UNLV Theses, Dissertations, Professional Papers, and Capstones

8-1-2012

Frequency-specificity and pattern-specificity of the buildup of auditory stream segregation

David Michael Weintraub

University of Nevada, Las Vegas, weintra2@unlv.nevada.edu

Follow this and additional works at: https://digitalscholarship.unlv.edu/thesesdissertations

Part of the Cognition and Perception Commons, Cognitive Psychology Commons, and the Neurosciences Commons

FREQUENCY-SPECIFICITY AND PATTERN-SPECIFICITY OF THE BUILDUP OF
AUDITORY STREAM SEGREGATION

By

David M. Weintraub

Bachelor of Liberal Arts
University of Nevada, Las Vegas
2009

A thesis submitted in partial fulfillment
of the requirements for the

Master of Arts in Psychology

Department of Psychology
College of Liberal Arts
The Graduate College

University of Nevada, Las Vegas
August 2012

THE GRADUATE COLLEGE

We recommend the thesis prepared under our supervision by

David Weintraub

entitled

Frequency-Specificity and Pattern-Specificity of the Buildup of Auditory Stream Segregation

be accepted in partial fulfillment of the requirements for the degree of

Master of Arts in Psychology
Department of Psychology

Joel Snyder, Committee Chair

Erin Hannon, Committee Member

Jefferson Kinney, Committee Member

Pierre Lienard, Graduate College Representative

Ronald Smith, Ph. D., Vice President for Research and Graduate Studies
and Dean of the Graduate College

August 2012

ABSTRACT

FREQUENCY-SPECIFICITY AND PATTERN-SPECIFICITY OF THE BUILDUP OF
AUDITORY STREAM SEGREGATION

By

David M. Weintraub

Joel S. Snyder, Examination Committee Chair
Assistant Professor of Psychology
University of Nevada, Las Vegas

During repeating sequences of low (A) and high (B) tones in an "…ABAB…" pattern, the likelihood of hearing two separate streams ("streaming") increases with more repetitions of the patterns, a phenomenon referred to as "buildup". Previous studies have shown that buildup is frequency specific (Anstis & Saida, 1985) and that its biasing effects decays over several seconds (Beauvois & Meddis, 1997). No study has examined whether the frequency specificity of buildup persists for such a long duration. To address these issues, Experiment 1 tested the decay of frequency-specific and non-frequency specific buildup. The results revealed that (1) frequency-specific buildup effects were strongest during short decay intervals and decayed with longer intervals, (2) non-frequency-specific buildup showed weaker buildup effects and less decay, and (3) both types of buildup had significant effects compared to a silence baseline comparison even after long decay intervals. It is assumed non-frequency-specific buildup involved mechanisms in a high-level auditory area not finely tuned to frequency and sensitive to complex features. Therefore, Experiment 2 tested whether mechanisms subserving buildup occur in areas of the auditory pathway sensitive to rhythmic pattern. The main results revealed that (1) frequency-specific and non-frequency specific buildup effects

were both disrupted by rhythmic pattern irregularity given their effects were large

without such irregularity, and (2) replicated all other aspects of Experiment 1. The results

of both experiments confirmed the presence of a frequency-specific mechanism

subserving buildup that may be longer-lasting than previously recognized and further

supported the presence of non-frequency specific mechanisms that are also long-lasting.

Additionally, buildup appeared to involve mechanisms in high-level auditory areas

sensitive to rhythmic pattern. Taken together, this study demonstrated buildup is a

complex process that involves multiple levels of analysis along the auditory pathway.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Natural environments are typically composed of an array of sounds coming from different sources. These sounds then enter our ears as a complex input. The problem faced by our auditory system is to break up this input and form distinct auditory streams for each source. Such perceptual organization is important for the recognition of speech during a noisy cocktail party in the presence of competing speakers (Cherry, 1953) or the perception of a melody played by an instrument amongst an ensemble of performers. This process, known as *auditory scene analysis* (Bregman, 1990), is often studied by looking specifically at *auditory stream segregation* (Bregman & Campbell, 1971) or the separation of interleaved sounds into separate streams. Despite its importance in hearing and the subjective ease with which it occurs, the neural mechanisms subserving auditory stream segregation are not completely understood.

Understanding auditory stream segregation in human listeners informs important ecologically valid issues within hearing sciences including evolution and animal communication. Psychophysics discoveries of the conditions which facilitate auditory stream segregation are important, because they provide insights into the cues the auditory system has evolved to process in the analysis of acoustic scenes (Bee & Micheyl, 2008). For example, sounds produced by a given source share more acoustic properties in common than with the sounds produced by other sources (Bregman, 1990). The auditory system appears to have evolved to exploit such cues when analyzing acoustic scenes. Additionally, studies of auditory stream segregation in humans is important for understanding animal communication as stream segregation capabilities are conserved

across species (Bee & Micheyl, 2008; Fay & Popper, 2000). For example, female gray treefrogs use frequency cues, in a manner qualitatively similar to humans, to segregate and localize mating calls from conspecific males in noisy multi-species breeding aggregations (Nityananda & Bee, 2011).

In the laboratory, auditory stream segregation is typically studied using a paradigm introduced by Van Noorden (1975). Listeners are presented with a sequence of low (A) and high (B) frequency pure tones in an alternating "…ABAB…" pattern. When the tones are integrated into the same stream the sequence is heard as a coherent trill of alternating frequencies (termed coherence); however, when the tones are segregated into separate streams the sequence is heard as two metronomes of different frequency (termed segregation). Although stream segregation may occur when streams differ on almost any salient perceptual cue (Moore & Gockel, 2002), the strongest influences are the frequency separation (Δf) between tones (Hartmann & Johnson, 1991) and presentation rate (PR) of the sequence (Bregman & Campbell, 1971). Furthermore, listeners have a tendency to first hear a sequence as coherent and later segregate it after several seconds of exposure (Bregman, 1978), a phenomenon called buildup (Anstis & Saida, 1985).

Studies on auditory stream segregation are typically concerned with answering two questions. *Where* in the brain does it take place and *how* (Shamma & Micheyl, 2010)? It has been asserted that much of stream segregation can be explained by activity within peripheral and primary auditory areas. This conclusion comes from several animal and human studies using physiological and psychophysical methods. However, as reviewed here, these studies are not sufficient to explain additional evidence that suggests

stream segregation is a complex process that involves multiple mechanisms including those beyond primary auditory areas (Moore & Gockel, 2002; Snyder & Alain, 2007). The literature review below addresses both questions concerning auditory stream segregation. Afterward, two studies will be presented which examine whether buildup involves multiple distinct mechanisms within and beyond peripheral and/or primary auditory areas, respectively.

Organization of the Auditory Pathway. An understanding of the possible mechanisms underlying auditory stream segregation requires familiarization with the organization of the auditory pathway. Beginning in the auditory periphery, nerve fibers contain a characteristic frequency to which they are most responsive (Konig, Heil, Budinger, & Scheich, 2005; Moore, 2003). Importantly, these fibers are located in an organized manner such that fibers selective for high frequencies are located more peripherally within the auditory nerve bundle and fibers selective for lower frequencies are located more centrally. Thus, a gradient from low to high frequency-tuned fibers forms a tonotopic map within the auditory periphery. Ultra-high-resolution fMRI reveals a similar tonotopic organization exists up to the primary auditory cortex in humans and becomes less precise thereafter (Formisano et al., 2003). Much like the visual system, a hierarchy of feature selectivity occurs by which auditory areas beyond those tonotopically organized are responsive to complex features (Moore, 2003; Rauschecker & Scott, 2009)For example, upon receiving input from frequency-selective neurons, a neuron late in the auditory pathway may be selective for a particular speech sound. In summary, the organization of the auditory pathway can be thought of as a hierarchy in which low-level areas are tonotopically organized and selective for simple features (e.g.,

frequency) whereas high-level areas are less precisely tonotopically organized and selective for complex features (e.g., speech sound).

Low-Level Mechanisms: Effects of Δf and PR. Much of stream segregation is likely to be a consequence of the tonotopic organization of low-level auditory areas (Micheyl et al., 2007). An influential theory by Hartmann and Johnson (1991) says that sound sources that activate separate non-overlapping peripheral frequency channels will be perceived as arising from separate sources. For example, low (A) and high (B) frequency tones in an "…ABAB…" sequence with a large Δf will activate spatially distinct areas along the tonotopic map of peripheral auditory areas. The non-overlapping patterns of activation will facilitate a segregated percept. Computational models further argue that high rates of adaptation within low-level auditory areas during fast PRs reduce the likelihood that two sounds will activate overlapping neuronal populations (Beauvois & Meddis, 1996; McCabe & Denham, 1997). Again, the non-overlapping patterns of activation will give rise to a segregated percept. Thus, these models simply argue that the effects of Δf and PR on auditory stream segregation can be explained by the tonotopic organization and adaptation of low-level auditory areas leading to the spatial separation of neuronal populations representing different streams.

