2008

The temporal dynamics of auditory memory for static and dynamic sounds

Brandon Robert Abbs
University of Iowa

Recommended Citation

Abbs, Brandon Robert. "The temporal dynamics of auditory memory for static and dynamic sounds." PhD (Doctor of Philosophy) thesis, University of Iowa, 2008.
http://ir.uiowa.edu/etd/4.

THE TEMPORAL DYNAMICS OF AUDITORY MEMORY FOR STATIC AND DYNAMIC SOUNDS

by

Brandon Robert Abbs

An Abstract

Of a thesis submitted in partial fulfillment
of the requirements for the Doctor of
Philosophy degree in Psychology
in the Graduate College of
The University of Iowa

August 2008

Thesis Supervisor: Associate Professor Prahlad Gupta

ABSTRACT

A challenge to research on auditory and speech perception has been a lack of correlation between measures of each ability, despite their shared acoustical basis. The hypothesis presented here is that stimuli in the auditory domain stimuli generally have a static quality whereas stimuli in the speech domain generally have a dynamic quality, but attempts to correlate the two assume that static and dynamic stimuli are processed similarly by auditory short-term memory (ASTM). This thesis utilizes a change-detection task, which is popular for investigating the function of visual short-term memory, to more closely examine the function of ASTM for the purpose of evaluating the plausibility of this hypothesis. Across three experiments the temporal dynamics of ASTM are explored by manipulating the interstimulus interval (ISI), the stimulus duration, and the type of feature change in a change-detection task. This final manipulation (Experiment 3) also involves change-detection in a sequence of stimuli rather than just pairs of stimuli (as in Experiments 1 & 2). Individual differences in ASTM function are also evaluated in order to ensure that any differences found are present for all listeners and to further understand the function of ASTM. It is found that ASTM for static stimuli is consistently better than dynamic stimuli, even for listeners with high ASTM performance. Longer ISIs lead to poorer performance only for dynamic stimuli and only for participants who show poor ASTM function at long ISIs. Longer durations lead to better performance for all stimuli to a point, but performance for dynamic stimuli plateaus, and even decreases, at the longest durations, while the static stimuli continue to show gains. This result generally holds for all listeners. Lastly, changes in temporal order and changes to stimuli at the end of the sequence are detected more often than changes in frequency and large changes are detected more often than small changes. These first two effects hold for all stimuli while the second effect is more true of static stimuli. These results are discussed within the context of ASTM and the impact of each manipulation on the encoding, maintenance, and comparison processes is considered.

Abstract Approved: ______________________________

Thesis Supervisor

____

Title and Department

____

Date

THE TEMPORAL DYNAMICS OF AUDITORY MEMORY FOR STATIC AND DYNAMIC SOUNDS

by

Brandon Robert Abbs

A thesis submitted in partial fulfillment
of the requirements for the Doctor of
Philosophy degree in Psychology
in the Graduate College of
The University of Iowa

August 2008

Thesis Supervisor: Associate Professor Prahlad Gupta

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

PH.D. THESIS

This is to certify that the Ph.D. thesis of

Brandon Robert Abbs

has been approved by the Examining Committee
for the thesis requirement for the Doctor of Philosophy
degree in Psychology at the August 2008 graduation.

Thesis Committee: ______________________________
Prahlad Gupta, Thesis Supervisor

____
Bob McMurray

____
JB Tomblin

____
Gregg Oden

____
Shaun Vecera

To Lena,
who kept me going.
&
My Family,
who got me on my way.

ACKNOWLEDGMENTS

I would like to acknowledge Prahlad Gupta for his continued support, patience, and mentorship over the past five years.

I would like to acknowledge the members of my Thesis Committee, especially Bob McMurray for his knowledge, thought, and patience during the revision of this thesis. I would also like to acknowledge Paul Abbas, Chi-Wing Ng, Amy Poremba, Chris Turner, and members of the Language Discussion Group at the University of Iowa for helpful discussion of the motivation and design of these experiments.

I would like to acknowledge John Lipinski, Stephanie Packard, Jamie Tisdale, Libo Zhao, and the undergraduate members of the Language and Memory Lab at the University of Iowa for their support and assistance in the completion of these experiments.

I would like to acknowledge Nancy Andreasen, Hsin-Jen Hsu, Amanda Owen, and J Bruce Tomblin for additional mentorship and helpful discussion on the application of these and related ideas to my future study of disordered cognition in mental illness and language impairment.

ABSTRACT

A challenge to research on auditory and speech perception has been a lack of correlation between measures of each ability, despite their shared acoustical basis. The hypothesis presented here is that stimuli in the auditory domain stimuli generally have a static quality whereas stimuli in the speech domain generally have a dynamic quality, but attempts to correlate the two assume that static and dynamic stimuli are processed similarly by auditory short-term memory (ASTM). This thesis utilizes a change-detection task, which is popular for investigating the function of visual short-term memory, to more closely examine the function of ASTM for the purpose of evaluating the plausibility of this hypothesis. Across three experiments the temporal dynamics of ASTM are explored by manipulating the interstimulus interval (ISI), the stimulus duration, and the type of feature change in a change-detection task. This final manipulation (Experiment 3) also involves change-detection in a sequence of stimuli rather than just pairs of stimuli (as in Experiments 1 & 2). Individual differences in ASTM function are also evaluated in order to ensure that any differences found are present for all listeners and to further understand the function of ASTM. It is found that ASTM for static stimuli is consistently better than dynamic stimuli, even for listeners with high ASTM performance. Longer ISIs lead to poorer performance only for dynamic stimuli and only for participants who show poor ASTM function at long ISIs. Longer durations lead to better performance for all stimuli to a point, but performance for dynamic stimuli plateaus, and even decreases, at the longest durations, while the static stimuli continue to show gains. This result generally holds for all listeners. Lastly, changes in temporal order and changes to stimuli at the end of the sequence are detected more often than changes in frequency and large changes are detected more often than small changes. These first two effects hold for all stimuli while the second effect is more true of static stimuli. These results are discussed within the context of ASTM and the

impact of each manipulation on the encoding, maintenance, and comparison processes is considered.

TABLE OF CONTENTS

CHAPTER

APPENDIX

LIST OF TABLES

Table

LIST OF FIGURES

Figure

CHAPTER 1

THE IMPORTANCE OF DETECTING FEATURE CHANGES IN AUDITORY STIMULI

Auditory Processing of Speech and Non-Speech

What is the relationship between non-speech auditory processing, as typically studied by psychophysical methods, and speech perception? In the literature, the evidence on this is mixed. A number of studies have concluded that there is no direct processing relationship, based on a failure to find correlations between performance on non-speech stimuli in an auditory psychophysical task (hereafter psychoacoustics) and a speech perception task (Christopherson & Humes, 1992; Espinoza-Varas & Watson, 1988; Era, Jokela, Qvarnberg, and Heikkinen, 1986; Johnson, Watson, & Jensen, 1987; Karlin, 1942; Surprenant & Watson, 2001; Van Rooij & Plomp, 1990, 1992). On the other hand, a number of recent studies have found a correlation between performance in a psychoacoustic and a speech perception task (Henry, Turner & Behrens, 2005; Fu, 2002).

The earliest large-scale study of the relationship between psychoacoustic ability and speech was conducted by Karlin (1942). Karlin's goal was to empirically validate clinical anecdotes regarding the independence of auditory abilities (e.g., pitch discrimination, duration judgment, etc.). To this end, Karlin (1942) developed a number of tests, each of which was designed to test a particular ability using both simple and speech stimuli, and then ran a large number of listeners (n = 200) through the battery. This overall approach of presenting a number of different tasks to listeners and interpreting either the factor structure of a factor analysis or the simple correlations between tasks is the most frequent approach to measuring the relationship between psychoacoustic ability and speech perception.

While the investigation of Karlin (1942) introduced the idea of describing the fundamental perceptual abilities that are present the auditory domain and relating these abilities to more complex auditory capabilities (i.e., the perception of complex sounds), the precise relationship between these abilities and speech processing remained unclear. Specifically, the speech and non-speech tasks that correlated well involved the identification of fundamental aspects of speech stimuli (e.g., pitch or level), but did not ask participants to link these sounds to any lexical item or otherwise identify the linguistic content of the speech, including the explicit recognition of the categorical label for the phoneme that was presented. Those that did ask participants to overtly respond by reporting the meaning of the speech that had been presented, or even simply discriminate two phonemes (embedded within a word) varying along a particular acoustic dimension, bore a weaker relationship with the simpler stimuli (correlations generally < .300). Together these data suggest that the relationship between two abilities within the auditory domain rests on the correspondence between the task demands, specifically the demand to combine information across frequency bands and time.

