateralization from the on-going complex tone. When a series of these ITD changes were applied sequentially to different components of the on-going complex tone, listeners could easily follow the tone-like percepts, and so hear a musical scale. However, each phase change was applied instantaneously, and so each tone-like percept was accompanied by an audible "click". To ensure that this click itself did not capture attention to the specific phase-changed component, the phases of all the non-changed components were reset to 0 $^\circ$ at the same time as the component-specific phase shift. Note, however, that this did not change the fact that each note on the scale occurred alongside a click. A second consideration when evaluating this study is that each tone-like percept was created by the same phase shift. Therefore, when this phase shift was applied to higher frequency components, the resulting change in perceived lateralization would have been reduced. This was not an issue for the current experiment, as lateralization changes were evoked by a constant change in ITD (i.e., a constant delay across all frequencies).

In an earlier study, Kubovy et al. (1974) presented broadly similar stimuli but applied each phase change gradually rather than instantaneously. The complex tone was an eight-component chord (duration = 24 s) which was heard at one side of space (an ITD of +1 ms was applied to all components for the duration of the chord). Part-way through the on-going complex tone, an ITD change was briefly applied to, and then removed from, a single component (an ITD change of -2 ms, so that the component was heard at the opposite side of space to the complex tone). In contrast to Kubovy & Daniel's (1983) study, the phase changes were gradually imposed and then removed using 45-ms ramps (the duration of the steady-state portion varied). Using this configuration, Kubovy et al. (1974) applied a series of brief ITD changes to different components of the complex tone - as one ITD was removed from a component, a subsequent ITD was simultaneously applied to a different component. As each ITD change generated a clear tone-like percept, the melody presented ("Daisy, Daisy...") was easily recognized by listeners. Despite the relatively gradual application and removal of each ITD change, Kubovy et al. (1974) noted that the associated phase change remained noticeable under monaural listening (when listening through one ear only - so that no ITD cues could be heard). These authors noted that these monaural phase cues were heard as a faint perturbation to the on-going tone. These cues were not especially salient - indeed, subjects were unable to recognize the melody when the standard sequence was presented to one ear only. Factors concerning the salience of monaural phase cues are considered later in this section.

Kubovy & Daniel (1983) suggested that a *change* in ITD is responsible for the "hearing out" of a single component, as a static ITD does not result in a single component being heard as segregated from a complex tone. Kubovy et al. (1974) noted that the onset of an ITD-generated tone percept sounded particularly striking - and was heard like a chime. In contrast, tone offsets were not particularly salient (note a similarity to Bregman et al. (1994b), who showed that the onset order of four tones was easier to discriminate than the offset order). The role of ITD changes in the segregation of components has since been studied further by Culling (2000). Using broadly similar stimuli, he reported similar findings to those of Kubovy et al. (1974), namely that when different components of a complex tone were successively changed in ITD, a recognizable melody could be detected (Big Ben's chimes). However he found that an ITD transition only generated a tone-like percept when the "background" complex tone remained static in ITD. Tone-like percepts were not heard when the changed component was part of an otherwise coherently changing complex tone (where all components gradually changed in ITD for the entire duration of the tone). In other words, specific ITD changes relative to a changing background were inaudible. Like Kubovy & Daniel (1983), Culling (2000) also observed that very rapid ITD changes were effective at evoking tone-like percepts.

Based on these findings, Culling (2000) rejected Kubovy et al.'s (1974) suggestion that the perception of movement is responsible for the segregation of the changing component from the static complex tone. Rather, Culling (2000) proposed that abrupt ITD changes are salient because the auditory system cannot identify the target component's ITD at the time of the transition. The auditory system is thought to detect ITDs through a process of cross-correlating the input arriving at the two ears (Jeffress, 1948). However, rapidly changing ITDs or multiple ITDs occurring simultaneously can be poorly defined, as the binaural signal may become decorrelated (and the auditory system is highly sensitive to binaural decorrelation, see Saberi et al., 1998). Given that a very rapid, or even instantaneous, ITD change can generate a clear tone-like percept, target segregation cannot be explained in terms of perceived movement. However, Kubovy et al.'s (1974) listeners would have been sensitive to the local inter-aural decorrelation against a continuing static ITD on the other components. Furthermore, as already noted, Culling (2000) also demonstrated that a component-specific ITD change was not salient when the entire complex tone was gradually changing in ITD. For this stimulus, any perception of movement would be relatively unaffected, but if listeners were using inter-aural decorrelation to detect the component with an abrupt ITD change, performance would be poor, as all components would be decorrelated at all times. Therefore, there would be no cues for the segregation of the target component(s) from the background components. Based on these findings, Culling (2000) concluded that rapidly changing ITDs are salient due to component-specific inter-aural decorrelation, rather than as a consequence of a motion-detection system.

As previously noted, the process of applying or removing an ITD to a single component of a complex tone generates a localized phase disparity between components (at the ear of the phase change). These phase disparities may be noticeable even if they are presented uniformly to both ears (i.e., no ITD present, only localized phase changes). These monaural phase cues are especially salient in harmonic complex tones (Kubovy, 1981), but less so in non-harmonic stimuli (Kubovy & Daniel, 1983). Kubovy & Jordan (1979) suggested monaural phase cues may be salient due to the compressive, non-linear transfer characteristics of the auditory system. They calculated that a localized phase shift to a single component will generate a peak in the power spectrum at the same frequency as the phase-shifted component. In other words, a monaural phase shift may in effect produce a localized amplitude change. Whilst these phase cues cannot be avoided entirely when presenting ITD changes, steps were taken to minimize their perceptual salience in the current experiment.

9.4 Experiment 9

For Experiment 9, pure-tone-like percepts were created with a different perceived lateralization from an on-going complex tone by changing the ITD of a single component. By analogy with Experiment 8, a sequence of ITD changes was applied to different components of the on-going complex-tone. The grouping of these tone-like percepts was examined using the same procedure as the previous experiment (i.e., an adapted version of Bregman & Rudnicky's (1975) method). As well as creating tone-like percepts using ITD cues rather than amplitude-increment cues, the stimuli used for Experiment 9 also differed from those used for Experiment 8 in several other ways. To reduce the salience of monaural phase cues, a non-harmonic complex tone was used (Kubovy & Jordan, 1979). Monaural phase cues were also reduced by applying the ITD change with relatively gradual onset and offset ramps (in comparison with the amplitude ramps used in Experiment 8). This change had a slight effect on the relative timings of the abrupt changes during the sequence. Despite these measures, it is plausible that monaural phase cues may still have remained perceptually salient. Therefore, as a precaution, an additional control condition was included in which performance was measured when only monaural phase cues were present (i.e., the same phase changes were applied to both ears and so that there was no change in ITD).

9.4.1 Method

Participants

Twelve listeners contributed to the main experiment, all of whom reported having normal hearing. Ten of the participants who completed Experiment 9 had previously completed Experiment 8.

Stimuli

The background complex tone was non-harmonic; the tone contained seven components that were equally spaced on a logarithmic scale. The lowest frequency component was 160 Hz, and the frequencies of higher components were spaced in five-semitone steps (i.e., 160, 214, 285, 381, 508, 678 & 905 Hz). Note that all of the components were separated from their immediate neighbours by more than 1 ERB_N (Glasberg & Moore, 1990). The smallest ERB separation was between the 160 and 214 Hz components (where the separation = 1.13 ERB_N), whilst the largest separation was between the 678 and 905 Hz components (1.85 ERB_N). The frequencies of the components were reduced from those of Experiment 8 as ITD cues begin to become ambiguous for frequencies above about 750 Hz (see Moore, 2003). In the current experiment, no ITD was applied to the 160-Hz and 905-Hz components. The total duration of the complex tone was 1.5 s, including linear onset and offset ramps of 10 ms each. Each of the components started in sine phase and was presented at 65 dB SPL. The complex tone had no ITD, and so was heard to occur straight ahead.

From this complex tone, a series of tone-like percepts were created (the number and timings of these percepts varied with the different experimental conditions). Each tone-like percept was created in exactly the same way - by temporarily applying an ITD to a single component of the complex tone. The component led at the right ear for 60 ms, although the change in ITD was gradually applied and then removed using 20-ms ramps. To reduce the salience of monaural phase cues, these ramps were twice as long as the equivalent amplitude ramps used in Experiment 7. Nonetheless, the resulting ITD changes were still perceived as relatively abrupt. Over the course of each 60-ms perturbation, the component first underwent a linear transition in ITD from 0 ms to 0.5 ms over 20 ms. The component was then held at an ITD of 0.5 ms for a further 20 ms, and in the final 20 ms it was returned to its original ITD value through a linear transition in ITD from 0.5 ms to 0 ms.

Conditions

For convenience, the 60-ms component-specific changes in ITD are referred to as 'notes'. The target notes occurred on the 5th (508 Hz) and the 6th (678 Hz) components. This was a 5-ST difference between the target notes. Note that this difference was larger than that used in both Experiment 8 (approximately 3.2 ST), and the 1.5-ST difference used by Bregman & Rudnicky (1975). Their onset times were 1000 ms and 1065 ms after the onset of the complex tone. This resulted in a 5-ms interval between the offset of the first target note and the onset of the second target note, during which no ITD was present on any of the components. This change was made as a pilot study indicated that a brief interval between the target notes assisted performance. The target notes could be arranged as either ascending or descending in frequency. Five conditions were tested, four of which were analogous to those used in Experiment 1. For the targets-only condition, the two target notes were unaccompanied by any other notes. For the distractors condition, two distracting notes on the 4th component (318 Hz) flanked the target notes. The onset times of these notes were 935 ms and 1130 ms after the onset of the complex tone. As for the two target notes, there was a 5-ms interval between each distractor and the neighbouring target note. In the matched-captors condition, five captor notes formed an isochronous sequence with the distractor notes (onset-to-onset time = 195 ms). As for Experiment 8, four captors occurred before the first distractor (155, 350, 545 and 740 ms after the onset of the complex tone), and one captor occurred after the second distractor (1325 ms after the onset of the complex tone). For the matched-captors condition, the captor notes were heard on the same component as the distractors (the 4th component, 318 Hz). For the remote-captors condition, the captor notes were placed on the 2nd component (214 Hz), but were otherwise identical to the captor notes previously described. Note that the frequency difference between the distractors and the remote-captors was less than the one octave reduction used in Experiment 8. An illustration of the remote-captors condition is shown in Figure 9.4.

