Stream Segregation on a Single Electrode
as a Function of Pulse Rate
in Cochlear Implant Listeners

by
Sara I. Duran

Department of Electrical and Computer Engineering
Duke University

Date: ______________________

Approved:

Dr. Leslie M. Collins, Supervisor

Dr. Lianne Cartee

Dr. Loren W. Nolte

Thesis submitted in partial fulfillment of the requirements
for the degree of Master of Science
in the Department of Electrical and Computer Engineering
in the Graduate School of Duke University
2010
Abstract
(Electrical Engineering)

Stream Segregation on a Single Electrode as a Function of Pulse Rate in Cochlear Implant Listeners

by

Sara I. Duran

Department of Electrical and Computer Engineering
Duke University

Date: ________________________

Approved:

Dr. Leslie M. Collins, Supervisor

Dr. Lianne Cartee

Dr. Loren W. Nolte

An abstract of a thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in the Department of Electrical and Computer Engineering in the Graduate School of Duke University
2010
Abstract

While cochlear implants usually provide a high level of speech recognition in quiet, speech recognition in noise and music appreciation remain challenging. In response to these issues, several studies have proposed increasing the number of channels of information through multiple pulse rate strategies. For the selection of pulse rates, studies of multi-rate strategies have considered implementation issues such as harmonics, pitch saturation, and tonotopic order but have not considered the fundamental perceptual question of whether two pulse rates can provide independent channels of information on a single electrode. This study measures stream segregation as an indicator of whether different pulse rates on the same electrode can be perceived independently. This approach differs from that of previous stream segregation studies which focused on stimulation of alternating electrodes, with the motivation of determining a relationship between electrode stream segregation and speech perception in challenging noisy environments.

Stream segregation in this study was measured using two stimulus sequences following an A-B-A-B structure where A and B were different pulse rates stimulating the same electrode. The timing between A and B was controlled to provide either a regular or irregular gap between the two pulse trains. The threshold at which subjects could distinguish a regular rhythm from an irregular rhythm was used as an estimate of stream segregation since detecting an irregular rhythm is an easier task when the streams are fused. Stream segregation in cochlear implant users, as with normal
hearing listeners, was hypothesized to be influenced by factors such as frequency and
the relative timing between tones. To attempt to assess the relationship between
these and stream segregation, subjects’ rate discrimination and gap detection abilities
were also measured. The results of this study indicate that stream segregation can
occur for two pulse rates on a single electrode; thus, it may be possible to use pulse
rates to create additional channels of information. Further, the stream segregation
results were not strongly correlated with the gap detection or rate discrimination
results. The lack of correlation with the gap detection results suggests that the task
was measuring a separate perceptual phenomenon rather than providing another
measure of gap detection. The lack of correlation with the rate discrimination results
suggests that discriminability may not be a limiting factor in selecting rates for
segregation. These results may have implications for the future design of multi-rate
speech processing strategies.
4 Rate Discrimination Experiment

4.1 Subjects, Stimuli and Equipment ........................................ 31
4.2 Experimental Task ........................................................... 32
4.3 Results ........................................................................... 33
4.4 Discussion ....................................................................... 34

5 Gap Detection Experiment ........................................................ 38

5.1 Subjects ........................................................................... 38
5.2 Experimental Task ........................................................... 39
5.3 Results ........................................................................... 40
5.4 Discussion ....................................................................... 44

6 Conclusions ....................................................................... 47

Bibliography ....................................................................... 50
List of Tables

3.1 Demographic information for implanted subjects. . . . . . . . . . . . . 21
3.2 Stimulation rates used in testing. . . . . . . . . . . . . . . . . . . . 22
# List of Figures

1.1 Auditory stream segregation diagram. ........................................... 2
2.1 Diagram (not to scale) of the human ear ................................. 6
2.2 Position of maximum vibration on the basilar membrane ............. 6
2.3 Diagram of an implanted ear. .................................................. 7
2.4 Block diagram of the ACE processing strategy. ......................... 8
2.5 Stream segregation as a function of frequency separation and TRT . 11
3.1 Example stimulus burst (not to scale). ................................. 23
3.2 Diagram of stimuli used in rhythm discrimination task ............... 24
3.3 Screen capture of graphical user interface used for stream segregation task. .......................................................... 25
3.4 Minimum detectable delay results grouped by subjects. ............. 26
3.5 Normalized minimum detectable delay results. ......................... 27
4.1 Screen capture of graphical user interface used for rate discrimination task. .......................................................... 32
4.2 Stream segregation and rate difference limen results grouped by subject. .......................................................... 36
4.3 Rate difference limens grouped by base rate. ......................... 37
4.4 Log of median rate difference limens and slopes of linear fits to minimum detectable delay data for versus base rate ............. 37
5.1 Stream segregation and gap detection threshold results grouped by subject. .......................................................... 41
5.2 Gap detection threshold results grouped by base rate. ............. 42
5.3 Slopes of linear fits to gap detection threshold data and log of median rate difference limens versus base rate .................................................. 43

5.4 Slopes of linear fits to gap detection threshold data and slopes of linear fits to MDD data versus base rate .................................................. 45
While cochlear implant users can achieve high speech recognition scores in quiet, more complex acoustic scenes, such as speech in noise or music, remain challenging to parse. Multi-rate speech processing strategies have been proposed as a possible way to overcome these issues by increasing the number of channels of information on a single electrode (e.g. [1], [2], [3], [4]). Some previously developed multi-rate strategies assumed that a higher number of rates on a single electrode would result in a higher number of pitches elicited on that electrode [1], [3] while other strategies used a discrete number of rates that were perceived as discriminable [2], [4]. However, to date multi-rate strategies have not investigated whether the rates chosen cause perceptual segregation or integration of the information contained in the two rates presented on a single electrode.

Auditory stream segregation is the process whereby a listener groups sounds together according to the percept they generate (e.g. [5],[6]). For instance, a listener can perceive a musical piece played by an orchestra as a single fused stream or attend selectively to the flute or the violin, segregating one instrument’s score from the rest. Similarly, if an individual is presented with a stimulus sequence following
Figure 1.1: Pictorial example of two percepts that may occur in auditory stream segregation. Lower left illustrates a fused percept and lower right illustrates a segregated percept.

an $A - B - A - B...$ structure, where A represents a tone at one frequency, B represents a tone at a different frequency, and "-" represents silence (Figure 1.1, top), he or she could perceive the sequence as a single sound with oscillating pitch (Figure 1.1, bottom left), which can create a "galloping" percept, or as two separate and alternating streams of tones (Figure 1.1, bottom right). In the first case the listener perceives a single integrated or fused stream of sound that could produce an oscillating percept if the two tones in the sequence have sufficiently different sound qualities. In the second case fission is perceived; the two streams are perceived as independent or segregated.

The phenomenon of stream segregation has been studied at length in normal hearing subjects (e.g. [5], [6], [7], [8]). Sequences of tones following an $A - B - A - B...$ or $ABA - ABA...$ structure are commonly used to test subjects’ stream segregation abilities because they comprise a simple and easily controlled stream segregation problem. Factors such as the frequency separation between the tones, the difference in onset between successive tones and subject attention have been found to influence the perception of one or two streams in normal hearing subjects.

Stream segregation has also been studied in electric hearing. For implanted lis-
teners, the possible variables affecting stream segregation include whether pitch is being controlled through place (electrode) or rate of stimulation. Previous research in cochlear implant subjects has focused mainly experimental paradigms where stimuli are presented on multiple electrodes, thus primarily considering place-based stream segregation (e.g. [9],[10], [11], [12]). Stream segregation on a single electrode as a function of pulse rate has not yet been investigated. In the study described in this document, stream segregation is proposed as a method to investigate whether different pulse rates on the same electrode can be perceived as independent channels of information, or segregated streams. The results of this study may determine when fission (segregation) or fusion (integration) are occurring and may guide the selection of stimulation rates for multi-rate speech processing strategies. A more informed rate selection could improve the performance of multi-rate processing strategies by increasing the number of independent channels of information available to the user. This focus differs from that of previous studies of stream segregation with cochlear implant users that focused on stimulation of alternating electrodes, with the motivation of determining a relationship between electrode stream segregation and speech perception in challenging noisy conditions such as multi-talker babble (e.g. [13], [10]).