Later studies approaching the same question yielded similar results (Elliott, Riach, Sheposh, & Trahiotis, 1966; Harris, 1964; Johnson et al., 1987; McLeish, 1950; Stankov & Horn, 1980). In Johnson et al. (1987) listeners were presented with a number of different tasks including ones that was designed to measure the temporal precision and sequencing ability of the auditory system. This task represents a novel contribution to this literature and was designed for use in previous studies (Watson, Kelly, & Wroton, 1976; Watson, Wroton, Kelly, & Benbassat, 1975) as an easily-controlled, non-linguistic, analogue to speech stimuli. Like Karlin (1942), the goal of Johnson, et al. (1987) was not necessarily to relate abilities on this task to speech perception, rather it was to assess performance on psychophysical tasks with little training and to assess the correlations between these psychophysical measures themselves. An additional goal of Johnson et al. and the studies of Watson et al. (1975;1976), upon which it was based, was to explore the processes underlying the perception of speech analogues. As will be demonstrated later, the present thesis will take a similar approach as it is an important, but often neglected, area of psychoacoustics research.

Overall it was found that performance on the sequencing task did relate to the discrimination of changes to frequency and, to a lesser extent, intensity, duration, and gap size. This relationship could be attenuated by a reduction in the memory components of the two tasks. This was demonstrated by a weakening of the relationship following the change from 3IFC to same-different and following a change from a 4-stimulus sequence to a 2-stimulus sequence, both of which reduce the role of memory in these tasks. Thus, evidence is presented that a more ‘peripheral’ psychophysical ability and a more general ‘central processing ability’ (the latter of which would be used for combining information across frequency bands and time) may be separate contributors to the overall ability to perceive speech.

Another central contributor to the overall ability to perceive speech is the ability to ‘overcome’ noise and distortion of the speech signal to identify speech, an ability that can be somewhat independent of one’s auditory perceptual ability (Stankov & Horn, 1980; Van Rooij & Plomp, 1990, 1992). In some listeners the effect of noise can be so large as to be deemed pathological (Middelweerd, Festen, & Plomp, 1990; Rodriguez, DiSarno, Hardiman, 1990; Rupp & Phillips, 1969). These listeners are said to have “normal fragile ear” (Rupp & Phillips, 1969) or “central auditory processing disorder” (Rodriguez et al., 1990). These deficits may include the inability to combine information across different bands that is exaggerated when the central auditory system is taxed by noise and the signals arising from different bands are incomplete or imperfect (Surprenant & Watson, 2001).

The ability to overcome distortion and interference was investigated systematically by Surprenant and Watson (2001) who tried to relate speech in noise tasks to threshold estimation for simple sounds and temporal order judgments. The speech stimuli included CVs, nonsense syllables, one-syllable words, and two to thirteen word sentences (spoken by a male), all of which were presented against either “cafeteria” noise (nonsense syllables) or speech-shaped noise (all others). For the CVs and sentences, the task was an identification task, whereas for the words the task was a 4AFC task with the alternatives differing from the target by one feature. This feature lead to changes in either the initial consonant, the final consonant, or the vowel. First, substantial individual differences were found in the speech tasks. Performance for the group varied by +/- 15 % around the mean performance on average across the three tasks. However, this variance only correlated within the set of speech tasks.

A factor analysis of these tasks showed one factor each for the discrimination of pitch, duration, and intensity of simple tones in various contexts; a factor for the speech identification; and a factor for temporal order judgments. There were small correlations (< .30) between the identification of words in noise and the measures of pitch discrimination, duration, intensity, and temporal order; however, these correlations were not high enough to consist of a unified factor. This result is similar to the results of Karlin (1942) and again seems to reflect a contribution of many different types of ability to overall speech ability; even if some abilities contribute more than others. The separate factor for speech stimuli in this experiment seems to be related to the ability to separate the speech signal from noise; an ability that is likely dependent on the combination of information across the periphery (Karlin, 1942), but is much more specific.

While these studies all failed to find a direct relationship between very simple and well-defined psychoacoustic abilities, such as pitch discrimination, and speech perception, two recent studies have had success in this realm (Fu, 2002; Henry, Turner, & Behrens, 2005). Henry et al. (2005) completed an investigation into the relationship between frequency selectivity and speech perception not by measuring pitch discrimination (which the authors point out can be affected by audibility, the effect of which may explain even small correlations between pitch discrimination and speech perception), but by measuring the ability to detect changes in the frequency position of spectral peaks in rippled noise. Rippled noise selectively attenuates frequency bands of broadband noise and the size of these ripples can be manipulated. Listeners must detect when these stimuli are completely inverted and the measure of spectral resolution comes from the minimum size of the ripples necessary to detect inversion. Speech stimuli were consonants presented in an /aCa/ context and vowels presented in an /hVd/ context using multiple talkers. Listeners in this study were normal hearing, hearing impaired, and cochlear implant listeners; which ensured a large degree of variability in performance from which correlations may be detected.

Overall, the relationship between spectral peak resolution and speech recognition was nonlinear, but highly predictive. Spectral peak resolution explained 64% of the variance in the perception of vowels and 66% of the perception of consonants. The nonlinear nature of the relationship was due mostly to the high performance of the normal hearing listeners. For cochlear implant and hearing impaired subjects, the relationship was linear and again approximately the same amount of variance was explained for vowels and consonants. Here, spectral resolution is found to be important for speech recognition when measured appropriately and this correspondence is likely to be based on the processing of *sets of frequencies* in both cases. More on this point will be introduced later.

Fu (2002) allows us to draw similar conclusions for the processing of *sets of modulated* frequencies in the case of consonants. In this experiment, only cochlear implant listeners were used and they were presented with temporally modulated noise. For the noise stimuli, the measure of performance was the threshold in terms of the depth of modulation needed to detect the modulation. The speech stimuli used in this experiment were consonants presented in an /aCa/ context and vowels presented in an /hVd/ context; these stimuli again utilized multiple talkers.

For these listeners, the correspondence between detection of modulation and speech recognition was extremely high and differed for vowels and consonants. The ability to detect modulation explained 72% of the variance for the perception of vowels and 97% of the variance for consonants. This relationship may be inflated given the importance of temporal modulations to the detection of speech in cochlear implant users, however this result indicates that a significant contributor to the relationship between processing of non-linguistic auditory stimuli and speech lies in the exact type of non-linguistic stimuli.

Why have some studies found a relationship whereas other have not? The answer may lie in a variable that has been overlooked in these studies: the difference between *static* auditory stimuli (whose frequency value does not change across presentation) and *dynamic* auditory stimuli (whose frequency is modulated during presentation). The findings of the studies described above can be partitioned based on whether they employed static or dynamic tones. The earlier studies of Karlin (1942) and those that followed him in utilizing a psychometric approach to study auditory and speech perception found data that indicated that the ability to accumulate auditory information across frequencies and across time were important to processing speech, however none considered the implications for these findings in the type of stimuli that are traditionally used in psychoacoustic investigations. That is, they all continued to use static tones in psychoacoustic measures despite the inherently dynamic nature of speech stimuli. This dynamic nature necessarily requires the accumulation of information across frequency bands and across time in order to efficiently process these types of stimuli. Validation of the plausibility of this hypothesis comes in the form of the results of Henry et al. (2005) and Fu (2002). Both of these studies asked listeners to utilize information across frequency domains or across time in their psychoacoustic measure and both found a correlation between this measure and speech perception.

In fact, another tack that has been taken for examining similarities between general auditory ability and speech processing is to examine similarities in the processing of different stimulus types utilizing the same task. Here, the variable of interest is the stimulus that is presented and these stimuli generally consist of speech stimuli and speech analogs, or sounds containing key features of speech. In contrast to the first set of studies reviewed above, the line of research that focuses on stimulus properties has been quite successful in demonstrating similarities between the processing of speech and non-speech stimuli. Many have focused on categorical perception of speech and non-speech and have found similar response profiles (i.e., categorical perception) for violins when attack time is manipulated (Cutting & Rosner, 1974), duplex stimuli when the onset time of one tone is manipulated (Triesman, Faulkner, Naish, & Rosner, 1995), and harmonic stimuli when the onset time of the lower hamonics is manipulated (Jusczyk, Rosner, Reed, & Kennedy, 1989). This last study was conducted with infants, which was used to argue that domain-general mechanisms of perceptual discrimination are used to build phonological competency while acquiring language. The important contribution of these studies (and others like them) is their demonstration that the manipulation of different types of acoustic stimuli, which are analogous to certain acoustic properties of speech, leads to similar response profiles as when the analogous properties are manipulated in speech itself.