Evidence in support of the models described above was provided by studies on the auditory cortex of monkeys (Fishman, Arezzo, & Steinschneider, 2004; Fishman, Reser, Arezzo, & Steinschneider, 2001) and songbird forebrain (Bee & Klump, 2004, 2005). While recording from a primary auditory cortical area whose best frequency (BF) response corresponded to the A tone of an "…ABAB…" sequence, the response to the intervening B tone at this site decreased with increasing Δf and faster PR. The B-tone

response reduction with increasing Δf likely reflected the spatial separation of neuronal populations activated by the two tones. Furthermore, forward suppression at the BF A-tone site caused by the preceding A tone temporally inhibited the response to the B tone at this site, and this suppression was stronger with faster PR and smaller Δf. This, in effect, ensured that the BF A-tone site only responded to the A tone, consequently increasing the spatial separation of neural populations activated by the two tones. Note that the same forward suppression likely occurred at the BF B-tone site causing a similar suppression to A-tone activity at this site. Unfortunately, these studies did not measure perception of streaming during recording making it difficult to confirm the findings as a neural correlate of auditory stream segregation. Nonetheless they provided convincing evidence consistent with the theories described above that the effects of Δf and PR on auditory stream segregation can be explained by the spatial separation of neuronal populations representing different streams within primary auditory areas or earlier.

Low-Level Mechanisms: Buildup. A series of experiments by Anstis and Saida (1985) argued for a low-level mechanism underlying buildup. They showed that the rate of buildup causing a continuous frequency-modulated (FM) tone to be heard as split (segregated) rather than continuous (coherent) increased with faster modulation rate (Experiment 1). This evidence was used to argue that buildup reflected stimulus-driven adaptation. Furthermore, adapting to one frequency did not produce buildup in other frequencies outside a small range (Experiment 4). Therefore, buildup occurred within frequency-tuned tonotopically-organized auditory areas. Finally, adapting to one ear did not produce buildup in the other ear (Experiment 5). Therefore, buildup had an early peripheral origin before information from the two ears was integrated. These results were

later reproduced by computational modeling based on similar principles (McCabe &
Denham, 1997). Additionally, consistent with the argument that buildup reflects
stimulus-driven adaptation, Rogers & Bregman (1993) showed that the amount of
buildup increased with the number of tone onsets (i.e., event density) when using discrete
pure tones.

Single- and multi-unit recording in animals confirmed that buildup reflects the
adaptation of frequency-tuned neurons within primary auditory cortex of awake rhesus
monkeys (Micheyl, Tian, Carlyon, & Rauschecker, 2005) and the cochlear nucleus of
anesthetized guinea pigs (Pressnitzer, Sayles, Micheyl, & Winter, 2008). In particular, the
likelihood that a neuron tuned to the A-tone frequency in an "…ABA_..." sequence ('_'
denotes a silent gap) also partially responded to the intervening B-tone decreased
following sufficient exposure time. Consequently, the spatial separation of neuronal
populations representing the two tones produced a segregated percept. Indeed,
computational modeling using the animal physiology data closely replicated perceptual
reports of stream segregation in humans using similar stimulus parameters (Micheyl, et
al., 2005; Pressnitzer, et al., 2008).

In summary, the studies discussed above provide converging evidence that
processes within low-level auditory areas can account for much of auditory stream
segregation. In particular, the effect of Δf can be accounted for by the tonotopic
organization of peripheral and primary auditory areas. The effect of PR and buildup can
be accounted for by adaptation within these areas. These studies argued that the
perception of streaming occurred when neuronal populations representing different
streams formed spatially distinct non-overlapping patterns of activation within low-level

auditory areas. However, as discussed below, this argument is insufficient to explain evidence for more complex mechanisms underlying auditory stream segregation (Moore & Gockel, 2002; Snyder & Alain, 2007).

Effects of Attentional Modulation. Psychophysical studies have shown that buildup of stream segregation was modulated by attention, and therefore suggested the involvement of high-level mechanisms in perception of streaming. In these studies, participants were presented with an "…ABA_..." pattern to one ear. The role attention plays in auditory stream segregation was examined by assessing the buildup of streaming while participants were engaged in a separate auditory, visual, or non-sensory task in which participants counted backwards (Carlyon, Cusack, Foxton, & Robertson, 2001; Carlyon, Plack, Fantini, & Cusack, 2003; Cusack, Deeks, Aikman, & Carlyon, 2004; Thompson, Carlyon, & Cusack, 2011). By having participants engaged in one of these other tasks, attention was diverted away from the "…ABA_..." pattern. When attending to the "…ABA_..." pattern, participants showed a typical pattern of buildup. However, when attending to the other task for first part of the "…ABA_..." pattern, participants failed to show any sign of buildup when they switched their attention. Thus, buildup either did not occur while attention was diverted to the other task or it was reset following the brief switch in attention (Cusack, et al., 2004), a distinction that has been quite difficult to resolve using psychophysical measurements. These effects occurred regardless of the task used to capture attention (Carlyon, et al., 2003), suggesting that buildup involves mechanisms within central auditory areas, multimodal pathways, and/or in peripheral areas that can be influenced in a top-down fashion by attention.

To explain these results, Cusack et al. (2004) proposed a hierarchical model of stream segregation. According to this model, preattentive mechanisms segregate streams based on acoustic features (e.g., Δf) and attention-dependent buildup mechanisms further break down outputs (streams) of this earlier process that are attended to. For example, when talking to a friend at a concert, low-level processes automatically segregate the friend's voice from the music. However, since attention is allocated to the friend's voice and not the concert, buildup processes do not further decompose the music into its constituent parts (e.g., guitar, drums, bass, etc.).

Consistent with this model, Snyder et al. (2006) provided event-related potential (ERP) evidence for at least two mechanisms contributing to stream segregation: a preattentive segregation mechanism and an attention-dependent buildup mechanism. In particular, auditory cortical responses (P2 and N1c) to an ABA_ pattern increased in amplitude with increasing Δf and correlated with behavioral measures of streaming; this enhancement occurred even when attention was directed away from the ABA_ pattern. Additionally, a temporally broad enhancement following the onset of an ABA_ pattern progressively increased in positivity throughout the course of the pattern. The time course of this progressive increase indicated a strong link with the buildup of streaming. Importantly, this enhancement was diminished when participant's attention was directed away from the ABA_ pattern. These findings support the existence of an attention-dependent buildup mechanism in addition to a preattentive segregation mechanism. Also, since buildup-related processes were measured during passive listening, these findings are more consistent with an effect of sustained attention as opposed to the possibility that buildup is simply reset following brief switches in attention [cf. (Cusack, et al., 2004)].

Importantly, because neural processes in secondary auditory areas are more sensitive to attentional modulation than primary auditory areas (Okamoto, Stracke, Bermudez, & Pantev, 2011; Petkov et al., 2004), it seems likely that preattentive segregation is mediated by mechanisms lower in the auditory pathway than those mediating attention-dependent buildup (Snyder, et al., 2006).

Effects of Context. Evidence for high-level mechanisms, in addition to those described above, has been provided by a series of studies showing that perception of streaming varies depending on the preceding context. In particular, stream segregation is modulated by previous stimuli and percepts, respectively (Snyder, Carter, Hannon, & Alain, 2009; Snyder, Carter, Lee, Hannon, & Alain, 2008; Snyder, Holder, Weintraub, Carter, & Alain, 2009; Snyder & Weintraub, 2011). For example, an effect of prior stimulus occurs when an "…ABA_…" sequence with an intermediate Δf is more likely to be heard as segregated when it is preceded by a small-Δf sequence and coherent when preceded by a large-Δf sequence. An effect of prior perception occurs when perception during the previous sequence is maintained such that the same perception is later heard during the subsequent sequence when the two sequences have the same Δf. Importantly, ERP work suggests prior Δf and prior perception involve distinct mechanisms from those involved with current Δf (Snyder, Holder, et al., 2009). For the effect of prior Δf, these mechanisms likely occur in areas of the auditory pathway not finely tuned to frequency (Snyder, Carter, et al., 2009) and sensitive to more complex features such as rhythmic pattern (Snyder & Weintraub, 2011). Together with the findings of Snyder et al. (2006), these studies suggest the involvement of multiple mechanisms subserving auditory stream segregation, at least some of which occur beyond low-level auditory areas.

Outside Auditory Cortex. Given that the studies described above suggest that mechanisms subserving stream segregation are not constrained entirely within low-level auditory areas, it seems possible that brain areas outside of auditory cortex are also involved in the perception of streaming. Indeed, recent fMRI studies have provided converging evidence that, in addition to auditory cortex (Hill, Bishop, Yadav, & Miller, 2011), areas within the intraparietal sulcus (IPS) were more active during perception of two streams compared to one (Cusack, 2005; Hill, et al., 2011; Teki, Chait, Kumar, von Kriegstein, & Griffiths, 2011). However, given the poor temporal resolution of fMRI, it is not clear whether activity within IPS reflected modulation of streaming itself or the output of segregating mechanisms within auditory cortex. Additionally, a recent intracranial electroencephalography study showed that evoked-potential neural correlates of Δf were found over widespread brain areas including posterior superior temporal gyrus, middle temporal gyrus, pre- and post-central gyri, inferior and middle frontal gyri, and the supra-marginal gyrus (Dykstra et al., 2011). Finally, feedforward and feedback connections between the medial geniculate body and auditory cortex were recruited during perceptual reversals (e.g., one to two streams, two to one stream) (Kondo & Kashino, 2009; Schadwinkel & Gutschalk, 2011). Taken together, these studies suggest that stream segregation involves multiple levels of analysis in brain areas within and outside of auditory cortex.