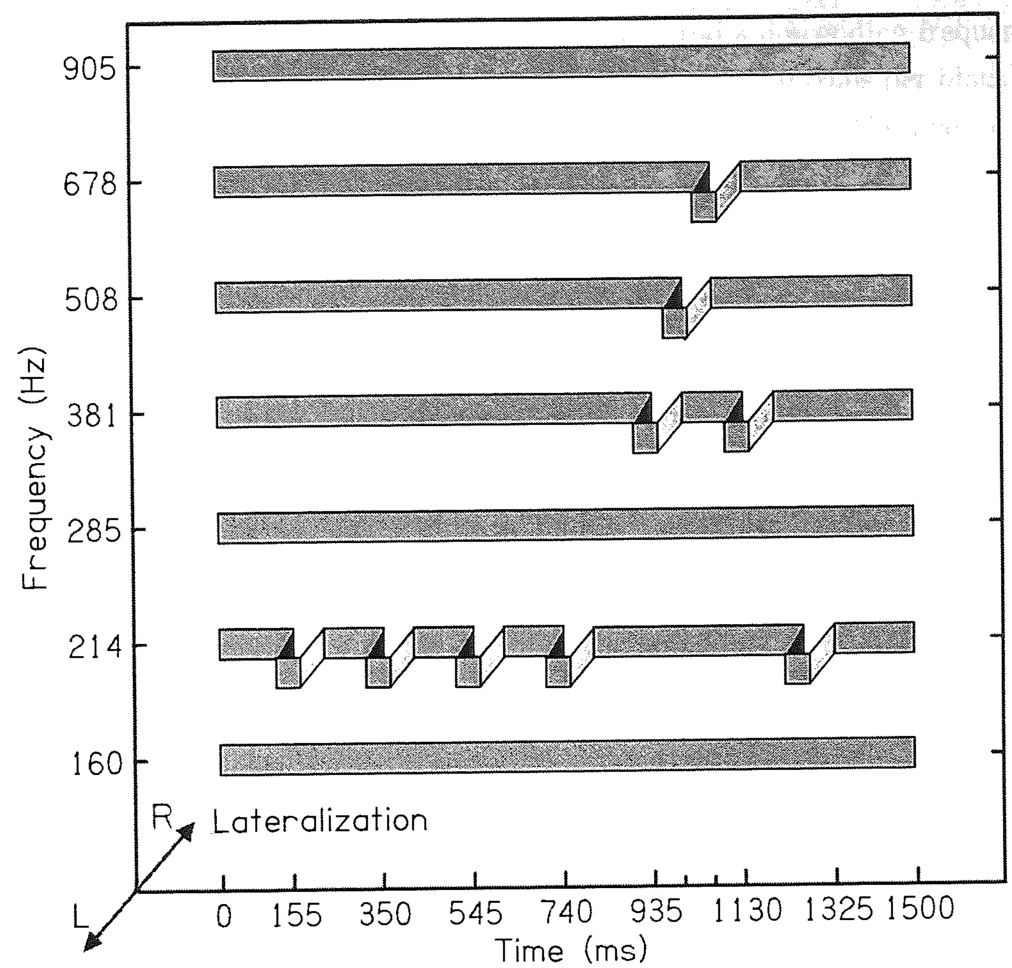

Figure 9.4: **Illustration of the remote-captors condition tested in Experiment 9. Each tone-like percept is created by the application of a 60-ms ITD to a single component of the complex tone (complex tone heard to occur straight ahead, tone-like percepts heard at the left side of space). The two target increments are on the 5th and 6th components (508 & 678 Hz), the two distractors are on the 4thcomponent (381 Hz), and the five captors are on the 2nd component (214 Hz).**

For the targets-monaural condition, the phase changes used to generate the ITD cues were applied to both ears in unison. For this condition, the local monaural phase changes associated with the ITD cues were present, but there was no difference in the signal sent to each ear. Aside from this difference, the targets-monaural condition was identical to the targets-only condition. The targets-monaural condition was included to assess the perceptual salience of the monaural phase cues. As the stimulus was designed to minimize these cues, it was predicted that performance for this condition would be poor.

Procedure

The training/screening method, and the general procedure used, were the same as those for Experiment 8. The task was to report the presentation order of the two target notes as either ascending or descending in pitch in a 2AFC procedure. Out of 13 listeners, only one person failed the training. For the main experiment, a trial block comprised one presentation of each condition in both an ascending- and a descending-frequency target configuration (i.e., 5 conditions $\times$ 2 target configurations = 10 trials per block). For the main experiment, 20 blocks were presented in total (200 trials). The main experiment lasted about 40 minutes, which each listener completed in one session.

Apparatus

The apparatus was identical to that described in Chapter 2, except that the sampling rate was increased from 20 kHz to 40 kHz. This change was made to reduce further the salience of monaural phase cues by doubling the number of steps over which the ITD was introduced and removed during the 20-ms ramps. The maximum ITD change used here was 0.5 ms, which corresponds to 20 samples (applied as a cumulative inter-aural delay of one sample per ms) at this higher sampling rate.

9.4.2 Results

For each condition, results from all listeners were averaged to give an overall percentage of correct responses. For the targets-only, distractors, matched-captors, remote-captors, and targets-monaural conditions, these means were 85.7%, 67.9%, 81.9%, 69.2% and 57.9%, respectively. These means are displayed with corresponding inter-subject standard errors in Figure 9.5. The results are as predicted - performance was best for the targets-only and matched-captors conditions, and reduced for the distractors and remote-captors conditions. The worst performance was observed in the targets-monaural condition. These results were assessed using a one-way, repeated-measures ANOVA. This analysis showed a significant main effect of condition [$F(4, 44) = 22.736$, $p < 0.001$, $\eta_p^2 = 0.674$].

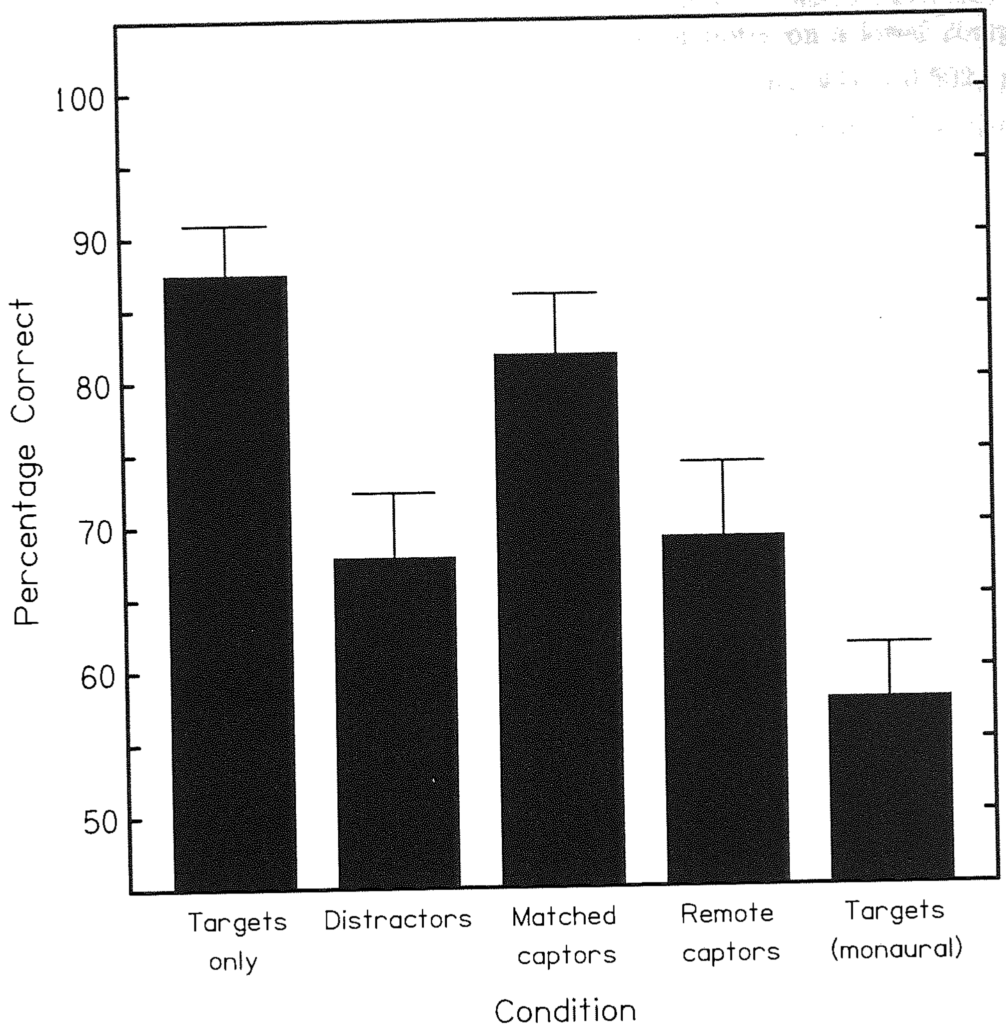

Figure 9.5: **Results from Experiment 9. Performance for each condition is expressed as an overall percentage of correct responses (from 12 participants). The error bars display +1 inter-subject standard error.**

Multiple comparisons were conducted using the Student-Newman-Keuls test[1]. These statistics are presented along with a measure of the difference in performance between the two conditions (expressed as a percentage). There was a significant difference in performance between the targets-only and distractors conditions [difference = 19.6%, $q(k = 4, df_{error} = 44) = 7.886$, $p < 0.01$], indicating that the presence of the distractors greatly impaired performance. Adding captor notes on the same component as the distractors (matched-captors condition) improved performance considerably [difference = 14.0%, $q(3, 44) = 5.621$, $p < 0.01$] and restored performance to that for the targets-only condition, so that the remaining difference was non-significant [5.6%,

[1] From the Studentized range table (Glantz, 2005; Sheskin, 2004), critical q values were were unavailable for $df_{error} = 44$ Instead, these values were obtained from the nearest approximation of $df_{error} = 40$. As for Experiment 8, q values are only compared to the 0.05 and the 0.01 significance levels.

$q(2, 44) = 2.265$, $p > 0.05$]. In contrast, adding captor notes on a lower component to the distractors did not aid performance [difference = 1.3%, $q(2, 44) = 0.503$, $p > 0.05$], and performance in the remote-captors condition was significantly below that for the targets-only condition [difference = 18.3%, $q(3, 44) = 7.382$, $p < 0.01$].

Finally, performance for the targets-monaural condition was poor. The percentage of correct responses for this condition was significantly lower than that for the equivalent condition in which the targets were presented with an ITD difference associated with the same monaural phase cues [29.6%, $q(5, 44) = 11.912$, $p < 0.01$]. Indeed, performance for the targets-monaural condition was not greatly above chance level. This was reflected in a one sample t-test, which found that performance in the targets-monaural condition was not significantly different from a comparison value of 50% [7.9%, $t(11) = 2.168$, $p > 0.05$]. When assessed using the same measure, all of the other (ITD based) conditions were significantly above chance (comparison value = 50%, $p < 0.05$ in all cases).

9.4.3 Discussion

Experiment 9 examined whether tone-like percepts evoked by applying brief ITD changes to single components of a complex tone are organized into perceptual streams. Such an ITD change can create a highly salient pure-tone-like percept which is heard at a separate location from the on-going complex tone (Kubovy et al., 1974). To examine whether these tone-like stimuli are organized into perceptual streams, the current study largely replicated Bregman & Rudnicky's (1975) design. The results from Experiment 9 are largely consistent with those of the Bregman & Rudnicky (1975) study, which used discrete pure tones, and with those of Experiment 8, which used pure-tone-like percepts evoked by sudden changes in amplitude. Performance in the targets-only condition was good (see Figure 9.5). This indicates that the tone-like percepts were perceived clearly, and that listeners were sensitive to the presentation order of the two target notes. Performance was impaired considerably when the target notes were flanked by two distractor notes. This is consistent with Bregman & Rudnicky's (1975) results, and is taken as evidence that the targets and the distractors grouped to form a single perceptual object (for which the order of the interior targets was less salient). Performance was restored when the distractors were embedded within a sequence of frequency-matched captor notes, indicating that the captors and distractors grouped together so that the targets were heard separately from this on-going stream. However, captor notes on a lower component, spectrally remote from the distractors, did not aid performance. This is taken as evidence that captor/distractor grouping is frequency specific. These results are entirely consistent with those observed by Bregman & Rudnicky (1975) and with those of Experiment 8. Therefore, it is concluded that tone-like percepts arising from an abrupt change in the level or lateralization of a component are organized into streams, in a way that is consistent with the organisation of isolated pure tones.