Together, the data discussed so far suggest that as the acoustic features of auditory stimuli more closely resemble speech, the behavioral profile of these tasks more closely resembles that of speech tasks, and perception in these tasks better predicts perceptiond of speech. This thesis raises the possibility that processing of *dynamic* non-speech stimuli may be related to the ability to recognize speech, even if the processing of *static* non-speech stimuli is not. The primary goal of this thesis is to examine whether processing differs systematically for static and dynamic stimuli. However, the goal of the thesis is not to relate processing of dynamic non-speech stimuli to speech processing, rather it is to examine the impact of static and dynamic stimuli on *auditory memory* in order to understand (1) whether differences exist in memory for static and dynamic stimuli and (2) how we might characterize the function of the memory system responsible for performance on these tasks (i.e., a desire to understand auditory memory function in and of itself). These goals are established so that we might develop hypotheses as to the underlying mechanism(s) of performance on speech and non-speech tasks and the variables that might affect its/their performance. Elaboration on the mechanisms underlying auditory discrimination is an area of theoretical development that has been lacking in comparison with the many studies investigating the ways in which discrimination ability for both speech and non-speech can be affected (Schouten, Gerrits, & van Hessen, 2003). Further, these studies have neglected the impact that individual differences in this mechanism may have on performance and how this variance may or may not relate to variance in a speech context. However, the general purpose of examining the correspondence between the properties of speech and non-speech stimuli have a long history (e.g., Johnson et al., 1984; Jusczyk et al., 1989; Treisman et al., 1995; Watson et al., 1975; 1976).

Given these data and the goals of the present thesis, what method should we adopt for examining processing differences between static and dynamic stimuli in the types of tasks discussed above? I have chosen a *change-detection task* for examining these differences given both its application to questions of perception and memory in a variety of sensory modalities as well as its correspondence to previous tasks that have undertaking similar questions as the present thesis. In a change-detection task, a perceiver is presented first with one stimulus, and then, after a delay, with a second stimulus, and must judge whether the two stimuli are the same or different. That is, the perceiver must judge whether or not the second stimulus is changed from the first, and the task therefore requires *change detection*. The studies cited above are, in effect, all change-detection studies, although they have not always been cast as such.

In a psychophysical task, listeners are generally presented with a tone, a delay, and a second tone. On “same” trials the second tone is exactly the same as the first whereas on “different” trials a feature (e.g., frequency or amplitude) may change. On each trial listeners must decide if the tone has changed from the first presentation to the second. Alternatively, listeners can be presented with three stimuli in succession in which just one stimulus is different, in which case listeners must detect the “different” stimulus on every trials. Using different masking techniques, including simultaneous (Egan & Hake, 1950; Shower & Biddulp, 1931; Wegel & Lane, 1924), temporal (Harris & Dallos, 1979; Relkin & Turner, 1988; Smith, 1977), or notched-noise masking (Wright et al., 1997; Viemeister, 1974; 1983) in conjunction with a change-detection task one can measure perceptual tuning curves that describe when listeners are able to detect changes in relation to the frequency of the ‘target’ tone as well as irrelevant, interfering tones that are played alongside the ‘target’. These data have traditionally been used to describe how different locations on the basilar membrane within the cochlea react to perceptual information and how this information is represented in the auditory nerve. These techniques have provided information about the *resolution* of the human auditory system (i.e., how finely feature information can be represented in the auditory system).

Similarly, in a speech perception task listeners are often times presented with a change detection task. As is the case in psychophysics, the second sound of a trial may be the “same” or “different” and listeners must decide if the tone has changed from the first presentation to the second. By manipulating the phonological features of a speech stimulus (voice onset time (VOT), formant frequencies, formant ratios, etc.) it is possible to study what features of a stimulus are critical to its recognition and how variation in these features affect recognition (e.g., Carney, Widin, & Viemeister, 1977; Liberman, Harris, Hoffman, & Griffith, 1957; Schouten, et al., 2003[1]).

---

[1] While not all of these authors utilized a simple AX discrimination procedure (as befits the present description of a change detection task), each of these tasks includes an element that requires the listener to detect a change in order to respond in the task.

Change detection ability in a speech context has recently been found to be predictive of the ability to repeat novel sequences of linguistic information, so-called nonwords (Abbs, Gupta, Tomblin, & Lipinski, 2007). Importantly, nonword repetition ability has itself been related to vocabulary development and is therefore seen as a proxy for language acquisition ability (Baddeley, Gathercole, & Papagno, 1998; Gathercole, 2006) and word learning in adults (Gupta, 2003). Others consider change detection *within* a phoneme as critically important for processing features of the stimulus, such as formant transitions, and enhancing the contrast of temporally-contiguous sounds (Kluender, Coady, & Kieft, 2003). While these two perspectives (Abbs et al., 2007 & Kluender et al., 2003) consider change detection that occurs on different time scales and, potentially, through different mechanisms, they both suggest that change detection may be an important ability for language processing and acquisition.

Static and Dynamic Auditory Patterns in Speech

The distinction between static and dynamic stimuli is of interest not only as a variable that may have been overlooked in certain prior studies, but in its own right. For speech sounds, this distinction corresponds roughly with the distinction between vowels and consonants.

Of the eight major classes of sounds that humans produce, three are the easiest to describe acoustically: vowels, fricatives, and nasals (Kent, & Read, 2002, pp. 17) because they are relatively static[2]. For all speech sounds, the frequency information generated by the larynx passes through the vocal tract and produces resonant frequencies known as formants. These formants are theoretically infinite in number, but in practice only the first four (F1 – F4) need to be used to categorize the speech sounds produced by humans

---

[2] Compared to other classes of stimuli, the frequency components of vowels, fricatives, and nasals do not change much over time, but dynamic properties of these classes are used by listeners in some contexts.

(cf. Plomp, 1964). Vowels can thus be defined as sets of frequencies (i.e., formants) spaced far apart in systematic patterns (Fant, 1960; 1970). For English vowels, the high-low and front-back position of the tongue in the oral cavity are the two dimensions that matter for distinguishing vowels and they systematically relate to the frequency of the first and second formant (F1 and F2). The reasons for the differences in formant frequencies between vowels as a function of vocal tract shape are beyond the scope of the present discussion; suffice it to say that vowels can be thought of as relatively static auditory patterns that maintain a systematic relationship between formants.

As I briefly discussed earlier, the correspondence between change detection for rippled noise and speech seen by Henry et al. (2005) is likely due to the correspondence between the type of information in the psychoacoustic and speech stimuli. In both types of stimuli, the identification of and memory for spectral peaks of static stimuli are critical. Research on the response of auditory cortex to vowels and rippled noise in non-humans have found a close correspondence to the neural representation of both types of stimuli, which lends further evidence to the reasons for their correlation (Versnel & Shamma, 1994). In the present (simplified) view of speech perception, such static auditory patterns can be modified to yield fricatives and nasals, two other classes of speech that contain relatively static frequency information. I turn now to more dynamic aspects of speech to learn what information is conveyed by *modulations* of frequency and *sequences* of auditory patterns.

The auditory patterns of other classes of sounds are more complex in that features of the stimulus are modulated and these modulations are used to recognize speech (Kent & Read, 2002, pp. 140 - 188). Stop consonants, such as /t/ or /d/, are produced by completely closing the vocal tract (i.e., both the oral and nasal cavities). Most consonants are considered in relation to a vowel that either follows (e.g., /da/) or precedes (e.g., /ad/) the consonant's articulation. These are called prevocalic and postvocalic stops, respectively. In the case of a prevocalic stop, the following sequence of sounds is produced: 1.) a period of silence lasting < 100 ms (i.e., a closure), 2.) a 5-40 ms burst of air as the obstruction is removed (i.e., a release), and 3.) a transition to a static sound lasting < 50 ms. A postvocalic stop consonant begins with a short transition *from* a static sound, followed by the closure and an *optional* release. If a sound is voiced, then a low-frequency laryngeal energy may be present during each of these sounds, otherwise the onset of voicing in a prevocalic consonant will be delayed by 25 - 100 ms until the articulation of the vowel begins. Again, the specifics of the entire class of sounds known as consonants is beyond the scope of this thesis, but do note the general idea that most consonants consist of a quick successions of sounds, each of which are particularly short (< 100 ms) in duration and many of which are dynamic.