Current Study. Previous studies have shown neural correlates of buildup occur in at least two levels of the auditory pathway, the cochlear nucleus (Pressnitzer, et al., 2008) and primary auditory cortex (Micheyl, et al., 2005). However, no study has shown that buildup occurs beyond these levels in parts of the auditory pathway not finely tuned to

frequency. It seems likely given evidence that stream segregation involves multiple mechanisms including those in high-level auditory areas and non-auditory areas. One study that tested the frequency-specificity of buildup showed that the effects of adapting to one frequency did not transfer to other frequencies outside a small range (Anstis & Saida, 1985). Therefore, they concluded that buildup occurred within frequency-tuned auditory areas. However, since no baseline measure was included, it could only be concluded that adapting to the same frequency produced more buildup than adapting to frequencies outside a small range. It is not clear in this latter case whether no buildup occurred at all or a weaker buildup occurred that was still above a baseline level.

Experiment 1 examined whether buildup involves mechanisms in areas of the auditory pathway not finely tuned to frequency, in addition to those in frequency-tuned areas, using a paradigm borrowed from Beauvois and Meddis (1997). Listeners were presented with a 10-second "…AAAA…" induction sequence designed to buildup the tendency to hear a subsequent 1.44-second "…ABAB…" test sequence as segregated. Buildup was measured as the strength with which the induction sequence induced streaming during the test. Importantly, the frequency of the induction tones varied such that they either matched or mismatched the frequency of the test sequence tones. We tested whether the amount of buildup produced by these non-frequency-matching induction tones was larger compared to a silent induction sequence (i.e., baseline measure). If so, this would provide evidence for mechanisms subserving buildup within auditory areas not finely tuned to frequency. Furthermore, we measured the decay of buildup by varying the silent interval duration between the induction and test sequences. For frequency-matching induction sequences, buildup decays to an asymptotic level by

11

about 4 seconds (Beauvois & Meddis, 1997). However, given that the suppressive effects of adaptation last longer higher in the auditory pathway (Fitzpatrick, Kuwada, Kim, Parham, & Batra, 1999; Harms & Melcher, 2002), buildup during non-frequency-matching induction sequences was expected to display less decay compared to frequency-matching induction sequences. If so, this would provide evidence that the two types of buildup are subserved by distinct mechanisms with different rates of decay. Finally, we tested whether buildup completely decayed by 4 seconds as proposed by Beauvois and Meddis (1997). Since their study included no baseline measure, it is not clear whether buildup decayed to a level equivalent to no buildup at all or stabilized at a level above baseline.

Experiment 2 was designed to further constrain where non-frequency-matching buildup occurs within the auditory pathway. In particular, we tested whether buildup within auditory areas not finely tuned to frequency occurs within areas sensitive to rhythmic pattern. Rogers and Bregman (1993) showed that rhythmic pattern regularity between an induction and test sequence played no role in stream segregation. However, if buildup produced by non-frequency-matching induction sequences involves mechanisms within high-levels of the auditory pathway, these mechanisms may be localized within an area sensitive to complex features such as rhythmic pattern. Indeed, there is growing evidence that rhythmic pattern is important for facilitating the effect of Δf (Jones, Kidd, & Wetzel, 1981) and prior Δf (Snyder & Weintraub, 2011), segregating sounds from background noise that overlaps in frequency (Teki, et al., 2011), suppressing a distracter stream from a stream-of-interest (Devergie, Grimault, Tillmann, & Berthommier, 2010), and stabilizing perception onto a segregated stream (Bendixen, Denham, Gyimesi, &

12

Winkler, 2010). Therefore, it seems likely that rhythmic pattern is also important in the buildup of auditory stream segregation. If manipulating rhythmic pattern regularity between the induction and test sequences have different effects on frequency-matching and non-frequency-matching buildup, this would provide further evidence that the two are subserved by distinct mechanisms. It would also provide evidence that buildup involves mechanisms in parts of the auditory pathway sensitive to rhythmic pattern.

CHAPTER 2

EXPERIMENT 1

Method

Participants

Forty-six undergraduates (30 females, mean age=21.48, age rage=18-47) from the

University of Nevada, Las Vegas participated in this study for course credit. All

participants reported having normal hearing. Informed consent was obtained from all

participants prior to the start of the experiment according to a protocol approved by the

University's Office for the Protection of Research Subjects.

Apparatus

Stimulus presentation and behavioral responses were collected using a custom

program written for Presentation (Neurobehavioral Systems, Inc., Albany, CA) running

on a Pentium 4 computer with an SB X-Fi sound card (Creative Technology, Ltd.).

Auditory stimuli were delivered through Sennheiser HD 280 headphones (Sennheiser

Electronic Corporation, Old Lyme, CT). Behavioral responses were made on a computer

keyboard.

Stimuli

Auditory stimuli were generated off-line in Matlab (The MathWorks, Inc., Natick,

MA) and consisted of pure tones (50ms in duration, including 5ms rise/fall times)

presented binaurally at around 70 dB SPL. All trials consisted of an induction phase

followed by a subsequent test phase (Figure 1). For all trials, the test phase was fixed

such that it always consisted of a 1.44-second sequence of four repetitions of an ABAB

pattern. A refers to a low frequency 1000 Hz tone and B refers to a high frequency 1420

CHAPTER 2

EXPERIMENT 1

Method

Participants

Forty-six undergraduates (30 females, mean age=21.48, age rage=18-47) from the

University of Nevada, Las Vegas participated in this study for course credit. All

participants reported having normal hearing. Informed consent was obtained from all

participants prior to the start of the experiment according to a protocol approved by the

University's Office for the Protection of Research Subjects.

Apparatus

Stimulus presentation and behavioral responses were collected using a custom

program written for Presentation (Neurobehavioral Systems, Inc., Albany, CA) running

on a Pentium 4 computer with an SB X-Fi sound card (Creative Technology, Ltd.).

Auditory stimuli were delivered through Sennheiser HD 280 headphones (Sennheiser

Electronic Corporation, Old Lyme, CT). Behavioral responses were made on a computer

keyboard.

Stimuli

Auditory stimuli were generated off-line in Matlab (The MathWorks, Inc., Natick,

MA) and consisted of pure tones (50ms in duration, including 5ms rise/fall times)

presented binaurally at around 70 dB SPL. All trials consisted of an induction phase

followed by a subsequent test phase (Figure 1). For all trials, the test phase was fixed

such that it always consisted of a 1.44-second sequence of four repetitions of an ABAB

pattern. A refers to a low frequency 1000 Hz tone and B refers to a high frequency 1420

Hz tone. The frequency separation (Δf) between the tones corresponded to a musical interval of 6 semitones. The interstimulus interval (ISI) between tones was 40ms and the stimulus-onset asynchrony (SOA) was 90ms. When the test phase is played in isolation, it is usually heard as a trill (coherent) (Van Noorden, 1975), however, when preceded by a frequency-matching induction sequence (as described below), it can be heard as two metronomes (segregated) (Beauvois & Meddis, 1997). The induction phase consisted of 10 seconds of either silence (i.e., baseline condition) or 111 isofrequency tones (i.e., "…BBBB…") that were 1420 Hz, 1690 Hz, or 2840 Hz with a 40ms ISI and 90ms SOA. Thus, the Δf between the non-silent induction phase and the high tone of the test phase was 0, 3, or 12 semitones, respectively. The silent interval duration between the induction and test phases was 0, 1, 4, or 8 seconds long. Each induction type was paired with each silent duration interval making a total of 16 trial types. The intertrial interval (ITI) was 5 seconds.
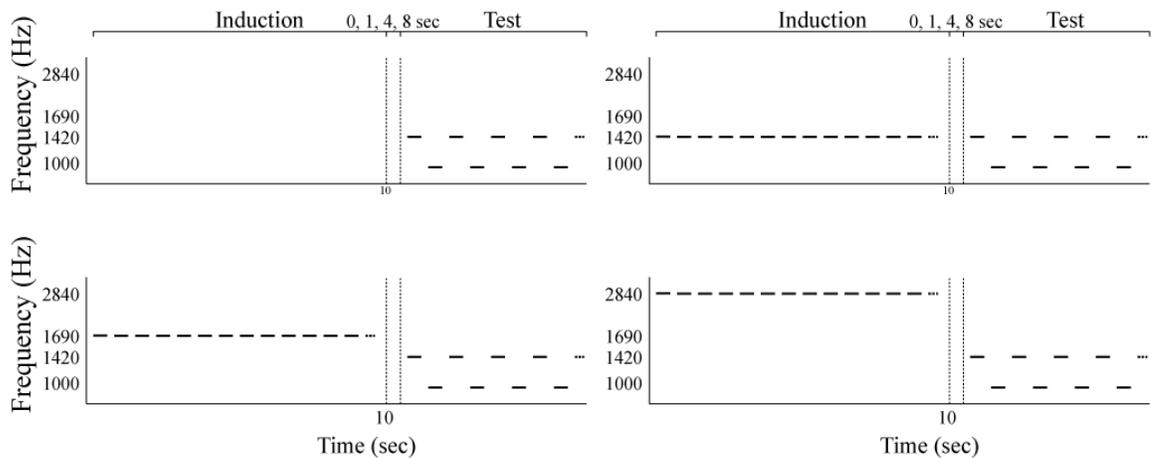

Figure 1: Visual diagram of all trial types. Diagrams depict silent (top left), 1420 Hz (top right), 1690 Hz (bottom left), and 2840 Hz (bottom right) induction types. Within each diagram the silent duration period between induction and test is either 0, 1, 4, or 8 seconds long.