The present study considers stream segregation on a single electrode as a function of stimulation rate. As noted above, stream segregation is subjective and it is therefore hard to measure directly. However, previous research has utilized objective tasks that become harder or easier depending on whether the subject perceives the streams as integrated or segregated (e.g. [8], [10]). The study presented in this document uses one such task, in which subjects were asked to detect relative timing differences between two different sequences of tones. Subjects were presented with two sequences following an $A - B - A - B...$ structure; one had a regular rhythm throughout and the other one had an irregular rhythm imposed on tone $B$. Subjects
were asked to identify the stimulus containing the irregular rhythm. This is a more difficult task if the streams are perceived as segregated because the subjects cannot as easily compare the rhythm discontinuity in B to the regularity of A [8]. Thus, performance on a rhythm discrimination task can be taken as an indirect measurement of stream segregation abilities.

It was hypothesized that, as with normal hearing listeners, factors such as frequency and relative timing between tones would influence stream segregation results. To attempt to assess the relationship between these, subjects’ rate discrimination and gap detection abilities were measured. The rate discrimination results provide an estimate of subjects’ rate-pitch perception abilities while the gap detection results are a measure of subjects’ temporal resolution. The results of these experiments provide a multifaceted view of stream segregation. Additionally, these results may indicate when and why fission or fusion are perceived and could aid in the selection of stimulation rates to provide multiple channels of information on a single electrode.
2.1 Cochlear Implants

Cochlear implants are auditory prostheses best suited for patients whose hearing loss is due to damaged hair cells (Figure 2.1). Hair cells are responsible for translating the mechanical vibrations of the basilar membrane, located inside the cochlea, into neural information. When bent by the movement of the basilar membrane, they release neurotransmitters that cause neurons to fire. These neural firings are relayed to the brain via the auditory nerve, resulting in the perception of sound. The frequency of the impinging sound determines which group of hair cells are maximally deflected. The basilar membrane’s frequency sensitivity is ordered tonotopically, or from high to low. The apex vibrates maximally at low frequencies and the base does so at high frequencies (see Figure 2.2). The basilar membrane can thus be thought of as a frequency analyzer. The tonotopic arrangement is maintained at higher levels of the auditory system.

The death of hair cells causes the loss of this mechanical to chemical transduction and is one of the most common causes of deafness. Sensorineural hearing loss can
Figure 2.1: Diagram of the human ear, not to scale. The trajectory of sound, from the external ear to the central nervous system, is shown. Reprinted from [14] with permission of the IEEE. ©1988 IEEE.

Figure 2.2: The position of maximum vibration on the basilar membrane in response to sinusoids at indicated frequencies is indicated from low, at the apex, to high, at the base. Reprinted from [15] with permission of the IEEE. ©1999 IEEE.

have a variety of causes, including but not limited to aging, toxicity caused by certain medications, autoimmune illness, genetics or physical trauma. The purpose of a cochlear implant is to replace the function of damaged hair cells and stimulate the nerve fibers directly. A diagram of an implanted ear is provided in Figure 2.3. The speech processor is typically located in a small, hook-shaped package that sits behind the ear, indicated by (1) in Figure 2.3. The speech processor specifies which electrodes to stimulate depending on the spectral and temporal characteristics of
Figure 2.3: Diagram of an ear with a cochlear implant. The basic components of the cochlear implant system, both internal and external, are shown: speech processor (1), transcutaneous radio frequency link between external and internal receiver/transmitters (2), electrode array (3), auditory nerve fibers (4) (Cochlear Corporation, www.cochlear.com).

The sound arriving at the microphone and on the speech processing strategy of the particular device. This information travels from the external processor to the external antenna which in turn transmits the information to the internal part of the device via radio frequency communication (2). The external receiver/transmitter of the implant is held in place by a magnet. The internal receiver/transmitter of the implant transmits this information to the electrode array (3), which stimulates the auditory nerve fibers directly (4).

As previously stated, it is the speech processing strategy that determines how the electrodes are stimulated given a certain incoming sound. A variety of strategies have been developed and their use depends on the device manufacturers. The subjects in this study used the Advanced Combinational Encoder (ACE) strategy (e.g. [16],[17]; see Figure 2.4), a variation on both the Continuous Interleaved Sam-
Figure 2.4: Block diagram the Advanced Combinational Encoder (ACE) strategy. Each processing stage is indicated.

In ACE, the audio signal arrives at the microphone and is then pre-amplified to simulate a high-pass filtering process that the human auditory system performs naturally [20]. This signal is then sent through a bank of bandpass filters which approximates the tonotopic arrangement of the cochlea. The frequency range of the bank of filters approximately spans 200 Hz to 8000 Hz. The lower limit was chosen to capture the fundamental frequencies of speech [21] while the upper limit is bound by the sampling frequency of the codec. Although the frequency range could be extended, this would provide lower frequency resolution for a fixed number of filters or channels, but this could be detrimental to speech which generally lies in the aforementioned range. The envelope of the signal in each channel is calculated by rectifying and low pass filtering the signal. Once the envelope has been extracted the energy content of each channel is calculated for a given analysis window. The $N$ highest energy channels out of the $M$ available channels are selected for stimulation. Before the electrodes are stimulated, amplitude compression is necessary to account for the reduced dynamic range in electric hearing compared to normal hearing [22].
The compressed envelope corresponding to each channel is then used to amplitude modulate a biphasic pulse train. These modulated pulse trains are sent to the electrodes in an interleaved fashion such that only one electrode is stimulated at a given time in order to minimize interactions between electrodes [18]. The time difference between the onset of pulses on a single electrode is referred to as stimulation rate. Most clinically used speech processing algorithms only use one stimulation rate, common on all electrodes. Recently however, multi-rate speech processing strategies have been developed in an effort to increase the information presented on each electrode and therefore improve speech understanding in adverse conditions.

2.2 Multi-rate Speech Processing

Speech processing strategies used in current devices, such as ACE or CIS, use a single stimulation rate for all active electrodes. For this reason, it is often assumed that the number of pitches that can be evoked is tied to the number of stimulation sites or electrodes. Previous research has focused on providing subjects with intermediate pitch percepts without modifying the design of currently available electrode arrays. One method to generate these intermediate pitch percepts is current steering; the proportion of current delivered simultaneously to two or three electrodes is varied, resulting in stimulation of neural populations between the stimulated electrodes (e.g. [23], [24], [25]). Multi-rate speech processing strategies also seek to improve the frequency resolution of cochlear implant users by using multiple rates per electrode (e.g. [1], [2], [3], [4]). Previous research has shown that increasing stimulation rate up to approximately 300-500 pulses per second (pps) typically results in a monotonic increase in pitch (e.g. [26], [27], [28]), with some subjects perceiving changes in pitch for rates up to 500-1000 pps (e.g. [29], [30]). Therefore, presenting multiple rates on a single electrode is expected to result in additional frequency channels, or additional pitch percepts.
Some previously developed multi-rate speech processing algorithms continuously varied stimulation rate depending on the frequency content of the incoming signal, implicitly assuming that using a large number of rates on a single electrode would result in a larger number of pitches elicited on that electrode [1], [3]. Other multi-rate strategies proposed the use of a discrete number of discriminable rates on each electrode [2], [4]. While some multi-rate speech processing strategies have shown significant improvements with respect to clinically implemented strategies in acoustic models (e.g. [3], [4]), these results were not as promising for cochlear implant patients. However, to date, multi-rate strategies have not considered whether the rates chosen cause stream segregation or integration on a single electrode. It is possible that certain rates would cause two streams on a single electrode to segregate, while other rates may cause streams to integrate. In certain cases, such as vowel perception, it may be preferable to integrate streams across frequency. In other situations, such as speech perception in noise, subjects may benefit from being able to segregate the target stream from the competing stream. The results of the stream segregation study presented in Chapter 3 may aid in the selection of rates for multi-rate speech processing strategies by indicating whether a certain pair of rates evokes an integrated or segregated percept.