For the current experiment, the tone-like percepts were created by applying, briefly maintaining, and then removing a phase change to a single component of the complex tone. However, such phase changes may be audible under monaural listening conditions (when the same phase change is presented to both ears, so that no ITD is heard - Kubovy & Jordan, 1979). For the current experiment, several steps were taken to reduce the salience of these monaural phase cues. These steps were largely successful, as performance in the targets-monaural condition was not significantly above chance. This indicates that these monaural cues were not of any real benefit in performing the task. Therefore, the differences in performance observed for the four main (ITD-based) conditions must primarily reflect differences in the ability of listeners to hear the salient ITD changes, as opposed to the less salient monaural phase cues.

Overall, it is concluded that tone-like percepts arising from ITD cues are organized into perceptual streams in essentially the same way as for sequences of isolated pure tones. This conclusion is consistent with that of Experiment 8, which demonstrated streaming using tone-like percepts that were created by applying amplitude increments to single components of a complex tone. Given that stream segregation has been demonstrated for both types of stimulus, it seems likely that sequences of any form of tone-like percept may be organized into perceptual streams. This conclusion is consistent with Moore & Gockel's (2002) hypothesis that stream segregation may occur on the basis of *any* salient perceptual differences between sounds. Furthermore, the current experiments have shown that tone-like percepts can be organized into streams either in the presence of a harmonic or a non-harmonic background tone (experiment 8 and 9 respectively). This further demonstrates that the effect is robust enough to be observed in a variety of sequence arrangements.

9.5 Experiment 10

Experiment 10 had two general aims. The first aim was to examine whether resetting of the build-up of stream segregation can be demonstrated in a sequence of tone-like percepts evoked using amplitude increments applied to components of a complex tone (Bregman et al., 1994a, and Experiment 8). If so, the second aim was to develop an objective measure of the extent of resetting more efficient than the temporal-discrimination threshold measure used in Experiment 2. The approach was to introduce a further modification of Bregman & Rudnicky's (1975) design to that already used in Experiments 8 and 9.

The experiments presented in this thesis have consistently demonstrated that a single, noticeable change to an on-going sequence can substantially reset reported stream segregation (see Experiment 1, for an example). So far, this active resetting effect has only been demonstrated for sequences of discrete pure tones. As tone-like percepts are organized into perceptual streams (Experiments 8 and 9), one would predict that an abrupt change would also have a resetting effect on these types of stimuli. Experiment 10 tested this hypothesis for sequences of tone-like percepts created from brief, component-specific, amplitude increments.

Previous experiments have clearly demonstrated active resetting using a specific type of tone sequence (see Experiments 1, 3a, 3b, 5 and 7). For these experiments, resetting was observed when the final tone of a same-frequency induction sequence (10 L-tone repetitions) was noticeably altered, and streaming was measured in a subsequent test sequence (3 LHL- repetitions, except for Experiment 7). It has been proposed that a same-frequency induction sequence may promote stream segregation by capturing one subset of test-sequence tones into an on-going stream (for the previous studies, this refers to the L tones). This capturing effect may cause the second subset of tones to be heard as segregated from the established stream. This hypothesis was initially suggested by Rogers & Bregman (1993).

This argument also relates to the study of Bregman & Rudnicky (1975). For this experiment, the captor tones may have aided performance by capturing the distractors into an on-going stream, and so caused the segregation of the targets from the task-irrelevant distractors. If this explanation is correct, one would expect that an abrupt change to a single captor tone would tend to disrupt the grouping of the captors with the distractors in exactly the same way that a change at the induction/test boundary can reduce the tendency for subsequent stream segregation of the test sequence. For the current design, any such resetting effect should result in a reduced tendency for the targets to segregate from the captors, and so should have an adverse effect on performance.

Experiment 4b demonstrated that resetting was only apparent when the final tone of a ten-tone induction sequence was altered (changes to the 4^{th} or the 7^{th} tone had no discernible resetting effect). This finding was taken as evidence that, for a same-frequency induction sequence, build-up rapidly (re-)occurred after a changed tone. To ensure that resetting was observable in the current experiment, the deviant change was applied to the last captor note before the distractors and targets. This captor is referred to as the *linking captor*, to avoid any confusion associated with the "final" captor note which occurred after the targets and distractors.

This experimental design also offers an objective measure of resetting, in contrast to the majority of the previous experiments which have demonstrated resetting using a subjective report of perception. The exception to this was Experiment 2, which measured resetting in a temporal-discrimination task (Roberts et al., 2002, 2008). Whilst the results from Experiment 2 were broadly compatible with the subjective experiments, the tendency for resetting was less apparent in the temporal-discrimination task and data collection was much more time consuming for the listeners. By using an alternative method, Experiment 10 aimed to demonstrate unambiguously resetting in a performance-based task. If achieved, this finding would strengthen the argument that resetting is a robust effect which can be observed in a variety of stimuli, and using a range of different measures.

9.5.1 Method

The method and reference stimuli were like those used in Experiment 8. Specifically, tone-like percepts were created by applying rapid amplitude increments to a single component of an on-going complex tone, and the task was to determine the presentation order of two target tones as either ascending or descending in frequency. The only difference from Experiment 8 was that one condition was removed (the remote-captors case), and two additional conditions were included to provide a measure of the extent of resetting.

Participants

Twelve listeners took part in Experiment 10. All reported having normal hearing. Nine of these participants had previously taken part in Experiments 8 and/or 9.

Stimuli

The general properties of the stimuli were identical to those used in Experiment 8. On each trial, the listener heard a complex tone comprising the first seven harmonics of a 300-Hz fundamental. The complex tone lasted for 1.2 seconds (including linear onset and offset ramps of 10 ms each). All of the harmonics started in sine phase, and each was set to 65 dB SPL. From this complex tone, a sequence of pure-tone-like percepts was created by applying a series of amplitude increments to individual harmonics. The properties of the sequence varied for each of the five conditions tested, but each individual increment shared the same properties. Specifically, a 6-dB amplitude increment was applied to an individual harmonic for 50 ms (including 10 ms linear onset and offset ramps).

Conditions

The properties of the background harmonic complex tone were the same as that used in Experiment 8 (seven components, $F0 = 300$ Hz). As previously, two target increments occurred in succession on the 5th (1500 Hz) and the 6th (1800 Hz) harmonics (see section 9.2.1 for a complete description). Three conditions were re-used from Experiment 8; these were the targets-only, distractors, and matched-captors conditions. Again, the reader is referred to section 9.2.1 for a full description of these conditions. As Experiment 8 clearly established that spectrally remote captors (added on the 2nd harmonic) did not aid performance, no such condition was included in Experiment 10.

Resetting was measured using two experimental conditions, both of which were modifications of the standard matched-captors condition. For the synchronous-global condition, amplitude increments were applied to all the other harmonics at the same time as the linking captor (the captor that occurred prior to the distractors, at 600 ms after the onset of the complex tone). The increments used were identical to those described in the stimuli section. For this condition, the linking captor was still physically present, but it was not perceptually salient (as a pure-tone-like percept). Instead, as the entire complex tone was briefly increased in level in unison, no individual component was heard as a segregated percept, and so the linking captor was effectively absent (deviant condition). It was predicted that this change from the established captor-increment sequence would disrupt the grouping of the captors with the distractors and so should impair performance at the task (as the distractors would become more likely to group with the targets). An illustration of the synchronous-global condition is provided in Figure 9.6.

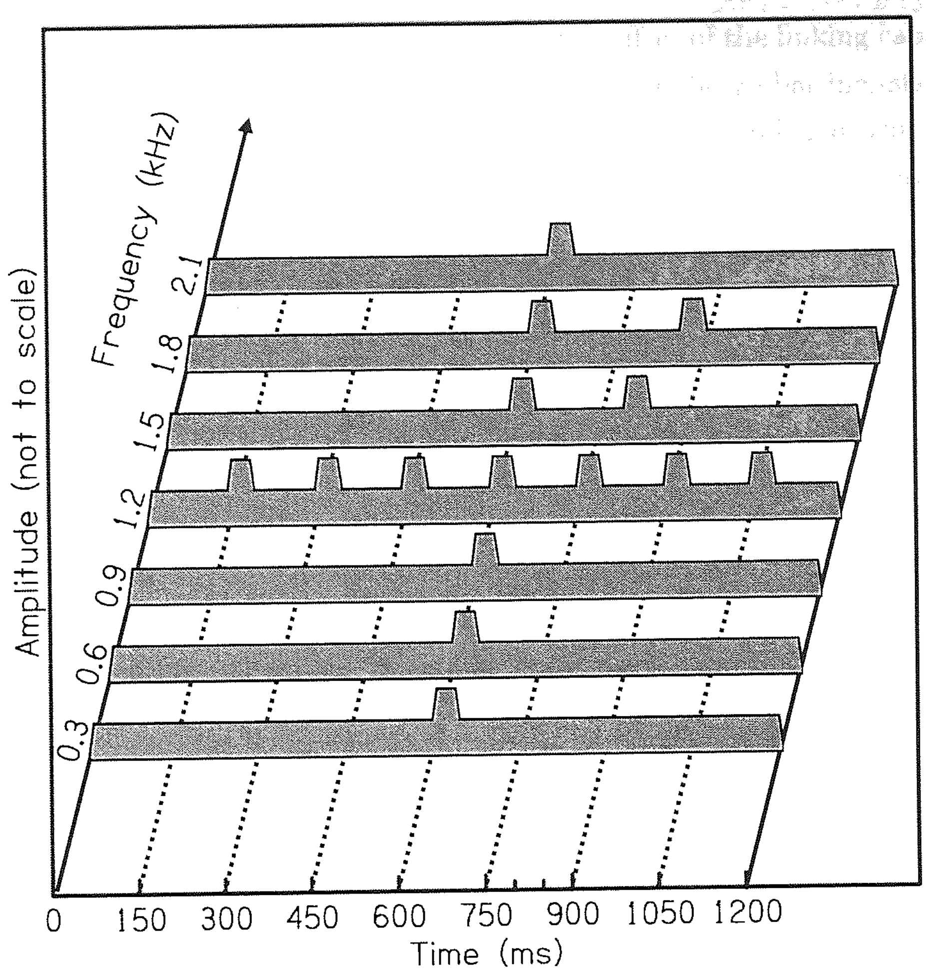

Figure 9.6: **Illustration of the synchronous-global condition tested in Experiment 10. Each tone-like percept is created by the application of an amplitude increment to a single component of the harmonic complex tone. The two target increments are on the 5th and 6th harmonics (1500 & 1800 Hz), and the two distractors on the 4th harmonic (1200 Hz) begin 750 and 900 ms after the onset of the complex tone. Five captors are also present on the 4th harmonic, four of which occur before the distractors. A global amplitude increment is applied to all components at the same time as the "linking captor" (i.e., the last captor before the distractors and targets, 600 ms after the onset of the complex tone). Note that this configuration constitutes the deviant condition, because the synchronous increments on the other harmonics form a simultaneous group with the linking captor and therefore effectively remove it from the captor sequence.**

Any reduction in performance for the synchronous-global condition could be taken as evidence that the change reset the build-up in the tendency for stream segregation induced by the captor sequence. However, a plausible alternative explanation would be that the unexpected global increment applied to all of the components merely distracted participants from the main task. This issue was addressed in the delayed-global condition. For this condition, amplitude increments were applied to all of the harmonics at the mid-point in the 100-ms interval between the offset of the linking captor and the onset of the first distractor (i.e., a global increment onset = 675 ms after the onset of the complex tone). This corresponds to an onset asynchrony of 75 ms with respect to the linking captor, and results in 25-ms intervals between the offset of the linking captor and the onset of the global increment, and between the offset of the global increment and the onset of the first distractor. As for the synchronous-global condition, this change was perceived as an increase in level for the entire complex tone, and no component-specific tone-like percepts were heard (other than for the prior linking captor itself). The linking captor and the first distractor should be clearly heard as separate from the non-overlapping global increment. Previous studies have shown that an asynchrony of about 30 ms is sufficient to cause the segregation of complex-tone components (Rasch 1978; Dannenbring & Bregman 1978, see also Roberts & Moore, 1991; Turgeon et al, 2005).

