

**HUMAN SENSITIVITY TO MINIMAL AUDITORY
TEMPORAL IRREGULARITIES**

by

Laura S. Ovens

Submitted in partial fulfillment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
July 2014

© Copyright by Laura S. Ovens, 2014

TABLE OF CONTENTS

LIST OF FIGURES	iii
ABSTRACT	iv
LIST OF ABBREVIATIONS USED	v
ACKNOWLEDGEMENTS	vi
CHAPTER 1: INTRODUCTION	1
1.1 OVERVIEW OF AUDITORY PERCEPTION	1
1.2 AUDITORY TEMPORAL PROCESSING	5
1.3 THE AGING AUDITORY SYSTEM	7
1.4 AUDITORY NEUROPATHY	12
1.5 THE JITTER PARADIGM	14
1.6 CURRENT STUDY	18
CHAPTER 2: METHODS	21
2.1 SUBJECTS	21
2.2 STIMULI AND APPARATUS	21
2.3 PROCEDURE	24
2.4 ANALYSIS	25
CHAPTER 3: RESULTS	26
3.1 OUTLIER ANALYSIS	26
3.2 EFFECTS OF STIMULUS CONDITION AND ICI	29
3.3 BROADBAND CONDITIONS	31
3.4 HIGH PASS CONDITIONS	33
3.5 COMPARISON TO UNIFORM JITTER PARADIGM	35
CHAPTER 4: DISCUSSION	38
4.1 INTERPRETATION OF RESULTS	38
4.2 METHODOLOGICAL LIMITATIONS	44
4.3 IMPLICATIONS AND DIRECTIONS FOR FUTURE RESEARCH	46
4.4 CONCLUSION	50
REFERENCES	52
APPENDIX: TABLES	55

LIST OF FIGURES

Figure 1.1	Jitter thresholds (expressed as % of base ICI) as a function of ICI for five subjects, from Phillips et al. (2012).....	17
Figure 2.1	Visual depictions of standard and jittered click train stimuli.	23
Figure 3.1	Normal (a) and abnormal (b) staircase plots.....	27
Figure 3.2	Jitter thresholds (measured as % of base ICI) as a function of ICI, presented on a log scale.....	30
Figure 3.3	Jitter thresholds (measured as % of base ICI) as a function of ICI, presented on a log scale, for (a) jitter type 1, (b) jitter type 2, (c) jitter type 3, and (d) jitter type 4, with a line for each stimulus condition.....	31
Figure 3.4	Jitter thresholds (measured as % of base ICI), as a function of ICI, presented on a log scale, for (a) broadband stimuli, and (b) high pass stimuli, with a line for each jitter type.....	33
Figure 3.5	Jitter thresholds (measured as % of base ICI) as a function of ICI, presented on a log scale, for (a) broadband, and (b) high pass stimuli, with data from Phillips et al. (2012)	37

ABSTRACT

The objective of the present study was to characterize the temporal processing abilities of healthy adults using a modified jitter paradigm in which only a single click in 25-click train stimuli had been displaced. It was hypothesized that thresholds would be higher than those observed for wholly jittered stimuli. It was also expected that the data would support the ‘dual mechanism hypothesis’, i.e., that the auditory perceptual system employs different strategies across different ranges of inter-click intervals (ICIs) to detect irregularities. Using 48 stimulus conditions (four jitter types, two masking conditions, six ICIs), it was hypothesized that performance would be superior on tasks with two aberrant ICIs compared to one. Fifteen subjects were tested, and repeated measures ANOVAs were performed to assess the significance of the perceptual effects observed. The data were supportive of all stated hypotheses. They provide a greater understanding of how the auditory system detects temporal irregularities.

LIST OF ABBREVIATIONS USED

ABR	Auditory brainstem response
AN	Auditory neuropathy
ANOVA	Analysis of variance
BB	Broadband
dB	Decibel
ERP	Event-related potential
HP	High pass
ICI	Inter-click interval
MMN	Mismatch negativity
N1	Negative-going evoked auditory response with a latency near 100 ms
P1	Positive-going evoked auditory response with a latency near 200 ms
PI-PB	Performance intensity-phonetically balanced
VOT	Voice-onset time

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. Phillips, for all of the encouragement, reassurance, and patience over the past three years. Thanks are also due to Susan Hall for her ongoing assistance with programming and technical questions. I would also like to express my gratitude to Dr. Semba, Dr. Crowder, and Dr. Aiken for serving on my committee. And finally to my family, friends, and to Jonny for the unending support.

CHAPTER 1: INTRODUCTION

1.1 Overview Of Auditory Perception

As with all sensory systems, the purpose of the auditory system is to translate information from the outside world into a useful form represented in the central nervous system and that is ultimately elaborated in conscious perception. In the auditory system, this processing begins with the transduction of mechanical information (variations in air pressure) to electrical information (action potentials in the cochlear nerves).

This process begins with the peripheral auditory system, encompassing the outer, middle, and inner ear. Sound waves first arrive at the outer ear and are funneled through the ear canal, which serves as a broadly tuned filter, making us more sensitive to frequencies in the 1000-6000 Hz range (Plack, 2005). The changes in air pressure result in displacement of the *tympanic membrane* (or ear drum), which in turn causes movement of the *ossicles*, the three bones of the middle ear. By acting as a lever system, and by concentrating the force received at the tympanic membrane onto the much smaller surface area of the oval window, the middle ear bones serve to increase the strength of the vibrations being transferred to the inner ear. This is essential, as it is difficult for vibrations to travel from an air medium to a fluid one (as in the cochlea).

The cochlea is the inner ear structure in which transduction occurs; it is responsible for encoding the various properties of sound stimuli. It is a coiled tube that is widest at the base (adjacent to the oval window) and narrowest at the apex. The tube is divided into three fluid-filled compartments. The *scala media*, bordered on either side by *Reissner's membrane* and the *basilar membrane*, is where the structures of interest are located. The basilar membrane is narrow and stiff at the base of the cochlea, and wide

and flexible at the apex (Plack, 2005). The *organ of Corti* rests on the basilar membrane, and within it are embedded hair cells, rows of which extend the length of the cochlea. Hair cells are so named for the stereocilia projecting out of their superior surfaces. There is a single row of inner hair cells, as well as three rows of outer hair cells whose longest row of stereocilia are embedded in the *tectorial membrane* above them.

As vibrations enter the cochlea, the basilar membrane serves to separate the frequency components of the sound. The membrane moves up and down in response to vibrations moving through the fluid. The narrow and stiff end moves in response to high frequency sounds (fast vibrations), while the wide and loose end moves in response to low frequency sounds (slow vibrations). The properties of the outer hair cells cause them to contract, pulling up on the basilar membrane and resulting in enhanced movement (Plack, 2005). The movements of the basilar membrane through the fluid of the sub-tectorial space cause deflections of the stereocilia of the inner hair cells, which causes the inner hair cells to respond with alterations of their membrane potentials.

The selective responsiveness of the basilar membrane is an essential component to our sense of hearing; it acts to mechanically separate the frequency components of a sound. Each section of the membrane has a 'characteristic frequency' to which it responds most strongly. When complex sounds enter the cochlea, the basilar membrane effectively performs a spectral decomposition, with sections closer to the base responding to the higher frequency components, and sections closer to the apex responding to lower frequency components of the sound. Although each place on the membrane has a characteristic frequency, in reality they each respond to a range of frequencies, with the strongest response reserved for the characteristic frequency (Plack, 2005). The

narrowness of the tuning of the basilar membrane is aided by the action of the outer hair cells, which amplify movement at the location of strongest response.

As mentioned above, deflections of the stereocilia on the inner hair cells result in their depolarization. Inner hair cells form synapses with auditory nerve cells, which carry information out of the cochlea. This is where transduction occurs; the translation of mechanical information into electrical activity that can be interpreted by the brain. The inner hair cells release a neurotransmitter when stereocilia are deflected in one direction, and stop release when they are deflected in the other direction. Stronger vibrations result in larger deflections and more neurotransmitter being released. Auditory nerve cells each collect information from only one inner hair cell, and are therefore only carrying information about the movement of the basilar membrane at one location (Plack, 2005). This configuration ensures that auditory nerve cells have the same characteristic frequency as the inner hair cell from which they receive information. This tonotopic organization is preserved throughout the auditory system to the primary auditory cortex, similarly to retinotopic organization in the visual system.

As the level (loudness) of a sound is increased, the rate of action potentials in the auditory nerve responding to that sound increases. The tonotopic organization of the cochlea and auditory nerves results in a place code; the auditory system can infer the qualities of a sound stimulus by comparing the firing rates of auditory nerve fibers with different characteristic frequencies. This is not the only strategy for extracting the frequency information of sound stimuli, however. Within a certain range of frequencies, the inner hair cells respond to sounds with an action called ‘phase-locking’. This occurs because the release of neurotransmitter is triggered only when the stereocilia are

deflected in one direction, which consistently occurs during the same phase of a sound. The resulting action potentials will therefore always occur with intervals equal to integer multiples of the period of the sound (Plack, 2005). For example, when a 50 Hz sound is played, the basilar membrane moves up and down at a rate of 50 times per second, or with 20 ms periods. The action potentials triggered by these movements will occur every 20 ms, or at multiples thereof (40 ms, 60 ms, etc). Even at frequencies which approach or surpass the maximum rate at which a given neuron can fire action potentials (around 200 Hz), information from several neurons each firing in response to the stimulus can be combined to extract frequency information. The usefulness of this mechanism is limited to frequencies below approximately 5000 Hz, above which auditory nerve fibers do not consistently fire at the same phase of a stimulus repeatedly (Plack, 2005). Within that range, however, this time-based coding plays an important role in pitch perception and sound localization.

From the cochlea, the auditory nerve carries information to a group of nuclei in the brainstem. These nuclei contain neurons with widely varying response properties, and some of the nuclei receive bilateral input. The integration of input from both ears is essential for processing of certain sound properties, such as sound source location. Localization of sound sources is important for guiding behaviour and orienting attention. There are two main mechanisms by which information from the two ears is compared in order to locate a sound source, one involving time cues, and one involving level cues (Plack, 2005). Interaural time difference refers to the difference in arrival time of the sound wave between the two ears, or the difference in phase of the sound wave when it arrives at the two ears. Interaural time differences are most useful for localizing low

frequency sounds, whose wavelengths are long relative to the size of the head. Interaural level cues occur because of the ‘sound shadow’ that is cast by the head, when a sound arrives from a lateral direction. Sounds arriving from, for example, the right side, will be reduced in energy when they reach the left ear. This cue is most useful for localizing high frequency sounds, which do not diffract around the head as easily as low frequency sounds. These types of binaural comparisons can be accomplished in the auditory brainstem.

The last stop in the auditory brainstem is the inferior colliculus, which projects to the medial geniculate body of the thalamus. From there, auditory information is sent to the primary auditory cortex in the temporal lobe, for higher order processing.

1.2 Auditory Temporal Processing

Behavioural temporal acuity depends on the fidelity with which both the temporal fine structure and the amplitude envelopes of sound stimuli are represented when they are encoded in the auditory system. Slow or irregular responses in the auditory system will lead to a compromised internal representation of sound, and therefore affect perceptions that rely on good temporal acuity. The healthy human auditory system is incredibly temporally sensitive, and this sensitivity relies on several levels of synchrony in neural firing. The first is the ability of auditory nerve cells to synchronize their firing to events in a stimulus (in other words, phase locking). The firing of a nerve cell should have a mean period identical to that of the frequency it is encoding; departures from this will result in a distorted pitch percept and lost information. Auditory nerve cells responding to the same stimulus must also be able to synchronize their activity to each other; input to

the auditory brainstem that is delayed with respect to other inputs will compromise the neural representation of the stimulus.

Temporal processing has long been a topic of interest to researchers, and many methods of measuring acuity have been devised. The variety of temporal processes used by the auditory system is large, as illustrated even by one of the more analytic or “reduced” tasks, commonly called the gap detection paradigm. A gap detection task measures the listener’s ability to differentiate between two stimuli, one of which is a continuous noise, the other of which contains a brief gap. When the two noises delimiting the gap are of the same frequency, the task is referred to as ‘within-channel’ gap detection (Phillips, Taylor, Hall, Carr & Mossop, 1997). When the two noises are spectrally dissimilar, it has been referred to as a ‘between-channel’ gap detection task. Performance on these two tasks differs greatly; the smallest gap detectable in within-channel tasks has often been found to be less than 5-6 ms for some listeners, while between-channel gap thresholds are invariably much higher, and increase with the disparity of the frequencies bordering the gap (Phillips et al., 1997). A proposed explanation for this is that the within-channel task amounts to a very simple ‘discontinuity detection’ within one ‘perceptual channel’, which can be performed at the level of the auditory nerve. When the two noises are of different frequencies, they stimulate different areas of the basilar membrane, and therefore different perceptual channels. The detection of a gap must then be performed centrally, as a comparison between the offset of activity in one channel, and the onset in another. This type of comparison has been referred to as a ‘relative timing operation’ (Phillips et al., 1997). The more complex relative timing operation relies on temporally coordinated neural

activity at more levels than the simple discontinuity detection. Firing must be synchronized both within channels and between them, in order to accurately perceive a gap between widely disparate frequencies.