Procedure

After signing a consent form and filling out a brief demographics questionnaire, participants were seated in a quiet room and given verbal instructions by the primary investigator or a trained research assistant. Instructions also included examples of extended 5.75-second test sequences with a Δf of 3, 6, or 12 semitones. This was intended to give participants clear examples of the different possible perceptions. In order to control for visual attentional focus, participants were instructed to maintain fixation on a white cross centered on a black background on a computer screen throughout the experiment. The cross remained white during the induction phase and ITI and turned red 2 seconds prior to the onset of the test phase. It remained red until the end of the test phase. This was intended to notify participants of the onset of the test phase. At the end of each trial, participants indicated whether they heard the test phase as a trill (coherent) or two metronomes (segregated) by pressing the '1' or '2' button on a computer keyboard number pad, respectively. Participants were instructed to listen attentively to the induction phase, but that they were making no judgments about it. Participants were encouraged not to bias their perception. Prior to beginning the main experiment, participants were given six practice trials (and additional trials if at first they had a poor understanding of the procedures) selected randomly from all possible trial types.

During the experiment, trials were presented in four different blocks. Each block contained 40 trials all with the same silent interval duration between induction and test phases. The order of block presentation was randomized using a Latin square design. Of the 40 trials, 10 of each of the four induction types were presented. Thus, a total of 160 trials, 10 of each trial type, were presented. Between blocks, participants were

encouraged to break for as long as they liked. The experiment lasted approximately 60 minutes.

<div align="center">Results</div>

For each participant, the proportion of segregated responses was averaged across all 10 trials of each trial type. Figure 2 displays the average proportion of segregated responses for each trial type. These proportions were then entered into a 4 (induction type) x 4 (silent interval duration) repeated-measures analysis of variances (ANOVA). The degrees of freedom were adjusted using the Greenhouse-Geisser ε, a conservative estimate of p. All reported probability estimates based on the reduced degrees of freedom; however, no multiple-comparison corrections were applied for post-hoc analyses. There was a significant main effect of induction type, $F(3,135)=26.05$, $p<.001$, such that the 1420 and 1690 Hz inductions produced significantly more segregated responses than the silent induction [1420 Hz: $F(1,45)=72.87$, $p<.001$; 1690 Hz: $F(1,45)=41.94$, $p<.001$]. The 2840 Hz induction produced more segregated responses than the silent induction, but this difference was only marginally significant, $F(1,45)=3.95$, $p=.053$. Other significant differences between induction types can be summarized as 1420 Hz > 1690 Hz > 2840 Hz [1420 Hz > 1690 Hz: $F(1,45)=8.49$, $p<.01$; 1420 Hz > 2840 Hz: $F(1,45)=24.0$, $p<.001$; 1690 Hz > 2840 Hz: $F(1,45)=17.0$, $p<.001$]. The main effect of induction type remained significant even at the 8-second silent interval duration, $F(3,135)=10.55$, $p<.001$, such that the 1420 and 1690 Hz induction types each produced significantly more segregated responses compared to the silent induction [1420 Hz: $F(1,45)=22.11$, $p<.001$; 1690 Hz: $F(1,45)=6.72$, $p<.05$]. Here,

the 2840 Hz and silent induction types produced similar amounts of segregated responses, $F(1,45)=.01$, $p=.94$.

Though the main effect of silent interval duration was non-significant, $F(3,135)=1.68$, $p=.17$, there was a significant induction type x silent interval duration interaction, $F(9,405)=2.36$, $p<.05$. This interaction became marginally significant when the silent induction was excluded from analysis, $F(6,270)=2.06$, $p=.07$. The interaction reflected the reduced effect of induction type with increasing silent interval duration, because the induction types were differently affected by the silent interval duration manipulation. Follow-up analyses showed that while the effect of the silent interval duration within the 1420 Hz induction was non-significant, $F(3,135)=1.70$, $p=.17$, the effect did show a significant linear trend, $F(1,45)=4.06$, $p<.05$, such that fewer segregated responses were reported with increasing silent interval duration with the largest difference between 0 and 1 second. The effect of silent interval duration within the 1690 Hz induction type was also non-significant, $F(3,135)=1.46$, $p=.23$, and showed a marginally significant linear trend, $F(1,45)=3.18$, $p=.08$. The 2840 Hz induction typed show a significant effect of the silent interval duration, $F(3,135)=4.35$, $p<.01$, and showed a significant quadratic trend, $F(1,45)=12.88$, $p<.001$, such that fewer segregated responses were reported for the 0- and 8-second silent interval durations. Finally, the effect of the silent interval duration within the silent induction type was non-significant, $F(3,135)=1.19$, $p=.32$, and showed neither a linear nor quadratic trend [linear: $F(1,45)=2.40$, $p=.13$; quadratic: $F(1,45)=.01$, $p=.91$].
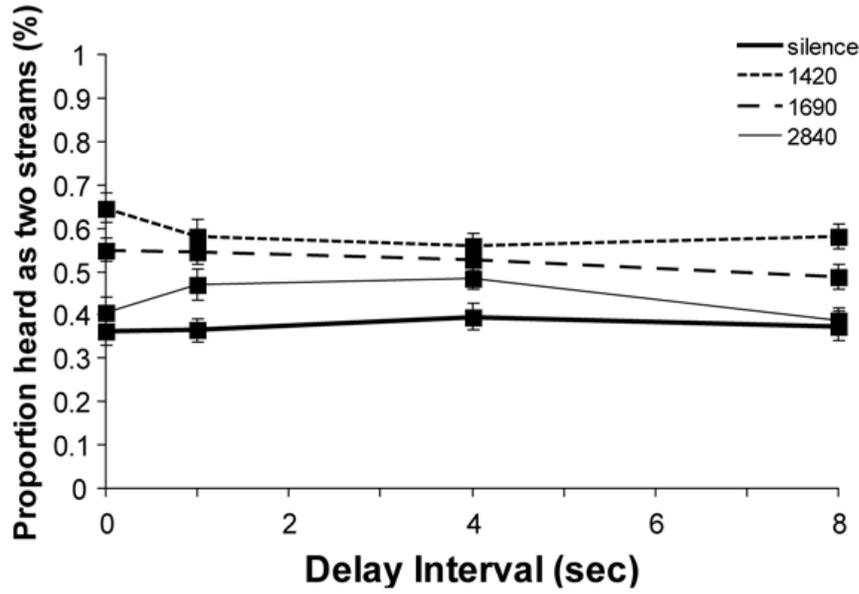
Figure 2: Results from Experiment 1. Error bars based on within-subjects confidence intervals (Cousineau, 2005).

Note that compared to Beauvois & Meddis (1997), our 1420 Hz induction demonstrated little decay of segregated responses across the silent interval durations. For example, between the 0- and 1-second silent interval durations, we found about a 5% reduction in responses, whereas, Beauvois & Meddis (1997) found about a 15% reduction. To address this issue, we tested the hypothesis that participants who displayed a weak buildup effect would also display little decay of segregated responses. For each participant, we calculated the buildup effect size as the difference in segregated responses between the 1420 Hz and silent induction types at the 0-second silent interval duration. The decay of segregated responses was calculated as the slope of segregated responses within the 1420 Hz induction as a function of the silent interval duration. In this latter measure, the more negative the slope, the more decay of segregated responses that participant displayed. A correlation analysis revealed that the two measures were significantly negatively correlated, $r(46)=-.43$, $p<.01$, such that participants who

displayed a large buildup effect also displayed more decay of segregated responses (Figure 3). It is possible that, for those participants that displayed little decay of buildup, attention was not sufficiently directly towards the induction sequences. This would explain the weak buildup effect size in these participants given buildup does not occur in the absence of attention (Carlyon, et al., 2001; Carlyon, et al., 2003; Cusack, et al., 2004; Thompson, et al., 2011). Therefore, we applied a median split between our participants based on their buildup effect size (*M*=.37) and repeated the analyses on just those participants above the median. Note that, when averaged together, the decay of segregated responses for this group of participants was more similar to Beauvois & Meddis (1997) (11% vs. 15%, respectively). Furthermore, groups of participants above and below the median split did not significantly differ on any of our demographic variables such as gender (χ^2=1.07, *p*=.59), age (*t*=-.227, *p*=.82), handedness (χ^2=3.20, *p*=.20), or musical training (*t*=-1.39, *p*=.169).
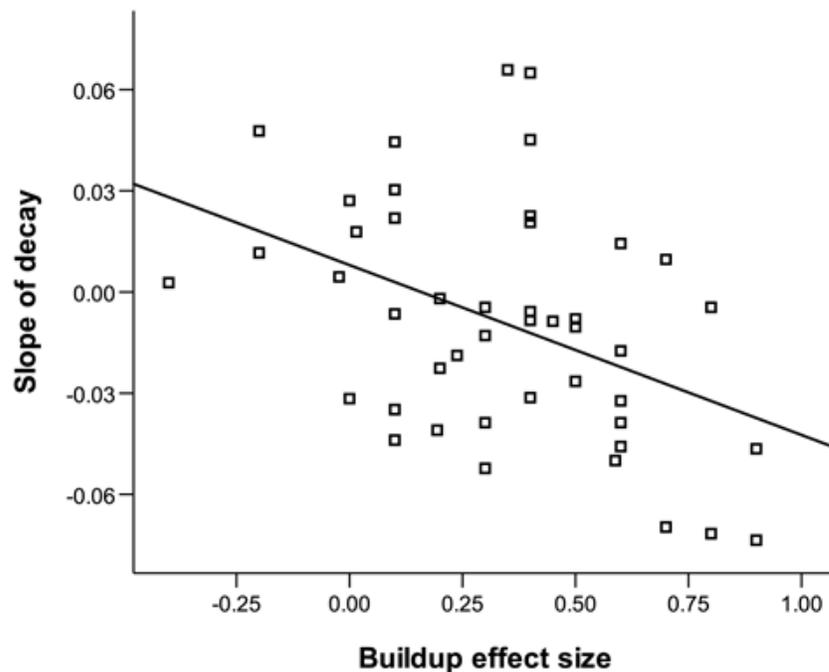

Figure 3: Scatter plot showing the relationship between the slope of segregated responses and buildup effect size. The linear line depicts the line of best fit.