If it is the case that any broadband increment will distract listeners from the main task, performance for the delayed-global condition should be poor. However, if such a change only affects performance when it directly disrupts the captor stream, then the delayed increments should not cause any impairment, as the salience of the linking captor and distractor increments should be unaffected by this global change.

Procedure and Apparatus

The general procedure used was identical to that described for Experiment 8 (see section 9.2.1). Out of the 12 participants, none failed the training/screening session. The main experiment was divided into trial blocks, comprising one presentation of each condition in both an ascending- and a descending-frequency target configuration (i.e., 5 conditions × 2 target configurations = 10 trials per block). 20 blocks were presented in the main experiment (200 trials in total). The main experiment lasted about 40 minutes, and each listener completed this in one session. The apparatus used was identical to that described in Chapter 2.

9.5.2 Results

For each condition, results from all listeners were averaged to give an overall percentage of correct responses. For the targets-only, distractors, matched-captors, synchronous-global, and the delayed-global conditions, these means were 91.0%, 70.0%, 87.5%, 79.0% and 87.3%, respectively. These means are displayed with corresponding inter-subject standard errors in Figure 9.7. These results were assessed using a one-way, repeated-measures ANOVA. This analysis confirmed a significant difference between conditions, $[F(4, 44) = 14.042,\ p < 0.001,\ \eta_p^2 = 0.561]$. The results for the reference conditions were as predicted - performance was best for the targets-only and matched-captors conditions, and reduced for the distractors condition. .

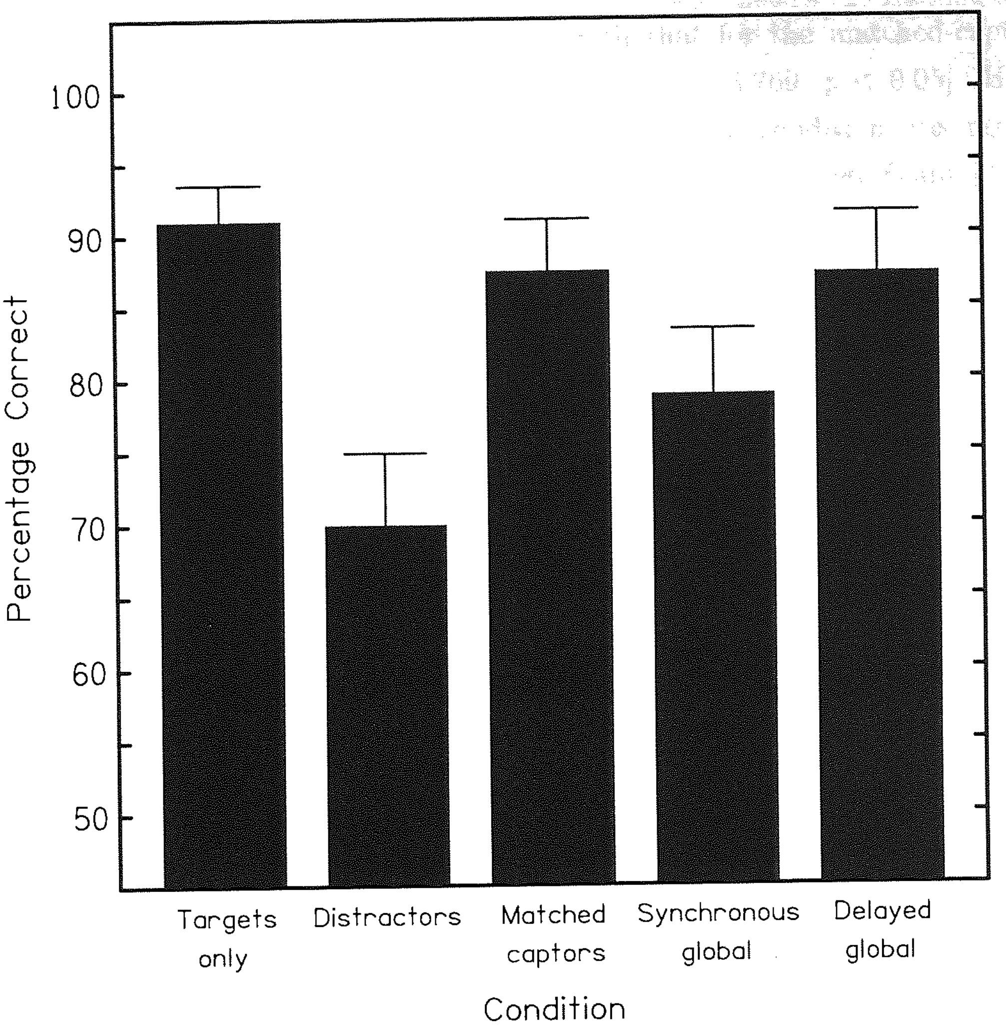

Figure 9.7: Results from Experiment 10. Performance for each condition is expressed as an overall percentage of correct responses (from 12 participants). The error bars display +1 inter-subject standard error. Note that the synchronous-global configuration constitutes the deviant condition, because the synchronous increments on the other harmonics form a simultaneous group with the linking captor and therefore effectively remove it from the captor sequence. The delayed-global configuration is not deviant, because the linking captor does not group with the (non-simultaneous) increments on the other harmonics.

The Student-Newman-Keuls test was used for multiple comparisons[2].There was a large and significant difference in performance for the targets-only and distractors conditions [difference = 21.0%, $q(k = 5,\ df_{error} = 44) = 9.282$, $p < 0.01$]. The presence of the matched captor increments largely restored performance to that for the targets-only condition [17.5%, t($q(4, 44) = 7.719$, $p < 0.01$], and the remaining difference was not significant [3.5%, $q(2, 44) = 1.562$, $p > 0.05$]. Performance for the synchronous-

2

From the Studentized range table (Glantz, 2005; Sheskin, 2004), critical q values were were unavailable for $df_{error} = 44$ Instead, these values were obtained from the nearest approximation of $df_{error} = 40$.. As for experiments 8 and 9, q values are only compared to the 0.05 and the 0.01 significance levels.

global (deviant) condition was reduced compared with that for the matched-captors condition and this difference was significant [8.5%, $q(3, 44) = 3.769$, $p < 0.05$]. However, the reduction in performance for the synchronous-global condition was not as great as that for the distractors condition, and this difference was significant [9.0%, $q(2, 44) = 3.952$, $p < 0.01$]. This is taken as evidence that the global amplitude increment (see Figure 9.6) partially disrupted the captor-distractor stream, and so increased the likelihood of grouping between the targets and the distractors. In contrast, a broad-band amplitude increment that did not coincide with any of the captor or distractor notes did not impair performance. There was no significant difference between the delayed-global condition and the matched-captors condition [0.2%, $q(2, 44) = 0.092$, $p > 0.05$]. Furthermore, the synchronous-global and delayed-global conditions differed significantly [8.3%, $q(2, 44) = 3.676$, $p < 0.05$]. Performance in all conditions was above chance level, as all conditions were significantly different from a comparison value of 50% when assessed using one-sample t-tests ($p > 0.05$ in all cases).

These findings are taken as evidence that the addition of a global increment which coincided with the linking captor note reset the stream-formation process. This resetting effect is probably due to the perceptual loss of the linking captor. In some ways, resetting due to the perceptual loss of the linking captor is analogous to the resetting of the deviant conditions demonstrated in the earlier studies[3] (e.g., replacing the final induction tone with an equivalent duration silent interval, see experiments 1, 2, and 4a). When expressed as a proportional shift away from the matched-captors condition (defined as 0.00) towards the distractors condition (defined as 1.00), this resetting effect reduced performance by 0.49.

9.5.3 Discussion

The general pattern of results is consistent with previous studies (Bregman & Rudnicky, 1975, see also Experiments 8 and 9). Performance was good for the targets-only condition, but significantly impaired when the targets were flanked by two distractors. Performance was restored when the distractors were embedded within a sequence of

[3] One would expect to find further evidence of resetting in a condition for which no increment was applied to any component at the time of the linking captor. Similar to the synchronous-global condition, this arrangement would have been analogous to the 'last tone replaced with silence' deviant conditions tested in experiments 1, 2 and 4a. Therefore, this proposed condition would be expected to further facilitate the comparison between the current results and the resetting effects of a silent interval shown in the earlier experiments. In hindsight, the current experiment may have benefitted from the inclusion such a condition.

captor increments of the same frequency as the distractors. These findings suggest that the distractors were captured into an on-going stream, which aided the segregation of the targets from the distractors. The results from these three conditions are in good accord with those of Experiment 8, which measured performance for these conditions using exactly the same stimuli.

Two additional conditions were included to examine whether a sudden change to the sequence of amplitude increments could reset the stream-formation process, and so increase the likelihood of the targets grouping with the distractors. When all of the harmonics were incremented at the same time as the linking captor note, performance was impaired in relation to the matched-captors case (a shift towards performance in the distractors condition corresponding to a resetting effect of about 0.49). This is taken as evidence that the salient change to the on-going captor-increment sequence reset the stream-formation process, so that the distractors became more likely to group with the targets. Indeed, the extent of resetting is largely consistent with that observed for the previous subjective experiments, which used sequences of discrete pure tones. For example, a range of different deviant tones was tested in Experiment 1, and it was found that, on average, these changes resulted in a 0.45 extent of resetting (as a proportional shift away from the standard induction condition towards the no-induction case - see section 3.2).

There was, however, an important consideration for the current experiment - namely that the global amplitude increment may simply have distracted listeners from the main task (as opposed to a decline in performance arising specifically from a resetting effect). This hypothesis was tested in the delayed-global condition, for which the global amplitude increment did not coincide with any of the harmonic-specific captor or distractor increments, but still remained close in time to the target increments (indeed, 75 ms closer than for the synchronous-global case in terms of the offset-to-onset time between the global increment and the first target). If the broadband increment merely distracted a listener from the main task, performance for the delayed-global condition would also have been relatively poor. Instead, it was found that performance was unaffected for the delayed-global condition (i.e., similar to that for the matched-captors condition). This evidence for resetting has two implications. First, that resetting can be observed in a sequence of tone-like percepts - as discussed, this resetting effect appears to be largely consistent with that observed for discrete pure-tone stimuli. Second, that resetting can be demonstrated effectively and efficiently using a performance measure in the current experimental design. This may relate to the similar function of the current captor increments to the same-frequency induction tones used in Experiments 1 - 7. It has been proposed that both of these sequences affect subsequent perception by capturing one specific tonal subset into an on-going stream, and so increase the tendency to hear stream segregation when tones of a different frequency are introduced (whether these are the H tones of a test sequence, see Experiments 1 - 7, or the target tones used here). The finding that resetting can be demonstrated in both types of experimental design is in good accord with the proposal that both types of captor/induction sequence have a similar effect on subsequent perception.

9.6 General Discussion

The current experiments examined whether tone-like percepts created by applying brief changes to a single component of a complex tone are organized into perceptual streams. For Experiment 8, clear evidence was found that tone percepts arising from amplitude increments (Kubovy & Daniel, 1983; Bregman et al., 1994a) are organized into streams. Experiment 9 demonstrated similar results using tone-like percepts created using ITD changes (Kubovy et al., 1974). These two sets of findings appear to indicate that any sequence of salient tone-like percepts will adhere to the same grouping principles as those for discrete pure tones (Bregman & Rudnicky, 1975). This argument was supported in Experiment 10, which demonstrated that active resetting of build-up could also be observed in a sequence of these tone-like percepts.