A comprehensive understanding of human temporal processing ability is important because impairments in temporal resolution are implicated in numerous pathologies associated with the auditory system. Insight into the functioning of the healthy, normal-hearing auditory system could shed light on the mechanisms and manifestations of disorders that impair auditory perception. A brief review of two of these hearing related disorders (age-related hearing loss and auditory neuropathy) will provide a better understanding of how perception is impaired, and how temporal processing abilities are specifically implicated as the likely cause of these impairments.

1.3 The Aging Auditory System

Age-related hearing difficulty (or presbycusis) is extremely common, affecting 50% of individuals over the age of 75 (Bance, 2007). In the elderly, hearing loss has historically been attributed to two sources; a decrease in absolute hearing sensitivity (especially at high frequencies), and to cognitive decline. Decreases in absolute hearing sensitivity are generally understood to be caused by the degeneration of peripheral auditory structures, including cochlear cells (sensory neurons as well as supporting cells), as well as outer and middle ear problems (reviewed in detail by Chisolm, Willott & Lister, 2003). Additionally, it is commonly understood that the aging process involves a general slowing of cognitive and mental processes, which could impact perceptual processes indiscriminately (Salthouse, 2000).

Although these factors are undoubtedly involved in some (if not most) cases of age-related hearing difficulty, many elderly individuals experience deficits beyond what would be expected from these factors alone. Although teasing apart the contributions of different factors can be difficult, there is growing evidence that a decline in central auditory processes may contribute to age-related deficits, specifically a decline in temporal processing ability (see Fitzgibbons & Gordon-Salant, 1996, for a review of early evidence).

Lister, Besing and Koehnke (2002) employed a gap detection paradigm to measure temporal processing abilities in listeners of three age groups, all of whom had normal hearing thresholds. Their finding that older listeners unaffected by hearing sensitivity loss, performed significantly worse, especially on ‘between-channel’ gap detection tasks, suggests a central origin for these deficits. Snell (1997) similarly attributed higher gap detection thresholds in older listeners with normal hearing sensitivity to a deficit in temporal resolution. Stuart and Phillips (1996) employed a “word recognition in noise” paradigm to investigate the temporal acuity of young normal hearing, older normal hearing, and older presbycusis listeners. The procedure entails testing participants on their word recognition performance in the presence of both continuous and interrupted noise maskers, while varying the signal to noise ratio. Participants all performed better in the interrupted noise condition, as their auditory systems make use of the brief ‘glimpses’ of the speech stimuli during interruptions in the masker. The finding of interest is that the superiority of performance in the interrupted noise condition declined with age for both normal hearing and presbycusis subjects, while this decline was not apparent in the continuous noise condition. The reduced ability

to make use of the brief silent periods implicates a decline in temporal resolution with age.

In addition to behavioural and psychoacoustic measures of auditory temporal processing, evidence from electrophysiological studies has implicated impaired temporal acuity in aging. Since psychoacoustic tasks rely in part on the attention and motivation of the subject, electrophysiological tests can be useful for obtaining more direct measures of auditory processing. One of these tests is called the ‘mismatch negativity’ (MMN) test. It measures perception at a pre-attentive level of a deviant event embedded within a series of standard stimuli (Bertoli, Smurzynski & Probst, 2002). It is a component of the auditory event-related potential (ERP), and can be observed in response to a change in any one of a number of stimulus properties, such as duration, frequency, or intensity. Pekkonen (2000) reviewed several studies, and found evidence of a reduced MMN amplitude in elderly listeners in response to deviations in duration, but not to deviations in frequency, suggesting a deficit in temporal resolution specifically. Bertoli et al. (2002) performed a study with older, normal hearing listeners in order to investigate the hypothesis of a specific temporal processing deficit in the elderly. The authors measured gap detection thresholds behaviourally, as well as measuring the MMN in response to deviant gap durations, and concluded that temporal resolution is indeed reduced.

A common experience for elderly listeners is difficulty in speech comprehension, which is exacerbated in noisy environments (Pichora-Fuller, Schneider, MacDonald, Pass & Brown, 2007). Accurate perception of speech sounds requires discrimination of brief acoustic cues, and there is an abundance of research implicating temporal processing deficits in difficulties with speech comprehension. Some studies have used time-

compressed speech (speech stimuli that have been digitally sped up) in order to shed light on these issues. Versfeld and Dreschler (2002) related the speech reception threshold (a measure of speech intelligibility in noise) to a time-compression threshold (a measure for temporal acuity), demonstrating their correlation and the worsened performance of elderly listeners, even after correcting for hearing loss. Gordon-Salant and Fitzgibbons (2001) used time-compressed speech stimuli to identify whether age-related difficulties in speech comprehension could be attributed to a general limitation in speed of information processing, or a specific limitation in processing transient acoustic cues (such as consonant bursts and transitions). Their finding that performance was poorer when consonant cues were selectively compressed suggests that a specific auditory temporal processing deficit may be responsible for these difficulties. Tremblay, Piskosz and Souza (2003) used both behavioural and electrophysiological measures to examine the perception and neural coding of voice-onset-time (VOT). VOT is a temporal cue in speech that distinguishes voiced from unvoiced consonants, for example, the /b/ sound from the /p/ sound. Older listeners with and without hearing loss exhibited difficulty in discriminating 10 ms VOT contrasts compared to younger listeners, as well as prolonged N1 and P2 latencies in response to longer VOT durations. The N1 is thought to reflect synchronous neural activity in response to acoustic change (Tremblay et al., 2003).

This abundance of evidence suggests that some of the perceptual deficits experienced by older adults can be attributed to degradation of the temporal response properties of the central auditory system. Further, there is evidence to support that these age-related changes take the form of disrupted neural synchrony, or temporal ‘jitter’ in neural firing (that is, the ability of auditory nerves to synchronize to the phase of a sound

stimulus). Miranda and Pichora-Fuller (2002) demonstrated how neural dyssynchrony may contribute to degraded perception by introducing temporal jitter into speech stimuli and determining the effect on ‘performance intensity-phonetically balanced’ (PI-PB) rollover. PI-PB rollover is the presence of a speech comprehension deficit at high presentation levels, and is often seen in elderly populations. By simulating neural jitter in young, normal-hearing listeners (achieved by temporally “jittering” the stimuli), they were able to recreate a PI-PB rollover effect that subjects did not exhibit in response to normal, intact speech stimuli. Pichora-Fuller et al. (2007) extended this work, seeking to specifically demonstrate that simulated neural dyssynchrony in young listeners would produce speech comprehension deficits similar to those seen in older listeners. Speech stimuli were again manipulated in order to reduce periodicity. In describing how temporal jitter was imposed on the stimulus, the authors first explain how normal phase-locked responses typically operate. A perfectly phase-locked response to a 500 Hz tone would have a mean period of 2 ms, though the actual inter-spike intervals would likely be distributed around this mean, with a relatively small standard deviation. By artificially increasing the standard deviation of this distribution, the authors can effectively simulate a loss of synchrony in healthy auditory systems. The authors note that manipulating the temporal fine structure of the stimulus could have the unwanted effect of introducing spectral splatter, or broadening the pattern of activation on the basilar membrane in response to the stimulus. To control for this frequency effect, and isolate the effects of temporal jitter, they included a control condition with similar levels of spectral splatter but no temporal distortion. Their results confirmed that simulated neural dyssynchrony resulted in reduced speech comprehension in noise that was not observed with spectral

distortion. The deficits exhibited by their subjects were similar to deficits observed in older adults with normal hearing thresholds.

The evidence that aging impacts neural synchrony and temporal processing abilities is compelling, and the perceptual deficits experienced by some elderly individuals, especially in speech comprehension, can be severe. The jitter hypothesis of aging is a useful theory to account for these effects, and evidence in support of it is mounting.

1.4 Auditory Neuropathy

Auditory neuropathy (AN) is another disorder related to timing in the auditory system. Identified and named by Starr, Picton, Sininger, Hood and Berlin (1996), the disorder is characterized by absent or abnormal auditory brainstem response (ABR), while normal otoacoustic emissions and pure tone thresholds are preserved. Otoacoustic emissions, which can be detected with a microphone placed in the ear canal, suggest normal outer hair cell function. Auditory brainstem potentials are a reflection of the electrical activity, measured by scalp electrodes, of cranial nerve VIII and the auditory brainstem pathway (Bess & Humes, 1995). A normal ABR relies on synchronized neural activity in the auditory brainstem, and an absent or abnormal ABR can be indicative of disrupted auditory nerve activity (Miranda & Pichora-Fuller, 2002).

The deficits displayed by patients with auditory neuropathy are associated with perceptions that rely on neural timing (Zeng, Kong, Michalewski & Starr, 2005). The most common perceptual deficit is poor speech comprehension, disproportionate to what would be expected based on pure tone thresholds. A patient described by Starr et al., (1996), who had progressive hearing difficulties beginning at age 15, could recognize

speech sounds but could not identify words, and had to rely on lip-reading to communicate. Zeng et al., (2005) extensively documented the perceptual deficits of a group of 21 AN patients across a broad age range. Eight of these patients had previously been found to have temporal processing deficits that correlated in degree with the severity of their speech comprehension deficits (Zeng, Oba, Garde, Sininger & Starr, 1999). It was found that AN subjects generally had normal intensity perception, exhibiting no difference compared to controls in intensity discrimination tasks, or in sound localization tasks using interaural level difference cues. In contrast, they exhibited poor performance relative to controls on temporal processing tasks, exhibiting elevated gap detection thresholds, inability to localize sound using interaural time difference cues, and poor performance on both forward and backward masking tasks, which indicate an inability to distinguish sounds occurring successively (Zeng et al., 2005). Sound localization ability was assessed by progressively increasing the phase difference between the sounds arriving at each ear, which for normal controls resulted in the perception of the sound source moving to one direction, while AN patients did not perceive a change in the position of the sound source. The authors make note of the difference between AN and other hearing disorders, which normally involve cochlear damage, and typically result in impairments in intensity-related perception.

The deficits observed in AN patients are all in line with what would be expected with disrupted synchrony in auditory nerve firing, which is consistent with an absent or abnormal ABR. Despite this, it was suggested that a reduction of overall activity of the auditory nerve, and resulting reduced neural input, could be an alternative cause of the AN pathology (Starr et al., 1996). This question appears to have been resolved by

Cowper-Smith, Dingle, Guo, Burkard and Phillips (2010), who reviewed evidence from the carboplatin-treated chinchilla to investigate the effects of reduced cochlear output. Carboplatin treatment results in a loss of inner hair cells and possibly of auditory nerve cells, but does not result in impaired temporal responses in the auditory nerve or inferior colliculus, suggesting that reduced cochlear output alone is not the cause of the disordered ABR observed in auditory neuropathy patients.

The locus of pathology in AN could exist at one of several locations along the auditory pathway; the inner hair cells, the auditory nerve, the synapse between them, or a combination of the above (Starr et al., 1996). Desynchronized activity could be attributable to demyelination of the auditory nerve fibers. The AN patient population is not uniform, with possible etiologies including genetic, infectious (such as measles and mumps), or other causes (Starr et al., 1996). Eight of ten patients described by Starr et al. (1996) exhibited evidence of peripheral neuropathy, which the authors suggest could have impacted their auditory nerve function. The finding that cochlear implants are effective in remediating hearing difficulties in some but not all patients is indicative of variable pathologies (Cowper-Smith et al., 2010).

Taken together, the evidence from the auditory neuropathy research makes clear that impaired neural synchrony results in severe perceptual deficits, somewhat similar to those observed in elderly individuals.

1.5 The Jitter Paradigm

The foregoing raises the question of just how sensitive normal listeners are to temporal jitter that exists in a stimulus. Until recently, little research had been done into sensitivity to departures from temporal regularity. Phillips, Dingle, Hall and Jang (2012)

decided to investigate this aspect of auditory temporal processing, and developed a jitter paradigm in order to do so. The jitter paradigm functions similarly to other auditory psychophysics tasks, such as gap detection; listeners are subject to a two-alternative forced decision task, embedded within an adaptive staircase. The stimuli consisted of two trains of 25 click noises each, one in which the clicks are regularly spaced in time, and one in which 'jitter' has been imposed on the stimulus, resulting in irregular intervals between all of the clicks. The amount of jitter was progressively reduced every time the listener accurately indicated which was the target stimulus, until their performance reached stable levels to stimuli with little jitter. A threshold value was calculated for each trial, measured as a percentage of the base inter-click interval (ICI), similar to Weber fractions.