Stop consonants are categorized based mainly on the presence of low-frequency voicing during the period of silence, the frequency information in the burst (Liberman, Delattre, & Cooper, 1952), and the formant transition to the new sound during the final phase of articulation. Each of these cues can be thought of as contributing to a complex auditory pattern that is extended in and modulated over time. Other classes of consonants, such as affricates, liquids, and semivowels, utilize differences in the exact acoustic features of each phase of articulation to distinguish individual phonemes. Each of these dynamic sounds can be thought of as auditory patterns that depend more on *the modulation of frequency information* and other features related to the duration of frequency information. *Findings of processing differences between static and dynamic stimuli will thus have implications for processing differences between consonants and vowels*.

Each of the classes of speech described above contains useful information that spans a few milliseconds to many seconds. This information may include short-term information such as the rise or fall time of a particular formant or long-term information such as the exact sequence of features. Thus, theories of speech perception have attempted to describe the recognition process on many time scales. Some posit serial processes that recognize individual features, phonemes, and words in both an interactive (McClelland & Elman, 1986) and non-interactive (Oden & Massaro, 1978) framework. Others accumulate information on short time scales, but ultimately recognize speech based on the more long-term characteristics of the spectrum (Klatt, 1979; Goldinger, 1998). However, the common thread that unites each of these theories is their reliance on some sort of short-term memory system for initially storing information, even if that information is not immediately utilized in the recognition process. Importantly, this information comes from multiple, time-varying, signals and it needs to be maintained across time scales that are longer than the ones that were just described. Further, there is no evidence that as development proceeds, the auditory system that contributes to language learning has any a priori information as to what time scales will ultimately be important for recognizing one's native language. Thus, the memory system that contributes to speech processing must ultimately be one that can maintain information across a wide range of time scales and researchers must ultimately study how auditory information is processed across a wide range of time scales.

Method and Aims

The task best suited for the study of memory for auditory information across a wide range of time scales is the change-detection method. Further, this method has been extensively utilized within the field of psychoacoustics and speech perception as well as other sensory domains (most notably, vision). In vision research, change detection tasks are thought to measure processing capabilities of visual short-term memory (VSTM). The task that is used in this area of research is derived from that of Phillips' (1974) in which viewers are presented with a number of distinct visual objects each defined by a unique set of low-level features at a unique spatial location. This array of stimuli is presented briefly, removed, and after a short delay is replaced with a second array. On some trials the display stays the “same” and on others one feature of on object can be “different”.

Using this change-detection paradigm, Luck and Vogel (1997) presented a number of colored squares and varied the number of objects presented in the array and it was found that for most viewers the task could be completed fairly well with four or fewer objects. Beyond this number participants were left guessing. These authors concluded that the inability to detect changes when more than four objects were present indicated a capacity limitation in VSTM of about four objects. Extensions of this paradigm have had participants judge changes in different sets of features and found that it was not the number of features, but the number of objects that was important for determining how much information can be represented in VSTM at any one time (Luck & Vogel, 1997; Vogel, Woodman, & Luck, 2001; Zhang & Luck, 2008; but see Alvarez & Cavanagh, 2004). Other extensions (Woodman, Vecera, & Luck, 2003) have found that the grouping of multiple objects through gestalt cues leads to the storage of these objects as a single representation in VSTM. Together these investigations have revealed fundamental information about the way that low-level visual information is represented in VSTM and the present thesis strives to utilize a change detection task to begin to understand how auditory short-term memory (ASTM) processes static and dynamic stimuli and whether there are any differences in how these types of stimuli are processed.

The temporal parameters of the change detection tasks employed in this thesis (such as stimulus duration and inter-stimulus delay) are sometimes longer than those employed in psychoacoustic change detection tasks. Further, they are not within the ranges that would apply to change detection *within* a phoneme or for detecting changes between features of a stimulus such as formant transitions, or enhancing the contrast of temporally-contiguous sounds (Kluender, et al., 2003). Some of them, however, are within a range that might apply to memory *within a word* if lexical entries are stored as largely auditory exemplars (Goldinger, 1998; Klatt, 1978). Any differences that are obtained in performance on static versus dynamic stimuli will therefore not speak directly to the conflicting results regarding correlations between change detection for speech and nonspeech stimuli. They would, however, indicate that this question needs re-examination. They would set the stage for a systematic re-examination of this relationship, with the static/dynamic nature of the testing stimulus now being examined as a potentially important variable. As noted earlier, determining whether there are such differences between static and dynamic stimuli is therefore the primary goal of this work.

The temporal parameters employed are, however, in the realm of those studied in language processing and language development. Before further discussion of this, it is worth examining the nature of processing during a change-detection task. The processing underlying performance in this type of task must minimally include the ability to *encode*, *store*, and *maintain* the first stimulus across some delay, as well as *encode* and *compare* the second stimulus of the trial in order that response systems can respond in an appropriate fashion (i.e., give a same or different response) on each trial. Thus, the use of the change detection methodology introduces a memory component to the task. Again, in the visual domain, this kind of memory has been termed visual short-term memory (VSTM) and by analogy, I will utilize the term auditory short-term memory (ASTM) to describe this mechanism in the auditory change detection tasks examined here. For now this term can be thought of as a placeholder for the memory system responsible for performance on an auditory change-detection task, but as will be demonstrated it is a much more complex system than this singular term suggests.

In the change detection tasks studied in this thesis, the temporal parameters of this memory are in the range of those studied in language processing and language development. The acquisition of linguistic knowledge relies on a number of abilities most of which critically require some sort of memory component. For example, in the segmenting of words based on sentential stress patterns, the stress components of a sentence must be remembered and compared in order to extract the metrical cues needed for segmentation. Further, when linking similar or distinguishing dissimilar words within sentences, words must be held in memory and compared in order to determine similarity between words and develop linguistic knowledge that relies upon these determinations. As I mentioned earlier, change detection ability using temporal parameters of the kind employed in this work is predictive of the ability to repeat nonwords (Abbs, et al., 2007). Again, nonword repetition ability has itself been related to vocabulary development and is therefore seen as a proxy for language acquisition ability (Baddeley, et al., 1998; Gathercole, 2006) and word learning in adults (Gupta, 2003). Thus the study of ASTM in a change-detection task has relevance for the study of language processing more generally.

Even without these empirical links, the integral role of memory in many different theories of speech perception (Goldinger, 1998; Klatt, 1979; McClelland & Elman, 1986; Oden & Massaro, 1978) suggests a more detailed examination of this memory system and its properties (Schouten et al., 2003). Thus, the secondary goal of this thesis is to examine properties of ASTM, and to examine whether variables expected to affect memory in general do have an impact on change detection, as well as whether they have a differential impact on processing of static and dynamic stimuli.

In each of the three experiments in this thesis, change detection performance is compared for static and dynamic stimuli. In each experiment, a second variable is manipulated that is hypothesized to affect the memory component of performance -- in Experiment 1 the second variable is interstimulus interval (ISI), in Experiment 2 it is stimulus duration, and in Experiment 3 it is the type of change (temporal order v. frequency value) of multiple to-be compared stimuli. Type of feature change is a memory-related variable because it manipulates the information needed from memory in order to detect a change. In each of these experiments, three questions are of interest:

1) Whether there is a main effect of the static/dynamic stimulus difference. This question relates directly to the primary goal of the thesis. In each experiment, it

examines whether a processing difference can be identified for static and dynamic stimuli.

2) Whether there is a main effect of the hypothetically memory-related manipulated variables. Again, in Experiment 1 this variable will be ISI, in Experiment 2 it will be duration, and in Experiment 3 it is type of feature change. This question relates directly to the secondary goal of the thesis. To the extent that this second variable is a memory variable, such a main effect speaks to properties of ASTM in the change detection task, in much the same way that the presence of ISI and stimulus duration effects in immediate serial list recall (ISR) has been taken to indicate that information in PSTM decays over time, and that serial position effects have been taken to indicate that maintenance of ordering information is also part of the functioning of PSTM.

3) Whether there is an interaction between type of stimulus (static/dynamic) and the other manipulated variable. This question relates to both the primary and secondary goals of this work. To the extent that static and dynamic stimuli are impacted differently by the second manipulated variable, this further examine processing differences between static and dynamic stimuli. Additionally, to the extent that the second variable is a memory variable, such an interaction indicates that some of the processing differences between static and dynamic stimuli may arise specifically in ASTM.

Additionally, a fourth question examined in Experiments 1 and 2 is whether any differences between performance on static and dynamic stimuli, and any effects of the second manipulated variable, are the same for all listeners, thus inquiring into individual differences in the processing of static/dynamic stimuli and in ASTM.