Overall, these findings indicate that a sequence of tone-like percepts will be organized into perceptual streams. This is taken as evidence that sequential grouping still occurs against a background of other broadband sounds which have extensive spectral overlap with the sequence of sounds in question. Future research might complement these findings by using a subjective measure of streaming, i.e., a direct report of perception. As these tone-like percepts are highly salient, one would expect streaming could be demonstrated subjectively in these sequences. Also of interest would be to explore more systematically the build-up of stream segregation for sequences of these tone-like percepts, for example by adapting the approach used by Anstis & Saida (1985).

Chapter 10

General Discussion

10.1 Introduction

The thirteen experiments presented in this thesis were designed to examine the dynamics of auditory stream segregation. A sequence of alternating low- and high-frequency tones can be perceived as either integrated into a single perceptual stream or segregated into two separate streams, depending on the sequence properties and also on listening instructions and attentional set (Miller & Heise, 1950; van Noorden, 1975). It has previously been demonstrated that, for an unchanging and on-going tone sequence, the tendency to hear stream segregation increases gradually and progressively over time (Bregman, 1978; Anstis & Saida, 1985). Similarly, a preceding same-frequency induction sequence (e.g., a repetition of an identical tone burst) can also promote segregation in a subsequent alternating-frequency sequence (Rogers & Bregman, 1993; Beauvois & Meddis, 1997; Roberts et al., 2008). Despite the apparent similarity between the segregation-promoting effect of a same-frequency induction sequence and the build-up which occurs in an alternating-frequency sequence, there has been little investigation into exactly how alike these two processes are.

An abrupt change to the properties of an on-going sequence can substantially reset the build-up of stream segregation, so that perception reverts back to integration. This resetting effect has been demonstrated following a variety of sudden changes when listeners directly reported their perception (Anstis & Saida, 1985; Rogers & Bregman, 1993, 1998), and also in a temporal-discrimination task (Roberts et al., 2008). Whilst resetting has been clearly demonstrated in a variety of different sequence arrangements and experimental procedures, there remain several potential explanations for the effect. The experiments presented in this thesis had three general aims - to explore further the resetting of the build-up of stream segregation, to develop a greater understanding of the segregation-promoting effect of a same-frequency induction sequence, and to explore the dynamics of stream segregation against a background of continuous sounds.

10.2 Summary and Conclusions

10.2.1 The Resetting Effect of a Single Deviant Tone

The build-up of stream segregation in an alternating-frequency sequence (e.g., repeating LHL- or LH cycles) is thought to reflect a gradual accumulation of evidence in favour of two separate sources being present (Bregman, 1978, 1990) and it has been previously established that an abrupt change to an on-going sequence can have a substantial resetting effect on build-up (e.g., a change in frequency or ear of presentation (Anstis & Saida, 1985); a change in level or perceived location (Rogers & Bregman, 1998)). Rogers & Bregman (1998) offered two explanations for resetting, both of which relate to Bregman's (1978, 1990) evidence-accumulation hypothesis. If the evidence-accumulation process is specific to the sequence being heard, any increased tendency to hear stream segregation should not transfer to a subsequent sequence comprising tones with different characteristics (a failure to transfer). Although this is a plausible explanation, Rogers & Bregman (1998) found that a sudden increase in level (+12 dB) at the inducer/test boundary had a substantial resetting effect, but that an equivalent decrease in level (-12 dB) did not. As the directional properties of a transition itself appear to influence resetting, the magnitude of a difference in sequence characteristics alone cannot fully explain the effect. Based on this, Rogers & Bregman (1998) offered an alternative explanation - that resetting occurs when a new acoustic event is perceived. Certain abrupt changes may indicate a new event, and so cause a re-analysis of the entire auditory scene. This may result in a resetting of the evidence-accumulation process. Rogers & Bregman (1998) attempted to evaluate the importance of sudden changes by comparing reported segregation following either an abrupt or a gradual change in sequence properties, and found that only sudden changes resulted in substantial resetting. However, the properties of the induction sequence were substantially altered in the gradual change conditions, resulting in mean tonal properties that were much more similar to those of the test sequence. Given this confound, evidence from the gradual change conditions is unable to differentiate between the two explanations of resetting (failure to transfer or active resetting).

Several of the current experiments distinguished between these two explanations of resetting by taking a novel approach. This was to present an induction sequence (usually 10 L- repetitions, where L = a 100 ms, 1 kHz pure tone, and - = a 100 ms silence) in which only a single (the final) inducer was altered in comparison with the standard L-tones. Listeners were required to report their perception of a brief test sequence as either integrated or segregated. The test sequence comprised three LHL- triplets, for which the L-tones and silences shared identical properties to those of the induction sequence, and the H-tones were varied across a range of frequencies (see Chapter 2). In this arrangement, any transferral of build-up from the prior induction tones to the test sequence should be largely unaffected by the presence of only a single deviant tone.

For Experiment 1, reported segregation in the test sequence was often substantially reduced for the deviant-tone conditions, and so it was concluded that *an abrupt change to a sequence can have an active resetting effect.* A broadly equivalent resetting effect was observed for a variety of deviant tones (tone altered in either frequency, level, duration, or replaced with silence), although this reduction was a non-significant trend in the case of the level change. For Experiment 2, two deviant-tone induction conditions were replicated from Experiment 1 (increase in level or replacement with silence), and test-sequence streaming was measured in an adaptive temporal-discrimination task, similar to that used by Roberts et al. (2008). As for Experiment 1, replacing the last tone with silence had a substantial resetting effect, although resetting following the level change was small (and non-significant). Despite being less apparent in the temporal-discrimination task, both experiments 1 and 2 offered broadly consistent evidence of the resetting effect of a single deviant tone. Consistent evidence of resetting was also observed in the temporal-discrimination task of Experiment 10, as discussed in section 10.2.4.

The resetting effect of a single deviant tone is consistent with Rogers & Bregman's (1998) suggestion that resetting may occur when a new sound source is perceived. To further test this hypothesis, Experiment 3a examined whether the extent of resetting varied with the size of a frequency change applied to the final induction tone (range = -0.5 to -12 ST from the L-tone frequency of 1 kHz). Using the same subjective procedure as Experiment 1, Experiment 3a found a moderate and consistent increase in resetting as the size of the frequency change was increased. The magnitude of the extent of resetting was linearly related to the (log) size of the frequency change applied to the deviant (see Figure 4.4). Experiment 3b examined the resetting effect of smaller frequency changes (range = -2.5 to -20 Hz wrt. 1 kHz). Changes of -10 and -20 Hz both had a slight, but significant resetting effect. Given that even small changes had a partial resetting effect once the change in frequency was suprathreshold (see Moore, 2003) and that the trend towards increased resetting with increased size of deviant frequency change was largely non-significant, it was concluded that resetting is linked to the perceptual salience of a change, and that any noticeable change can have an active resetting effect. This argument contrasts with Rogers & Bregman's (1998) proposal that resetting only occurs when a new acoustic event is perceived. An exception to the current argument was observed in experiments 1 & 2, where a +12 dB deviant tone failed to produce significant resetting. The reason for this is unclear, especially as this level change was clearly salient and previous studies have demonstrated resetting when a level change was applied to an on-going sequence (Rogers & Bregman, 1998; Roberts et al., 2008). However, it is worth noting that for experiments 1 & 2, the +12 dB deviant tone did result in a clear trend towards resetting.

An alternative explanation for the reduction in segregation associated with a single deviant tone is simply that one fewer identical tone contributes to the cumulative build-up occurring during an induction sequence. This hypothesis was tested in Experiment 4a, where the serial position of a deviant tone was manipulated (created by replacing the tone with an equivalent-duration silence). Resetting of build-up was observed only when the final induction tone was replaced; the same change applied either to the 4th or the 7th tone did not reduce reported segregation in the test sequence. As each of these induction sequences comprised nine tones, but only a change to the final tone had any effect on test-sequence streaming judgments, it was concluded that a deviant at the boundary between the induction and test sequences must actively reset build-up. This finding was explained in Experiment 4b, where it was demonstrated that build-up of the tendency for stream segregation is very rapid during a same-frequency induction sequence. Indeed, build-up appeared to be complete after only three induction L-tones. However, listeners were asked only to report their perception of the *final* LHL- triplet of the test sequence (the test sequence comprised 3 LHL- repetitions). It is plausible that the build-up of segregation may have continued through both the induction sequence *and the initial portion of the test sequence.* This additional period of build-up may have contributed to the observed results. Nevertheless, any build-up during the test sequence would have been present in every condition tested, and reported segregation following the three-inducers condition remained substantially higher than that for the no-induction sequence control condition. The rapid build-up observed following a same-frequency induction sequence is discussed further in section 10.2.3.

The Decay of Stream Segregation in a Silent Interval

By manipulating the silent interval between tone sequences, previous studies have shown that the build-up of stream segregation gradually decays during a subsequent silence (Bregman, 1978; Beauvois & Meddis, 1997). These studies found that the decay of segregation was typically most rapid over the first few seconds of a silence, but was still apparent after four, or even eight, seconds of silence (at least for the musically trained listeners in the Beauvois & Meddis study). The current experiments found that reported segregation was greatly reduced when the final induction tone was replaced with an equivalent-duration silent interval (i.e., only a 300-ms silence between the 9th

induction tone and the 1st test-sequence tone, see experiments 1, 2, and 4a). Whilst it is difficult to compare this finding directly with earlier studies, it is clear that the loss of build-up for this brief silence was much greater than the decay of the tendency for stream segregation previously associated with a silent interval of similar duration (Bregman, 1978; Beauvois & Meddis, 1997). Therefore, the currently observed effect of a brief silent interval has been interpreted primarily as rapid, active resetting, as opposed to a slower, passive decay. Note that the resetting effect of a brief silence was also demonstrated in the temporal-discrimination task of Experiment 2. The current findings are somewhat consistent with a study by Cusack et al. (2004). These authors measured loss of build-up after a range of silent intervals (1 s - 10 s), which were inserted into an on-going LHL- sequence. All of the intervals tested resulted in a largely similar and very substantial reduction in reported stream segregation. This suggests that their results for the shortest interval that they tested (1 s) were influenced by active resetting as well as by decay.

One might speculate that differences between studies in the time course of loss of build-up may reflect the fact that listening context affects the relative contributions of (rapid) active resetting and (more gradual) decay of build-up. A review of the literature suggests that the current Experiment 2 provides the only objective measure of loss of build-up for which only the effect of a single, very brief, silence was measured. It may prove illuminating to adapt Experiment 2 to study the time-course of decay through longer silences using an objective, temporal-discrimination measure.

10.2.2 The Role of Attention in Stream Segregation

Experiments 3a and 3b demonstrated that relatively small, but noticeable, frequency changes applied to the final inducer can have a partial resetting effect. This may suggest that resetting is linked to attentional factors, and this issue relates to the wider debate concerning the role of attention in stream segregation. Although this debate remains on-going, there is a body of 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. 2003a), and also that stream segregation builds up even in an unattended sequence (e.g., Sussman et al., 2007). Consistent evidence has also been reported in behavioural studies that have measured the effect of an unattended sound sequence on performance in a non-auditory task (the irrelevant sound effect - see Jones, 1995; Jones et al., 1999; Macken et al., 2003).