2.3 Stream Segregation

2.3.1 Stream Segregation in Normal Hearing

Stream segregation has been studied in normal hearing listeners for over 30 years (e.g. [5], [6], [7], [8], [31],[32]). During that time, several factors have been found that influence this process. These include, but are not limited to, frequency separation between tones, tonal repetition time (TRT), length of stimulus sequence, and amplitude modulation rate of the tones.

Van Noorden studied stream segregation as a function of the difference in fre-
Figure 2.5: Perceptual grouping of streams of pure tones as a function of frequency separation and tone repetition time [6]. Sequences with parameters defined by the space above the temporal coherence boundary generally result in the perception of two streams while the opposite is true for sequences with parameters below the fission boundary. Reprinted from [11] with permission from Elsevier.

Frequency between two successive tones as well as the TRT, defined as the difference between the onset of successive tones [6]. He established the notion of the temporal coherence boundary (TCB) and the fission boundary (FB). He suggested that it is impossible to hear one stream for stimuli with parameters defined by the space above the TCB (see Figure 2.5). Conversely, he proposed that only one stream can be heard for stimuli with parameters below the FB. The combination of these two attention-independent phenomena are referred to as automatic stream segregation. In the area between these two boundaries one can hear either one or two streams. What an individual perceives when in this ambiguous region is heavily dependent on attention. For instance, a listener could at first perceive a single stream but then, by focusing on the higher pitch sound and bring it to the foreground, thereby relegating the lower pitch sound to the background, be able to perceive two streams.

An individual’s stream perception does not necessarily remain constant through-
out the presentation of a sequence. It is possible for the number of streams a listener perceives to flip back and forth during the presentation of a single stimulus sequence. Additionally, it has been suggested that the auditory system is biased towards integration and that it takes time for the brain to accumulate the information needed to build up a segregated percept [33], [34]. For instance, if a listener is presented with a sequence he or she may hear one stream to begin with, then after listening for a length of time hear two streams, then hear one stream again.

It has also been proposed that listeners can segregate two streams based on the envelope of a signal, in the absence of fine temporal structure and place cues [7]. In this study, the authors created stimuli with $ABA-ABA-\ldots$ structure where $A$ and $B$ were bursts of amplitude modulated white noise. This noise, therefore, had no fine temporal or spectral cues. These results indicate that individuals are not only able to segregate streams when the tones in the sequence excite different auditory filters; stream segregation can also occur within overlapping frequency channels as a function of modulation rate.

Previous research indicates that stream segregation in acoustic hearing is affected by temporal and frequency cues, the perception of which is inherently different in electric hearing. In normal hearing, place and rate of stimulation are tied together: the position of maximum vibration on the basilar membrane depends on the incoming frequency. In electric hearing place and rate of stimulation can be manipulated independently. While normal hearing subjects are able to resolve thousands of different frequencies, the number of pitches that a cochlear implanted patient can resolve is tied to the number of stimulation sites in their implant, typically 12 to 22 for current devices. Changes in rate below the rate-pitch saturation threshold have also been shown to affect perceived pitch (e.g. [26], [27], [28]). The representation of temporal cues in electric hearing is also degraded. The pulsatile stimulation method used in most currently available cochlear implants cannot capture instantaneous changes in
the envelope of the incoming signal, it can only sample its envelope. The consideration of these additional factors makes stream segregation in electric hearing a more complex issue than its normal hearing counterpart.

2.3.2 Methods for Measuring Stream Segregation

Stream segregation is a multidimensional phenomenon which depends on frequency and temporal cues. Electrical stimulation makes the perception of these cues inherently different. Additionally, there is no parameter set for which nearly all cochlear implant listeners perceive one stream or two. For this reason special attention must be paid to how stream segregation is measured in electric hearing.

The are two main methods for measuring stream segregation. The first method relies on a subject reporting whether they hear one or two streams when presented with a certain stimulus sequence (e.g. [9], [11]). This subjective method requires subjects to have an understanding of what one or two streams sound like. For normal hearing subjects, there exists a set of parameters for which one or two streams will almost certainly be perceived [6]. It is therefore easier to present subjects with an example of a sequence that produces a one or two stream percept prior to a subjective task. Since such a parameter set does not exist for electric hearing, subjects must rely on verbal descriptions of one and two stream percepts. It is possible that the phrasing of such a description could affect the outcome of the study.

The second method of measuring stream segregation is to present the subject with a task where performance is positively or negatively affected by stream segregation. It has been shown that subjects can more easily identify a target melody which is interleaved with distractor tones if the streams are perceived as segregated (e.g. [35], [12]). If an individual can focus on each stream individually the target melody will no longer be confounded by the distractors and therefore will become easier to identify. This method requires a subject to either recognize a popular melody or
to memorize a target musical sequence. Therefore, the accuracy of this method is dependent on the individual being able to remember the target melody or on his or her knowledge of certain popular melodies. Conversely, subjects seem to have more difficulty in detecting relative timing differences between two tone sequences when the streams are perceived as separate (e.g. [8], [10]). If the streams are perceived as segregated, it becomes harder to use one sequence as a reference to detect a rhythm discontinuity in a second sequence. The performance on these objective tasks does not require the subject to have any prior knowledge of what constitutes a one or two stream percept.

Both of these methods have been recently used to investigate stream segregation in electric hearing. However, the study described in this document uses a rhythm detection task in order to obtain the most objective results possible.

2.3.3 Stream Segregation in Electric Hearing

As previously discussed, the processes that govern auditory stream segregation have been studied more extensively for normal hearing than for electric hearing. Unilateral implant users cannot rely on frequency or time of arrival differences between ears to localize a sound source or to perceptually segregate it from the acoustic background. In order to study this complex problem a simplified case must be examined. Previous research in both normal and electric hearing has used tonal sequences with an $A - B - A - B - A...$ or $A - B - A - B$ structure to measure subjects’ stream segregation abilities. The focus of stream segregation in electric hearing in the literature to date has been to investigate the effects of place of stimulation and tonal repetition time (TRT), or difference in onset between successive tones.

In general, in electric hearing, subjects report streams as segregated more frequently when the stimulation sites corresponding to tones $A$ and $B$ were more separated [9],[10], [11], [12]. This is likely due the increase in perceived pitch between the
tones as distance between electrodes increases and as such is consistent with results obtained in normal hearing individuals where greater frequency separation between sequences facilitates stream segregation [6]. In normal hearing listeners the temporal coherence boundary (TCB), or the boundary in the TRT versus frequency difference space above which only two streams can be perceived, increases with respect to TRT. Stream segregation results in electric hearing do not show significant effects of TRT on how often subjects perceive two streams [11]. Because this differs from results obtained in normal hearing, this suggests that stream segregation in electric hearing may not be governed by the same mechanisms as in normal hearing.

However, there are other variables affecting stream segregation in normal hearing that have also been examined for cochlear implant listeners. Previous studies in normal hearing suggest that it takes time for a subject to perceive two streams as segregated; longer stimulus sequences allow more time for the brain to accumulate the information needed to determine that two streams are distinct (e.g. [33], [34]). While some previous results suggest that this phenomenon does not occur in electric hearing [9], [12], others indicate that subjects report streams as segregated more frequently when the tone sequences are longer (e.g. [36], [10]), thus a consensus on this matter has not been reached. These conflicting results serve as an indicator of subject to subject variability in cochlear implants.

In similar experiments to those performed with normal hearing subjects [7], the ability of cochlear implant listeners to segregate streams based on the envelope modulation has also been examined [9]. For a given stimulation site and stimulation rate, streams were reported as segregated more often as difference in modulation frequency increased. An increase in modulation depth also produced an increase in the number of two stream responses. These results suggest that individuals can segregate streams of tones with the same stimulation rate if the differences in their modulation envelope are sufficiently large.
The aforementioned studies generally focused on stream segregation across electrodes. While envelope modulation rate and depth have been shown to produce stream segregation on a single electrode, the same has not been shown for stimulation rate. Thus, the potential uses of rate-based stream segregation as a mechanism to present multiple channels of information on a single electrode has not yet been explored. In the study presented in Chapter 3, a rhythm detection task was used to obtain an objective measure of stream segregation (e.g. [8], [10], [12]). In this experiment, subjects are asked to detect relative timing changes between two streams, which becomes a harder task when the two streams are perceived as segregated.