CHAPTER 2

DIFFERENCES IN DETECTING CHANGES TO STATIC AND DYNAMIC AUDITORY STIMULI AT SHORT AND LONG MEMORY DELAYS

Experiment 1

Experiment 1 examines the ability to detect changes between two auditory tone stimuli that are both either static (i.e., each stimulus consists of a single frequency value throughout its duration) or dynamic (i.e., each stimulus incorporates a rising frequency value across its duration), under a variety of interstimulus interval (ISI) conditions. This is done in order to answer the primary question of the present thesis: Is change detection for static and dynamic sounds equivalent? The manipulation of ISI allows us to ask corollary questions of: Is there an effect of delay on change detection performance, are the effects of delay the same for static and dynamic stimuli, and are the effects of delay the same for all listeners? In this way we can address the main question of differences between processing of static and dynamic stimuli and more nuanced questions of how ASTM functions in a change detection task and how it is affected by variables of the task.

The levels of the ISI that will be utilized are meant to probe a wide range of possible values, rather than correspond to any particular meaningful value of speech and language. From a developmental perspective, utilization of ASTM for detecting changes in the linguistic environment will not be chained to a single temporal interval, rather the intervals between potentially relevant information may be unknown and information from a wide range of intervals may be important for learning the structure of language. Thus, in answering the question of whether delay affects memory, it is also important to estimate the performance function across a wide range of ISI values.

Figure 2.1 is a schematic of all possible types of change-detection trials at the shortest (100 ms) and longest (1800 ms) ISI for static stimuli; Figure 2.2 is the same schematic for dynamic stimuli. In both figures, the 100 ms ISI example is in the top panel and the 1800 ms ISI is in the bottom panel. Time moves from left to right in the figure, with specific time values plotted on the x-axis. Thus, stimulus one is on the left side of the panel and stimulus two is on the right side of the panel. The frequency of the stimulus at any point in time is plotted on the y-axis. A static stimulus consists of *a single frequency value* that extends over a period of time (*D*, the duration of the stimulus). A dynamic stimulus consists of *a set of frequency values that are monotonically increasing over a period of time* (also *D,* which was always the same as for static stimuli).

Two measures express this fundamental difference between static and dynamic stimuli. First, the *magnitude of the frequency change* that occurs through the duration of a particular stimulus was zero for static stimuli and nonzero for dynamic stimuli. Second, the *slope of the frequency change* (as a function of time) incorporated in a particular stimulus was zero for static stimuli and nonzero for dynamic stimuli. The two measures are, of course, related. The *slope of the frequency change* incorporated in a particular stimulus is a function of the magnitude of the frequency change and the duration of the stimulus.

The *magnitude of the frequency change* incorporated in a particular stimulus can be expressed either as the magnitude of the range covered by the starting and ending frequency, or as a proportion of the starting frequency of the stimulus. Likewise, The *slope of the frequency change* incorporated in a particular stimulus can be expressed as Hz/ms or as a percentage of the increase per millisecond from the starting frequency of the stimulus.

Although these two measures necessarily differed for static and dynamic stimuli, stimuli were constructed so that all dynamic stimuli had the same value as each other on these two measures. Expressed as a proportion of starting frequency, the *magnitude of the frequency change* was 0.0595 (i.e., 5.95%) for every dynamic stimulus (although the starting frequencies of different dynamic stimuli differed, as shown in Figure 2.2, and as will be described in detail below). Expressed as a percentage of the starting frequency, the slope of the frequency change was 0.01%/ms for every dynamic stimulus. (Having the same magnitude of frequency change for all dynamic stimuli necessarily ensured that all slopes would be the same given that all dynamic stimuli had equal duration; but the latter would not necessarily be true if the durations of different dynamic stimuli could differ).

For every static stimulus, the starting frequency was held constant throughout the duration of the stimulus, and thus the magnitude of frequency change was 0% for every static stimulus, although again, the starting frequencies of different static stimuli differed. For all static stimuli, the slope was 0%/ms, because frequency did not change throughout the stimulus's duration.

Thus static and dynamic stimuli differed in three ways: (1) *static* consist of a single frequency value while *dynamic* consist of a range of monotonically increasing frequency values; (2) *static* have a magnitude of frequency change of 0% while *dynamic* change 6%; and (3) *static* have a slope of frequency change of 0%/ms while *dynamic* have a slope of 0.01%/ms. However, within a change-detection trial, the two stimuli did not differ in any of these ways. Recall that each change-detection trial incorporated two stimuli that were either both static or both dynamic. For "same" trials, in which the first and second stimuli were identical, the two stimuli of course did not differ at all. For "different" trials, in which the two stimuli were in fact different, the difference would lie only in *what* the specific single frequency value was (for static stimuli) or what the set of frequency values was (for dynamic stimuli).

Turning to a consideration of task performance, the information that must be encoded and maintained about the first stimulus in a change-detection trial differs for static and dynamic stimuli. For static stimuli, information about a single frequency value is all that needs to be encoded and maintained. For dynamic stimuli, while there are a number of ways an abstract form of a dynamic stimulus could be represented (e.g., a mean frequency or a starting and ending frequency along with a trajectory) each of these schemes is ultimately derived from information about the range and order of frequency values. While it may not be theoretically necessary to encode the order of frequency values in order to extract a mean frequency for a dynamic stimulus, it is necessary for extracting the slope and the direction of the frequency change, an ability that humans undeniable possess. Thus, this range and order must be encoded and maintained at some level even if it is not the only way information about the dynamic stimulus is represented.

On a change-detection trial involving static stimuli, the information about the single frequency value of the first stimulus must later be compared with information about the single frequency value of the second stimulus. On a change-detection trial involving dynamic stimuli, the information about the ordered set of frequency values constituting the first stimulus must later be compared with information about the ordered set of frequency values constituting the second stimulus.

The motivation for the manipulation of stimulus type lies in the gross analogy of static and dynamic stimuli to vowels and consonants respectively, as discussed in the Introduction. A main effect of stimulus type would provide evidence for the hypothesis that in the literature, the findings of lack of correlation between the processing of non-speech and speech stimuli might be due to studies having focused primarily on static stimuli. It would also indicate that ASTM is sensitive to the differences between such stimuli.

The four specific questions asked in Experiment 1 are as follows: 1.) ***Is there a main effect of stimulus type (static vs. dynamic) on change detection performance?*** This question directly addresses the primary goal of this thesis. A main effect of stimulus type would support the hypothesis that in the literature, the findings of lack of correlation between the processing of non-speech and speech stimuli might be due to studies having focused primarily on static stimuli.

2.) ***Is there a main effect of ISI on change detection performance?*** This question addresses the secondary goal of the thesis. Manipulating ISI requires listeners to compare stimulus representations at different delays beginning with almost no delay (100 ms) and ending with a long delay (1800 ms). In order to discriminate two stimuli, the acoustic features from the first stimulus must be encoded in ASTM when it is first presented, and be made available for comparison (either through maintenance across the delay or reactivation) when the second stimulus is presented. An effect of ISI would suggest that ASTM is subject to decay over time, in much the same way that the detrimental effect of longer ISIs in ISR have been taken to indicate that PSTM is subject to trace decay (Baddeley & Lewis, 1984).

3.) ***Is there an interaction between stimulus type and ISI?*** This question addresses both the primary and secondary goals of the thesis. If there is no interaction, then any main effect of stimulus type would likely reflect a difference in the representation of static vs.dynamic stimuli – one that is present even at the 100 ms ISI and that does not change with ISI. An interaction would indicate, however, that memory demands affect static and dynamic stimuli differentially, providing further evidence of differential processing of static and dynamic stimuli as well as further information about the properties of ASTM.

4.) ***Are the effects of ISI and stimulus type equivalent for all listeners?*** If it is the case that all listeners do show effects of ISI and/or stimulus type, then this will indicate that there is something intrinsic to the variable that makes it functionally incompatible with ASTM (i.e., ASTM has difficulty processing because its capabilities and the demands of the stimulus are incompatible). On the other hand, if it is the case that certain groups of listeners do not show effects of ISI and/or stimulus type then the effects of these variables cannot be intrinsic to the variable themselves. Rather, a more accurate description of the effect will be one that depends on the function of the particular ASTM that is processing the stimulus.

For example, in a population of monolingual speakers there will be individual differences in the exact words present in each speaker's vocabulary thus some speakers will have difficulty recognizing a word not in their vocabulary, but other speakers may have no problem at all. Words from a foreign language cannot be processed by any speakers however (assuming the language is dissimilar enough from the native language), as the language system will (theoretically) not have a way of recognizing these words. Thus, native language recognition reflects variation in the individual vocabularies of the population while foreign language recognition reflects variation in the group's vocabulary and reflects an intrinsic incompatibility between the word and the vocabulary of the group.