Related to these findings, there is some evidence from behavioural studies which suggests that a switch in attention during an on-going sequence may have a resetting effect. For example, Cusack et al. (2004) prompted listeners to switch temporarily their attention away from an on-going tone sequence. They found that the build-up of stream segregation was substantially reset following such a brief shift in attention. Similarly, Carlyon et al. (2001) presented a relatively long tone sequence, but diverted attention to another, separate auditory task for the initial half of the sequence. They found that when listeners began to report their perception of the tone sequence, there was no evidence of any build-up having occurred during the unattended portion of the sequence (see also Carlyon et al., 2003). One interpretation of these results is that the switch in attentional focus may have reset any prior build-up.

This evidence contributed to Cusack et al.'s (2004) hierarchical decomposition model. This model states that there may be some form of automatic stream segregation in a multi-source listening environment, but that only an attended source may be elaborated into further, more specific perceptual streams (see also Brochard et al., 1999). Based on this, the model proposes that build-up may occur only in an attended tone sequence (although there is some more recent evidence from EEG studies which suggests otherwise, see Sussman et al., 2007). This issue notwithstanding, Cusack et al. (2004) proposed that even a brief diversion of attention may reset any detailed elaboration (stream segregation) of a previously attended sequence. Therefore, when attention is switched back to the sequence, it is heard as a single, non-elaborated (i.e., integrated) percept. Based on this proposal, one might speculate that the currently observed resetting effect of a single deviant tone may have been due to the deviant tone temporarily diverting attention away from the on-going sequence.

If resetting of build-up occurs because of a switch in attention, one would predict that any individual tone which does *not* attract attention should not have any resetting effect. Experiment 5 attempted to test this hypothesis by presenting three different induction sequences, for each of which the final inducer was fixed at -2 ST from the standard L-tone frequency of 1 kHz. Consistent with the earlier studies, a resetting effect was observed when this tone was heard as a deviation from the preceding nine induction tones (which were all set to 1 kHz - the single-deviant condition). When all of the induction tones were lowered by 2 ST, a near-complete loss of build-up was observed[1]. The critical condition examined the effect of sequence predictability by randomizing the frequencies of the first nine inducers (tones varied in frequency across a four-semitone

[1] This indicates that whilst the properties of the final inducer can strongly influence subsequent stream segregation, the preceding induction tones also affect the segregation-promoting effect of an induction sequence.

range centred on 1 kHz). As for the single-deviant condition, the final inducer was fixed at -2 ST, but it should not have been heard as a deviation from an established pattern. Given this, it was predicted that the final inducer would be unlikely to capture a listener's attention, and therefore should not have any resetting effect. However, for this condition, reported segregation was broadly similar to that for the single-deviant condition. One possible explanation for this is that the spread of frequency excitation in the unpredictable induction sequence may have confounded any effect of sequence predictability on resetting. To compensate for this, a future study could make the sequence unpredictable across a different dimension, such as time (i.e., tone onset times and tone durations). This would remove any spread of frequency excitation, but still allow for a test of the relationship between sequence predictability and resetting.

10.2.3 The Segregation-Promoting Effect of a Same-Frequency Induction Sequence

For the current studies, test-sequence segregation was typically promoted by a prior same-frequency induction sequence (e.g., usually 10 L- repetitions). Several previous studies have used similar sequence arrangements, and have also found that a same-frequency induction sequence is highly effective at promoting subsequent segregation (Rogers & Bregman, 1993; Beauvois & Meddis, 1997; Roberts et al., 2008). Despite this, it has remained unclear exactly how alike the segregation-promoting effect of a same-frequency induction sequence is to the build-up of stream segregation which occurs during an alternating-frequency tone sequence.

During a same-frequency induction sequence, only a repetition of an identical tone is heard. Given this, Rogers & Bregman (1993) proposed that these tones will always be heard as single integrated stream. If the induction tones then continue into the test sequence, the corresponding stream is maintained, but the new subset of tones (i.e., the H tones in the current experiments) may be excluded from this on-going stream so that stream segregation is heard. In other words, at the onset of a test sequence, an L-tone-only stream has already been established, and this may increase the likelihood of the H-tones being heard as a segregated stream. This explanation can also be used to interpret the findings of Bregman & Rudnicky (1975), who demonstrated that an initial same-frequency tone sequence can aid the subsequent segregation of distractor and target tones (see chapters 5 and 9).

Several of the current experiments produced results consistent with this hypothesis. In Experiment 4b, it was found that induction sequences comprising either 3, 6, or 10 L-tones all had a broadly equivalent segregation-promoting effect. Therefore, build-up during a same-frequency sequence must be very rapid, and seems to be complete after only three L-tone repetitions. This suggests that an induction-tone stream is firmly established after three tone repetitions, and that the strength of this percept does not increase appreciably with further induction tones. This explanation seems plausible, as there are no competing perceptual organizations during a same-frequency induction sequence.

In Experiment 7, the segregation-promoting effect of a same-frequency induction sequence was directly compared with the build-up occurring during an alternating-frequency sequence. Unlike the earlier experiments, listeners were required to indicate continuously their perception of a long test sequence (duration = 20 s). As predicted, for most conditions, the tendency to report segregation built-up over the course of this test sequence. It was found that a 2-s same-frequency induction sequence promoted substantially more stream segregation than did the equivalent-duration alternating-frequency induction case. Indeed, for test-sequence frequency separations of 6 and 9 ST, reported segregation was observed to be greatest immediately after the same-frequency induction sequence, and then to decay over the initial portion of the test sequence (this decay was significant for the 9-ST case). In other words, a same-frequency induction sequence is able to promote stream segregation to a greater extent than that usually heard after the same time in an equivalent alternating-frequency sequence. The capturing effect of a same-frequency induction sequence may relate to attentional factors. During the currently tested same-frequency induction sequence, a listener will attend the L-tones, and this attentional bias may continue into the test sequence, resulting in the exclusion of the H-tones from the attended L-tone subset. This hypothesis relates to the hierarchical decomposition model (Cusack et al., 2004), which states that only attended auditory stimuli are subject to detailed perceptual representation.

Actual and Perceived Continuity in an Induction Sequence

In Chapter 7, the effects of perceived and actual continuity during an induction sequence were measured on streaming judgments in a test sequence. In experiments 6a and 6b, it was found that an induction sequence which comprised a single extended L-tone (duration = 1100 ms), *which ended 100 ms before the onset of the test sequence*, was largely ineffective at promoting test-sequence stream segregation. Consistent findings have been reported previously by Rogers & Bregman (1993) and Roberts et al. (2008). Because of this, Rogers & Bregman (1993) proposed that sequence similarity (i.e., in terms of a similar rate of tone onsets between the induction and test sequence), rather than tone density in the induction sequence *per se*, strongly influences the segregation-promoting effect of an induction sequence.

Experiment 6b measured the effect of a fully continuous inducer (i.e., the first test-sequence tone was extended to create a single, *uninterrupted* 1300-ms tone). This fully continuous modification of the first test-sequence tone served as the (arbitrary) induction sequence. This induction sequence arrangement was highly effective at promoting subsequent judgments of stream segregation. It would seem likely that the increase in reported segregation following a fully continuous tone was due to the disruption of the first LHL- triplet, caused by the extension of the first L-tone. Such a change would weaken the galloping rhythm associated with an integrated percept, which listeners may have been using as a cue to report integration.

This issue aside, experiments 6a and 6b were primarily designed to investigate the effect of perceived continuity in an induction sequence on streaming judgments in the test sequence. To my knowledge, no previous study has examined the effect of perceived continuity on *subsequent* stream segregation. In Experiment 6a, one induction condition presented six discrete, standard induction tones which were heard to continue through interleaved noise-bursts (so that a single extended tone was perceived). In this circumstance, the induction sequence did not promote test-sequence streaming, instead, reported segregation was similar to that for the actual extended-tone condition. In Experiment 6b, it was shown that an extended tone which was perceived to continue through a subsequent louder noise-burst, (so that a fully continuous tone was heard) was effective at promoting subsequent segregation. Together, these experiments demonstrate that actual and perceived continuity within an induction sequence have the same effect as one another, either promoting subsequent stream segregation or not, depending on context.

10.2.4 Tone Percepts Evoked by the Segregation of Components from a Complex Tone are Organized into Perceptual Streams

In experiments 1 - 7, stream segregation was studied for sequences of isolated, discrete pure tones. Indeed, the majority of previous research into sequential grouping has also used this kind of stimulus arrangement. However, natural listening environments commonly contain a mixture of both sequential and simultaneous grouping cues, and so the auditory system must be capable of using a combination of both types of cue to build an accurate perceptual representation of the auditory environment (Bregman, 1990). One example of this was studied in experiments 8 - 10, which measured the sequential organization of tone-like percepts evoked by changes in simultaneous grouping. An abrupt change to a single component of a complex tone may cause it to be perceived separately from the others, so that the changed component is heard as a pure-tone-like percept, distinct from the on-going complex tone. Previous studies have demonstrated the perceptual salience of tone-like percepts evoked from sudden changes in amplitude (Kubovy & Daniel, 1983; Bregman et al., 1994a), or interaural time difference (Kubovy et al., 1974) to an individual component of a complex tone.

The stream segregation of tone-like percepts was studied by adapting a temporal-order measure of streaming originally used by Bregman & Rudnicky (1975). Listeners were required to report the order of two target tone-like percepts as either ascending or descending in frequency. In the experimental conditions, performance was taken as an index of the grouping of task-irrelevant distractor tone percepts with additional captor tone percepts. Poor performance was taken as evidence that the targets and distractors grouped together, making the task more difficult, whereas good performance was taken as an indication that the captors and distractors were grouped into an on-going perceptual stream, from which the targets were excluded. Experiment 8 clearly demonstrated that tone-like percepts evoked from amplitude increments (Bregman et al., 1994a) are organized into distinct perceptual streams. Experiment 9 found consistent evidence of the stream segregation of tone-like percepts evoked from ITD changes (Kubovy et al., 1974). It was concluded that any form of salient tone-like percept is governed by the same sequential grouping principles as those for discrete pure tones (Bregman & Rudnicky, 1975). This is evidence that sequential grouping can occur against a background of other broadband sounds, even if these sounds occupy a similar frequency region to the sounds which are organized into perceptual streams.

Experiment 10 demonstrated that a single change to a sequence of tone-like percepts evoked from amplitude increments had an active resetting effect. The deviant tested involved replacing a component-specific amplitude increment with a synchronized global increment applied to each component of the on-going complex tone (so that no pure-tone-like percept was heard). This change was applied to the captor immediately prior to distractor and target tone-percepts (somewhat analogous to an alteration to the final inducer before a test sequence, as tested in the previous experiments). Resetting was evident as the distractors became more likely to group with the targets (indexed by impaired performance). This experiment clearly demonstrated active resetting using a temporal-discrimination measure, which strengthens the argument that active resetting is a robust effect that can be observed using a variety of different stimuli and response measures.

10.3 Future Questions for Research

Active Resetting in an Alternating-Frequency Sequence

For the current experiments, active resetting was studied primarily in a context where a deviant tone was inserted into a same-frequency induction sequence. The one exception to this was Experiment 7, which measured deviant-tone resetting in both a same- and an alternating-frequency sequence. For this experiment, a tone made deviant by altering in duration produced, an apparent (albeit non-significant) trend towards resetting when inserted into a same-frequency induction sequence; the same change had a significant resetting effect in Experiment 1. Generally, there was no apparent resetting effect when the same change was applied to the final L-tone of an alternating-frequency (LHL-) induction sequence, and it was speculated that the lack of resetting may have been due to the perceptual representation of the alternating-frequency sequence (i.e., integrated or segregated) prior to the deviant tone. More specifically, the deviation in tone duration tested may not have been especially salient when inserted into an alternating-frequency sequence which was perceived as integrated. Clearly, there is need for further study to establish whether deviant-tone resetting can be observed in an on-going alternating-frequency sequence, and whether this is affected by the perceptual representation of the tone sequence.