2.4 Pitch Perception in Electric Hearing: Rate and Place Cues

Previous research has shown that stream segregation in electric hearing is dependent on the perceived differences between the two sets of tones comprising the stimulus. Previous research has primarily focused on stream segregation as a function of electrode separation, thereby exploring the place-pitch impact on stream segregation. In normal hearing, the basilar membrane vibrates maximally at the apex of the cochlea for low frequencies and at the base for high frequencies (see Figure 2.2). Analogously, electrical stimulation of the fibers that innervate the base of the basilar membrane generally causes a higher pitch percept than stimulation of the nerve fibers at the apex [37], [38], [39].

The site of stimulation is not the only method of encoding pitch in electric hearing. An increase in stimulation rate at a given stimulation site has been shown to produce a monotonic increase in pitch up to approximately 300-500 pulses per second (pps) (e.g. [26], [27], [28]), with a small number of subjects being able to perceive changes in pitch for rates up to 500-1000 pps (e.g. [29], [30]). The different patterns of action potentials that occur as a result of stimulation at different rates are interpreted by higher levels of the auditory system at different pitches [40].
stimulated electrically, neurons fire synchronously (e.g. [41],[42]) at a rate limited by the maximum firing rate of a single neuron. The rate at which pitch saturates in most subjects is likely related to the maximum firing rate of a single neuron, measured at approximately 400 Hz in cats with sensorineural hearing loss [43]. In contrast to electrically stimulated neurons, acoustically stimulated neurons fire somewhat stochastically [44]; however, a population of nerve fibers as a whole can lock to a particular phase of a stimulus and fire simultaneously for up to about 5000 Hz [45]. This phase locking limit is likely tied to normal hearing listeners’ general inability to perceive changes in pitch for pure tones above 5000 Hz [40].

It has been proven that increasing stimulation rate can change pitch up to a certain rate threshold. It is therefore possible to present various pitches on a single electrode by selecting the appropriate rates; thereby presenting more information per channel than if a single rate had been used. In electric hearing, sequences of tones that were sufficiently separated in pitch have caused the subject to perceive two separate streams [9],[10], [11], [12]. The study presented in this document investigates whether it is possible to generate an analogous stream segregation percept by presenting a sequence of tones to a single electrode that are perceived as being different in pitch due to a difference in rate.

2.5 Gap Detection in Electric Hearing

In normal hearing, stream segregation has been shown to depend on both frequency and temporal cues [6]. In electric hearing, increasing electrode separation has been shown to encourage the perception of two streams [9], [10], [11], [12]. This study in this document investigates stream segregation on a single electrode as a function of rate. One dimension of the stream segregation phenomenon was further examined by conducting a rate discrimination experiment and comparing the results to subjects’ stream segregation results. In an attempt to investigate the second dimension of
subjects’ stream segregation abilities, a measure of their temporal resolution abilities was obtained via a gap detection experiment.

In a typical gap detection task a subject is asked to detect a silence gap in between two flanking tones or markers. The aim of such a task is to establish a gap detection threshold (GDT), or shortest detectable gap. Research has shown that gap detection thresholds for CI users are not significantly different from normal hearing users indicating that temporal resolution, and specifically how it contributes to gap detection, is not affected by sensorineural hearing loss [46], [47].

Place and rate of stimulation have both been found to affect GDTs. For within-channel gap detection results, no significant difference was found between the GDTs obtained by stimulating different locations along the cochlea [46], suggesting that subjects’ gap detection abilities are not dependent on absolute place-pitch. However, as the difference in rate between the markers presented to the same electrode increases, a deterioration of gap detection abilities is observed [48]. In contrast, GDTs have been shown to increase as the electrode separation between markers increases [49], [48], [50]. It is hypothesized that as dissimilarity between markers increases, in either place or rate of stimulation, a greater discontinuity is perceived even in the absence of a gap. This perceived discontinuity can cause subjects to perceive a gap in between dissimilar markers even when one is not present. Across channel gap detection has been suggested as a measure of electrode interaction, with higher GDTs possibly indicating greater distance between the neural populations stimulated by each marker [49].

A measurable effect of loudness on GDTs has also been observed. When markers were similar, in place of stimulation and rate, loudness was found to have a significant effect on GDTs [46], [49]. As loudness increased, GDTs decreased, indicating that listeners found it easier to detect a gap in between the two tones. However, this effect becomes less pronounced as the difference between the maskers increases.
Gap detection and rate discrimination describe the two dimensions of the stream segregation task presented in Chapter 3. Previous research suggests that these two dimensions of stream segregation are not independent; increasing the difference in rate between the markers has been shown to yield lower GDTs [48]. Thus, in addition to examining the relationship between stream segregation and each of these phenomena, this study compared subjects gap detection and rate discrimination abilities. In order to avoid loudness effects on GDT, the stimuli used for the gap detection task described in Chapter 5 were loudness balanced.
Multi-rate processing strategies aim to improve on clinically available speech processing strategies such as CIS or ACE by increasing the amount of information presented on each electrode. However, to date little consideration has been given to whether the selected rates would cause stream integration or segregation. Which one of these two percepts, fusion or fission, on a single electrode would most benefit speech recognition has not been determined. The purpose of this experiment was to investigate stream segregation on a single electrode as a function of stimulation rate, and more specifically, whether different rates on the same electrode could be perceived as independent channels of information (or segregated streams). The results of this experiment may be used to determine whether specific relationships between rates generally cause an integrated percept or whether some cause a segregated percept.

3.1 Subjects

Table 3.1 contains demographic information for the eight subjects that participated in the study, seven of whom were postlingually deaf. All subjects were users of Cochlear Corporation’s CI24 family of devices and used monopolar 1+2 (MP1+2) stimulation
Table 3.1: Demographic information for implanted subjects

<table>
<thead>
<tr>
<th>Subject ID</th>
<th>Gender</th>
<th>Age (years)</th>
<th>Age at onset of deafness (years)</th>
<th>Age at implantation (years)</th>
<th>Implant type</th>
<th>Mode of stimulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>M</td>
<td>49</td>
<td>0</td>
<td>48</td>
<td>CI24RE</td>
<td>MP1+2</td>
</tr>
<tr>
<td>S2</td>
<td>M</td>
<td>33</td>
<td>4</td>
<td>31</td>
<td>CI24RE</td>
<td>MP1+2</td>
</tr>
<tr>
<td>S3</td>
<td>M</td>
<td>58</td>
<td>55</td>
<td>56</td>
<td>CI24RE</td>
<td>MP1+2</td>
</tr>
<tr>
<td>S4</td>
<td>M</td>
<td>57</td>
<td>15</td>
<td>53</td>
<td>CI24RE</td>
<td>MP1+2</td>
</tr>
<tr>
<td>S5</td>
<td>M</td>
<td>54</td>
<td>48</td>
<td>49</td>
<td>CI24R</td>
<td>MP1+2</td>
</tr>
<tr>
<td>S6</td>
<td>F</td>
<td>74</td>
<td>46</td>
<td>66</td>
<td>CI24M</td>
<td>MP1+2</td>
</tr>
<tr>
<td>S7</td>
<td>M</td>
<td>66</td>
<td>20</td>
<td>61</td>
<td>CI24R</td>
<td>MP1+2</td>
</tr>
<tr>
<td>S8</td>
<td>F</td>
<td>45</td>
<td>10</td>
<td>41</td>
<td>CI24RE</td>
<td>MP1+2</td>
</tr>
</tbody>
</table>

mode, where both extra-cochlear electrodes (numbered 1 and 2) were used as ground for stimulation. The experiment was completed in two to six testing sessions lasting one and a half to four hours each. Except for subject S5, who volunteered his time, subjects were compensated for their participation. The use of human subjects in the experiments described in the following sections was approved by the Institutional Review Board of Duke University.