This procedure was carried out at seven different base inter-click intervals: 5, 10, 20, 40, 60, 80, and 100 ms. At short ICIs, it is impossible to distinguish the individual clicks within the stimulus, and the stimuli are perceived as a brief 'buzz' noise, while the jittered version has a 'rougher' quality to it. The procedure was also carried out using two different stimulus conditions. The first was a 'naturalistic' experiment, in which the sound stimuli used for the clicks were 'broadband' (containing a broad range of frequencies). It was speculated that spectral pitch cues might have contributed to performance on the task for ICIs <40ms (after Krumbholz, Patterson & Pressnitzer, 2000). To address this, a second experiment was done using a high-pass filter (so the click noises were of a higher frequency), presented against a low frequency noise masker. The noise masker is included because click stimuli of high frequency may still generate low frequency spectral elements that can be the basis of spectral pitch percepts. To

understand this mechanism, it helps to remember that click noises played with 5 ms inter-click intervals occur 200 times per second, in other words, with a frequency of 200 Hz. The auditory system may perceive a low frequency sound despite the fact that the sound energy within the clicks are designed to contain only high frequencies. This can come about in either or both of two ways. First, the pitch percept may reflect purely temporal processing, based on central analysis of the intervals between the clicks. Second, the transduction of the stimulus by the earphones (or by the cochlea) can introduce spectral elements with a frequency equal to the repetition rate of the clicks. Thus, by using high-pass clicks in the presence of a low-frequency masker, the task becomes a purely temporal one.

Five adult listeners completed these procedures. Based on the results of these experiments, Phillips et al. (2012) theorized that performance was dictated by differing mechanisms across different ranges of inter-click intervals. This is evident in the threshold sensitivities to broadband stimuli; performance was better on short ICIs (5-20 ms), and poorest performance was at 40 ms. Generally, performance improved again on longer ICIs, resulting in an inverted U-shaped function across ICIs (figure 1.1). In contrast, for high-pass, masked stimuli, performance was poorest at short ICIs, with thresholds much higher than those for the broadband stimuli. Performance improved slightly at longer ICIs, with no inverted U-shape (figure 1.1). This suggests that subjects were indeed experiencing a pitch percept based on spectral distortions when listening to broadband stimuli with short ICIs. This would be accomplished by comparing the salience of the pitch percept between the standard and target stimulus. The filter and masker successfully prevented the generation of a pitch percept, resulting in poorer

performance on the task. Absolute jitter thresholds were typically less than ten percent of the base ICI. These findings align with those of Krumbholz et al. (2000), who determined that pitch percepts based on temporal information were generated only for stimuli with intervals less than approximately 33 ms.

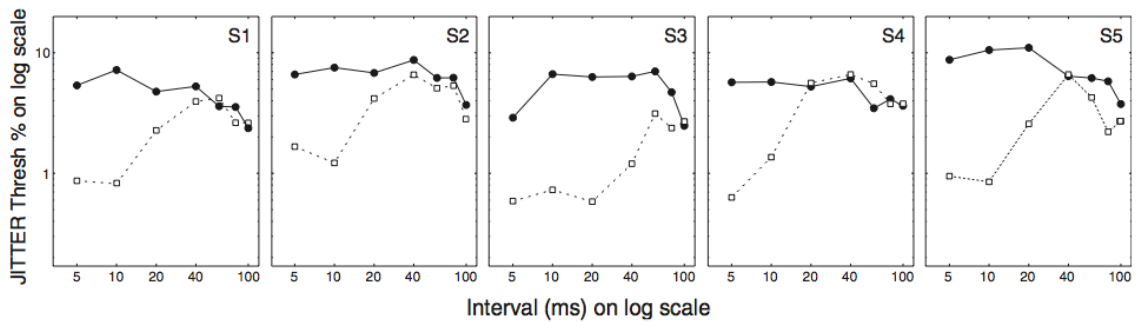


Figure 1.1. Jitter thresholds (expressed as % of base ICI) as a function of ICI for five subjects. Results are for both broadband clicks (dashed lines), and high pass, masked clicks (solid lines). Reproduced with permission of the senior author from Phillips et al. (2012).

At longer ICIs, different perceptual strategies must be employed, as a pitch percept becomes very weak. While the short ICI stimuli have a single emergent quality upon which listeners can base their judgment, the long ICI stimuli do not. The clicks are far enough apart in time that each inter-click interval must be successively and consciously compared, in what is called a relative timing operation, in order to identify the target stimulus. This type of operation is much more cognitively demanding. It was hypothesized that at the longest ICIs tested, the stimuli may have generated a rhythm percept that aided listeners in their judgment. Following this reasoning, it is likely that performance was poorest in the 40-60 ms range because neither pitch nor rhythm cues were available, resulting in the inverted U-shaped function observed for broadband stimuli.

Accurate performance on jitter tasks relies heavily on neural synchrony. Consider the neural mechanisms responsible for perception of the short ICI regular click trains; auditory nerves will fire with a mean interval equal to the inter-click interval of the stimulus, resulting in a temporally-based pitch percept. Reduced neural synchrony will impact the salience of this percept, and thus the differentiation between the standard and jittered trains is made more difficult. At longer ICIs, no pitch percept is generated, but the precision with which auditory nerve cells synchronize to the stimulus events is still important in determining performance. Temporally imprecise firing would result in a standard stimulus being perceived as slightly jittered, making the comparison between the two stimuli more difficult.

1.6 Current Study

The current study is designed to build on the work of Phillips et al. (2012). In order to further characterize human sensitivity to departures from temporal regularity, the jitter paradigm was again employed, in a modified form. Rather than imposing jitter on *all* clicks of the target stimulus, a single click in the train was displaced, resulting in minimal irregularity in each train. It was hypothesized that threshold sensitivity to these temporal irregularities would be higher than was found in the 2012 study, as listeners have only one opportunity to detect the jitter, as opposed to many.

It was hypothesized that the performance function for broadband (henceforth referred to as BB) stimuli across inter-click intervals would be similar to that found in 2012, namely an inverted U-shaped function, reflecting the dual mechanisms purportedly in use across different ranges of ICIs. Specifically, it was expected that a spectral pitch cue would result in superior performance (i.e., lower thresholds) at short ICIs.

Following the reasoning of the 2012 study, both naturalistic broadband stimuli and high-pass with masker stimuli were used (henceforth referred to as HP), to further explore the dual mechanism hypothesis (these are referred to as ‘stimulus conditions’). It was hypothesized that we would observe a greater influence of stimulus condition on performance at short inter-click intervals (5-20 ms) than at long inter-click intervals. In other words, the high pass filter and masking noise would negatively impact performance at short ICIs more strongly due to the removal of spectral pitch cues that may aid in decision-making for BB stimuli. This effect was not expected to be as strong at longer inter-click intervals, where pitch cues are unlikely to contribute to performance.

The stimuli of the current study were further manipulated in one of four ways, resulting in four jitter types (referred to as types 1-4). Types one and two each had one click in the train displaced, with all other clicks unaffected. The result of this manipulation is two aberrant inter-click intervals, one before the click and one after. Type one had the displaced click moved ‘backwards’ (see figure 2.1b for a visual depiction). The terminology employed here can be ambiguous; we are using ‘backwards’ to mean ‘to the left’ (in reference to a visual depiction of the click train where the first click is on the left, see figure 2.1), or ‘earlier in time’. Type two had the displaced click moved forward (to the right), or played later in time (figure 2.1c). These two manipulations should result in spectrally identical stimuli, but it is an empirical question as to whether listeners respond to the two manipulations equivalently. Types three and four differed from one and two in that the subsequent clicks in the train following the displaced click were ‘block-shifted’, so that the inter-click interval following the displaced click was of normal duration (figures 2.1d and 2.1e are useful for comprehension of the stimuli). In

type three, the displaced click was moved backwards (played earlier), resulting in one aberrant inter-click interval that was shorter in duration than all the others in the train. In type four, the displaced click was moved forwards (played later), resulting in one longer than normal inter-click interval in the train.

It was further hypothesized that the extra irregular inter-click interval in jitter types one and two compared to types three and four would positively impact performance, resulting in lower jitter thresholds for those types.

CHAPTER 2: METHODS

Methods were adapted from procedures developed by Phillips et al. (2012), and were approved by the Dalhousie University Health Sciences Human Research Human Research Ethics Board (protocol #2013-3082).

2.1 Subjects

Fifteen subjects (ten females, five males) were recruited through word of mouth. The subjects were predominantly university-aged (21-25 years), although two older adults were included as well (aged 48 and 58 years). None of them reported a history of hearing loss at frequencies less than 4 kHz.

2.2 Stimuli And Apparatus

Stimuli were trains of 25, 45.4 μ s-duration clicks, presented diotically (i.e. to both ears simultaneously). Two forms of stimuli were used; unfiltered (broadband) clicks, and high-pass clicks (>2.0 kHz), presented against a digitally generated low frequency noise masker. The noise masker used was $1/f$ “pink noise” <2.0 kHz. Filtering of the clicks and the noise masker was accomplished through digital elimination of unwanted Fourier components. The position of the displaced click in the target click train was drawn randomly on each trial from a distribution of positions 10-18 (inclusive). This ensured that participants could not anticipate the position of the jittered click.

Procedures were carried out at six different inter-click intervals (ICIs): 5 ms, 10 ms, 20 ms, 40 ms, 80 ms, and 160 ms. The ICI duration determines the overall duration of the stimulus. The initial level of jitter (click displacement) was set at an arbitrarily high level to help ensure that it would be perceptible to all subjects.

Four different jitter manipulations were performed (henceforth referred to as jitter types 1, 2, 3, and 4). Figure 2.1 depicts all four stimulus manipulations. In type 1, the jittered click was displaced ‘backward’ (that is, it was played earlier than normal). This resulted in two aberrant inter-click intervals; one short one before the displaced click, and one long one after it. Type 2 was similar, although the click was displaced ‘forward’, or played later than normal (still resulting in two aberrant inter-click intervals). Type 3 contained a click again displaced backward, although subsequent clicks were block shifted backwards as well, resulting in only one aberrant inter-click interval (the one before the displaced click, see figure 2.1d). This ICI was shorter than normal. Type 4 contained a click displaced forward (played later), again with subsequent clicks block shifted, resulting in only one aberrant inter-click interval before the target click (this time longer than normal).

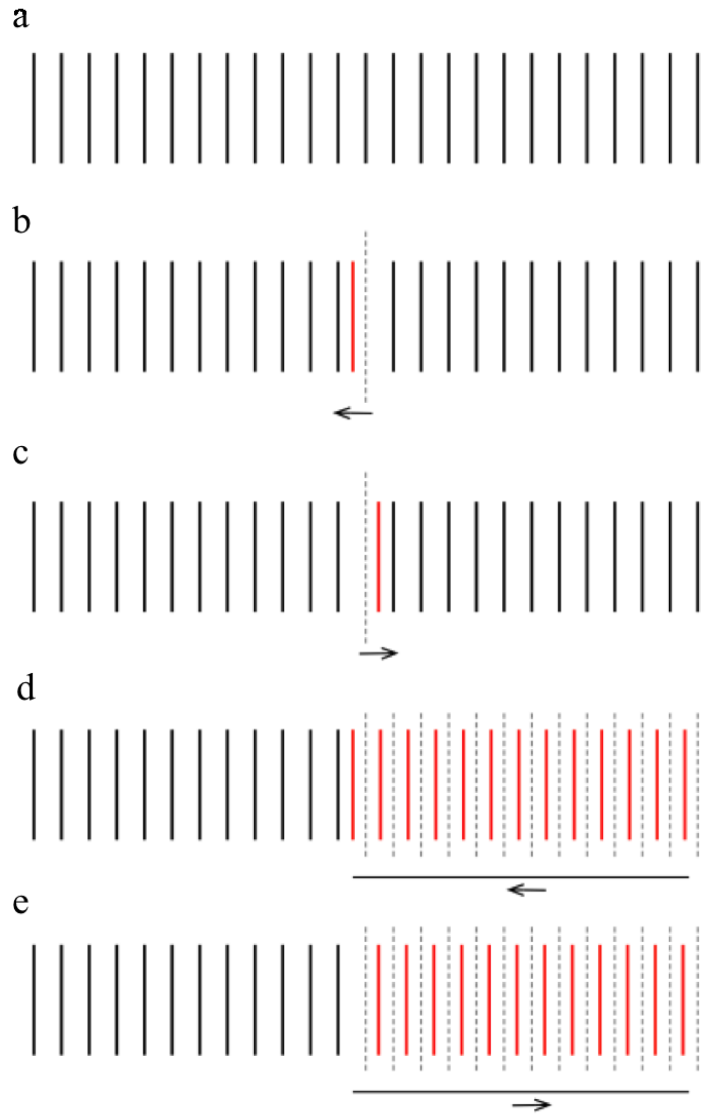


Figure 2.1. Visual depictions of standard and jittered click train stimuli. (a) standard stimulus with regular intervals between clicks (b) type 1 jitter; single click displaced backwards, resulting in two aberrant ICIs, (c) type 2 jitter; single click displaced forwards, (d) type 3 jitter; block-shift backwards of all clicks subsequent to the jitter, resulting in a single aberrant ICI (short in duration), and (e) type 4 jitter; block-shift forwards, resulting in a single (longer in duration) aberrant ICI.