20

Figure 4 displays the average proportion of segregated responses for each trial type for groups above and below the median split separately. Again, the following statistics characterize only those participants above the median. There was a significant main effect of induction type, $F(3,66)=26.12$, $p<.001$, such that the 1420, 1690, and 2840 Hz inductions all produced more segregated responses compared to the silent induction [1420 Hz: $F(1,22)=106.87$, $p<.001$; 1690 Hz: $F(1,22)=53.28$, $p<.001$; 2840 Hz: $F(1,22)=11.81$, $p<.01$]. Other differences between induction types can again be summarized as 1420 Hz > 1690 Hz > 2840 Hz [1420 Hz > 1690 Hz: $F(1,22)=11.24$, $p<.01$; 1420 Hz > 2840 Hz: $F(1,22)=10.05$, $p<.01$; 1690 Hz > 2840 Hz: $F(1,22)=3.89$, $p=.06$]. The main effect of induction type remained significant even at the 8-second silent interval duration, $F(3,66)=12.87$, $p<.001$, such that each induction type produced a significantly larger amount of segregated responses compared to the silent induction [1420 Hz: $F(1,22)=22.62$, $p<.001$; 1690 Hz: $F(1,22)=11.88$, $p<.01$; 2840 Hz, $F(1,22)=10.09$, $p<.01$].

Again, though the main effect of silent interval duration was non-significant, $F(3,66)=1.41$, $p=.25$, there was a significant induction type x silent interval duration interaction, $F(9,198)=3.70$, $p<.01$. This interaction remained significant when the silent induction type was excluded from analysis, $F(6,132)=4.43$, $p<.01$. Follow up analyses revealed a significant effect of silent interval duration within the 1420 Hz induction type $F(3,66)=3.41$, $p<.05$, and showed a significant linear trend, $F(1,22)=6.04$, $p<.05$. The effect of silent interval duration within the 1690 Hz induction type was non-significant, $F(3,66)=1.19$, $p=.32$, and showed neither a linear nor quadratic trend such that similar

21

amounts of segregated responses were reported across all silent interval durations [linear: $F(1,22)=2.25$, $p=.15$; quadratic trend: $F(1,22)=.25$, $p=.62$]. The 2840 Hz induction type revealed a significant effect of the silent interval duration, $F(3,66)=4.20$, $p<.05$, and showed a significant quadratic trend, $F(1,22)=8.94$, $p<.01$. Finally, unlike before, the silent induction type revealed a significant effect of silent interval duration, $F(3,66)=3.09$, $p<.05$, and showed a significant quadratic trend, $F(1,22)=6.79$, $p<.05$.
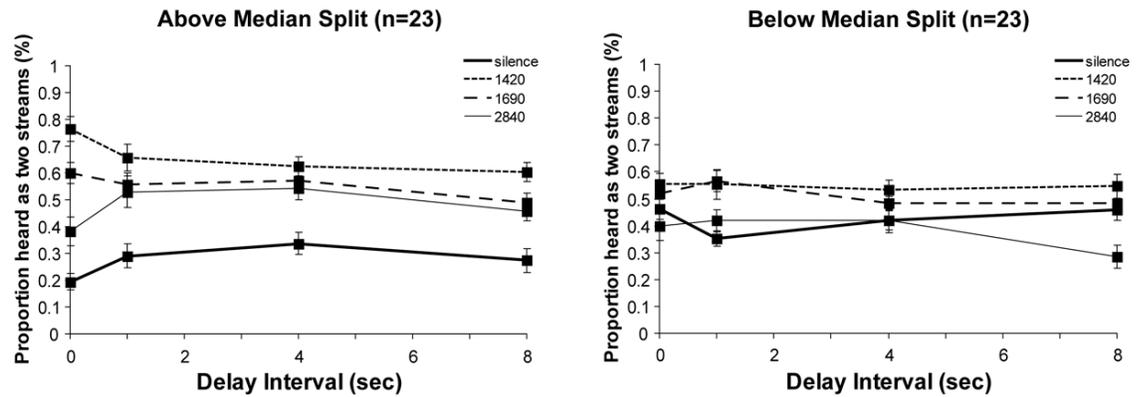

Figure 4: Results from Experiment 1 separately for participants above (left) and below (right) the median split. Error bars based on within-subjects confidence intervals (Cousineau, 2005).

CHAPTER 3

EXPERIMENT 2

Method

Participants

Thirty-five undergraduates (15 females, mean age = 22.0, age rage = 18-34) from

the University of Nevada, Las Vegas participated in this study for course credit. All

participants reported having normal hearing. Informed consent was obtained from all

participants prior to the start of the experiment according to a protocol approved by the

University's Office for the Protection of Research Subjects.

Apparatus

The apparatus was the same as in Experiment 1 with the following exception.

Behavioral responses were made on a RB-830 button box (Cedrus Corporation, San

Predro, CA).

Stimuli

Stimuli were the same as in Experiment 1 with the following exceptions. A 2530

Hz induction type (10 semitones above the high tone of the test phase) was included to

examine whether any pattern of buildup for the 2840 Hz induction type in Experiment 1

was due its octave (i.e., 12 semitones) relation with the high-frequency tone of the test

phase. The rhythmic pattern of each non-silent induction type (1420 Hz, 1690 Hz, 2530

Hz, 2840 Hz) was either isochronous (i.e., same as in Experiment 1) or non-isochronous.

Non-isochronous induction types consisted of triplets of tones with a 10ms ISI and 60ms

SOA between tones within a triplet and 100ms ISI and 270ms SOA between triplets. This

pattern takes on a galloping rhythm. Importantly, the induction length and number of

tones was held constant between pattern types. Given that the number of induction types increased from four in Experiment 1 to nine in the current experiment, only the 0- and 4-second silent interval durations were included. Thus, there were a total of nine induction types and two silent interval durations making a total of 18 trial types.

<div align="center">Procedure</div>

Procedures were the same as in Experiment 1 with the following exceptions. Participants were seated in a single-walled sound-attenuated room (Industrial Acoustic Corp, Bronx, NY). Prior to beginning the main experiment, participants were given 10 practice trials (and additional trials if at first they had a poor understanding of the procedures) selected randomly from all possible trial types. At the end of each trial, participants pressed different buttons on a button box depending on whether they heard the test phase as a trill or two metronomes. During the experiment, trials were presented in four different blocks. A block contained 45 trials with the same silent duration between induction and test phases. Of the 45 trials, five of each nine induction types (silent, 1420 Hz ISO/NON-ISO, 1690 Hz ISO/NON-ISO, 2530 Hz ISO/NON-ISO, 2840 Hz ISO/NON-ISO) were presented. Thus, a total of 180 trials, 10 of each trial type were presented. The order of blocks presented was randomized.

<div align="center">Results</div>

For each participant, the proportion of segregated responses was averaged across all 10 trials of each trial type. We first tested the hypothesis that participants who displayed a weak buildup effect would also show little decay of segregated responses. For each participant, we calculated the buildup effect as the difference in segregated responses between the isochronous 1420 Hz and silent induction types at the 0-second

silent interval duration. Furthermore, the decay of segregated responses was calculated as the slope of segregated responses within the isochronous 1420 Hz induction as a function of the silent interval duration. A correlation analysis revealed that the two measures were significantly negatively correlated, $r(35)=-.58$, $p<.001$, such that participants who displayed a large buildup effect also displayed more decay of segregated responses with increasing silent interval duration. Therefore, a median split was applied between our subjects based on their buildup effect size ($M=.30$). All reported statistical analyses are only on those subjects above the median ($n=18$). Groups of participants above and below the median split did not significantly differ in gender distribution ($\chi^2=.77$, $p=.38$), age ($t=-1.25$, $p=.22$), or handedness ($\chi^2=2.44$, $p=.30$); however, the group below the median split had significantly more years of musical training than the group above the median ($t=2.16$, $p<.05$). Beauvois and Meddis (1997) found that musicians exhibit more decay of segregated responses across time compared to non-musicians. Therefore, differences in response patterns between the two groups in the current study are unlikely due to differences in musical training.

Figure 5 (left) displays the average proportion of segregated responses for the isochronous trial types. To look at the effect of induction type and silent interval duration, as tested in Experiment 1, proportion of segregated responses were entered into a 5 (induction type) x 2 (silent interval duration) repeated measured ANOVA using only the isochronous and silent induction types. The degrees of freedom were adjusted using the Greenhouse-Geisser ε, a conservative estimate of p. All reported probability estimates based on the reduced degrees of freedom; however, no multiple-comparison corrections were applied for post-hoc analyses. There was a significant main effect of induction type,

$F(4,68)=33.81$, $p<.001$, such that all induction types produced significantly more segregated responses than the silent induction [1420 Hz: $F(1,17)=106.49$, $p<.001$; 1690 Hz: $F(1,17)=86.47$, $p<.001$; 2530 Hz: $F(1,17)=4.64$, $p<.05$; 2840 Hz: $F(1,17)=5.38$, $p<.05$]. This effect remained significant at the 4-second silent interval duration, $F(4,68)=14.03$, $p<.001$. Importantly, there was a significant induction type x silent interval duration interaction, $F(4,68)=4.70$, $p<.01$, such that the effect of induction type was smaller at the 4-second silent interval duration. This interaction remained significant even after removing the silent induction from the analysis, $F(3,51)=3.97$, $p<.05$. As in Experiment 1, this interaction could be interpreted by the differential effect of the silent interval duration between induction types. Indeed, only the 1420 Hz induction showed a significant effect of the silent interval duration, $F(1,17)=10.23$, $p<.01$, and this effect was non-significant for all other induction types [Silent: $F(1,17)=3.64$, $p=.07$; 1690 Hz: $F(1,17)=3.14$, $p=.09$; 2530 Hz: $F(1,17)=1.26$, $p=.28$; 2840 Hz: $F(1,17)=.02$, $p=.90$]. Taken together, these results replicated the findings of Experiment 1.