3.2 Stimuli and Equipment

Two different sets of tones, labeled $A$ and $B$, were used in this study. Both sets of stimuli consisted of biphasic pulse trains with 25 $\mu$s pulse width and an 8 $\mu$s interpulse gap, as seen in Figure 3.1. Tones $A$ and $B$ were 60 ms-long except for S5, S6 and S7 who could not hear 60 ms tones. For these subjects testing was completed using 100 ms-long tones. The stimulation rate of tones of type $A$ was set to one of three base rates (BRs): 200, 300 or 800 pps. The stimulation rate of the $B$ tones was set to one of three Weber fractions of the rate assigned to tone $A$: 0, 0.5, or 1 (see Table 3.2). Weber fraction is defined as the ratio of the difference in stimulation rate between tones $A$ and $B$ to the stimulation rate of tone $A$. 

Table 3.2: Stimulation rates used in stream segregation experiment: base rates and their corresponding Weber fractions

<table>
<thead>
<tr>
<th>Weber Fractions</th>
<th>Base rates (pps)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.5</td>
<td>300</td>
</tr>
<tr>
<td>1</td>
<td>400</td>
</tr>
<tr>
<td>200</td>
<td>300</td>
</tr>
<tr>
<td>300</td>
<td>450</td>
</tr>
<tr>
<td>800</td>
<td>1200</td>
</tr>
<tr>
<td>1200</td>
<td>1600</td>
</tr>
</tbody>
</table>

Prior to every testing session, an estimate of the subjects’ dynamic range was obtained. Since higher rates generally produce louder percepts than lower rates [51], subjects’ thresholds (Ts) were measured at each of the BRs, and maximum comfortable levels (MCLs) were measured at a Weber fraction of 1 (see Table 3.2). These measurements were taken using the method of adjustment: the subjects set the loudness of the stimuli to the requested level using a Griffin PowerMate™ USB knob. The Nucleus Implant Communicator (NIC v2) was used to stream all stimuli from a PC. Graphical user interfaces (GUIs) designed in MATLAB were used for all of the psychophysical tasks.

The amplitudes of $A$ and $B$ were loudness balanced using the method of adjustment. In this loudness balancing task, the three BRs were loudness balanced to each other and then each Weber fraction was loudness balanced to its corresponding base rate. This method of loudness balancing, in which each measurement depends on another, is referred to as the adjacent method [52]. Previous research has shown that the adjacent method provides results that are highly correlated to those obtained using the reference method, in which each measurement is independent [53].

3.3 Experimental Task

Subjects were presented with two stimulus sequences of biphasic pulse trains with $A - B - A - B...$ structure via direct stimulation of electrode 11, located in the middle of the implanted array. $A$ and $B$ were 60 or 100 ms-long pulse trains with
Figure 3.1: Example stimulus burst (not to scale). Duration, pulse width, inter-pulse gap and amplitude are indicated.

different stimulation rates. They were separated by a short period of silence denoted by (-). Each stimulus sequence was 2.4 s-long when 60 ms tones were used and 3.36 s-long when 100 ms tones were used. In the reference sequence, A and B maintained a regular rhythm throughout, therefore, the silence in between tones was always 40 ms (see panel a of Figure 3.2). As in [10], in the target stimulus sequence the rhythm of A was held constant while an accumulating delay was imposed on B starting at the 7th cycle (see panel b of Figure 3.2). In the subsequent cycles, the delay was incremented by \( d \). However, the accumulated delay of the 11th and 12th cycles was equal to that of the 10th cycle, \( 4d \). If the subjects perceive the two streams as fused this irregular-rhythm sequence is expected to create a discontinuity in the “galloping” percept with respect to the regular-rhythm reference stimulus.

A screen capture of the graphical user interface subjects used to perform the task is provided in Figure 3.3. Subjects were presented with one regular and one irregular rhythm sequence and asked to identify the one that contained the irregularity by clicking on the corresponding button with a computer mouse. Subjects were not informed that the goal of the task was to obtain a measure of their stream segregation abilities, or how often they perceived streams as separate, until the experiment was completed. A box above the response button corresponding to each interval provided subjects with a visual cue during the presentation of each stimulus interval. Subjects were asked to click on the Ready button before every trial. No feedback on subjects’
responses was provided. The goal of the experiment was to obtain an estimate of the minimum detectable accumulated delay (MDD), a predictor of the subjects’ auditory stream segregation capabilities [8]. Estimates of the MDD were obtained for nine different pairs of rates (see Table 3.2). For the first run of each estimate, the total accumulated delay was set to 32 ms. For each MDD estimate, the stimulation rate of A was fixed at one of the BRs and that of B was fixed at a Weber fraction of the stimulation rate of A (see Table 3.2). MDD estimates were obtained via a three-down one-up adaptive procedure, in order to converge on the 79.4% point on the psychometric function [54]. If the subject’s answer was correct three consecutive times, the delay was decreased, resulting in a more difficult task; if it was incorrect once, the delay was increased, resulting in an more simple task. The step size was 8 ms for the first two reversals and 4 ms for the next four reversals. The MDD
Figure 3.3: Screen capture of graphical user interface used for stream segregation task. Subjects indicated which interval contained the target stimulus sequence by clicking on button 1 or 2. The color of the boxes above the buttons was changed to blue during the stimulus presentation. Subjects clicked on the Ready button before every trial.

was taken to be the mean of the last 4 reversals. If the value of MDD is plotted versus experimental trial, a reversal is the point in the staircase where the direction is reversed. Five estimates of the MDD were obtained per pair of stimulation rates and averaged to report results.

3.4 Results

Figure 3.4 shows the MDD results for the three BRs, 200, 300 and 800 pps, and their corresponding Weber fractions of 0, 0.5, and 1 gathered from the 8 subjects that completed testing. Each point represents the average of 5 MDD estimates for that particular subject. A larger slope was taken to mean more stream segregation was occurring because as Weber Fraction increased, the MDDs were increasing, indicating that subjects were perceiving the streams as segregated more often. Subject S6’s
results illustrate the expected trend: more stream segregation at lower pulse rates and less stream segregation at higher pulse rates. Conversely, subject S2’s stream segregation capabilities seem to be relatively invariant to pulse rate. Cochlear implant subject variability is much greater than for normal hearing listeners, therefore these differences in performance are expected.

In addition to having different stream segregation abilities, it is possible that the subjects also had varying rhythm detection skills. In order to eliminate the effect of the latter, all values were normalized by subtracting the average MDD at a Weber fraction of 0 for their corresponding base rate [10]. These results can be seen in

Figure 3.4: MDD results grouped by subject. Results for the three base rates and their corresponding Weber fractions are shown. The blue line with square markers corresponds to 200 pps, green with circular markers corresponds to 300 pps and red with diamond markers corresponds to 300 pps. Each point is the average of 5 estimates of the MDD.
Figure 3.5: Normalized minimum detectable delay results. The results of each subject were normalized by subtracting the within-subject average MDD at a Weber fraction of 0 for the corresponding base rate. For the individual subjects, each point is the average of 5 normalized estimates of the MDD. Error bars indicate one standard deviation above and below the mean across subjects. Dotted line indicates best linear fit, in the least squared sense, to the subject average.

Figure 3.5. Normalization was also performed by dividing all results by the average MDD at a Weber fraction of 0 yielded similar results (not shown). Each point on the subject-specific lines is the average of 5 normalized MDD estimates. The average across subjects is represented by the black line, with error bars indicating one standard deviation above and below the across-subject mean. The normalized MDD indicates the degree of difficulty with respect to the baseline (at a Weber fraction of 0) when the stimulation rate of $B$ is increased. In order to measure this relative difficulty a straight line with a y-intercept of 0 was fit, in the least-squares sense, to the across-subject average results for each base frequency (see black lines
in Figure 3.5). The slope of the fitted line can be used to quantify the amount of stream segregation, that is how often A and B are perceived as two separate streams. Higher slopes indicate an increase in stream segregation as Weber fraction increases. The average slopes for each of the three BRs, 200, 300 and 800 pps are 10.87, 6.82 and 4.72 respectively. According to these results, subjects perceived streams as segregated most often when the base rate was 200 pps and least often when it was 800 pps. Linear fits were also found for the subject-specific MDD data shown in Figure 3.5, where subject data is plotted for all BRs. The difference between the slopes of the subject-specific linear fits for 200 and 300 pps, as well as the difference between the slopes of the linear fits for 200 and 800 pps were shown to be statistically significant (t-test, 95% confidence, \( p = 0.0236 \) and \( p = 0.0413 \) respectively).