2.3 Procedure

Fifteen subjects were recruited through word of mouth, and each took part in all conditions. They were tested individually over several one-hour sessions in an Eckel sound-attenuating booth. They wore Sennheiser HD590 headphones, seated in front of an Apple iMac computer monitor and keyboard, and mouse. Total testing time was approximately eight hours per subject.

The objective of each task was to measure the smallest departure from regularity (jitter) that a subject could perceive. Thresholds of detectable jitter were determined through the use of a two-alternative forced choice decision task embedded within an adaptive staircase. Subjects were played a reference train and a target train, in random order, separated by 1000 ms of silence. For HP stimuli, the masking noise began 350 ms before the click train, and terminated 150 ms after it. Broadband stimuli were presented at 63.5 dB SPL, high-pass clicks at 61.3 dB, and the masker at 71.3 dB (A-weighted: Extech model 407750 digital sound level meter). They were asked to indicate which of the two trains contained a jittered click by pressing the correct button with the mouse. Visual feedback was provided indicating correct and incorrect answers. Trials were self-paced. Up to the first incorrect response, the program automatically reduced the amount of jitter for every correct response. Following the first incorrect response, the adaptive staircase was a two-down, one-up design, with a step factor of 1.2. This type of procedure tracks 70.7% correct performance (Levitt, 1971). After eight reversals in direction of the adaptive step, the staircase was automatically terminated, and the threshold value was calculated as a geometric average the values of the last six adaptive step reversals. The threshold was expressed as a percentage of the base inter-click interval (in ms).

Each condition was tested three times in each participant, and an arithmetic mean of all three values was calculated for use in analysis. Conditions were performed in random order, chosen by the participants, who filled out a chart of all condition types in order to keep track of which tasks had been completed and to avoid repetition.

2.4 Analysis

Data collected were the arithmetic means of the three thresholds on each of 48 conditions (2 stimulus conditions (i.e. BB vs. HP) x 4 jitter types x 6 ICIs), except where outliers were excluded (see Results). A series of two-way repeated measures analyses of variance (ANOVAs) were performed in order to investigate the hypothesized effects, and post hoc tests were conducted to further characterize the conditions under which each effect was significant.

Data were also compared to those collected by Phillips et al. (2012), where applicable, to investigate the effect of minimal temporal irregularities vs. multiple irregularities. Jitter types 1 and 2 were combined for the purposes of that comparison.

CHAPTER 3: RESULTS

3.1 Outlier Analysis

Before any statistical analyses were performed, the data were examined for outliers. Visual inspection of the data revealed many outlying thresholds; out of 720 total stimulus conditions (2 stimulus conditions x 4 jitter types x 6 inter-click intervals x 15 subjects), 41 were found to have outlying values. Visual examination of the staircase plots for each of these outlying values confirmed that performance was irregular and that the thresholds were likely not valid. Examples of a typical staircase plot, as well as a plot depicting random or irregular performance, are provided in figure 3.1. A typical adaptive staircase is an exponential decay function, reflecting progressively reduced jitter, which eventually plateaus around the threshold sensitivity of the subject. Irregular performance is easy to distinguish from regular performance using the staircase plot because the staircase does not have the asymptotic form. The outlying values were then removed, and the remaining two values for that condition were used in calculations. In two instances, two out of three values were irregular, and only one data point was entered for those conditions. In many instances, more than three values were recorded. This was likely due to participants losing track of which conditions had been completed, and accidentally repeating a condition. Additionally, participants were encouraged to re-do a particular task if they felt they had not performed well due to lack of focus. In cases where more than three values were available, the last three (non-outlying) values were used for the purpose of analyses.

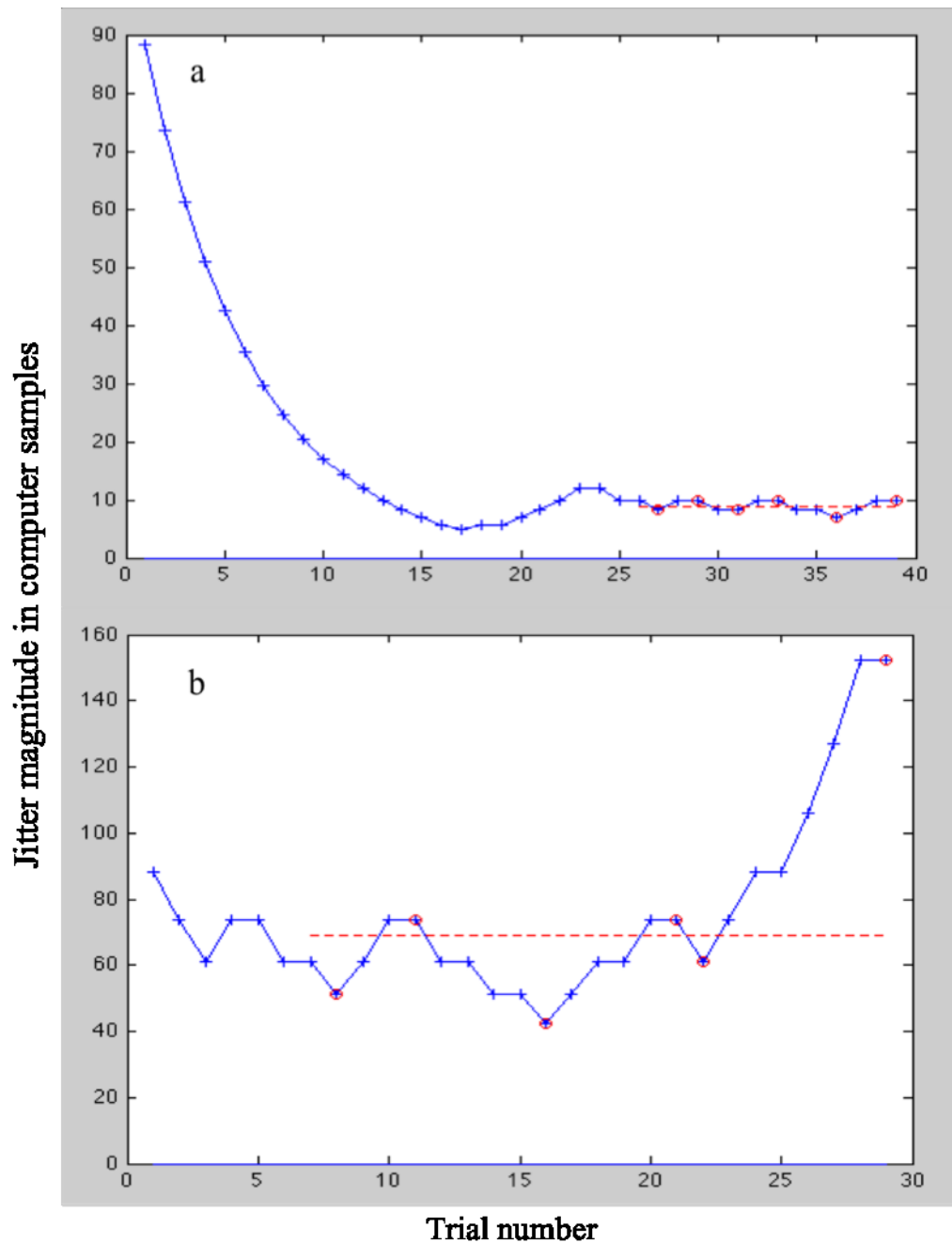


Figure 3.1. Normal (a) and abnormal (b) staircase plots. Normal plots exhibit an exponential decay function, reflecting correct responses at high levels of jitter, which level off around the threshold of the listener. Crosses indicate responses, red circles are reversals in direction of the staircase, and the red line indicates the mean of the last six reversals in direction, or the calculated threshold value. Values on the y-axis are measured in computer samples, where one sample = 23 μ s; in other words, there are about 44 samples / ms.

There are several possible reasons for the presence of outliers. The jitter task is cognitively demanding, and can be mentally exhausting over long periods, which impacts performance. Participants were encouraged to end a session when they felt they could no longer focus, even if a full hour had not elapsed. Despite this, subjects occasionally reported that their attention had drifted during a session, and that certain thresholds may not be valid.

In the case of one condition, the outliers were likely due to difficulty in perceiving the jitter in the stimulus even at the initial (high) level of jitter. This was the type 3, 5 ms, HP stimulus condition. It is perhaps unsurprising that type 3 was the most difficult condition, due to the nature of the stimulus manipulation resulting in only one irregular inter-click interval of briefer than normal duration. It was also hypothesized that, due to the removal of spectral pitch cues in the masked conditions, performance would suffer at short ICIs. For these reasons, it is not surprising that there would be subjects who would have difficulty perceiving the jitter in the stimulus, and that some many not be able to perceive it at all. Three separate subjects had no meaningful values for this condition (in other words, performance was random and the results were not reflective of true threshold sensitivities). For the purposes of statistical analyses, the mean threshold value of the other twelve subjects was imputed as values for these subjects. Unfortunately, this has the effect of artificially reducing the variance of this group. As described below, steps were taken to mitigate the consequences of this decision and ensure the validity of our results.

3.2 Effects Of Stimulus Condition And ICI

Table 1 provides group means and standard deviation values for all 48 conditions. Figure 3.2 depicts threshold values (measured as % of base ICI) as a function of ICI for both stimulus conditions, with all four jitter types collapsed together. The graph illustrates the superior performance of subjects on BB stimuli, with a larger disparity between stimulus conditions at low ICIs (<40 ms). Figure 3.3 (a-d) demonstrates the consistency of this pattern across all four jitter types. We performed a series of ANOVAs in order to investigate the effects discussed in our hypotheses. First, jitter types one to four were collapsed together and a two-way repeated measures ANOVA was performed in order to investigate the effects of ICI and stimulus condition (broadband vs. high pass and masked). There was a significant main effect of stimulus condition, $F(1,14) = 136.77$, $p < 0.001$. The overall group means, standard error, and 95% confidence intervals for the stimulus conditions (i.e. with all ICIs collapsed) are presented in table 2. There was a significant main effect of ICI, $F(2.91,40.73) = 5.99$, $p = 0.002$. The degrees of freedom were corrected using the Greenhouse-Geisser estimates of sphericity ($\epsilon = 0.58$), as Mauchly's test indicated that the assumption of sphericity had been violated ($\chi^2(14) = 26.87$, $p = 0.02$). Group means, standard error, and 95% confidence intervals for the six ICI groups (5, 10, 20, 40, 80, and 160 ms) are provided in table 3. The effect of ICI will be explored in more detail below, within each stimulus condition independently. There was also a significant interaction between stimulus condition and ICI, $F(5,70) = 35.31$, $p < 0.001$. Inspection of figure 3.2 is helpful in interpreting these results. Review of the graph suggests that the source of the interaction effect is that the effect of stimulus condition is much larger at short ICIs (5 and 10 ms) than at longer ICIs. A series of

paired-sample t-tests was performed to investigate this, with stimulus conditions being compared individually for each ICI (e.g. 5 ms broadband compared to 5 ms high pass/masked). The results of these tests are provided in table 4. All six of these comparisons were significant, but the higher t-values in the low ICI comparisons (see 5 and 10 ms comparisons in table 4 and figure 3.2) suggest a larger effect size at these ICIs. Interestingly, three of the five subjects tested by Phillips et al. (2012) showed the same effect, i.e., higher thresholds for HP stimuli across *all* ICIs, but with the greatest disparity between the two conditions at short ICIs.

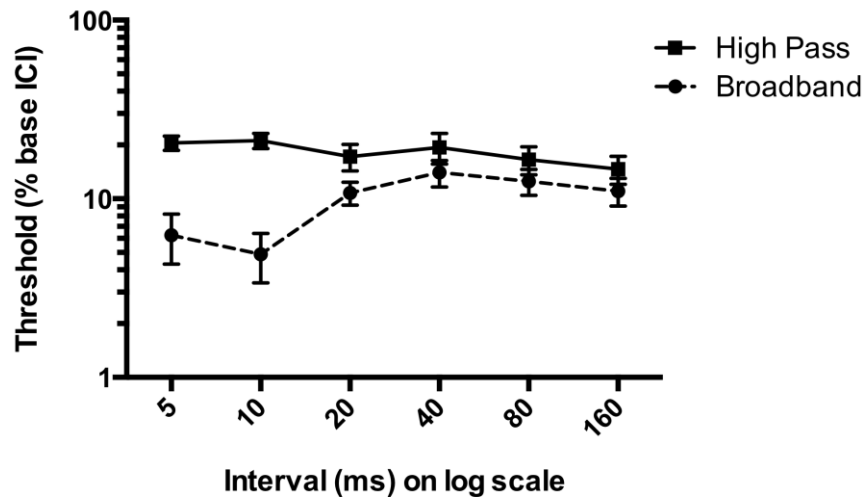


Figure 3.2. Jitter thresholds (measured as % of base ICI) as a function of ICI, presented on a log scale. The graph represents performance on all four jitter types, with a line for each stimulus condition. There was a significant main effect of stimulus condition ($p < 0.001$), of ICI ($p = 0.002$), as well as a significant interaction ($p < 0.001$). Paired samples t-tests indicated significant differences between stimulus conditions at each ICI (see table 4).