Stream Segregation as a Bi-Stable Percept

There has been some evidence to suggest that stream segregation may be a bi-stable percept (Pressnitzer & Hupé, 2006; Denham & Winkler, 2006). By calculating the average duration of each successive percept reported in a tracking procedure, these authors found that segregation was never exclusively heard even in a long, four-minute sequence. In other words, the build-up of stream segregation does not appear to continue up to a point were *only* segregation is heard. Pressnitzer & Hupé (2006) showed that there was no long-term trend in the average durations of each percept (integrated or segregated), except that the initial integrated percept lasted substantially longer than all subsequent percepts. Based on this finding, Pressnitzer & Hupé (2006) argued that the build-up effect reported by Anstis & Saida (1985) was a byproduct of the method of data analysis, and that this effect was primarily driven by the increased duration of the initial integrated percept. The same conclusion was reached by Denham & Winkler (2006), who showed that the average duration of the initial percept declined as the frequency separation between L- and H-tones was increased. This finding can be related to Anstis & Saida's (1985) study, as a shorter initial integrated percept should result in a more rapid build-up of segregation. Despite this, there has been no conclusive evidence to show how the bi-stability hypothesis can explain the long-term build-up over the course of a 30 s or 60 s tone sequence, or the observation that the presentation rate required for segregation decreased over the course of a long tone sequence (Anstis & Saida, 1985).

There is considerable scope to examine further the suggestion that streaming is bi-stable. A rigorous test of this hypothesis could employ some form of objective measure of stream segregation. If streaming is bi-stable after an initial prolonged integrated percept, then performance in a temporal-discrimination task should remain constant as the duration of a prior tone sequence is increased.

The Neural Mechanisms for Resetting

Recent research has attempted to identify the neural basis of auditory streaming. There is now a body of evidence to suggest a relationship between neural responses in the primary auditory cortex (A1) and stream segregation (Fishman et al., 2001, 2004; Kanwal et al., 2003; Bee & Klump, 2004, 2005; see Micheyl et al., 2007, for a review). Using awake macaque monkeys, Fishman et al. (2001) recorded single-unit responses in the A1 to alternating-frequency pure-tone sequences (a repeating AB arrangement). The frequency of the A-tones was fixed at the best-response frequency (BF) of the recording site, whilst the frequency of the B-tones was shifted from this value (from 1.65 to 7 ST, either above or below the frequency of the A-tones). The presentation rate of the sequence was also manipulated (from 5 to 25 tone onsets per second). The unit became less likely to respond to the B-tones as either the frequency separation or presentation rate of the sequence was increased. This finding correlates with the known effect of these parameters on stream segregation, as observed using psychophysical methods (e.g., van Noorden, 1975).

Single-unit recording data from A1 also correlates with the known build-up of segregation over time (Bregman, 1978; Anstis & Saida, 1985). Micheyl et al. (2005) demonstrated that the suppression of responses to (non-BF) B-tones increased over the course of a 10-s tone sequence (again, recorded from awake macaque monkeys). It could prove interesting to examine how a same-frequency induction sequence of BF tones would affect neural responses to a subsequent sequence comprising both BF and non-BF tones. As a same-frequency induction sequence is known to promote stream segregation, one might predict that such an induction sequence would result in highly suppressed responses to subsequent non-BF tones. Furthermore, an adaptation of such an experiment may also offer some insight into the neural basis of active resetting.

The previously described studies all employed single-unit recording from animals to examine the neural basis of stream segregation. A second approach has been to collect EEG or MEG data from humans. Data obtained from such methods reflect more global responses in the auditory cortex, as recording will encompass multiple BF sites. Gutschalk et al. (2005) measured auditory evoked responses to sequences of low- and high-frequency tones (an ABA- arrangement, where B = 1 kHz, A = +2 to +12 ST, duration = 100 ms, and within-triplet tone offset-to-onset time = 50 ms or 200 ms). MEG measurements were employed to record the magnitude of the neural response that coincided temporally with the B-tones (specifically the P_1m and the N_1m components). The magnitude of these components varied with the physical sequence properties. More specifically, larger responses were observed at greater A-B frequency separations and at faster sequence presentation rates. Therefore, the magnitude of neural responses at the time of the B-tones closely correlated with the known effects of frequency separation and tone-repetition time on stream segregation (van Noorden, 1975). Larger neural responses occurred when a listener would be likely to hear segregation, and Gutschalk et al. (2005) also collected psychophysical data to support this assertion. These results are thought to reflect to a form of feature-specific forward suppression (Gutschalk et al., 2005, 2007; Micheyl et al., 2007). This argument proposes that the neural response in the auditory cortex towards one sound (e.g., A-tones) can suppress the responses to subsequent sounds (e.g., B-tones), but that the degree of suppression decreases as the frequency separation or the presentation rate between tones is increased.

Furthermore, neural responses coinciding with the B-tones have been demonstrated to correlate with changes in the *perceptual* representation of a sequence. In a second experiment, Gutschalk et al. (2005) presented a 25 s repeating ABA- tone sequence (where B = 1 kHz, A = +4 or +6 ST, tone duration = 50 ms). For this sequence arrangement, perception typically fluctuated between integration and segregation. As perceptual responses were recorded concurrently with MEG data, it was found that neural responses coinciding with the B-tones were consistently of a larger magnitude when the subject reported hearing segregation - as opposed to smaller responses when integration was heard. In a related study, Snyder et al. (2006) presented somewhat similar ABA- sequences (A = 500 Hz, B = +4, +7 or +12 ST, tone duration = 20 ms, tone offset-to-onset time = 80 ms). These authors measured EEG responses at different time points during an on-going (10.8 s) sequence, and found that the magnitude of neural responses coinciding with the B-tone increased over the course of the trial. As a larger neural response is thought to correlate with stream segregation, Snyder et al.'s (2006) observation is consistent with the evidence that there is a build-up of streaming over time (Bregman, 1978).

In summary, there appears to be a strong correlation between the proposed forward suppression of neural activity coinciding with the B-tone and the likelihood of perceiving integration (Gutschalk et al., 2005; Snyder et al., 2006). From this, one could make two predictions based on the evidence provided in this thesis. Firstly, as a repeating same-frequency induction sequence strongly promotes subsequent test-sequence segregation, one might expect that the L-tone-only induction sequence would *reduce* any forward suppression of the H-tones in the test sequence. This prediction is congruent with Snyder et al.'s (2006) observation that forward suppression decreased throughout the duration of a 10.8 s ABA- sequence. Secondly, the current experiments have demonstrated that a single abrupt change at the induction/test sequence boundary can trigger resetting (i.e., promote subsequent integration). Based on this, one would predict that the presence of an abrupt change would *promote* the suppression of subsequent neural responses to the H-tones (in comparison with a standard induction case). As discussed in section 10.2.2., it is plausible that the resetting effect of a deviant tone may be linked to a temporary switch in attention away from the on-going sequence. It may be the case that any brief withdrawal of attention may result in a greater forward suppression of responses to the subsequent test-sequence H-tones. Consistent with this speculation, Snyder et al. (2006) found strong forward suppression of responses coinciding with the B-tones when listeners were instructed to ignore the tone sequence (in comparison to less forward suppression when the sequence was attended).

Build-up and resetting in normal listening enviroments

Bregman (1978, 1990) has offered a functional explanation of the build-up of segregation. He proposed that at the onset of a sound sequence, the default assumption is that all acoustic elements have arisen from the same acoustic event. Only over time does the stream-formation mechanism become increasingly likely to segregate sounds on the basis of their continued difference from one and other. This evidence-accumulation process is relatively slow, and as a result changes in grouping are reasonably conservative. Bregman argued that the conservative nature of the system is beneficial because it prevents excessive fluctuations in the perceptual organization of a sound sequence.

Indeed, it is clearly desirable for our auditory system to form stable perceptual representations of the acoustic events occurring around us. Based on this, it seems somewhat counterintuitive that relatively small changes to an on-going sequence can have a substantial effect on subsequent stream segregation (i.e., resetting effects). It is quite uncommon to hear a repeating and unchanging sound in normal listening environments, and we are evidently able to form accurately perceptual streams that correspond to fluctuating sound signals (following a melody or speech for example). Therefore, it cannot be the case that any salient change between two within-stream sounds will substantially reset the stream-formation process.

Instead it seems more likely that resetting may occur only when a deviant sound is perceived as a clear violation of an on-going stream. In the current experiments the standard induction sequence comprised ten repetitions of an identical L-tone. Because of this, any perceivable change in tonal properties would likely have been heard as a salient deviation from the pre-established properties of the stream. Perhaps in more naturalistic acoustic streams comprising varying sound signals, the auditory system may require a larger deviation before resetting will occur. There is considerable scope for studying further the build-up and resetting of stream segregation in more naturalistic sound sequences.

10.4 Concluding Remarks

In conclusion, the principal findings of this thesis are threefold. A salient, single change to an on-going tone sequence can actively reset the build-up of segregation, and this effect can often be substantial. Second, the segregation-promoting effect of a same-frequency induction sequence is not analogous with the build-up of segregation which occurs during an alternating-frequency sequence. Instead, a same-frequency induction sequence appears to promote subsequent stream segregation by capturing one sub-set of tones into an on-going perceptual stream. Finally, a sequence of tone-like percepts evoked from changes to individual components of an on-going complex tone can be organized into perceptual streams.

References

Ahad, P. A. (2000). Media Control Functions (MCF) (version 2.94) [computer software]. Montreal: Canada. http://www.digivox.ca

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

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

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

Boehnke, S. E., & Phillips, D. P. (2005). The relation between auditory temporal interval processing and sequential stream segregation examined with stimulus laterality differences. *Perception and Psychophysics, 67*, 1088-1101.

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

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

Bregman, A. S., Ahad, P., & Kim, J. (1994b). Resetting the pitch analysis system. 2: Role of sudden onsets and offsets in the perception of individual components in a cluster of overlapping tones. *The Journal of the Acoustical Society of America, 96*, 2694-2703.

Bregman, A. S., Ahad, P., Kim, J., & Melnerich, L. (1994a). Resetting the pitch analysis system. 1: Effects of rise time of tones in noise backgrounds or of harmonics in a complex tone. *Perception and Psychophysics, 56*, 155-162.

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., Colantonio, C., & Ahad, P. A. (1999). Is a common grouping mechanism involved in the phenomena of illusory continuity and stream segregation? *Perception and Psychophysics, 61*, 195-205.

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

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

Carlyon, R. P. (2004). How the brain separates sounds. *Trends in Cognitive Sciences, 8*, 465-471.

Carlyon, R. P., Cusack, R., Foxton, J. M., & Robertson, I. H. (2001). Effects of attention and unilateral neglect on auditory stream segregation. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 115-127.

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

Ciocca, V., & Bregman, A. S. (1987). Perceived continuity of gliding and steady-state tones through interrupting noise. *Perception and Psychophysics, 42*, 476-484.

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.

Culling, J. F. (2000). Auditory Motion Segregation: A Limited Analogy With Vision. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 1760-1769.

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

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

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

Dannenbring, G. L. (1976). Perceived auditory continuity with alternately rising and falling frequency transitions. *Canadian Journal of Psychology*, *30*, 99-114.