In the same way, individual variation in a change detection task reflects individual differences in ASTM, but any group-level differences (e.g., similar differences in change-detection ability for static and dynamic stimuli for all listeners) would reflect an inherent limitation of ASTM. Processing of dynamic stimuli can only be related to speech perception if individual differences in the ability to process dynamic stimuli are found (the previously reviewed studies inform us that individual differences in speech perception do exist). Thus, studying individual differences will provide information as to the natural variation of change detection ability in randomly selected listeners, which in turn provides information for future studies on the relationship between abilities on these tasks and speech perception.

Performance on the change detection task can be evaluated in three ways. First, overall accuracy in the form of a proportion correct score evaluates how well participants can detect changes on all trials (i.e., on both "same" and "different"). Second, the Hit Rate can evaluate performance just on "different" trials, where a change was actually made to the stimulus. This measures the participant's ability to accurately detect changes when changes are made to the stimulus. Third, the False Alarm Rate can evaluate performance just on "same" trials, where no change was made to the stimulus. This measures the participant's overall propensity toward responding with a "change" response. Each of these three measures will be used in Experiment 1 in order to fully understand performance in this task.

Method

Participants

Seventy undergraduates (40 = Female; 30 = Male) from the University of Iowa participated in this experiment in exchange for course credit.

Stimuli

Stimuli were sine wave tones, rising tones, and masking white noise. Sine waves and noise were generated using MATLAB® (The MathWorks, Inc., 2007, Natick, MA). Rising tones were generated using Tone Generator (NCH Swift Sound, 2005, Canberra, Australia) by specifying a start frequency, an end frequency, and a duration.

Frequency values for the stimuli were determined by choosing a starting frequency that was not associated with a note of the chromatic musical scale, but was within or near the range of frequencies that are most efficiently transmitted by the ear, 500 – 4000 Hz (Moore, 1997, p.18). From an arbitrarily chosen starting frequency of 503 Hz, 43 additional tones were created by increasing the frequency of the tones in 80 Hz increments. This procedure resulted in a large set of stimuli, ranging from 503 Hz – 3943 Hz, whose members were a constant frequency distance from and had no harmonic relationship with one another (see Table A.1 in Appendix A). Also listed in Table A.1 is the "Bark Value" for each stimulus. The bark value refers to the critical band along the basilar membrane that is thought to process a particular frequency based on our knowledge of basilar membrane function. Each frequency is transformed according to the equation:

(1) $$Bark = 13 \bullet \arctan\left(\frac{0.0076f}{1000}\right) + 3.5 \bullet \arctan\left(\frac{f}{7500}\right)^2$$

where f is the frequency being transformed. A stimulus is thought to be processed by a different critical band if its bark value is a different integer value (i.e., 3.9 v. 4.0) than another frequency.

For the static condition, the presented tones started at the values given in Table A.1 and maintained this frequency across a 500 ms presentation duration. For the dynamic condition, the tones started one-half step, or 50 cents, lower than the corresponding static stimulus and then increased one step on a musical scale, or 100 cents, across the presentation. Thus, a dynamic stimulus starts lower than the corresponding static stimulus, passes through the frequency of the static stimulus halfway through the presentation, and then ends at a higher frequency value (for the resulting frequency values of these stimuli see Table A.2). Since the magnitude of the change is always 100 cents, the increase in frequency is normalized by adjusting for the logarithmic relationship between low-and high-frequency sounds (and thus, discrimination should follow Weber's law and no change should be perceived in the magnitude of two stimuli with different starting frequencies). There will be larger magnitudes in terms of frequency range for high frequency sounds, but in terms of the percentage of the starting frequency (5.95%), the magnitude will be the same. Further, the slope value of these increases (in Hz/ms) will be larger for higher frequency sounds, but not in terms of percentage of the starting frequency (0.01% per ms). These values are also listed in Table A.2 and again it should be pointed out that the magnitude of frequency change (5.95%) and the slope of the frequency change (0.01%/ms) was the same for all dynamic stimuli, so that the only cue to change detection between the two stimuli on a change-detection trial involving dynamic stimuli was the particular set of frequency values comprising the two stimuli – just as on trials involving static stimuli. For the masking noise, there were

44 different white noise stimuli corresponding to 44 different random number vectors used during stimulus generation.

The static and dynamic stimuli were presented at approximately 75 dB. Participants were permitted to adjust the level of the stimuli if they found the starting level uncomfortable. Such an adjustment did not change the signal-to-noise ratio, however, which was determined by the experimental presentation software, PsyScope X (Cohen, MacWhinney, Flatt & Provost, 1993). This software presented the noise at full volume (255 units) and the tones at a reduced (150 units) volume. The reduced volume of the tones is approximately 60% lower than the white noise.

Design and Procedure

The main manipulation of ISI was made by block, with participants completing 100 change-detection trials within a single block at a single ISI value. The ISI values at which participants were tested were 100, 225, 450, 900, and 1800 ms. Within the 100 trials in a block, participants were randomly presented with 50 trials that involved static tones and 50 trials that involved dynamic tones. Within these 50 trials per stimulus type, 25 trials involved a random change in the frequency of a stimulus (i.e., "different" trials) and 25 trials involved the same stimulus presentation (i.e., "same" trials). Participants were given a mandatory 30 second break every 20 trials. Participants were also given 10 practice trials at the beginning of each block. Sequential effects of block were controlled through a Latin-Squares design.

To reduce the stimulus specificity of any effects seen in this design, a different sequence of trials, utilizing randomly selected frequencies, was created for each of the counterbalancing conditions, yielding five *stimulus sets*. Each of these sets can additionally be analyzed as a function of ISI, where again random selection means that participants are getting a different sequence of trials utilizing randomly selected frequencies. In the analysis to follow, it will be important to ensure that any effect of *stimulus set* (the specific combination of frequencies and differences between frequencies a participant received) has been removed from the performance of participants.

In order to utilize the whole stimulus set on "different" trials and reduce the predictability of the change component of this task, the amount of frequency change for each "different" trial was randomly determined. To ensure that these differences are not driving any subsequent differences between the variables of interest that might be found, a 2 (stimulus type) X 5 (ISI) analysis of variance (ANOVA) was conducted on the frequency differences on "different" trials of the 5 stimulus sets and found no significant difference between the amount of change as a function of stimulus type, F (1, 4) < 1, or as a function of ISI, F (4, 16) < 1, and there was no interaction between the two F (4, 16) < 1. On average the frequency change for the dynamic stimuli was approximately 1200 Hz and the change for the static stimuli was approximately 1160 Hz, a difference of less than 80 Hz. The average amount of change as a function of ISI varied between approximately 1120 and 1210, a difference of approximately 90 Hz. The large changes in frequency ensures easy comparison of stimuli if the stimulus is encoded and maintained in ASTM.

Given that the beginning frequency values of the dynamic stimuli were also randomly determined there exists the possibility that consecutive stimuli on a given trial simply overlapped and this overlap could drive differences between the stimulus variable of interest. This confound will be addressed in the analysis presented below.

The procedure for the analysis of Experiment 1 will be to first examine overall performance as a function of ISI and stimulus type as measured by proportion correct. Second, we will eliminate explanations of these results that would rest on particular characteristics of the data collected (i.e., ceiling effects and stimulus effects that are specific to the present experiment). Third, we will verify that the same effects occur for both hit rate and false alarms to understand the role of decision processes in these data. Fourth, we will address the frequency overlap issue described above. Fifth, we will examine individual differences in performance on this task as measured by proportion correct.

Results and Discussion

Overall Behavioral Performance

Overall performance was high in this task, especially in the case of static stimuli; however there was a significant difference between static and dynamic stimuli and a significant effect of ISI (Figure 2.3). A 2 (Stimulus Type) X 5 (ISI) Repeated-Measures ANOVA conducted on the proportion correct measure (see Figure 2.3) revealed a main effect of stimulus type, F (1, 69) = 88.95, $p < .05$, $\eta^2 = .56$. This result indicates that change detection for static and dynamic sounds is not equivalent. Further, a main effect of ISI, F (4, 276) = 6.22, $p < .05$, $\eta^2 = .08$, and a significant interaction, F (4, 276) = 4.61, $p < .05$, $\eta^2 = .06$ indicate that there is a negative effect of delay on change detection performance (i.e., the effect of ISI is to further degrade this memory) and that the effect is not the same for both types of stimuli[3].