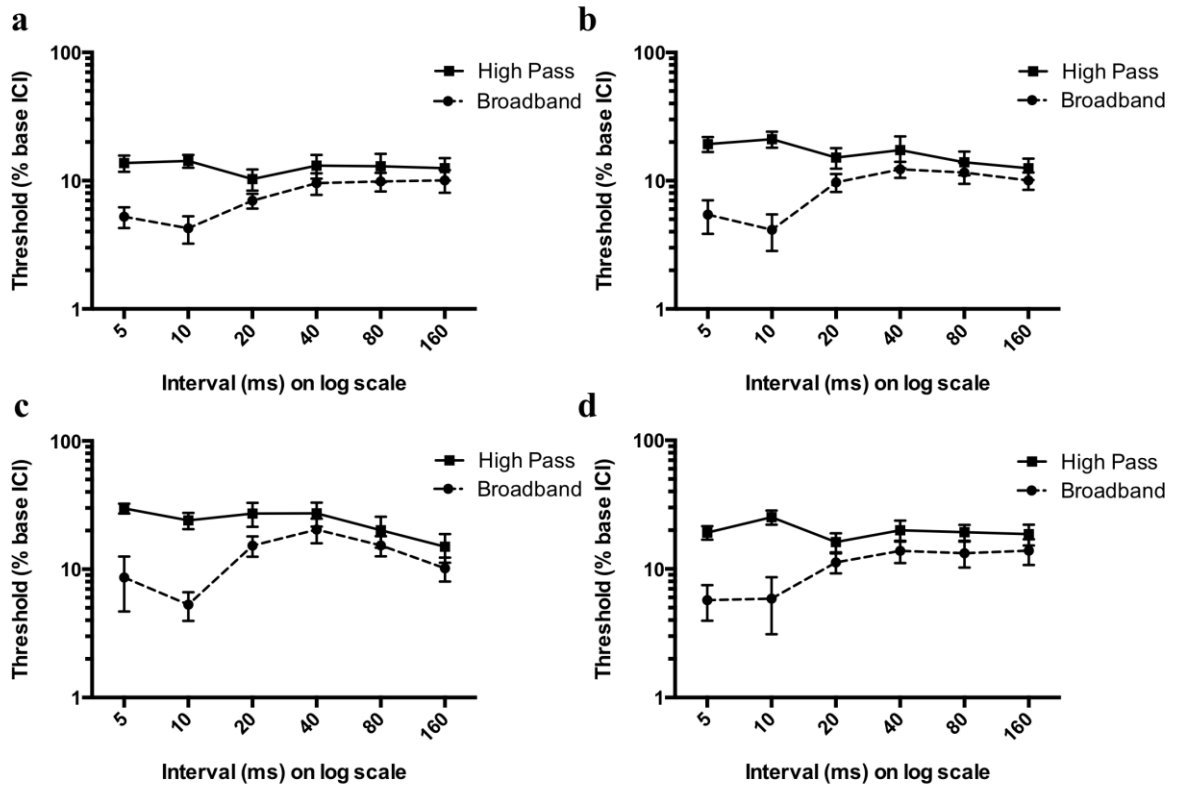


Figure 3.3. Jitter thresholds (measured as % of base ICI) as a function of ICI, presented on a log scale, for (a) jitter type 1, (b) jitter type 2, (c) jitter type 3, and (d) jitter type 4, with a line for each stimulus condition.

3.3 Broadband conditions

Following this, two 2-way repeated measures ANOVAs were performed, one for each stimulus condition, in order to investigate the effects of jitter type and ICI. The results of the ANOVA for broadband stimuli will be reported first. Figure 3.4a depicts the jitter thresholds for all four types as a function of ICI, for broadband stimuli. The graph demonstrates consistent differences between jitter types across the ICIs, although the differences appear to be greater at some ICIs than others. There was a significant main effect of jitter type, $F(3,42) = 42.431$, $p < 0.001$. Group means, standard error, and 95% confidence intervals for the four jitter types are presented in table 5, and show that type 3

generated the highest threshold values (corresponding to poorest performance), followed by types 4, then 2, then 1. Pairwise comparisons were subsequently performed, with a Sidak correction for multiple comparisons applied, to further explore this effect. All comparisons were found to be significant, that is, each jitter type was significantly different from every other type, and these results are presented in table 6. The main effect of ICI is reported again here, as the results reported above include both stimulus conditions in the analysis of the effect. There was a significant main effect of ICI, $F(2.37, 33.12) = 28.01$, $p < 0.001$. The Greenhouse-Geisser correction ($\epsilon = 0.47$) was again applied due to a violation of the assumption of sphericity according to Mauchly's test ($\chi^2(14) = 33.85$, $p = 0.003$). Group means, standard error and 95% confidence intervals are provided in table 7. Pairwise comparisons were subsequently conducted to determine which ICIs differed significantly from each other, using a Sidak correction for multiple comparisons. These results, presented in table 8, suggest that thresholds for 5 and 10 ms ICIs were each significantly different (and lower) than thresholds for all higher (≥ 20 ms) ICIs, none of which were significantly different from each other (aside from a significant difference between ICIs of 40 and 160 ms, $p = 0.034$). Figure 3.4a depicts these effects clearly. There was a significant interaction between jitter type and ICI, $F(15, 210) = 8.54$, $p < 0.001$. Figure 3.4a appears to show that the effect of jitter type may have been larger at some ICIs (such as 40 ms, where there are greater apparent differences between the thresholds for jitter types), than others (such as 160 ms, where there is greater overlap between thresholds for different jitter types). This is likely the source of the significant interaction.

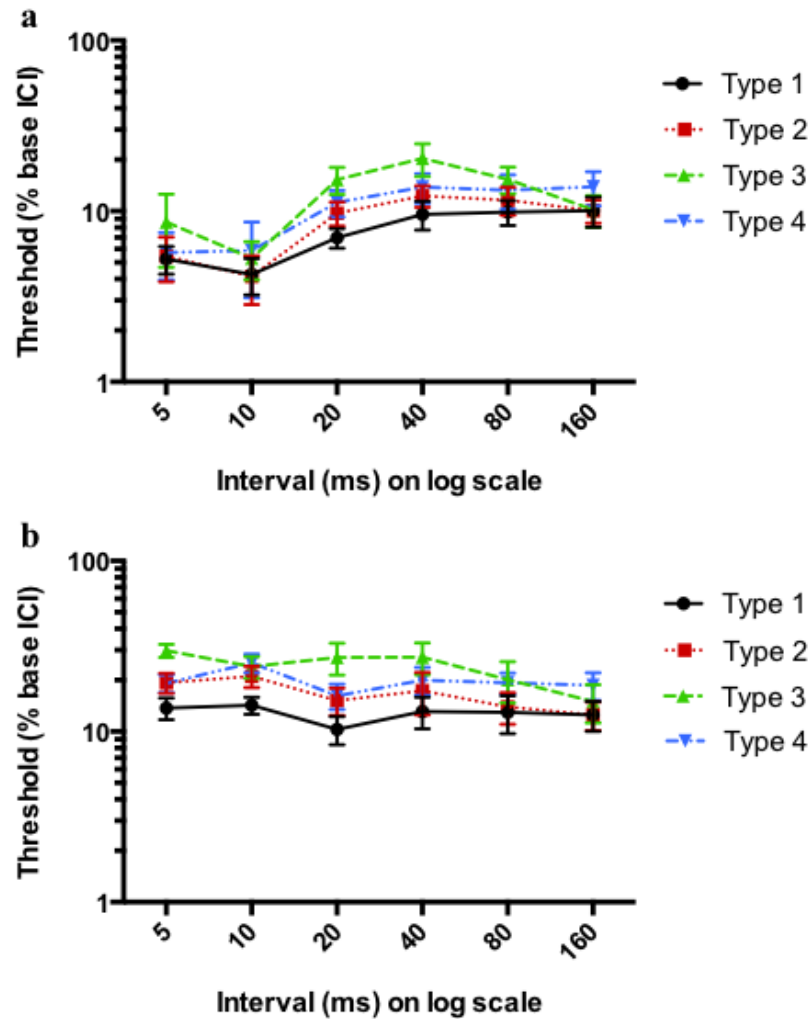


Figure 3.4: Jitter thresholds (measured as % of base ICI), as a function of ICI, presented on a log scale, for (a) broadband stimuli, and (b) high pass stimuli, with a line for each jitter type. For broadband stimuli, there was a significant main effect of jitter type ($p < 0.001$), of ICI ($p < 0.001$), as well as a significant interaction ($p < 0.001$). For high pass stimuli, there was a significant main effect of jitter type ($p < 0.001$), of ICI ($p < 0.001$), as well as a significant interaction ($p < 0.001$).

3.4 High Pass Conditions

Finally, the results from the two-way repeated measures ANOVA performed on data from the high pass/masked stimulus condition are as follows. Refer to figure 3.4b, which depicts jitter thresholds as a function of ICI for all four jitter types. The graph

appears to show a different pattern of results than for BB stimuli, which is apparent when comparing figures 3.4a and 3.4b. In comparison to figure 3.4a, which shows superior performance at low (<20 ms) ICIs for BB stimuli, figure 3.4b has much flatter lines, depicting more consistent performance across ICI. There was still a significant main effect of jitter type, $F(3,42) = 64.44$, $p < 0.001$, with type 3 having the highest threshold, followed by types 4, 2, and 1 (see table 9), as was the case for broadband stimuli. Sidak-corrected pairwise comparisons confirm significant differences between all jitter types (see table 10 for results). There was a significant main effect of ICI, $F(3.13,43.86) = 9.66$, $p < 0.001$. Degrees of freedom were again modified according to the Greenhouse-Geisser estimate ($\epsilon = 0.63$), as Mauchly's test indicated a violation of the assumption of sphericity ($\chi^2(14) = 27.27$, $p = 0.02$). Group means, standard error and 95% confidence intervals of each ICI can be found in table 11, while results of Sidak-corrected pairwise comparisons can be found in table 12. Unlike the case for broadband stimuli, the 5 ms ICI condition was only significantly different from ICIs of 80 and 160 ms ($p = 0.025$ and 0.003 , respectively), which is reflected in the flatter lines of figure 3.4b. There was a significant interaction between jitter type and ICI, $F(15,210) = 8.92$, $p < 0.001$. Examination of figure 3.4b suggests this interaction could again be due to the differential effect of jitter type at different ICIs. For example, at 20 ms, the range of thresholds represented is broader, with non-overlapping confidence intervals for some types, while the four types appear to converge more closely at 160 ms ICI.

For HP stimuli, one extra step was performed due to the decision to impute mean values in the type 3, 5 ms ICI condition. To evaluate the possibility that this lower-variance group may have been driving any significant effects, the entire 5 ms ICI

condition was removed, and the ANOVA repeated with only five ICI groups. There was still a significant main effect of type, although under this analysis Mauchly's test indicated a violation of the assumption of sphericity ($\chi^2(5) = 15.92, p = 0.007$), so degrees of freedom were adjusted with the Greenhouse-Geisser correction ($\epsilon = 0.64$). For the main effect of jitter type, $F(1.91, 26.67) = 39.90, p < 0.001$. A notable departure from the earlier analysis was that, according to the Sidak-corrected pairwise comparisons, there was no longer a significant difference between jitter types 3 and 4 ($p = 0.21$), which suggests that the 5 ms ICI condition was driving that specific effect. There was also still a significant main effect of ICI, $F(4, 56) = 10.76, p < 0.001$. The significant interaction between jitter type and ICI was also preserved, $F(12, 168) = 8.83, p < 0.001$. With this repeated analysis, we can be fairly confident that none of our hypothesized effects are being driven solely by the lower variance group as a result of our decision to impute mean values for three subjects.

3.5 Comparison To Uniform Jitter Paradigm

Finally, our results were compared to those of Phillips et al. (2012), in order to determine how the thresholds obtained here (using singular click dislocations) would compare to a paradigm in which listeners had many opportunities within each stimulus to detect jitter. Refer to figure 3.5, which depicts both groups' jitter thresholds, as a function of ICI, with each subject plotted individually. It is clear from the graph that jitter thresholds from the current study are systematically higher than those for stimuli in which all of the clicks were jittered. For this comparison, jitter types 1 and 2 were collapsed together, to create a more direct comparison between the two tasks (one in which listeners have 24 irregular ICIs available to detect jitter, one in which there is a

single displacement or two irregular ICIs). The ICIs employed in both experiments and which could therefore act as points of comparison were 5, 10, 20, 40, and 80 ms. Unfortunately, due to the large difference in group size (five subjects vs. fifteen), it was deemed that this comparison could not reliably be accomplished statistically. However, descriptive statistics (group means and standard deviations) are provided in tables 13 (for broadband conditions) and 14 (for high pass conditions). Although we have no significance tests to confirm our results, the almost entirely non-overlapping data sets from the two studies provide strong evidence that thresholds for single displacements in click trains are much higher than to all-jittered trains (figure 3.5). Furthermore, a finding of interest is that the ratio of the mean thresholds of one group to the other (in other words, the factor by which the current threshold sensitivities are higher than those reported in the 2012 study) are fairly consistent across ICIs, ranging from 2.19-5.68x for BB stimuli, and 1.87-2.82x for HP stimuli (see tables 13 and 14).