Dannenbring, G. L., & Bregman, A. S. (1978). Streaming vs. fusion of sinusoidal components of complex waves. *Perception and Psychophysics*, *24*, 369-376.

Darwin, C. J., & Carlyon, R. P. (1995). Auditory grouping. In B. C. J. Moore (Ed.), *Handbook of perception and cognition* (pp. 387-424). San Diego, California: Academic press.
Denham, S. L., & Winkler, I. (2006). The role of predictive models in the formation of auditory streams, *Journal of Physiology - Paris*, *100*, 154-170.

Dowling, W. J. (1973). The perception of interleaved melodies. *Cognitive Psychology*, *5*, 322-337.

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. *The 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.

French-St. George, M., & Bregman, A. S. (1989). Stream segregation as a function of predictability of frequency and timing. *Perception and Psychophysics*, *46*, 384-386.

Gardner, R. B. G., & Wilson, J. P. (1979). Evidence for direction-specific channels in the processing of frequency modulation *The Journal of the Acoustical Society of America*, *63*, 704-709.

Glantz, S.A. (2005). *Primer of biostatistics* (6^{th} ed.). McGraw-Hill.

Glasberg, B. R., & Moore, B. C. J. (1990). Derivation of auditory filter shapes from notched-noise data, *Hearing Research*, *47*, 103-138.

Green, G. G. R. , and Kay, R. H. (1973). The adequate stimuli for channels in the human auditory pathways concerned with the modulation present in frequency-modulated tones. *Journal of Physiology - London*, *234*, 50-52.

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

Hartmann, W. M., & Johnson, D. (1991). Stream segregation and peripheral channelling. *Music Perception*, *9*, 155-183.

Helmholtz, H. (1925). *Treatise on physiological optics.* New York: Optical Society of America.

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

Howell, D. C. (1992) *Statistical Methods for Psychology* (5th ed.). Wadsworth Publishing.

Jeffress, L. A. (1948). A place theory of sound localization. *Journal of Comparative and Physiological Psychology, 41*, 35-39.

Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review, 83*, 323-355.

Jones, D. M. (1995). The fate of the unattended stimulus: Irrelevant speech and cognition. *Applied Cognitive Psychology, 9*, 23-38.

Jones, D. M., Alford, D., Bridges, A., Tremblay, S., & Macken, W. J. (1999). Organizational factors in selective attention: The interplay of acoustic distinctiveness and auditory streaming in the irrelevant sound effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 464-473.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review, 96*, 459-491.

Jones, M. R., Kidd, G., & Wetzel, R. (1981). Evidence for rhythmic attention. *Journal of Experimental Psychology: Human Perception and Performance, 7*, 1059-1073.

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

Kay, R. H., and Matthews, D. R. (1972). On the existence in the human auditory pathway of channels selectively tuned to the modulation present in frequency-modulated tones. *Journal of Physiology - London*, 225, 657-677.

Koffka, K. (1935). *Principles of Gestalt psychology.* London: Lund Humphries.

Kubovy, M., (1981). Concurrent pitch segregation and the theory of indispensable attributes. In M. Kubovy, & J. R. Pomerantz (Eds.), *Perceptual Organization* (pp. 55-98). Hillsdale, New Jersey: Lawrence Erlbaum.

Kubovy, M., Cutting, J., & McGuire, R. M. (1974). Hearing with the third ear: Dichotic perception of a melody without monaural familiarity cues. *Science, 186*, 272-274.

Kubovy, M., & Daniel, J. E. (1983). Pitch segregation by interaural phase, by momentary amplitude disparity, and by monaural phase. *Journal of the Audio Engineering Society, 31*, 630-634.

Kubovy, M., & Jordan, R. (1979). Tone-segregation by phase: On the phase sensitivity of the single ear. *The Journal of the Acoustical Society of America, 66*, 100-106.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America, 49*, 467-477.

Macken, W. J., Tremblay, S., Houghton, R. J., Nicholls, A. P., & Jones, D. M. (2003). Does auditory streaming require attention? Evidence from attentional selectivity in short-term memory. *Journal of Experimental Psychology: Human Perception and Performance, 29*, 43-51.

McCabe, S. L., & Denham, M. J. (1997). A model of auditory streaming. *The Journal of the Acoustical Society of America, 101*, 1611-1621.

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

Micheyl, C., Tian, B., Carlyon, R. P., & Rauschecker, J. P. (2005). Perceptual organization of tone sequences in the auditory cortex of awake macaques. *Neuron, 48*, 139-148.

Miller, G. A., & Licklider, J. C. R. (1950). The intelligibility of interrupted speech. *The Journal of the Acoustical Society of America*, 22, 167-173.

Moore, B. C. J. (2003). *An introduction to the psychology of hearing* (5th ed.). London: Academic Press.

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

Miller, C. T., Dibble, E., & Hauser, M. D. (2001). Amodal completion of acoustic signals by a nonhuman primate. *Nature Neuroscience, 4*, 783-784.

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

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

Näätänen, R., & Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin, 125*, 826-859.

Necker, L. A., (1832). Observations on some remarkable optical phenomena seen in Switzerland; and on an optical phenomenon which occurs on viewing a figure of a crystal or geometrical solid. *The London and Edinburgh Philosophical Magazine and Journal of Science, 1*, 329-337.

Neff, D. L., Jesteadt, W., & Brown, E., (1982). The relation between gap discrimination and auditory stream segregation. *Perception and Psychophysics, 31*, 493-501.

Pastore, R. E., Harris, L. B., & Kaplan, J. K. (1982). Temporal order identification: Some parameter dependencies. *The Journal of the Acoustical Society of America, 71*, 430-436.

Petkov, C. I., O'Connor, K. N., & Sutter, M. L. (2003). Illusory sound perception in macaque monkeys. *The Journal of Neuroscience, 23*, 9155-9161.

Petkov, C. I., O'Connor, K. N., & Sutter, M. L. (2007). Encoding of Illusory Continuity in Primary Auditory Cortex. *Neuron, 54*, 153-165.

Plomp, R., & Mimpen, A. M. (1968). The ear as a frequency analyzer. *The Journal of the Acoustical Society of America, 43*, 764-767.

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

Rasch, R. (1978). The perception of simultaneous notes such as in polyphonic music. *Acustica, 40*, 21-33.

Regan, D., & Tansley, B. W. (1979). Selective adaptation to frequency-modulated tones: Evidence for an information-processing channel selectively sensitive to frequency changes. *The Journal of the Acoustical Society of America, 65*, 1249-1257.

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. *The 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 and Performance, 34*, 992-1006.

Roberts, B., & Moore, B. C. J. (1991). The influence of extraneous sounds on the perceptual estimation of first-format frequency in vowels under conditions of asynchrony. *The Journal of the Acoustical Society of America, 89*, 2922-2932.

Rogers, W. L., & Bregman, A. S. (1993). An experimental evaluation of three theories of auditory stream segregation. *Perception and 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 and Psychophysics*, *60*, 1216-1227.

Saberi, K., Takahashi, Y., Konishi, M., Albeck, Y., Arthur, B., & Farahbod, H. (1998). Effects of Interaural Decorrelation on Neural and Behavioral Detection of Spatial Cues. *Neuron*, *21*, 789-798.

Sek, A., & Moore, B. C. J. (1995). Detection of mixed modulation using correlated and uncorrelated noise modulators. *The Journal of the Acoustical Society of America*, *95*, 3511-3518.

Singh, P. G. (1987). Perceptual organization of complex tone sequences. *The Journal of the Acoustical Society of America*, *82*, 886-899.

Singh, P. G., & Bregman, A. S. (1997). Effect of different timbre attributes on the perceptual segregation of complex-tone sequences. *The Journal of the Acoustical Society of America*, *102*, 1943-1952.

Sheskin, D. (2004). *Handbook of parametric and nonparametric statistical procedures* (3^{rd} ed.). Chapman & Hall/CRC.

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

Snyder, J. S., Alain, C., & Picton, T. W. (2006). Effects of attention on neuroelectric correlates of auditory stream segregation. *The Journal of Cognitive Neuroscience, 18*, 1–13.

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

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 and Psychophysics*, *69*, 136-152.

Sussman, E., Ritter, W., & Vaughan, H. G., Jr. (1999). An investigation of the auditory streaming effect using event-related brain potentials. *Psychophysiology*, *36*, 22-34.

Thurlow, W. R. (1957). An auditory figure-ground effect. *American Journal of Psychology*, 70, 653-654.

Tremblay, S., & Jones, D. M. (1998). Role of habituation in the irrelevant sound effect: Evidence from the effects of token set size and rate of transition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 659-671.

Turgeon, M., Bregman, A. S., & Roberts, B. (2005). Rhythmic masking release: Effects of asynchrony, temporal overlap, harmonic relations, and source separation on crossspectral grouping. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 939-953.

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. *The Journal of the Acoustical Society of America*, *106*, 938-945.

Vliegen, J., & Oxenham, A. J. (1999). Sequential stream segregation in the absence of spectral cues. *The Journal of the Acoustical Society of America*, *105*, 339-346.

Warren, R. M. (1970). Perceptual Restoration of Missing Speech sounds. *Science*, *167*, 392-393.

Warren R. M. (1982). *Auditory perception: A new synthesis.* New York: Pergammon.

Warren, R. M., Obusek, C. J., & Ackroff, J. M. (1972). Auditory induction: perceptual synthesis of absent sounds. *Science*, *176*, 1149-1151.

Warren, R. M., Obusek, C. J., Farmer, R. M., & Warren, R. P. (1969). Auditory sequence: Confusion of patterns other than speech and music. *Science*, *164*, 586-587.

Watter, S., Geffen, G. M., & Geffen, L. B. (2001). The n-back as a dual-task: P300 morphology under divided attention. *Psychophysiology*, *38*, 998-1003.

Wessel, D. (1979). Timbre space as a musical control structure. *Computer Music Journal*, *3*, 45-52.

Winkler, I., Horváth, J., Teder-Sälejärvi, W. A., Näätänen, R., & Sussman, E. (2003a). 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.

Winkler, I., Sussman, E., Tervaniemi, M., Ritter, W., Horváth J., & Näätänen, R. (2003b). Pre-attentive auditory context effects. *Cognitive, Affective, and Behavioral Neuroscience*, *3*, 57-77.

Appendix

Many of the studies presented in this thesis have also contributed to posters presentations at various conferences. These were:

Haywood, N., & Roberts, B. (2006, September 14). *Exploring the build-up and resetting of auditory stream segregation.* Poster presented at the BSA Short Papers Meeting on Experimental Studies of Hearing and Deafness, Cambridge, UK.

Haywood, N., & Roberts, B. (2007, September 20). *Pure-tone percepts caused by brief changes in the level or lateralization of components in a complex tone are organized into streams.* Poster presented at the BSA Short Papers Meeting on Experimental Studies of Hearing and Deafness, London, UK.

Haywood, N., & Roberts, B. (2008, February 16). *Resetting effects of a single deviant tone on the build-up of auditory stream segregation.* Poster presented at the ARO Midwinter Research Meeting, Phoenix, USA[2].

Haywood, N., & Roberts, B. (2008, September 18). *The effects of perceived continuity and predictability on the build-up of auditory stream segregation.* Poster presented at the BSA Short Papers Meeting on Experimental Studies of Hearing and Deafness, York, UK.

Roberts, B., Haywood, N. R., Moore, B. C. J., & Glasberg B. R. (2007, August 17). *Exploring the build-up and resetting of auditory stream segregation.* Poster presented by Prof. Brian Roberts at the APA Convention 2007, San Fransisco, USA.

[2] Attendance to this conference was supported by a travel grant awarded from Deafness Research UK.