3.5 Discussion

Assuming that the greater the similarity in pitch between tones A and B, the more likely that a subject will fuse the two streams [6], the results obtained in this experiment are consistent with previous research that found the rate-pitch percept to saturate at approximately 300-500 pps (e.g. [27], [26]). These results are also consistent with previous stream segregation results for cochlear implant listeners which show that as the two tones are more separated in place-pitch, streams are perceived as segregated more often [10]; however, in this study the changes in pitch were caused by stimulation rate. For base rates below the saturation threshold, increasing the stimulation rate is likely to have increased the perceived pitch of tone B. As the difference in the perceived pitch of A and B increased, the ability of the subjects to segregate the two streams also increased. However, when the BR was 800 pps an increase in pulse rate probably did not elicit as large of a difference in pitch as rate increased, if it elicited a change in perceived pitch at all, thereby allowing the subjects to maintain both the A and B sequences grouped together as a single stream. It is
possible that increases in pulse rate above the rate-pitch saturation threshold causes changes in other qualities of the sound, such as timber; however, those changes may not be as salient as those caused by employing the rate-pitch mechanism at lower pulse rates. In order to test the hypothesis that changes in pitch are a significant factor in stream segregation, a rate discrimination experiment was conducted. The details of the experiment are provided in Chapter 4.

These results may provide a rough reference for rate selection in multi-rate speech processing strategies such as the Multiple Carrier Frequency Algorithm (MCFA) [4]. In previous research, the rates used for MCFA were selected using only the criteria that the rates be discriminable (therefore below the rate-pitch saturation threshold) and harmonic. No consideration was given to whether these rate variations on a single channel would be integrated or segregated. It is possible that one percept would be preferable over another. One method for investigating this would be to compare the speech recognition results of MCFA using a pair of rates that encouraged integration and another pair of rates that encourages segregation.
Rate Discrimination Experiment

The results of the stream segregation experiment suggest that for lower BRs there is a greater increase in stream segregation as the Weber Fraction between tones A and B is increased. This can be seen in Figure 3.5, where the average slope of the Weber fraction versus MDD lines is highest for the 200 pps BR and lowest for the 800 pps BR. It was hypothesized that a greater sensitivity to changes in pulse rate at lower base rates might be a factor in these results. An increase in the perceptual difference between the two tones was expected to result in an increase in stream segregation (and therefore MDD). In order to determine the relationship between pulse rate sensitivity and stream segregation, a rate discrimination experiment was conducted. The difference limens (DLs), or smallest detectable differences in rate, were measured with respect to the three BRs. The results of this experiment indicate how easily subjects can differentiate tones A and B based on rate.

Previous research has shown that increasing the stimulation rate above 300-500 pps typically does not result in an increase in perceived pitch (e.g. [26], [27], [28]). Therefore it is expected that DLs will be lowest for 200 pps and will increase as pulse rate increases.
4.1 Subjects, Stimuli and Equipment

Of the eight subjects that participated in the stream segregation experiment, S1, S4, S5, S7 and S8 also participated in the rate discrimination experiment. Their demographic and implant information is provided in Section 3.1. Subjects completed the experiment in two to three testing sessions lasting two to four hours each.

The stimuli used in this experiment consisted of 60 or 100 ms-long biphasic pulse trains similar to the tones used in the stream segregation experiment. Further details on the stimuli, T and MCL measurements, equipment and loudness balancing procedure used are included in Section 3.2. Because pitch perception depends on loudness, ideally every pair of rates to be compared would be loudness balanced. However, since the adaptive rate discrimination task can result in hundreds of unique rate pair comparisons this is not feasible. Instead, the loudness of the tones not included in Table 3.2 was set to that of the closest loudness balanced rate determined previously for the stream segregation task. Additionally, tones were loudness roved ±4 current steps in order to keep subjects from using any residual loudness cues to perform the task. A current step is a unit defined by Cochlear Corporation. The exact conversion between current in $\mu$A and current level (CL or number of current steps) for the CI24M/R and the CI24RE implants are provided in Equations 4.1 and 4.2. CL ranges from 0 to 255.

\[
I(\mu A) = 10e^{\frac{CL \cdot \ln(175)}{255}}
\]  

(4.1)

\[
I(\mu A) = 17.5 \cdot 100^{\frac{CL}{255}}
\]  

(4.2)
4.2 Experimental Task

A screen capture of the GUI subjects used to perform the rate discrimination task is provided in Figure 4.1. Subjects were presented with a four-interval two-alternative forced-choice task and were asked to identify which interval contained the tone that was different in pitch compared to the rest by clicking on the corresponding response button using a computer mouse. The color of the box above the response button corresponding to each interval was changed to blue during the stimulus presentation to provide subjects with a visual cue. The first and fourth interval always contained a tone at the base rate. The target was presented either in the second or third interval, with the remaining interval also containing a tone at the BR. The inter-stimulus interval was 0.5 s. No feedback was provided.

A multiplicative adaptive procedure was used to find the difference limen (DL)
for each BR. A one-down, one-up rule was used for the first 4 reversals. Each time the subject provided a correct response the difference in rate between the target and reference tones was decreased, resulting in a harder task. This difference was increased when the subject responded incorrectly, making the task easier. A two-down, one-up rule was used for the last 8 reversals. For all twelve reversals the step size was a factor of 1.4. The DL was taken to be the geometric mean of the last 8 reversals, corresponding to a 70.7% probability of correct detection [54]. Five DL estimates were obtained for each base rate.

4.3 Results

The top panel of Figure 4.2 shows the normalized MDD results grouped by subject for the 5 subjects who also completed the rate discrimination experiment. Normalization was performed by subtracting the average MDD at a Weber fraction of 0 in order to minimize the effect of subjects’ varying rhythm detection abilities. Higher MDDs indicate that more stream segregation is occurring. The rate discrimination results are shown in the bottom panel of Figure 4.2. Lower DLs indicate that subjects are able to detect a smaller change in rate with respect to the corresponding base rate. If, as hypothesized, a greater sensitivity to changes in pulse rate was a factor in the stream segregation results, subjects with higher MDD slopes should have lower rate DLs for the corresponding BRs.

Figure 4.3 shows the difference limen results for all 5 subjects. Pairwise t-tests confirmed that the DLs for all three base rates are statistically significantly different from each other (99% confidence, \( p < 0.01 \)). These results are consistent with those reported previous studies, which indicate a decreased sensitivity to changes in rate above 300 – 500 pps (e.g. [26], [27], [28]). Lastly, Figure 4.4 shows the both the median rate DLs and the slopes of the linear fits to the normalized MDD data as a function of base rate. Higher MDD slopes indicate more stream segregation. As
expected, DLs increased as base rate increased while the opposite trend was observed for the slopes of the linear fits to the MDD data. This trend suggests that as BR increases sensitivity to changes in rate decrease and stream segregation increases. However, this negative correlation is not statistically significant ($r = -0.29, p = 0.2934$).

4.4 Discussion

The results displayed in Figure 4.3 confirm results obtained by previous research which found the rate-pitch percept to saturate between 300 and 500 pps (e.g. [26]; [27], [28]). It was hypothesized that subjects who had higher MDD slopes, indicating more stream segregation abilities, would also have lower DLs. Examination of the the side-by-side results of the stream segregation and rate discrimination experiments for each subject, shown in Figure 4.2, does not appear to support this hypothesis. The results were also grouped for all subjects and examined as a function of BR. Figure 4.4 shows a plot of the log median DLs and MDD slopes versus BR. As BR increases, log median DLs increase, as does their spread, while the average MDD slopes slightly decrease. While the correlation between the two values is negative as expected, it is not significant ($r = -0.29, p = 0.2934$).

This lack of correlation between the measures could be due to the limited number of points, 15 for each variable, representing 3 median DLs or 3 MDD slopes for each of 5 subjects. If data for more subjects were available it is possible that a stronger relationship could have been observed. Additional data obtained by increasing the number of BRs or Weber fractions might also have resulted in a stronger relationship being measured.