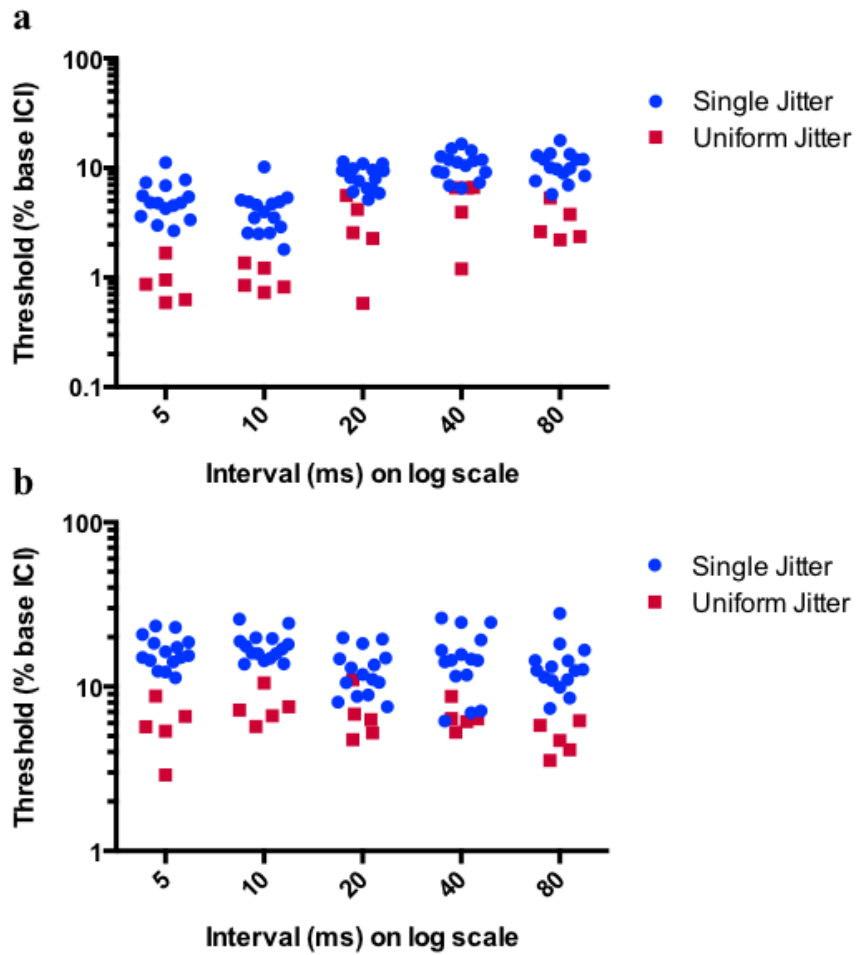


Figure 3.5. Jitter thresholds (measured as % of base ICI) as a function of ICI, presented on a log scale, for (a) broadband, and (b) high pass stimuli. Each point represents an individual subject, blue circles represent thresholds to single instances of jitter in click trains, while red squares represent thresholds to uniformly jitter stimuli. Data from performance on uniformly jittered stimuli collected from Phillips et al. (2012).

CHAPTER 4: DISCUSSION

4.1 Interpretation Of Results

The central objective of this experiment was to further characterize the sensitivity to auditory temporal irregularities of healthy, normal-hearing adults. By building on the work of Phillips et al. (2012) and employing a comprehensive new jitter paradigm, we were able to investigate several hypotheses. Ultimately, it was found that sensitivity to minimal temporal irregularities mirrors performance on tasks of uniform jitter, but with systematically higher thresholds. Furthermore, the data collected support all of our main hypotheses. The pattern of performance across short and long ICIs, and how it differs between BB and HP stimuli, provides compelling evidence for the dual mechanism hypothesis. The significant differences between types confirm that stimuli with two aberrant inter-click intervals provide a perceptual advantage for detecting jitter over stimuli with a single irregularity.

A primary finding of interest is the apparent confirmation of the dual mechanism hypothesis. Phillips et al. (2012) posited that the inverted U-shaped performance function that they observed across a range of ICIs (see figure 1.1) was due to the presence of different mechanisms available for perceiving jitter that work over different ranges of ICIs. Specifically, at short ICIs (<40 ms) and for broadband stimuli, it is believed that the click stimuli generate spectral pitch cues, which can then be used to more easily differentiate the standard and target stimuli. By comparing the salience of the pitch cues between the standard and target stimuli, listeners can base their judgment on information other than just the temporal properties of the sound. This results in an upward sloping performance function, reaching a peak of poor performance at 40 ms. It is upward-

sloping because the temporal pitch cue itself becomes very weak when ICIs are longer than about 40 ms. The peak of 40 ms aligns with data from Krumbholz et al. (2000), whose findings indicated that pitch percepts based on temporal information were generated only in response to stimuli with intervals of about 33 ms or less. When listeners are instead exposed to stimuli in which residual spectral cues have been removed or masked, this upward sloping performance disappears. Restricting the sound energy of the clicks to high frequencies, and playing a low-frequency masking noise to obscure any residual low frequency spectral pitch cues leaves subjects with purely temporal properties on which to base their judgments. This results in a large disparity in performance between BB and HP stimuli at short ICIs (figure 1.1).

This pattern of results was evident in the current study as well. Figure 3.2 shows a large disparity in performance between the two stimulus conditions at 5, 10, and to a lesser extent, 20 ms ICIs. Further, figure 3.3 makes clear that this pattern is present for all four jitter types. Paired samples t-tests confirm significant differences between the stimulus conditions at these ICIs (see table 4).

For longer ICIs (>40 ms), Phillips et al. (2012) observed that the lines representing performance for the two stimulus conditions converged. The large disparity in performance observed at short ICIs disappeared. This finding indicated that the superior performance exhibited at short ICIs for BB stimuli had indeed been based on pitch cues. As the ICIs increased in duration, spectral pitch cues were no longer available, and the BB stimuli provided no perceptual advantage over the filtered and masked clicks. This pattern of converging performance at longer ICIs is further evidence of dual mechanisms of jitter perception operating across different ranges of ICIs. Figure 3.2 from

the current study appears to confirm a similar pattern of performance. The lines representing the two stimulus conditions approach each other at the 20 ms ICI, and run in close proximity across the stimulus conditions with longer ICIs. Paired samples t-tests assessed whether differences between the stimulus conditions were significant at each ICI, and found all differences to be significant (see table 4). Despite this, the graph (figure 3.2) makes clear that the effect of stimulus condition (as manifested in the disparity in performance) is more pronounced at short ICIs, and the higher t-values at those conditions are also suggestive of a larger effect size. Once again, this pattern is replicated across all four jitter types (figure 3.3). The continued superior performance for broadband stimuli in the absence of useful pitch cues could be due to the higher cognitive demand of the masked and filtered click stimuli. Ignoring the masking noise and focusing attention on the clicks requires more effort, which may impact performance. Nevertheless, the disparity in performance between stimulus conditions is much more subtle at long ICIs where no pitch cues are available, as was hypothesized. It should be noted that three of five subjects from Phillips et al. (2012) also exhibited slightly superior performance on broadband stimuli at long ICIs, although the greatest disparity between the two conditions was at short ICIs.

Another effect observed by Phillips et al. (2012) was a downward sloping performance function as ICI increased above 40 ms, which, when considering the upward slope at short ICIs, resulted in an inverted U-shape. They reasoned that at longer ICIs, clicks in the train could be individuated, and that a regular click train may result in a rhythm cue. Jittered stimuli would not generate a similarly rhythmic percept, and therefore a judgment could be made identifying the jittered stimulus based upon the

salience of the perceived rhythm. It was also suggested that the peak of poor performance around 40 ms ICI could be due to the existence of a narrow range of ICIs in which neither perceptual strategy (pitch or rhythm cues) is readily available to the listener. In the current results, there does not appear to be as pronounced of a decline in threshold sensitivities at long ICIs. If we accept the above explanation as the reason for superior performance at long ICIs, it is perhaps not unexpected that we did not see a similar improvement in our own results. While a stimulus in which all clicks are jittered would be a striking violation of an established rhythm percept, a stimulus in which only one click is displaced would not provide a similarly dramatic perceptual difference. The majority of the stimulus would be similarly rhythmic, and only through conscious comparison of successive inter-click intervals could a subject identify the displaced click. Despite this, it appears that there is still a slight improvement in performance at long ICIs, which is more apparent for jitter types 2 and 3 (see figure 3.3b and 3.3c). For these two jitter types, the peak of poor performance on broadband stimuli is at 40 ms ICI, as expected.

The evidence from the current results is strongly supportive of the dual mechanism hypothesis, and extends the work of Phillips et al. (2012). It appears that the central auditory system has multiple perceptual strategies for detecting irregularities, which are employed over different ranges of inter-click intervals, and that accurate performance at short ICIs relies on perception of pitch cues.

It was also hypothesized that there would be an effect of jitter type. The various stimulus manipulations imposed on the click trains result in different perceptions that may impact listeners' sensitivity to the temporal irregularities. It was hypothesized that

thresholds for types 1 and 2 would be lower (i.e. the listener would be more sensitive), than for types 3 and 4, due to the presence of two aberrant inter-click intervals in types 1 and 2, as opposed to only one aberrant interval in types 3 and 4. This was confirmed to be the case. For broadband stimuli, there was a significant main effect of type ($F(3,42) = 42.431$, $p < 0.001$), and post hoc pairwise comparisons confirmed significant differences between all four types (see table 6), while the mean thresholds of each type confirm that type 3 and 4 thresholds are higher than those for types 1 and 2 (table 5). The pattern of results for the high pass and masked stimuli were similar ($F(3,42) = 64.44$, $p < 0.001$), the only difference being that the significant difference between types 3 and 4 disappeared when the 5 ms ICI condition was removed from the analysis ($p = 0.21$). However, the order of highest to lowest thresholds was the same in both BB and HP stimulus conditions (type 3, then 4, then 2, then 1).

It is unclear why performance on type 1 was consistently better than type 2, as both jitter types contained two aberrant inter-click intervals (one short and one long- see figure 2.1b and 2.1c), and the only difference was the order in which they occurred. Thus the two stimulus manipulations should result in spectrally identical stimuli, and we had no predictions about how listeners might respond to the two manipulations differently.

Another finding of interest is that that the type 3 jitter was consistently the most difficult, with higher thresholds. We believe that the single aberrant inter-click interval of shorter than normal duration may have been more difficult to perceive (especially at short ICIs), than the longer inter-click interval in type 4. Our data confirm that for both stimulus conditions, the type 3 jitter had the highest threshold values (see table 5 and 9), and was significantly different from the jitter type with the next highest thresholds (type

4) for BB stimuli ($p = 0.008$), and for HP stimuli when the 5 ms ICI condition was included ($p = 0.007$). It is possible that a longer than normal inter-click interval (as in the type 4 jitter) is easier to detect as it creates a perceptible pause, or discontinuity, in the stimulus.

Finally, it was hypothesized that threshold sensitivities to minimal temporal irregularities would be systematically higher than threshold sensitivities to jitter imposed on all clicks in a click train stimulus. Despite the absence of formal statistical analysis, this was overwhelmingly confirmed by our comparison to data collected by Phillips et al. (2012). Only five ICI conditions were used in both studies, and could therefore serve as points of comparison: 5, 10, 20, 40 and 80 ms. Figure 3.5 illustrates the dramatic difference in performance on the two types of tasks, emphasized by the almost entirely non-overlapping data points between the two groups. This is of course unsurprising, as a task in which all clicks are jittered provides 24 opportunities to detect irregularity (as there are 24 irregular inter-click intervals), which is considerably easier than detecting a single instance of jitter. Tables 13 and 14 provide the mean thresholds for each task at each ICI, and the ratio between the two tasks (i.e. the degree to which subjects performed better with uniform jitter than with single click displacements). The largest disparity in performance is for BB stimuli with 5 ms ICI, where thresholds for single displacements are 5.68x higher. The smallest difference in performance was observed for HP stimuli with 20 ms ICI, where thresholds were only 1.87x higher in the present study. Plack (2005) discusses the concept of ‘multiple looks’ and how the variability in our internal representation of a stimulus limits our ability to make discrimination judgments accurately. A stimulus of longer duration allows the auditory system ‘multiple looks’,

allowing the listener to overcome that variability (whether it is due to variations in firing rate, in phase locking precision, etc) to make more accurate discriminations. This concept could be relevant to the finding that threshold sensitivities are higher for minimal irregularities, as the central auditory system has no opportunity to integrate ‘multiple looks’ at the stimulus to form a more accurate perception on which to base decisions.