Figure 5 (right) displays the average proportion of segregated responses for the non-isochronous trial types. A second 5 (induction type) x 2 (silent interval duration) repeated measures ANOVA using only the non-isochronous and silent induction types revealed a significant main effect of induction type, $F(4,68)=13.51$, $p<.001$, and remained significant even at the 4-second silent interval duration, $F(4,68)=5.12$, $p<.01$. Follow-up analyses revealed that the 1420, 1690, and 2840 Hz inductions produced more segregated responses than the silent induction [1420 Hz: $F(1,17)=43.02$, $p<.001$, 1690 Hz: $F(1,17)=13.27$, $p<.01$; 2840 Hz: $F(1,17)=7.13$, $p<.05$]. There was no difference in the amount of segregated responses produced by the 2530 Hz and silent inductions,

$F(1,17)=2.92$, $p=.10$. Finally, the induction type x silent interval duration interaction was only marginally significant, $F(4,68)=2.81$, $p=.06$, and none of the induction types showed a significant effect of the silent interval duration [Silence: $F(1,17)=3.64$, $p=.07$; 1420 Hz: $F(1,17)=3.66$, $p=.07$; 1690 Hz: $F(1,17)=.59$, $p=.45$; 2530 Hz: $F(1,17)=1.10$, $p=.31$; 2840 Hz: $F(1,17)=.01$, $p=.92$].
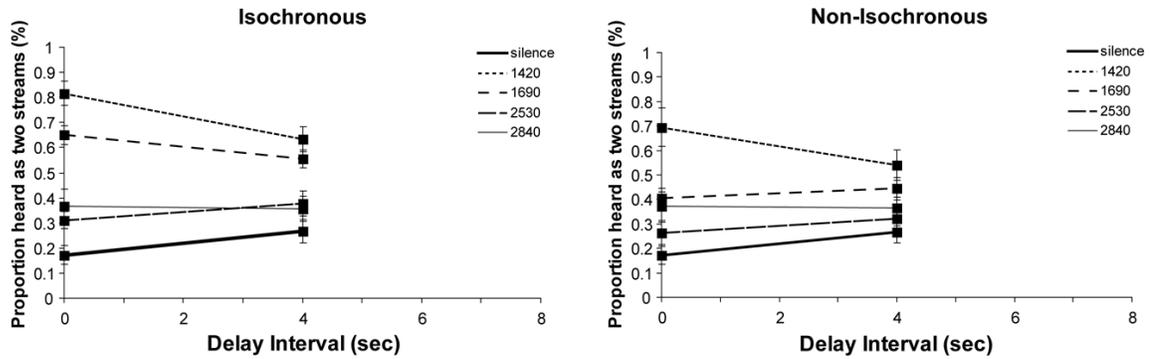

Figure 5: Results from Experiment 2. Error bars based on within-subjects confidence intervals (Cousineau, 2005).

To examine the effect of our pattern manipulation, the proportion of segregated responses were entered into a 2 (pattern type) x 4 (induction type) x 2 (silent interval duration) repeated-measures ANOVA. The silent induction type could not be included in this analysis as it could not change in pattern. There was a significant main effect of pattern type, $F(1,17)=5.08$, $p<.05$, such that the isochronous induction types produced more segregated responses than the non-isochronous induction types. Importantly, there was a significant pattern type x induction type interaction, $F(3,51)=6.60$, $p<.01$, such that the effect of induction type was larger for isochronous than non-isochronous pattern types. Another way to interpret the interaction is that the pattern manipulation had more of a disruptive effect on the two inductions producing the most segregated responses when isochronous (i.e., 1420 and 1690 Hz). Consequently, the amount of segregated

27

responses produced for these two conditions were reduced to a size more similar to the

non-isochronous 2530 and 2840 Hz induction types. Consistent with this interpretation,

both the 1420 and 1690 Hz induction types revealed a significant effect of pattern type

[1420 Hz: $F(1,34)=12.56$, $p<.01$; 1690 Hz: $F(1,34)=9.92$, $p<.01$]; however, neither the

2530 nor 2840 Hz induction types revealed an effect of pattern type [2530 Hz:

$F(1,34)=3.59$, $p=.07$; 2840 Hz: $F(1,34)=.02$, $p=.88$] (Table 1). Furthermore, the effect of

pattern type was not significantly different between the 1420 and 1690 Hz induction

types, $F(1,17)=2.17$, $p=.16$, nor was there a s significant pattern type x induction type x

silent interval duration interaction, $F(1,17)=1.16$, $p=.30$.

Table 1. Effect of pattern irregularity.*

	Delay Interval (sec)	
	0	**4**
1420	0.12 (.05)	0.09 (.04)
1690	0.24 (.05)	0.11 (.05)
2530	0.05 (.03)	0.05 (.04)
2840	0.00 (.04)	-0.01 (.04)

*=values represent difference in segregated responses between isochronous and non-isochronous induction types. Note that a positive score means that more segregated responses were reported for the isochronous than non-isochronous induction type. Values in parentheses represent within-subject confidence intervals (Cousineau, 2005).

Finally, to examine differences between our 2530 and 2840 Hz induction types,

these trial types were entered into a 2 (induction type) x 2 (pattern type) x 2 (silent

interval duration) repeated-measures ANOVA. There was a main effect of induction type,

$F(1,17)=4.63$, $p<.05$, such that the 2840 Hz induction produced more segregated

responses than the 2530 Hz induction. Though the effect of pattern type was non-

significant, $F(1,17)=.32$, $p=.58$, there was a significant induction type x pattern type

interaction, $F(1,17)=4.71$, $p<.05$. This interaction was driven by the large differences

between the two inductions when they were non-isochronous compared to isochronous.

CHAPTER 4

GENERAL DISCUSSION

Previous studies have shown that neural correlates of buildup occur in at least two levels of the auditory pathway that are frequency specific, the cochlear nucleus (Pressnitzer, et al., 2008) and primary auditory cortex (Micheyl, et al., 2005). The results of Experiments 1 and 2 suggest an additional level of analysis beyond primary auditory cortex in an area not finely tuned to frequency. This was evidenced in the increased tendency to hear an "…ABAB…" test sequence as segregated when preceded by an induction sequence, compared to a silent sequence, even when this sequence matched neither of the test-tone frequencies. Therefore, these findings challenge the previous assertion that buildup is frequency specific within a narrow range (Anstis & Saida, 1985) and instead suggest a much broader range up to an octave in width. Furthermore, non-frequency-matching buildup had a slower rate of decay compared to frequency-matching buildup, suggesting that the two involve distinct mechanisms. Consistent with a high-level basis for non-frequency-matching buildup, neurons late in the auditory pathway are not finely tuned to frequency (Konig, et al., 2005; Moore, 2003) and show a longer-lasting suppressive effect of stimulus-driven adaptation (Fitzpatrick, et al., 1999; Harms & Melcher, 2002). Finally, the current study strengthens the argument that auditory stream segregation is a rather complex process that involves multiple levels of analysis within and outside of auditory cortex (Snyder & Alain, 2007). However, since the current study only employed psychophysical methods, future physiology studies are needed to confirm that mechanisms subserving buildup are found beyond primary auditory cortex.

Experiment 1 also demonstrated that the effect of buildup was strikingly persistent and did not fully decay even by 8 seconds. This considerably updates previous descriptions that buildup fully decays by about 4 seconds (Beauvois & Meddis, 1997). Together with the findings of Beauvois and Meddis (1997), this experiment suggests mechanisms subserving buildup are associated with a long auditory store (Cowan, 1984). Furthermore, these findings contend that a previous model of auditory stream segregation, which relies on a "leaky integrator" function, needs to be considerably updated (Beauvois & Meddis, 1996). According to this model, buildup reflects the accumulation of stimulus-driven neural excitation within a leaky integrator. Importantly, excitation should increase exponentially during stimulus onset, but also decrease (decay) exponentially during stimulus offset. Beauvois & Meddis (1996, 1997) proposed a leaky integrator's time constant, which controls the accumulation and decay of excitation, is 4 seconds. However, because the current study has demonstrated that buildup does not fully decay even by 8 seconds, this time constant needs to be substantially increased. Additionally, the decay patterns of non-frequency-matching buildup were inconsistent with an exponential decay of excitation within a leaky integrator. For example, decay of the 1690 Hz condition was best described as linear without exponential decay. Additionally, decay of the 2840 Hz condition was quadratic such that there was an increase in buildup from 0 to 1 second. This pattern could have occurred if stimulus-driven neural excitation continued to accumulate within a leaky integrator during the additional 1 second of stimulus offset (i.e., silent interval duration). However, given that similar patterns were not replicated in Experiment 2, this finding requires replication. In

31

short, models of stream segregation that rely on a "leaky integrator" function are insufficient to account for the observed patterns of buildup.