These results suggest sensitivity to spectral changes does not have a strong relationship in stream segregation in electric hearing. However, another factor that has been shown to influence stream segregation in acoustic hearing is the temporal
characteristics of the stimuli. It is possible that subjects were attending only to the irregular silence gaps at the end of the target sequence, instead of attending to the whole stimulus and relying on the entire length of the stimulus sequences and the segregated or integrated percept they elicited, in order to perform the task. In order to investigate the relationship between sensitivity to temporal changes in short stimulus sequences and stream segregation, a gap detection task was conducted.
Figure 4.2: Top: Normalized MDD results plotted for all of the 5 subjects who also completed the rate discrimination experiment. For each subject, results were normalized by subtracting the average MDD at a Weber fraction of 0. Each point represents the mean of 5 MDD measurements. Error bars designate ±1σ. Small horizontal jitter was added for better visualization. Bottom: DL as a function of BR for 5 subjects. 5 estimates of DLs were obtained for each subject at each BR. DLs indicate smallest noticeable change from the corresponding base rate. Median, upper and lower quartiles and whiskers extending to 1.5 times the interquartile difference are shown. Box colors (blue, green, red) correspond to the three BRs (200, 300 and 800 pps) as in top plot. Plus symbol (+) indicates outliers.
Figure 4.3: Rate discrimination results grouped by BR for 5 subjects. DLs indicate smallest noticeable change from the corresponding base rate. 5 estimates of DLs were obtained for each subject at each BR. DLs indicate smallest noticeable change from the corresponding base rate. Median, upper and lower quartiles and whiskers extending to 1.5 times the interquartile difference are shown.

Figure 4.4: Log of median rate difference limens (represented by diamonds), in pulses per second, and slopes of linear fits to normalized MDD data (indicated by circles) plotted versus base rate. Data is shown for 5 subjects. The correlation between the log DLs and slopes of the linear fits is not significant ($r = -0.29$, $p = 0.2934$).
The results presented in Section 4.3 did not show a statistically significant relationship between rate DL and MDD. This was true when the data was examined on a subject by subject basis (Figure 4.2) as well as when the data for the group was analyzed in aggregate (Figure 4.4). Subjects could have been using the detection of the irregular silence gaps that occurred between tones A and B (Figure 3.2) at the end of the target sequence to perform the task in addition to or instead of using a perceived difference in pitch or other tonal quality between the tones. In order to test this hypothesis, a gap detection experiment was conducted. The gap detection threshold (GDT), defined as the smallest detectable gap, was estimated for all pairs of rates used in the stream segregation experiment (described in Chapter 3).

5.1 Subjects

Four of the five subjects that participated in the rate discrimination task, S4, S5, S7 and S8, also took part in the gap detection experiment. The study was completed in one to two testing sessions lasting two to four hours each. The subjects demographic information is provided in Table 3.1. Section 3.1 contains additional subject
information. Further details about the stimuli, T and MCL measurements and the equipment employed are provided in Section 3.2.

Previous studies have shown that louder markers, or stimuli flanking the gap, result in lower GDTs (e.g. [47], [46], [48]). In order to minimize these effects, all tones were loudness balanced as described in Section 3.2. As in the rate discrimination experiment, the stimuli were loudness roved ±4 current steps (see Section 4.1 for further information) to eliminate any residual loudness cues which could aid subjects in performing the task.

5.2 Experimental Task

Subjects were presented with a four-interval two-alternative forced-choice task and were asked to identify which interval contained the two tones that were separated by a silence gap by clicking on the response button corresponding to the target stimuli. The GUI used was practically identical to the one used for the rate discrimination task; a screen capture of the GUI is provided in Figure 4.1. A box above each response button provided a visual cue during the stimulus presentation. The first and fourth interval always contained two tones without a gap in between them, for reference. The target was presented either in the second or third interval, with the remaining interval also containing the reference. In each interval, the first tone was always presented at one of the three base rates and the second tone was presented at one of the three Weber fractions (see Table 3.2).

A multiplicative adaptive procedure was used to find the gap detection threshold (GDT) for each base rate. A one-down, one-up strategy and a multiplicative factor of 2 was employed for the first 4 reversals. For the last 8 reversals a two-down, one-up rule and a factor of 1.3 was used. The GDT was taken to be the geometric mean of the last 8 reversals, corresponding to a 70.7% probability of correct detection [54]. The parameters for the adaptive procedure were obtained from a previous gap
detection study [48]. The initial length of the gap between the two target tones was set to 40 ms. Five GDT estimates were obtained for each pair of rates.

5.3 Results

The MDD and GDT results for the 4 subjects who completed both experiments are shown in the top and bottom panels of Figure 5.1 respectively. Higher MDD slopes indicate an increase in stream segregation as Weber fraction increases. Conversely, lower GDTs denote better gap detection abilities. For a particular base rate, an increase in GDT as a function of Weber fraction suggests that the subject found detecting gaps between tones harder as rate increased. If a subject had been using gap detection to perform the rhythm discrimination task the normalized MDD and GDT would be positively correlated; that is if the slope of the linear fit to the MDD data was high, the slope of the linear fit to the GDT data would also be high and vice versa. This would indicate that the rhythm detection task was purely measuring subjects’ ability to detect relative changes in timing, not their stream segregation abilities.

Figure 5.2 shows the GDT results grouped by BR. Regression lines were fit, in the least squares sense, to each subject’s data for each BR. Pairwise t-tests showed that the GDT results across subjects at a Weber fraction of 0 are not statistically significantly different, suggesting that the subjects had similar baseline gap detection abilities, thus allowing their data to be considered in aggregate without normalization, which would have removed an offset in GDT across BR. Previous research suggests that as dissimilarity between the markers increases, gap detection performance worsens due to a perceived discontinuity between the flanking tones even in the absence of a gap [48]. It was therefore expected that GDT slopes would be higher for BRs below the rate-pitch saturation threshold of 300-500 pps (e.g. [26]; [27], [28]) because as Weber fraction increased subjects’ would perceive the makers to be
Figure 5.1: Top: Normalized MDD results by subject for the 4 subjects who also completed the gap detection experiment. For each subject, results were normalized by subtracting the average MDD at a Weber fraction of 0. Each point represents the mean of 5 MDD measurements. Error bars indicate $\pm 1\sigma$. Bottom: Mean of 5 GDT measurements as a function of Weber fraction for 4 subjects. Error bars indicate $\pm 1\sigma$. A small horizontal offset was added for better visualization.
Figure 5.2: Gap detection results by base rate. For the individual subjects, each point is the average of 5 estimates of the GDT. Dashed line indicates the average across subjects. Error bars indicate one standard deviation above and below the mean across subjects.

more dissimilar. For a BR of 800 pps, an increase in Weber fraction was not expected to produce a very noticeable increase in pitch, therefore the GDT slopes were expected to be lower than for the two other BRs. While the average GDT slope does decrease slightly as BR increases, this difference is not significant (pairwise t-test, 95% confidence, \( p > 0.05 \)).

Since stream segregation on a single electrode is a two dimensional problem, affected by subjects’ ability to discern the stimulation rate of the two tones in the sequence as well as the gaps between them, the relationship between the two dimensions was examined. The slopes of the linear fits to the GDT data and the log of the median DLs are plotted together versus BR in Figure 5.3. A positive GDT slope
Figure 5.3: Slopes of linear fits to GDT data (represented by squares) and log of median rate difference limens (indicated by diamonds), in pulses per second, versus base rate. Data is shown for 4 subjects. There exists a significant negative correlation between the GDT slopes and the log or the median DLs ($r = -0.6396, p = 0.0251$).

indicates an improvement in gap detection abilities as Weber fraction increases for a particular BR. Conversely, higher rate DLs indicate worse rate discrimination abilities. As BR increases, rate DLs increase indicating that subjects find it increasingly difficult to detect the difference between tones at the BRs and their three Weber fractions. On the other hand, GDT slopes decrease as BR increases. This negative correlation between the GDT slopes and the log of the median DLs was significant ($r = -0.6396, p = 0.0251$). These results indicate that as the BR increases subjects no longer find it easier to detect the gap between a tone at the BR and a tone at a Weber fraction of 1 than they do between two tones at the BR. This is likely due to differences in rates below the rate-pitch saturation threshold being more distinguishable than rates above the threshold and is consistent with previous results that indicate that as the perceptual difference between tones increases gap detection worsens [48].