The effect of ISI is marginally linear, F (1, 69) = 3.65, $p = .06$, $\eta^2 = .05$, with a large cubic component, F (1, 69) = 19.27, $p < .05$, $\eta^2 = .22$, owing to the inflection point introduced by performance at the 225 ms level. No quadratic component is present F (1, 69) < 1. This function does mean that a general decrease is expected with ISI; thus, while the decrease at 225 ms is contributing to the main effect of ISI, so is the decrease at longer ISI intervals.

---

[3] Unless otherwise reported, violations of sphericity, when present, did not change the significance level when corrected using the Greenhouse-Geisser correction and uncorrected statistics are presented.

Overall these results demonstrate that the answer to the main question of interest: **1.)** ***Is there a main effect of stimulus type (static vs. dynamic) on change detection performance?*** is a resounding "Yes". Change detection with dynamic stimuli is less accurate than change detection with static stimuli, no matter how short the delay interval (Figure 2.3)**.** This effect was large and it suggests that ASTM is sensitive to the difference between changing and unchanging stimulus information, at least for stimuli with a duration of 500 ms.

In terms of the manipulation of ISI, the answer to the main question here: **2.)** ***Is there a main effect of ISI on change detection performance?*** is also "Yes". The function of this decrease is not simply linear however and there was an unexpected drop in performance at the 225 ms level. Further, the answer to another question of the present study: **3.)** ***Is there an interaction between stimulus type and ISI?*** is also "Yes", providing a more nuanced view of how static and dynamic auditory stimuli are processed. Specifically, change detection for dynamic but not static stimuli decreases as a function of increasing ISI, at least in the ISI range 100-1800 ms. This finding suggests that degradation of information occurs in ASTM over a delay for dynamic but not for static stimuli. Even for dynamic stimuli, however, the memory degradation as a function of ISI does not emerge until the longest delay of 1800 ms. Before turning to the final question of individual differences, a number of different aspects of the present set of data need to be further examined in order to further validate any conclusions that might be drawn.

Ceiling Effects

Given that performance of listeners in this task is near ceiling, there is the chance that the extremity of these values is masking effects of ISI on change detection for static stimuli. To rule out this interpretation an arcsin transformation of the data was completed according to the formula, $\arcsin\left(\sqrt{p}\right)$, where p is accuracy in terms of proportion correct. A 2 X 5 Repeated Measures ANOVA on these transformed data yielded the same results.

That is, there was a significant main effect of stimulus, *F* (1, 69) = 88.78, $p < .01$, $\eta^2 = .56$, a main effect of ISI, *F* (4, 276) = 5.42, $p < .01$, $\eta^2 = .07$, and a significant interaction, *F* (4, 276) = 7.53, $p < .05$, $\eta^2 = .10$. For the trend analysis, there was again a significant cubic component, *F* (1, 69) = 20.53, $p < .05$, $\eta^2 = .23$, but no linear, *F* (1, 69) = 1.96, $p = .17$, $\eta^2 = .03$, or quadratic, *F* (1, 69) < 1.

Stimulus Set Effects

A visual inspection of the data plotted as a function of stimulus set showed three data points that are lower than performance at other levels of our experimental design: the static stimuli at 225 ms and 1800 ms as well as the dynamic stimuli at 1800 ms. As stated earlier, a stimulus set is a between-groups factor that denotes the exact combination of frequencies and differences between frequencies on change trials. It is important to validate that the effects seen so far are not simply an effect of a particular stimulus set, which is a nuisance variable in this experiment. Determining that this factor does not have an effect on performance will allow a conclusion that the effects seen here are reliable effects of stimulus type and ISI, the two variables of interest and not the particular stimuli presented in the respective conditions.

Two of these drops in performance were predictable (static and dynamic stimuli at 1800 ms) and one was not. In order to ensure the validity of both effects and establish that they are both truly a function of ISI and not a function of any stimulus-specific effects (i.e., the particular combination of frequencies and differences in frequency) that may be present in our design at particular ISIs, I initiated an analysis of performance as a function of ISI and stimulus set for each of our two stimulus types.

For this analysis, a 5 (Stimulus Set) X 5 (ISI) Mixed-Design ANOVA was conducted with proportion correct on static stimuli as the dependent variable (see Figure 2.4). There was no overall effect of stimulus set, *F* (4, 65) < 1, however there was an interaction between stimulus set and ISI, *F* (16, 260) = 6.40, $p < .05$, $\eta^2 = .28$. A 5 X 5

ANOVA for dynamic stimuli again revealed no overall effect of stimulus set, *F* (4, 65) = 1.60, $p > .05$, but an interaction between stimulus set and ISI, *F* (16, 260) = 6.21, $p < .05$, , $\eta^2 = .28$ (Figure 2.5). Stimulus set is not an overall factor in determining performance for static or dynamic stimuli, thus the overall difference as a function of ISI cannot be attributed to stimulus set.

In terms of the ISI by stimulus set interaction for static stimuli, a visual inspection of the data for the stimulus set X ISI interaction for static stimuli (see Figure 2.4) shows that the interaction is driven by set specific effects (i.e., just 1 stimulus set performing at a different level than the 4 others) at the 900 ms ISI and 1800 ms ISI. Thus, while these effects were marginally significant they do not seem reliable. At the 225 ms ISI, the ISI of main interest, 3 of the stimulus sets show lower performance while 2 of these sets show performance relatively in line with other stimulus sets at other ISIs. The sets with the smallest difference between them (sets B & D, difference = 5.71%) were significantly different according to a two-tailed independent samples *t*-test, *t* (26) = 2.57, $p < .05$, d = 1.00, suggesting a bimodal distribution between these two clusters of performance levels. However, 4 of 5 sets (including set B) show a decrease in performance from the shortest ISI, 100 ms, to the current ISI, 225 ms. Thus, while there are absolute differences between stimulus sets at particular ISIs, the relative differences (which control for individual differences between participants) suggest a consistent drop in performance at 225 ms across all stimulus sets.

For the ISI by stimulus set interaction for dynamic stimuli, the interaction appears more complicated, but also driven by single stimulus set differences at 225, 450, and 900 ms ISIs (see Figure 2.5). At the time point of main interest, 1800 ms, again 3 of 5 stimulus sets show performance at a similar level and 4 of 5 stimulus sets show a decrease in performance from the previous ISI level. However, a two-tailed independent samples *t*-test did not show a significant difference between the groups with the smallest difference between them (B & D, 4.00% difference), t (26) = 1.09, $p > .05$, d = .40. Thus, there is a wide range of performance levels, but not a bimodal distribution of scores.

The results of the cross-sectional analyses at both the 225 ms ISI with static stimuli and the 1800 ms ISI with dynamic stimuli suggest a consistent, reliable effect as a function of ISI and not particular stimulus sets or individual differences. The effects at 900 and 1800 ms for static stimuli do seem to be the result of a single particular stimulus set.

False Alarms

The number of "different" responses on "same" trials (i.e., false alarms), provides information as to how similarly stimuli are perceived when they are the same. These results are depicted in Figure 2.6. A 2 (Stimulus) x 5 (ISI) ANOVA revealed a main effect of stimulus, F (1, 69) = 7.61, $p < .05$, $\eta^2 = .10$, suggesting that overall two of the same dynamic stimuli are perceived as *more different* than two of the same static stimuli or there is greater decision noise for dynamic stimuli. There was also a main effect of ISI, F (4, 276) = 3.69, $p < .05$, $\eta^2 = .05$ and no significant interaction, F (4, 276) < 1 meaning that for both dynamic and static stimuli, the longer the interval between two stimuli, the *more different* they are perceived to be or there was greater noise in the decision process. Importantly, at no time are false alarms for dynamic stimuli lower than that of static stimuli, thus at no time are dynamic stimuli perceived as more similar than static stimuli. This indicates that the overall lower change-detection performance for dynamic stimuli is not simply an effect of perceptual similarity. Thus, the overall lower change-detection performance for dynamic stimuli may indicate poorer encoding/maintenance of dynamic (as compared to static) stimuli in ASTM.

Hit Rate

As shown in Figure 2.7, the hit rate results closely mimic those of proportion correct. The results of a 2 (Stimulus Type) x 5 (ISI) repeated measures ANOVA also mimic the results of proportion correct with a main effect of stimulus, F (1, 69) = 84.38, $p < .05$, $\eta^2 = .55$, main effect of ISI, F (4, 276) = 5.99, $p < .05$, $\eta^2 = .08$, and an interaction of stimulus and ISI, F (4, 276) = 4.44, $p < .05$, $\eta^2 = .06$. The means of the significant interaction mimic those of proportion correct, thus I may infer that the same analysis of performance in terms of proportion correct as a function of both stimulus and ISI holds in terms of the hit rate.