4.2 Methodological Limitations

The present study is not without methodological limitations or ambiguities, most prominently related to the arbitrariness with which some decisions about stimulus parameters must be made. A clear shortcoming is that three values had to be removed from the analysis and replaced with mean values from that condition. Although we attempted to set initial levels of jitter high enough that they would be easily perceptible to all subjects, evidently this was not the case for the HP type 3, 5 ms ICI condition. It is possible that a more thorough examination of the data being collected throughout the study (rather than at the end), may have led us to raise the initial level of jitter and instruct subjects to repeat those conditions in order to gather meaningful thresholds. It is important to keep in mind, however, that there is a limit to the displacement that can be imposed on a click in the stimulus, namely, with a 5 ms inter-click interval, the click can only be played <5 ms earlier than its ‘regular’ position in time. Beyond that, it would simply occupy the position of the preceding click. Moreover, the more that the click is displaced (using the type 3 manipulation), the more difficult it becomes to resolve it from the preceding click, so rather than making the task easier (by increasing jitter), it might be straining the temporal acuity of the listener. Nevertheless, it is likely that measures could have been taken to collect more reliable data for this condition.

Another arbitrary decision related to stimulus parameters that may have impacted performance was the relative levels (i.e. loudness) of the clicks and the masking noise in the high pass and masked stimulus condition. Although the clicks must be audible over the masking noise, the masker must be loud enough to obscure any residual low frequency spectral cues. In pursuit of this, it is possible that we overcompensated and set the level of the masking noise loud enough that it may have made it difficult to perceive the clicks well. At the very least, it requires more cognitive effort and attention in order to ‘tune out’ the masking noise and focus on the clicks, so lapses in attention have a more pronounced effect on performance than in the broadband noise condition.

A common limitation for psychophysical studies is the question of the reliability of data that is gathered from only three repetitions of a psychophysical task. There is often an improvement in performance observed from the first to the third threshold measurement, which may reflect a practice effect. Ideally, a listener’s threshold values will reflect the limits of their temporal resolution, and it is possible that further repetition of the tasks would lead to a plateau of consistent performance that is a truer representation of their temporal processing ability. Indeed, some researchers have subjects perform each task up to five times, treating the first two thresholds as practice and excluding them from analysis, a strategy employed by Phillips et al. (2012). For the purposes of this study with all of its conditions, it was considered impractical to require more than three repetitions of each task, as each subject already required six to ten sessions in order to complete all of the data collection. The possibility of treating the first of the three thresholds as practice and excluding it from our analyses was considered, but that approach would have left only one available data point on any conditions that

contained outliers, and ultimately it was decided that the conservative approach would be to include all three wherever possible. Nevertheless, it should be kept in mind that although the pattern of performance is likely a good representation of the characteristics of temporal processing abilities, the individual threshold values might not be precise reflections of the limit of each listener's temporal acuity.

Another consideration that deserves mention is the inextricable impact of attention on psychophysical tasks, and especially how variable attention spans between subjects may contribute to variability in the data that is not representative of disparate temporal processing abilities. This is an unavoidable consequence of the fact that conscious perception is linked to attention. We did our best to mitigate the effects of wavering attention. Subjects were encouraged to take breaks when necessary and to end sessions when they could no longer focus adequately on the task, and as mentioned, to repeat conditions when they felt they had lost focus and not performed well. It should be noted that attention might be a more important factor in performance on some conditions than others. On short ICI conditions where pitch cues produce a single emergent quality, performance relies less on close attention than on longer ICI conditions in which successive inter-click intervals must be consciously compared. Similarly, greater attention is required on HP conditions in order 'tune out' the masker and focus on solely the click train. Thus, the effects of attention on performance may be variable not just between subjects, but between conditions as well.

4.3 Implications And Directions For Future Research

Our results are a meaningful addition to a large body of research on temporal processing ability in normal-hearing adults. Temporal processing is an essential

component of auditory perception, and a thorough understanding of the properties and limits of temporal acuity is essential to a comprehensive understanding of sensory processes. The abundance of evidence related to deficits in temporal acuity from behavioural (e.g. Lister et al., 2002) and some electrophysiological measures (e.g. Bertoli et al., 2002) raises the question of how a loss of neural synchrony manifests physically in the nervous system. This is of course a question that cannot be easily addressed through research on human subjects. A study performed by Willott (1991) may provide some useful insight into these processes. The author used single and multi-unit recording to study the responses of cells in the central auditory system of young and old mice. It was found that many neurons in the inferior colliculus of older mice were ‘sluggish’ in their responses, and exhibited changes in their temporal discharge patterns. There was also an increase in the proportion of neurons that exhibited spontaneous activity, which the author suggested might have an impact on the internal signal-to-noise ratio and interfere with the ability to distinguish sounds from background neural activity. Although there is always a question of whether results from animal studies can be used to inform our understanding of the human nervous system, these findings could be relevant to our understanding of deficits in temporal processing. The finding of ‘sluggish’ responses in some (but not all) neurons could be consistent with selective demyelination, which would result in reduced synchrony of responses arriving at the auditory brainstem structures. A better understanding of the physical manifestations of auditory system pathologies could also aid in attempts to model these pathologies and simulate their consequences in healthy listeners, as Pichora-Fuller et al. (2007) have sought to do. Their paradigm imposes uniform jitter across sound stimuli, which would simulate jittered responses in

all neurons equivalently. Aging and other pathological processes may not work in that way. Moreover, if only a small proportion of auditory neurons develop sluggish responses, then their model may not be truly representative of jittered responses in the nervous system.

Understanding the baseline temporal processing abilities of healthy adults, as is the aim of our study, could have useful clinical implications if that knowledge could be used to diagnose related problems in those who are experiencing perceptual difficulties, or even to identify deficits in temporal acuity before they become perceptually apparent. There is some evidence that age-related deficits in temporal acuity can be observed long before difficulties in word recognition are exhibited (Snell & Frisina, 2000). Importantly, the ability to identify deficits in temporal processing is only truly useful if measures can be taken to improve on these abilities, prevent further deterioration, or to mitigate the resulting perceptual difficulties. This raises the question of whether these deficits, if identified early, are susceptible to treatment or intervention, or whether preventative measures could be taken that reduce the likelihood of developing perceptual difficulties associated with reductions in temporal resolution.

Evidence from musicians or those with musical training suggests that this is indeed the case. Musicians develop a variety of auditory skills related to practicing music, including, for example, the ability to isolate single instruments within an orchestra or arrangement. The ability to isolate sounds that overlap in pitch relies on accurate perception of timbre, or the temporal fine structure of fluctuating harmonics (Parbery-Clark, Tierney, Strait & Kraus, 2012). It appears that the refined auditory skills exhibited by musicians in practice are objectively measurable as improved temporal processing and

speech comprehension. A study performed by Parbery-Clark, Skoe, and Kraus (2009) found that musicians had superior performance on a Hearing in Noise Test (HINT). When musicians and controls' subcortical neurophysiological responses to speech in quiet and noise were compared, musicians exhibited enhanced neural timing, and a more robust subcortical representation of the stimulus, with less degradation of the response with the addition of background noise. Similar results have been found in musically trained children, who demonstrated less degradation of the auditory brainstem response in noise conditions, and superior speech-in-noise perception (Strait, Parbery-Clark, Hittner & Kraus, 2012). Parbery-Clark, Anderson, Hittner and Kraus (2012) sought to determine if musical training could offset the effect of aging on the encoding of complex stimuli like stop consonants, and found that musical experience does indeed mitigate the effects of aging, that is to say, musically trained adults did not exhibit the same deficits in neural timing as non-musicians.

The compelling evidence that musical training can positively impact central auditory processing suggests that these processes may be amenable to training, treatment, or preventative measures. The mechanism by which musical training improves temporal processing has not been elucidated. If the perceptual difficulties of elderly individuals are caused by selective demyelination that causes sluggish responses in some cells, it seems unlikely that the issue could be targeted directly. Perhaps, though, training would enable the auditory system to re-wire itself in ways advantageous to perception, in order to compensate for slow or unsynchronized responses. While it is clear that musical training in the absence of perceptual deficits can improve temporal processing, it may be the case that similar training would have little to no impact on listeners already experiencing

difficulties. It would nonetheless be valuable if treatment could prevent further deterioration and worsening symptoms.

Future research in this area could further investigate the performance of elderly individuals on jitter tasks, as well as simulating temporal processing deficits in healthy individuals in order to better understand their characteristics. Developing paradigms that more accurately model the deficits observed in those with temporal processing difficulties could provide further insight into the manifestations of central auditory pathologies. Jitter tasks could also be used to identify temporal processing deficits, and perhaps training protocols could be developed based on the observed benefits of musical experience in order to determine if temporal resolution could be improved. Ideally, those individuals with degraded temporal resolution could be identified before they experience meaningful perceptual deficits, and intervention could mitigate the effects.

Conclusion

The objective of the present study was to introduce a new jitter paradigm of minimal irregularity rather than uniform jitter, in order to measure the temporal processing of healthy adults and investigate several hypotheses related to the properties of these abilities in different conditions. The data collected are consistent with the dual mechanism hypothesis proposed by Phillips et al. (2012); that the central auditory system employs different perceptual strategies across different ranges of ICIs, and that pitch percepts generated at short ICIs aid in the discrimination of displaced clicks. Our pattern of results are as expected, with superior performance on broadband stimuli compared to masked and filtered clicks, especially at short ICIs. It was also confirmed that performance on tasks with two aberrant inter-click intervals (types 1 and 2) is superior to

those with a single irregularity (types 3 and 4), as hypothesized. Finally, through comparison to previous data collected by Phillips et al. (2012) using uniformly jittered stimuli, it was confirmed that performance suffers when listeners have only one opportunity to detect jitter, rather than many. Our research contributes to a better understanding of auditory temporal resolution, which helps to inform understanding of numerous auditory pathologies. Further research could modify the jitter paradigm to be used as a diagnostic tool, with clinical implications for how we recognize and treat central auditory pathologies.

REFERENCES

- Bance, M. (2007) Hearing and aging. *Canadian Medical Association Journal*, 176 (7), 925-927.
- Bertoli, S., Smurzynski, J., & Probst, R. (2002). Temporal resolution in young and elderly subjects as measured by mismatch negativity and a psychoacoustic gap detection task. *Clinical Neurophysiology* 113 (3), 396-406.
- Bess, F. H. & Humes, L.E. (1995) *Audiology: The fundamentals*. Baltimore: Williams & Wilkins.
- Chisolm, T. H., Willott, J. F., & Lister, J. J. (2003). The aging auditory system: anatomic and physiologic changes and implications for rehabilitation. *International Journal of Audiology* 42 (2), 2S3-2S10.
- Cowper-Smith, C. D., Dingle, R. N., Guo, Y., Burkard, R., & Phillips, D. P. (2010). Synchronous auditory nerve activity in the carboplatin-chinchilla model of auditory neuropathy. *The Journal of the Acoustical Society of America* 128 (1), EL56-62.
- Fitzgibbons, P. J., & Gordon-Salant, S. (1996). Auditory temporal processing in elderly listeners. *Journal of the American Academy of Audiology* 7 (3), 183-189.
- Gordon-Salant, S., & Fitzgibbons, P. J. (2001). Sources of age-related recognition difficulty for time-compressed speech. *Journal of Speech, Language, and Hearing Research* 44, 709-719.
- Krumbholz, K., Patterson, R. D., & Pressnitzer, D. (2000). The lower limit of pitch as determined by rate discrimination. *The Journal of the Acoustical Society of America* 108 (3 pt.1), 1170-1180.
- Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America* 49 (2), 467-477.
- Lister, J., Besing, J., & Koehnke, J. (2002). Effects of age and frequency disparity on gap discrimination. *The Journal of the Acoustical Society of America* 111 (6), 2793-2800.
- Miranda, T. T., & Pichora-Fuller, M. K. (2002). Temporally jittered speech produces performance intensity, phonetically balanced rollover in young normal-hearing listeners. *Journal of the American Academy of Audiology* 13 (1), 50-58.
- Parbery-Clark, A., Anderson, A., Hittner, E. & Kraus, N. (2012). Musical experience offsets age-related delays in neural timing. *Neurobiology of Aging* 33 (7), 1483e1-1483e4.