It was initially hypothesized that subjects might have been using the irregular
silence gaps between the tones in the stimulus sequences to perform the stream segregation task. Subjects could have been attending to the end of the stimulus sequences, where the irregularity was presented, in order to perform the rhythm detection task, thereby not relying on the entire length of the stimulus sequence and the stream segregation or integration percept elicited. In order to investigate this hypothesis, the relationship between the MDD results and the gap detection results was examined. A plot of the slopes of the linear fits to the GDT data and the slopes of the linear fits to the MDD data versus base rate is provided in Figure 5.4.

It can be seen that the average GDT slopes decrease slightly as BR increases. As previously mentioned, this decrease did not prove to be significant (pairwise t-test, 95% confidence, \( p > 0.05 \)). The MDD slopes also appear to decrease slightly. This decrease as a function of BR is not significant when the data from only 4 subjects is used (pairwise t-test, 95% confidence, \( p > 0.05 \)). However, when the data from all 8 subjects that participated in the experiment was used, a statistical difference was found between the slopes for BRs of 200 and 300 pps as well as between the slopes for 200 and 800 pps (t-test, 95% confidence, \( p = 0.0236 \) and \( p = 0.0413 \) respectively). Therefore, the lack of statistical difference between the slopes for the three BRs is potentially due to only 4 subjects being considered. The correlation between the two data sets is not significant (\( r = 0.1731, \ p = 0.5906 \)).

5.4 Discussion

The gap detection experiment was conducted to test the hypothesis that subjects who were better at detecting gaps between tones \( A \) and \( B \) also perceived streams as segregated more frequently. This would have indicated that subjects may have been using the gaps between the tones, instead of perceived differences in pitch or other tonal quality between them, to perform the rhythm detection task. The results obtained do not seem to support this hypothesis. However, the data obtained is
Figure 5.4: Slopes of linear fits to GDT data (represented by squares) and slopes of linear fits to MDD data (indicated by circles), in pulses per second, versus base rate. Data is shown for 4 subjects. The is no significant correlation between the two variables ($r = 0.1731$, $p = 0.5906$).

consistent with a previous study which found GDTs to be approximately 1-5 ms for pulsatile markers similar to those used in this study [46]. When the data from the stream segregation and gap detection experiments are compared for each subject (Figure 5.1), there appears to be no relationship between the two variables. The GDT and MDD slopes are plotted together in Figure 5.4. The lack of a statistical and observed relationship between the gap detection and stream segregation data suggests that subjects were not solely using the gaps in between the tones to perform the task.

The relationship between the GDTs and the rate DLs was also examined. Previous research suggested that subjects may find it harder to detect a gap in between two tones as the perceptual difference between the tones increases [48]. This is due to the fact that subjects may perceive a discontinuity in between two tones that have no gap in between them if the perceptual difference between the tones is large enough. A plot of the slopes of the linear fits to the GDT data and the log of the median
DLs are provided in Figure 5.3. As BR increases the log of the median DL increases, indicating that two tones with different rates become harder to distinguish as the BR increases. The opposite trend is observed for the GDT slopes. The correlation between the two data sets is negative and significant ($r = -0.6396$, $p = 0.0251$). These results seem to indicate that the easier tones are to tell apart based on their rate the harder it is for subjects to detect a gap in between them. For instance, when the BR is 200 pps and the Weber fraction is progressively incremented subjects find it increasingly difficult to detect gaps in between the two tones, presumably because at this BR the markers become progressively more distinct and create a perceptual discontinuity even in the absence of a gap.

The results of this experiment suggest that subjects were not exclusively focusing on the irregular gaps at the end of the stimulus sequence to perform the rhythm detection task; instead, it appears that subjects were employing a higher order cognitive ability. Although the results in Section 4.3 did not show a significant relationship between rate DLs and MDD slopes, it is still possible that subjects were partly relying on perceived differences in pitch to segregate streams. While no significant relationship was observed between stream segregation and the two variables hypothesized to affect it, rate DLs and GDTs, these two dimensions are significantly related to each other.
The results of this study suggest that subjects can experience stream segregation on a single electrode. Their stream segregation abilities were quantified by the slope of the linear fits to the MDD versus Weber fraction. The MDD slopes were significantly higher for 200 pps than for 300 and 800 pps BRs. Had subjects not been experiencing stream segregation, the MDD slopes would be approach zero for all BRs. The results of this study suggest that subjects were perceiving streams as segregated more frequently when they could distinguish the tones based on their rate. This hypothesis was formed with the knowledge that previous research had shown that increases in rate above 300-500 pps can no longer cause perceived increases in pitch (e.g. [26], [27], [28]). The results of the rate discrimination experiment that was performed to further understand the stream segregation data were also consistent with these previous studies: the rate DLs were found to increase significantly with BR. However, there was no correlation between the slopes of the linear fits to the MDD data and the log of the rate DLs. One reason for this could be the reduced number of data points. More data could have been obtained by having more subjects perform the stream segregation and rate discrimination experiments, by obtaining more than 5
measurements per pair of rate or by considering more Weber fractions. If one or several of these measures had been adopted it is possible that the relationship between the two data sets would have been significant. It is also possible that a subject’s stream segregation abilities are influenced by higher order cognitive processes, such as attention ([55], [10]). In this case, normalized MDDs would provide a measure of both a subject’s frequency resolution and cognitive source separation abilities. It is therefore possible that subjects with impaired frequency selectivity and very good cognitive processing could obtain high MDDs, indicating good stream segregation abilities [10].

A gap detection experiment was performed to investigate the temporal dimension of stream segregation. While the relationship between spectral discrimination and stream segregation was not significant, it is possible that the ability to perform rhythm detection task was more strongly related to temporal discriminability. In examining the GDT results, an increase in the slope of the linear fit to the GDT data was taken to suggest a worsening of gap detection abilities as a function of Weber fraction. The MDD and GDT slopes were not found to be correlated significantly, indicating that subjects’ stream segregation abilities are not correlated with their temporal resolution. These results were consistent with those obtained in a previous study [10] in which rhythm detection abilities were compared for long $A - B - A - B - \ldots$ sequences, such as those presented in this study, and short $A - B - A$ triplets. Subjects performed significantly worse when the short sequences were used. This suggest that subjects were not merely using the irregular silence gaps at the end of the stimulus sequence to perform the task and that they benefited from the build-up stream segregation effect. Lastly, the relationship between GDT and rate DL was examined. The correlation between the rate slopes of the linear fits to the GDT data and the logarithm of the rate DLs was found to be negative and significant. These results were consistent with previous studies that have indicated that as the
perceptual difference between the two tones increases (that is, as the DLs become smaller) the gap detection task becomes harder (the slope fit to the GDT is higher) [48].

The fact that stream segregation may be possible on a single electrode as a function of pulse rate has implications for multi-rate stimulation strategies. Multi-rate strategies have not considered stream segregation when selecting rates for stimulation, instead it was assumed that using more rates would lead to more useful channels of information on a single electrode. This may only be true if the rates are distinct enough to cause a segregated percept. The results of this study show that it is possible to achieve such a percept on a single electrode. However, it is possible that subjects may find it to disruptive to perceive independent channels on a single electrode. For instance, when listening to vowels it may be preferable for subjects to integrate across all frequencies. Thus, it may be preferable to create channels that are perceptually different but still produce a somewhat integrated percept. The results of this study suggest that this may be hard to achieve. Discriminability of rates was not strongly related to MDD, suggesting that even for rates that were indistinguishable, other factors caused stream segregation.

In electric hearing, the focus of stream segregation research has been on stimulating alternating electrodes. This study investigated whether stream segregation could be used to present multiple channels of information on a single electrode. The results of the experiments performed suggest that stream segregation is not strongly tied to pulse rate discrimination or temporal sensitivity, which is surprising given the results in the normal hearing literature. The results presented in this document may have implications for the design of multi-rate speech processing algorithms which may want to use pulse rate to create independent channels on a single electrode.
Bibliography