Experiment 2 showed that rhythmic pattern irregularity between induction and test sequences had a disruptive effect on buildup for those induction sequences that produced a high proportion of segregated responses otherwise (i.e., 1420Hz, 1690Hz). These findings suggest the involvement of brain areas sensitive to rhythmic pattern in auditory stream segregation. This likely includes high-level auditory areas, such as the planum temporale (Chen, Penhune, & Zatorre, 2008), where auditory information is integrated over relatively long temporal windows (Harms & Melcher, 2002; Ligeois-Chauvel, Peretz, Babai, Laguitton, & Chauvel, 1998). Additional areas outside of auditory cortex involved in rhythmic processing include prefrontal cortex, cerebellum, supplementary motor areas, and premotor cortex (Chen, et al., 2008). Furthermore, these findings are consistent with the theory of rhythmic attention in auditory stream segregation (Jones, et al., 1981). Rhythmic attention is assumed to be a time-dependent process that dynamically fluctuates in a periodic fashion between a high and low state (Large & Jones, 1999). According to this theory, rhythmic attention aids listeners in picking up relations between adjacent and nonadjacent events when they are nested in a common rhythm. Therefore, when stimuli have a regular periodic pattern, rhythmic attention detects sounds that do and do not belong to that stream.

Nevertheless, the findings of Experiment 2 are slightly inconsistent with our original hypothesis that rhythmic pattern irregularity would have a larger disruptive effect for non-frequency-matching induction sequences than frequency-matching sequences. Instead, the effect size was not statistically different between the 1420 Hz and 1690 Hz

conditions (although, there was a trend for the effect to be larger for the 1690 Hz condition) and was non-significant for the two non-frequency-matching sequences with the largest frequency deviations (2530 Hz, 2840 Hz). The significant effect of the rhythmic pattern manipulation on the 1420 Hz condition was inconsistent with Rogers and Bregman (1993) who reported rhythmic pattern irregularity had no disruptive effects on buildup during frequency-matching induction sequences. One possibility is that brains areas sensitive to rhythmic pattern are recruited for stream segregation only to the extent that low-level cues (e.g., frequency) sufficiently facilitate stream segregation. This would explain the lack of a significant rhythmic pattern manipulation on the two conditions that showed the least amount of buildup when rhythmically regular (i.e., 2530 Hz ISO, 2840 Hz ISO). Moreover, it may be that Rogers and Bregman (1993) failed to show a significant effect of rhythmic pattern manipulation, because a large portion of their conditions were chosen to promote a coherent perception. Future studies are needed to more closely assess the conditions under which rhythmic regularity plays a strong role in the buildup of auditory stream segregation.

The results of Experiment 2 also suggest that buildup involves mechanisms sensitive to complex relationships between sounds across a broad frequency range. This was evidenced in the significantly larger proportion of segregated responses for the 2840 Hz condition compared to the 2530 Hz condition. Importantly, 2840 Hz is one octave larger than the high-frequency tone of the test. However, it appears that these mechanisms are insensitive to rhythmic pattern. This would explain the larger difference between the two conditions when rhythmically irregular compared to regular. That is, for the isochronous patterns, the 2530 Hz and 2840 Hz conditions may have produced similar

proportion of segregated responses because of the former's rhythmic pattern similarity and the latter's octave relationship with the test. However, for non-isochronous patterns, the rhythmic pattern irregularity disrupted buildup produced during the 2530 Hz condition (note that there was a strong trend for the effect of pattern type to be significant for this condition, $p=.07$) but had little effect on the 2840 Hz condition, which still had its octave relationship with the test. This would, in effect, increase the difference between the two conditions only when they were non-isochronous and drive the significant induction type x pattern type interaction between them. In summary, the 2530 Hz and 2480 Hz conditions may involve distinct mechanisms, one sensitive to rhythmic pattern and another sensitive to complex frequency relationships, respectively.

Finally, it is possible that the segregated responses produced by the 2530 Hz and 2840 Hz conditions do not actually reflect effects of buildup and may instead reflect a more general auditory mechanism. This may explain the non-significant effect of the pattern manipulation and silent interval duration on these conditions. It has been shown that categorization of speech sounds are modulated in a contrastive-manner when preceded by non-speech sounds (Aravamudhan, Lotto, & Hawks, 2008; Holt, 2005; Holt & Lotto, 2002). For example, categorization of a phoneme within a /ga/-/da/ series is highly dependent on the frequency of its third formant (F3), such that a phoneme is more likely to be heard as /ga/ when its F3 is low and /da/ when its F3 is high. However, an ambiguous phoneme within a /ga/-/da/ series (i.e., intermediate F3) is more likely to be categorized as /ga/ when preceded by a high-frequency non-speech sound. In contrast, the same phoneme is more likely to be categorized as /da/ when preceded by a low-frequency non-speech sound. These findings suggest that speech categorization involves a general

auditory mechanism sensitive to speech and non-speech sounds. Furthermore, the outcome of this mechanism appears to enhance the contrast between adjacent and non-adjacent sounds (Aravamudhan, et al., 2008; Holt, 2005). Thus, it is possible that similar mechanisms are involved in auditory stream segregation. Notably, perception of streaming is highly modulated by the previous context in a contrastive-manner such that a sequence with an intermediate Δf is more likely to be heard as segregated when it is preceded by a small-Δf sequence and coherent when preceded by a large-Δf sequence (Snyder, Carter, et al., 2009; Snyder, et al., 2008; Snyder, Holder, et al., 2009; Snyder & Weintraub, 2011). In the current study, it may be that the absence of frequency separations within the induction sequence contrastively enhanced the frequency separations present within the test sequence. The frequency-separation enhancement in this latter case would have facilitated a segregated percept. This hypothesis predicts that a test sequence is more likely to be heard as segregated when preceded by any induction sequence that lacks frequency separations. Future work is needed to test this hypothesis and the involvement of general contrastive mechanisms in auditory stream segregation.

In conclusion, the results of this study suggest that buildup of auditory stream segregation involves distinct mechanisms in high-level auditory areas not finely tuned to frequency and sensitive to rhythmic pattern, in addition to those in the cochlear nucleus (Pressnitzer, et al., 2008) and primary auditory cortex (Micheyl, et al., 2005). Furthermore, this study also suggests that the effects of buildup are longer lasting than previously recognized (Beauvois & Meddis, 1997). Future physiology studies are needed to substantiate claims of this behavioral study. Additional studies are also needed to address the outstanding questions remaining from our results.

Social/Behavioral IRB – Expedited Review
Continuing Review Approved

NOTICE TO ALL RESEARCHERS:
Please be aware that a protocol violation (e.g., failure to submit a modification for <u>any</u> change) of an IRB approved protocol may result in mandatory remedial education, additional audits, re-consenting subjects, researcher probation, suspension of any research protocol at issue, suspension of additional existing research protocols, invalidation of all research conducted under the research protocol at issue, and further appropriate consequences as determined by the IRB and the Institutional Officer.

DATE: December 9, 2011

TO: **Dr. Joel Snyder**, Psychology

FROM: Office of Research Integrity – Human Subjects

RE: Notification of IRB Action by /Charles Rasmussen/ Dr. Charles Rasmussen, Co-Chair
Protocol Title: **Neural Mechanisms of Auditory and Visual Processing in Healthy Adults**
Protocol #: 0710-2518
Expiration Date: December 8, 2012

Continuing review of the protocol named above has been reviewed and approved.

This IRB action will reset your expiration date for this protocol. The protocol is approved for a period of one year from the date of IRB approval. The new expiration date for this protocol is December 8, 2012. If the above-referenced project has not been completed by this date you must request renewal by submitting a Continuing Review Request form 30 days before the expiration date.

PLEASE NOTE:
Upon approval, the research team is responsible for conducting the research as stated in the protocol most recently reviewed and approved by the IRB, which shall include using the most recently submitted Informed Consent/Assent forms and recruitment materials. The official versions of these forms are indicated by footer which contains current approval and expiration dates.

Should there be *any* change to the protocol, it will be necessary to submit a **Modification Form** through ORI - Human Subjects. No changes may be made to the existing protocol until modifications have been approved by the IRB. Modified versions of protocol materials must be used upon review and approval. Unanticipated problems, deviations to protocols, and adverse events must be reported to the ORI – HS within 10 days of occurrence.

If you have questions or require any assistance, please contact the Office of Research Integrity - Human Subjects at IRB@unlv.edu or call 895-2794.

BIBLIOGRAPHY

Anstis, S., & Saida, S. (1985). Adaptation to auditory streaming of frequency-modulated tones. *Journal of Experimental Psychology: Human Perception and Performance, 11*(3), 257-271.

Aravamudhan, R., Lotto, A. J., & Hawks, J. W. (2008). Perceptual context effects of speech and nonspeech sounds: The role of auditory categories. *Journal of the Acoustical Society of America, 124*(3), 1695-1703.

Beauvois, M. W., & Meddis, R. (1996). Computer simulation of auditory stream segregation in alternating-tone sequences. *Journal of the Acoustical Society of America, 99*(4), 2270-2280.

Beauvois, M. W., & Meddis, R. (1997). Time decay of auditory stream biasing. *Perception & Psychophysics, 59*(1), 81-86.

Bee, M. A., & Klump, G. M. (2004). Primitive auditory stream segregation: A neurophysiological study in the songbird forebrain. *Journal of Neurophysiology, 92*(2), 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*(3), 197-214.