- Parbery-Clark, A., Skoe, E. & Kraus, N. (2009). Musical experience limits the degradative effects of background noise on the neural processing of sound. *The Journal of Neuroscience* 29 (45), 14100-14107.
- Parbery-Clark, A., Tierney, A., Strait, D. L. & Kraus, N. (2012). Musicians have fine-tuned neural distinction of speech syllables. *Neuroscience* 219, 111-119.
- Pekkonen, E. (2000). Mismatch negativity in aging and in Alzheimer's and Parkinson's diseases. *Audiology and Neuro-Otology* 5, 216-224.
- Phillips, D. P., Taylor, T. L., Hall, S. E., Carr, M. M. & Mossop, J. E. (1997). Detection of silent intervals between noises activating different perceptual channels: some properties of "central" auditory gap detection. *The Journal of the Acoustical Society of America* 101 (6), 3694-3705.
- Phillips, D. P., Dingle, R. N., Hall, S. E. & Jang, M. (2012). Dual mechanisms in the perceptual processing of click train temporal regularity. *The Journal of the Acoustical Society of America* 132 (1), EL22-EL28.
- Pichora-Fuller, M. K., Schneider, B. A., MacDonald, E., Pass, H. E. & Brown, S. (2007). Temporal jitter disrupts speech intelligibility: a simulation of auditory aging. *Hearing Research* 223 (1-2), 114-121.
- Plack, C. J. (2005). *The sense of hearing*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Salthouse T. A. (2000) Aging and measures of processing speed. *Biological Psychology* 54 (1-3), 34-54.
- Snell, K. B. (1997). Age-related changes in temporal gap detection. *The Journal of the Acoustical Society of America* 101 (4), 2214-2220.
- Snell, K. B. & Frisina, D. R. (2000). Relationships among age-related differences in gap detection and word recognition. *The Journal of the Acoustical Society of America* 107 (3), 1615-1626.
- Starr, A., Picton, T. W., Sininger, Y., Hood, L. J. & Berlin, C. I. (1996). Auditory neuropathy. *Brain* 119, 741-753.
- Strait, D. L., Parbery-Clark, A., Hittmer, E. & Kraus, N. (2012). Musical training during early childhood enhances the neural encoding of speech in noise. *Brain & Language* 123 (3), 191-201.
- Stuart, A. & Phillips, D. P. (1996). Word recognition in continuous and interrupted broadband noise by young normal-hearing, older normal-hearing, and presbycusis listeners. *Ear and Hearing* 17 (6), 478-489.

- Tremblay, K. L., Piskosz, M. & Souza, P. (2003). Effects of age and age-related hearing loss on the neural representation of speech cues. *Clinical Neurophysiology* 114 (7), 1332-1343.
- Versfeld, N. J. & Dreschler, W. A. (2002). The relationship between the intelligibility of time-compressed speech and speech in noise in young and elderly listeners. *The Journal of the Acoustical Society of America* 111 (1), 401-408.
- Willott, J. F. (1991). Central physiological correlates of ageing and presbycusis in mice. *Acta Oto-Laryngologica* 476, 153-156.
- Zeng, F-G., Kong, Y-Y., Michalewski, H. J. & Starr, A. (2005). Perceptual consequences of disrupted neural activity. *Journal of Neurophysiology* 93 (6), 3050-3063.
- Zeng, F-G., Oba, S., Garde, S., Sininger, Y., & Starr, A. (1999). Temporal and speech processing deficits in auditory neuropathy. *Neuroreport* 10 (16), 3429-3435.

APPENDIX: TABLES

Table 1. Means and standard deviations of all 48 conditions

Jitter type	ICI (ms)	<u>Broadband</u>		<u>High pass</u>	
		Mean (% base ICI)	SD	Mean (% base ICI)	SD
1	5	5.23	1.73	13.71	3.56
	10	4.25	1.86	14.27	2.95
	20	6.97	1.65	10.31	3.52
	40	9.58	3.33	13.11	4.97
	80	9.88	2.95	12.95	5.83
	160	10.07	3.64	12.50	4.51
2	5	5.44	2.86	19.34	4.68
	10	4.15	2.37	21.23	5.43
	20	9.72	2.82	15.19	5.01
	40	12.29	3.20	17.33	8.68
	80	11.61	3.91	13.95	5.30
	160	10.05	2.82	12.50	4.31
3	5	8.63	7.12	29.81	4.75
	10	5.29	2.40	24.02	6.30
	20	15.25	4.96	27.19	10.43
	40	20.37	8.01	27.24	10.39
	80	15.35	4.93	20.17	9.86
	160	10.17	3.87	15.01	6.79
4	5	5.72	3.19	19.19	4.21
	10	5.87	4.99	25.24	5.73
	20	11.22	3.59	16.21	4.91
	40	13.84	4.92	20.01	6.70
	80	13.27	5.46	19.28	4.86
	160	13.86	5.64	18.66	6.18

Table 2. Group data for each stimulus condition (ICIs collapse, jitter types collapsed)

Stimulus condition	Mean (% base ICI)	Std. Error	<u>95% Confidence interval</u>	
			Lower bound	Upper bound
Broadband	9.92	0.65	8.53	11.32
High pass	18.26	1.06	15.99	20.53

Table 3. Group data for each ICI (stimulus conditions collapsed, jitter types collapsed)

ICI (ms)	Mean (% base ICI)	Std. Error	<u>95% Confidence interval</u>	
			Lower bound	Upper bound
5	13.38	0.66	11.98	14.79
10	13.03	0.64	11.66	14.40
20	14.01	0.93	12.01	16.00
40	16.72	1.31	13.90	19.54
80	14.56	1.09	12.21	16.90
160	12.85	0.98	10.74	14.96

Table 4. Paired samples *t*-tests comparing HP to BB at each ICI

ICI pair (ms)	Paired differences					t-value	Degrees of freedom	Sig. (2- tailed)
	Mean (% base ICI, HP-BB)	SD	Std. Error mean	<u>95% CI of the difference</u>				
				Lower bound	Upper bound			
5	14.25	4.67	1.21	11.67	16.84	11.81	14	0.000
10	16.27	4.23	1.09	13.93	18.62	14.90	14	0.000
20	6.43	4.44	1.15	3.97	8.89	5.61	14	0.000
40	5.40	5.05	1.30	2.60	8.20	4.14	14	0.001
80	4.06	3.61	0.93	2.06	6.06	4.36	14	0.001
160	3.63	3.44	0.89	1.73	5.53	4.09	14	0.001

Table 5. Group data for each jitter type, broadband conditions only (ICIs collapsed)

Jitter type	Mean	Std. Error	<u>95% Confidence interval</u>	
			Lower bound	Upper bound
1	7.67	0.46	6.67	8.66
2	8.88	0.53	7.74	10.01
3	12.51	0.93	10.52	14.50
4	10.63	0.81	8.89	12.37

Table 6. *Pairwise comparisons between jitter types, broadband conditions only (with Sidak adjustment for multiple comparisons)*

Types compared	Mean difference (% base ICI)	Std. Error	Sig.	<u>95% CI of the difference</u>	
				Lower bound	Upper bound
1 vs. 2	-1.21	0.28	0.004	-2.05	-0.37
1 vs. 3	-4.85	0.55	0.000	-6.51	-3.18
1 vs. 4	-2.97	0.42	0.000	-4.26	-1.68
2 vs. 3	-3.64	0.51	0.000	-5.19	-2.08
2 vs. 4	-1.76	0.48	0.016	-3.23	-0.28
3 vs. 4	1.88	0.47	0.008	0.44	3.32

Table 7. *Group data for each ICI, broadband conditions only (jitter types collapsed)*

ICI (ms)	Mean (% base ICI)	Std. Error	<u>95% Confidence interval</u>	
			Lower bound	Upper bound
5	6.26	0.91	4.32	8.20
10	4.89	0.71	3.38	6.40
20	10.79	0.75	9.20	12.39
40	14.02	1.10	11.65	16.39
80	12.53	0.96	10.46	14.59
160	11.04	0.91	0.09	12.99

Table 8. *Pairwise comparisons between ICIs, broadband conditions only (with Sidak adjustment for multiple comparisons)*

ICIs compared (ms)	Mean difference (% base ICI)	Std. Error	Sig.	95% CI of the difference	
				Lower bound	Upper bound
5 vs. 10	1.37	0.52	0.260	-0.46	3.19
5 vs. 20	-4.54	0.75	0.000	-7.17	-1.90
5 vs. 40	-7.76	1.35	0.001	-12.50	-3.03
5 vs. 80	-6.27	1.14	0.001	-10.27	-2.27
5 vs. 160	-4.78	1.05	0.007	-8.47	-1.10
10 vs. 20	-5.90	0.58	0.000	-7.95	-3.85
10 vs. 40	-9.13	1.22	0.000	-13.41	-4.85
10 vs. 80	-7.64	1.10	0.000	-11.49	-3.78
10 vs. 160	-6.15	1.03	0.001	-9.77	-2.52
20 vs. 40	-3.23	0.99	0.081	-6.70	0.24
20 vs. 80	-1.73	0.94	0.747	-5.06	1.59
20 vs. 160	-0.25	1.02	1.000	-3.85	3.36
40 vs. 80	1.49	0.85	0.801	-1.51	4.50
40 vs. 160	2.98	0.90	0.034	0.16	5.81
80 vs. 160	1.49	0.61	0.356	-0.66	3.64

Table 9. *Group data for each jitter type, high pass conditions only (ICIs collapsed)*

Jitter type	Mean	Std. Error	95% Confidence interval	
			Lower bound	Upper bound
1	12.81	0.90	10.87	14.75
2	16.57	1.08	14.27	18.88
3	23.91	1.62	20.43	27.38
4	19.76	0.96	17.70	21.83

Table 10. *Pairwise comparisons between jitter types, high pass conditions only (with Sidak adjustment for multiple comparisons)*

Types compared	Mean difference (% base ICI)	Std. Error	Sig.	<u>95% CI of the difference</u>	
				Lower bound	Upper bound
1 vs. 2	-3.77	0.78	0.002	-6.14	-1.39
1 vs. 3	-11.10	0.97	0.000	-14.07	-8.13
1 vs. 4	-6.96	0.44	0.000	-8.31	-5.60
2 vs. 3	-7.33	0.91	0.000	-10.10	-4.56
2 vs. 4	-3.19	0.72	0.003	-5.40	-0.99
3 vs. 4	4.14	1.03	0.007	1.00	7.29

Table 11. *Group means data for each ICI, high pass conditions only (jitter types collapsed)*

ICI (ms)	Mean (% base ICI)	Std. Error	<u>95% Confidence interval</u>	
			Lower bound	Upper bound
5	20.51	0.88	18.63	22.40
10	21.17	0.96	19.11	23.22
20	17.22	1.35	14.32	20.13
40	19.42	1.76	15.65	23.19
80	16.59	1.38	13.63	19.54
160	14.67	1.23	12.04	17.30

Table 12. *Pairwise comparisons between ICIs, high pass conditions only (with Sidak adjustment for multiple comparisons)*

ICIs compared (ms)	Mean difference (% base ICI)	Std. Error	Sig.	95% CI of the difference	
				Lower bound	Upper bound
5 vs. 10	-0.66	1.26	1.000	-5.08	3.77
5 vs. 20	3.29	1.11	0.143	-0.61	7.19
5 vs. 40	1.09	1.59	1.000	-4.50	6.68
5 vs. 80	3.92	1.02	0.025	0.35	7.50
5 vs. 160	5.84	1.19	0.003	1.67	10.01
10 vs. 20	3.95	1.22	0.088	-0.36	8.25
10 vs. 40	1.75	1.37	0.977	-3.07	6.56
10 vs. 80	4.58	1.43	0.092	-0.45	9.61
10 vs. 160	6.50	0.85	0.000	3.49	9.50
20 vs. 40	-2.20	0.79	0.203	-4.99	0.59
20 vs. 80	0.63	0.95	1.00	-2.69	3.96
20 vs. 160	2.55	0.76	0.067	-0.12	5.22
40 vs. 80	2.83	1.16	0.352	-1.25	6.92
40 vs. 160	4.75	1.09	0.010	0.92	8.58
80 vs. 160	1.92	1.05	0.753	-1.78	5.62

Table 13. *Descriptive statistics comparing uniform to single jitter (jitter types 1 and 2 collapsed), broadband conditions only*

ICI (ms)	<u>Uniform jitter (UJ)</u>		<u>Single jitter (SJ)</u>		Ratio (mean _{SJ} /mean _{UJ})
	Mean (% base ICI)	SD	Mean (% base ICI)	SD	
5	0.94	0.43	5.34	2.22	5.68
10	1.00	0.28	4.20	2.00	4.20
20	3.04	1.92	8.35	2.05	2.75
40	4.99	2.41	10.94	2.97	2.19
80	3.26	1.31	10.74	3.09	3.29

Table 14. *Descriptive statistics comparing uniform to single jitter (jitter types 1 and 2 collapsed), high pass conditions only*

ICI (ms)	<u>Uniform jitter (UJ)</u>		<u>Single jitter (SJ)</u>		Ratio (mean _{SJ} /mean _{UJ})
	Mean (% base ICI)	SD	Mean (% base ICI)	SD	
5	5.86	2.12	16.52	3.70	2.82
10	7.53	1.82	17.70	3.57	2.35
20	6.82	2.48	12.75	4.04	1.87
40	6.57	1.28	15.22	6.31	2.32
80	4.88	1.11	13.45	4.90	2.76