Bee, M. A., & Micheyl, C. (2008). The cocktail party problem: What is it? How can it be solved? And why should animal behaviorists study it? *Journal of Comparative Psychology, 122*(3), 235-251.

Bendixen, A., Denham, S. L., Gyimesi, K., & Winkler, I. (2010). Regular patterns stabilize auditory streams. *Journal of the Acoustical Society of America, 128*(6), 3658-3666.

Bregman, A. S. (1978). Auditory streaming is cumulative. *Journal of Experimental Psychology: Human Perception and Performance, 4*(3), 380-387.

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

Bregman, A. S., & Campbell, J. (1971). Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of Experimental Psychology, 89*(2), 244-249.

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*(1), 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*(11), 1393-1402.

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Listening to musical rhythms recruits motor regions of the brain. *Cerebral Cortex, 18*(12), 2844-2854.

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America, 25*(5), 975-979.

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology, 1*(1), 42-45.

Cowan, N. (1984). On short and long auditory stores. *Psychological Bulletin, 96*(2), 341-370.

Cusack, R. (2005). The intraparietal sulcus and perceptual organization. *Journal of Cognitive Neuroscience, 17*(4), 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*(4), 643-656.

Devergie, A., Grimault, N., Tillmann, B., & Berthommier, F. (2010). Effect of rhythmic attention on the segregation of interleaved melodies. *Journal of the Acoustical Society of America, 128*(1), EL1-EL7.

Dykstra, A. R., Halgren, E., Thesen, T., Carlson, C. E., Doyle, W., Madsen, J. R., et al. (2011). Widespread brain areas engaged during a classical auditory streaming task revealed by intracranial EEG. *Frontiers in Human Neuroscience, 5*, 74.

Fay, R. R., & Popper, A. N. (2000). Evolution of hearing in vertebrates: The inner ears and processing. *Hearing Research, 149*(1-2), 1-10.

Fishman, Y. I., Arezzo, J. C., & Steinschneider, M. (2004). Auditory stream segregation in monkey auditory cortex: Effects of frequency separation, presentation rate, and tone duration. *Journal of the Acoustical Society of America, 116*(3), 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*(1-2), 167-187.

Fitzpatrick, D. C., Kuwada, S., Kim, D. O., Parham, K., & Batra, R. (1999). Responses of

neurons to click-pairs as simulated echoes: Auditory nerve to auditory cortex.

Journal of the Acoustical Society of America, 106(6), 3460-3472.

Formisano, E., Kim, D. S., Di Salle, F., van de Moortele, P. F., Ugurbil, K., & Goebel, R.

(2003). Mirror-symmetric tonotopic maps in human primary auditory cortex.

Neuron, 40(4), 859-869.

Harms, M. P., & Melcher, J. R. (2002). Sound repetition rate in the human auditory

pathway: Representations in the waveshape and amplitude of fMRI activation.

Journal of Neurophysiology, 88(3), 1433-1450.

Hartmann, W. M., & Johnson, D. (1991). Stream segregation and peripheral channeling.

Music Perception, 9(2), 155-184.

Hill, K. T., Bishop, C. W., Yadav, D., & Miller, L. M. (2011). Pattern of BOLD signal in

auditory cortex relates acoustic response to perceptual streaming. *BMC

Neuroscience, 12*, 85.

Holt, L. L. (2005). Temporally nonadjacent nonlinguistic sounds affect speech

categorization. *Psychological Science, 16*(4), 305-312.

Holt, L. L., & Lotto, A. J. (2002). Behavioral examinations of the level of auditory

processing of speech context effects. *Hearing Research, 167*(1-2), 156-169.

Jones, M. R., Kidd, G., & Wetzel, R. (1981). Evidence for rhythmic attention. *Journal of

Experimental Psychology: Human Perception and Performance, 7*(5), 1059-1073.

Kondo, H. M., & Kashino, M. (2009). Involvement of the thalamocortical loop in the

spontaneous switching of percepts in auditory streaming. *Journal of

Neuroscience, 29*(40), 12695-12701.

Konig, R., Heil, P., Budinger, E., & Scheich, H. (2005). *The auditory cortex: A snythesis of human and animal research*. Mahwah, NJ: Lawrence Erlbaum Associates.

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review, 106*(1), 119-159.

Ligeois-Chauvel, C., Peretz, I., Babai, M., Laguitton, V., & Chauvel, P. (1998). Contribution of different cortical areas in the temporal lobes to music processing. *Brain, 121*, 1853-1867.

McCabe, S. L., & Denham, M. J. (1997). A model of auditory streaming. *Journal of the Acoustical Society of America, 101*(3), 1611-1621.

Micheyl, C., Carlyon, R. P., Gutschalk, A., Melcher, J. R., Oxenham, A. J., Rauschecker, J. P., et al. (2007). The role of auditory cortex in the formation of auditory streams. *Hearing Research, 229*(1-2), 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*(1), 139-148.

Moore, B. C. J. (2003). *An introduction to the psychology of hearing*. San Diego, CA: Academic Press.

Moore, B. C. J., & Gockel, H. (2002). Factors influencing sequential stream segregation. *Acta Acustica United with Acustica, 88*(3), 320-333.

Nityananda, V., & Bee, M. A. (2011). Finding your mate at a cocktail party: Frequency separation promotes auditory stream segregation of concurrent voices in multi-species frog choruses. *Plos One, 6*(6).

Okamoto, H., Stracke, H., Bermudez, P., & Pantev, C. (2011). Sound processing

hierarchy within human auditory cortex. *Journal of Cognitive Neuroscience,*

23(8), 1855-1863.

Petkov, C. I., Kang, X. J., Alho, K., Bertrand, O., Yund, E. W., & Woods, D. L. (2004).

Attentional modulation of human auditory cortex. *Nature Neuroscience, 7*(6),

658-663.

Pressnitzer, D., Sayles, M., Micheyl, C., & Winter, I. M. (2008). Perceptual organization

of sound begins in the auditory periphery. *Current Biology, 18*(15), 1124-1128.

Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex:

Nonhuman primates illuminate human speech processing. *Nature Neuroscience,*

12(6), 718-724.

Rogers, W. L., & Bregman, A. S. (1993). An experimental evaluation of three theories of

auditory stream segregation. *Perception & Psychophysics, 53*(2), 179-189.

Schadwinkel, S., & Gutschalk, A. (2011). Transient bold activity locked to perceptual

reversals of auditory streaming in human auditory cortex and inferior colliculus.

Journal of Neurophysiology, 105(5), 1977-1983.

Shamma, S. A., & Micheyl, C. (2010). Behind the scenes of auditory perception. *Current*

Opinion in Neurobiology, 20(3), 361-366.

Snyder, J. S., & Alain, C. (2007). Toward a neurophysiological theory of auditory stream

segregation. *Psychological Bulletin, 133*(5), 780-799.

Snyder, J. S., Alain, C., & Picton, T. W. (2006). Effects of attention on neuroelectric

correlates of auditory stream segregation. *Journal of Cognitive Neuroscience,*

18(1), 1-13.

Snyder, J. S., Carter, O. L., Hannon, E. E., & Alain, C. (2009). Adaptation reveals multiple levels of representation in auditory stream segregation. *Journal of Experimental Psychology: Human Perception and Performance, 35*(4), 1232-1244.

Snyder, J. S., Carter, O. L., Lee, S. K., Hannon, E. E., & Alain, C. (2008). Effects of context on auditory stream segregation. *Journal of Experimental Psychology: Human Perception and Performance, 34*(4), 1007-1016.

Snyder, J. S., Holder, W. T., Weintraub, D. M., Carter, O. L., & Alain, C. (2009). Effects of prior stimulus and prior perception on neural correlates of auditory stream segregation. *Psychophysiology, 46*(6), 1208-1215.

Snyder, J. S., & Weintraub, D. M. (2011). Pattern specificity in the effect of prior Δf on auditory stream segregation. *Journal of Experimental Psychology: Human Perception and Performance, 37*(5), 1649-1656.

Teki, S., Chait, M., Kumar, S., von Kriegstein, K., & Griffiths, T. D. (2011). Brain bases for auditory stimulus-driven figure-ground segregation. *Journal of Neuroscience, 31*(1), 164-171.

Thompson, S. K., Carlyon, R. P., & Cusack, R. (2011). An objective measurement of the build-up of auditory streaming and of its modulation by attention. *Journal of Experimental Psychology: Human Perception and Performance, 37*(4), 1253-1262.

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

VITA

Graduate College
University of Nevada, Las Vegas

David M. Weintraub

Degrees:
Bachelor of Liberal Arts, Psychology, 2009
University of Nevada, Las Vegas

Special Honors and Awards:
ARO Student Travel Award, 2011
Graduate & Professional Student Association Travel Grant, 2010, 2011, 2012

Publications:
Snyder, J.S., Gregg, M.K., Weintraub, D.M., & Alain, C. (2012). Attention, awareness, and the perception of auditory scenes. *Frontiers in Psychology*, *3*, 15.

Snyder, J.S. & Weintraub, D.M. (2011). Pattern specificity in the effect of prior Δf on auditory stream segregation. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 1649-1656.

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

Thesis Title: Frequency-specificity and pattern-specificity of the buildup of auditory stream segregation.

Thesis Examination Committee:
Chairperson, Joel Snyder, Ph. D.
Committee Member, Erin Hannon, Ph. D.
Committee Member, Jefferson Kinney, Ph. D.
Graduate Faculty Representative, Pierre Lienard, Ph. D.