Together, the hit rate and false alarm data suggest that the effect of stimulus type on proportion correct is due to differences in the memory for static and dynamic stimuli, not differences in the perceived similarity between stimuli. That is, **the answers given above to Questions 1-3 reflect differences in *the encoding, maintenance, and/or comparison* processes in change detection for dynamic vs. static stimuli, *not* simply greater perceived similarity of the dynamic over the static stimuli.**

Overlapping Stimuli

Dynamic stimuli had ranges of frequency values that could result in the overlap between the actual frequency values of two "different" stimuli on "different" trials. For example, if the stimulus that began at 1732 Hz and ended at 1835 Hz was given as the first dynamic stimulus and the second, "different", stimulus began at 1809 Hz and ended at 1918 Hz then there would be overlap in these stimuli from 1809 to 1835 Hz. Given that stimuli were randomly selected, this overlap was not controlled and could potentially account for any observed differences between static and dynamic stimuli. Thus, I analyzed the number of times such overlap occurred in Experiment 1 in order to rule out overlap as the sole cause of differences in accuracy between static and dynamic stimuli. On average, overlap occurs on approximately 4.16% of all "different" trials with dynamic stimuli. In Experiment 1, participants correctly detected changes to dynamic stimuli with a hit rate of 88.78% and detected changes to static stimuli with a hit rate of 94.67%. Thus, the 5.9% difference in hit rate cannot be *wholly* explained by the overlap of stimuli in the case of dynamic, but not static, stimuli. Further, an overlap explanation cannot explain the differences in false alarms found above.

In order to further eliminate the overlap of stimuli as a potential explanation for these data, I looked at individual conditions that had a large number of dynamic change-detection trials with overlap and those that had a small number of trials with overlap. In terms of high numbers, stimulus set E – 900 ms ISI, stimulus set B – 900 ms ISI, and stimulus set A – 1800 ms ISI, each had the highest proportion of overlapping trials (16, 12, and 12% respectively). In terms of low numbers, stimulus set E – 225, 450, and 1800 ms ISI, stimulus set D – 100, 225, and 900 ms ISI, stimulus set B – 450 and 1800 ms ISI, and stimulus set A – 100 ms condition, each had the lowest proportion of overlapping trials (0%). The average proportion correct for the "High" overlap set on change detection for dynamic stimuli was 93% and the average proportion correct for the "Low" overlap set was 93.29%. Looking just at "different" trials, the average hit rate for dynamic stimuli in the "High" set was 88.95% and the average hit rate for the "Low" set was 88.92%. These "Low" hit rates are biased towards including short ISIs, but eliminating scores from ISIs less than 900 ms only changes the hit rate to 89.81%. Thus, in none of the above analyses is performance on change detection for dynamic stimuli in the "High" overlap stimulus sets ever more than 1% lower than performance for "Low" overlap stimulus sets. I can thus conclude that the overlap of dynamic stimuli on "different" trials is not the sole cause of differences between accuracy in change detection for static and dynamic stimuli and likely plays little role in such differences.

Individual Differences

In order to further understand performance in this task and examine whether the findings of (1) more accurate change detection performance for static stimuli, (2) less accurate change detection performance with longer ISIs, and (3) less of a decrease in change detection performance as a function of ISI for dynamic than static stimuli hold for all subjects, individual differences in performance were assessed. Individual differences in performance were analyzed by sorting participants according to the proportion correct at the 1800 ms ISI condition averaging across stimulus type. Thus, I separated listeners based on who had the *best change-detection performance* for both static and dynamic stimuli. I wished to examine whether this individual difference would be predictive of performance differences at other ISIs. Those listeners whose score fell below the median (at the 1800 ms ISI) were classified as 'low sensitivity' and those whose score fell above the median at the same ISI were classified as 'high sensitivity'. The median for this sample was 94.5% (range = 83 – 100%) and a division based on this median resulted in 36 participants in the high-sensitivity group and 36 in the 'low-sensitivity group.

Differences between these groups' performance was assessed with a 2 (Sensitivity) X 2 (Stimulus Type) X 4 (ISI) Mixed-Design ANOVA, which left out the performance upon which these groups were sorted (i.e., 1800 ms; Figure 2.8). For this analysis, there was first a significant main effect of sensitivity, F (1, 68) = 25.66, $p < .05$, $\eta^2 = .27$. The main effect of sensitivity indicates that the selection of participants based on performance at the 1800 ISI did yield two groups who showed different discrimination abilities when performance was collapsed across all stimulus types and the remaining 3 ISI levels (see Figures 2.9 & 2.10). Even though the low-sensitivity group was selected based on their change-detection performance at long delays, they also show poorer change detection, more generally. Thus, separating the groups based on performance on 1800 ms ISI reveals a more general decrease in accuracy across ISIs and stimulus types.

Separating the groups in this way also selected for a group that was less affected by the manipulation of stimulus as there was a significant sensitivity by stimulus type interaction, F (1, 68) = 10.40, $p < .05$, $\eta^2 = .13$. Participants in the high-sensitivity group were not only more accurate with greater delays, but they were also more accurate on dynamic stimuli and showed less of a decrease in change detection performance as a function of stimulus type. Thus, there was a more general decrease in accuracy from high- to low-sensitivity listeners. The characterization of these differences as a more general decrease in accuracy between groups is further supported by a significant main effect of ISI, *F* (3, 204) = 5.86, $p < .05$, $\eta^2 = .08$, but no sensitivity by ISI interaction, *F* (3, 204) < 1, and no 3-way interaction, *F* (3, 204) < 1. (The main effect of stimulus type and the interaction between stimulus type and ISI was the same as reported earlier).

In order to explore the interaction between stimulus type and sensitivity, a *t*-test was completed on the average accuracy across the 100, 225, 450, and 900 ms ISIs for both high-sensitivity listeners. There was a significant difference for both high-, *t* (34) = 4.32, $p < .05$, d = .73 and low-sensitivity listeners, *t* (34) = 7.76, $p < .05$, 1.31. Thus, stimulus type does have an effect on both groups and the interaction points to a difference in the size of the effect between the two groups with an almost 2 x larger effect in low- compared with high-sensitivity listeners. These data indicate that the effect of stimulus type is an individual difference and not a functional incompatibility with ASTM. Thus, one's ability to process dynamic stimuli is an individual difference that could potentially relate to individual differences in language processing.

Again, because of concerns with ceiling effects, especially in the case of high-sensitivity listeners, an arcsine transform of the data was completed and the analysis was re-run. These results did not differ from the previous results: main effect of sensitivity, *F* (1, 68) = 34.91, $p < .05$, $\eta^2 = .34$, a main effect of ISI, *F* (3, 204) = 6.38, $p < .05$, $\eta^2 = .09$, a sensitivity by stimulus type interaction, *F* (1, 68) = 5.78, $p < .05$, $\eta^2 = .08$, but no sensitivity by ISI interaction or a 3-way interaction, *F*s (3, 204) < 1.

The answer to the final question of **4.)** ***Are the effects of ISI and stimulus type equivalent for all listeners?*** is "No" and "No". First, I was able to identify a subset of listeners that were significantly better than the remaining listeners at change detection with an 1800 ms ISI. Thus, ISI does not affect all listeners in the same way, some show less of a decrease in change detection performance with increasing ISI. Second, this analysis also revealed an interaction with stimulus type that indicates that the effect of stimulus type is not the same for all listeners. Together, the individual differences data demonstrate that the effects of ISI and stimulus type are not due to functional limitations of ASTM in general, rather they are related to individual differences in its functioning, which could theoretically contribute to individual differences in speech and language capabilities. Further, these individual differences create a relatively large range of performance levels that one could attempt to relate to speech perception in the future. One could also attempt to look for group differences on speech-related tasks after a larger sample has been split based on performance at the 1800 ms ISI level.

Since these groups were not randomly assigned to counterbalancing conditions, the potential effects of counterbalancing were examined. The counterbalancing conditions of the participants in each group were examined. Participants in the high-sensitivity group saw the 1800 ms condition in the 4th task position more often than the low-sensitivity group. Participants in the low-sensitivity group saw the 1800 ms condition in the 3rd task position more often. Given the simplicity of the task, it is unlikely that practice effects would be contributing to any differences between groups. Further, the stimulus sets corresponding to these particular orders, C & D, did not show any anomalous performance at the 1800 ms block (see Figure 2.4 & 2.5). Thus, the individual differences appear to be real differences in change detection and not artifacts of particular stimulus sets or particular task sequences.

Summary

Experiment 1 provides a concrete answer to our main goal of examining whether change detection for static and dynamic stimuli is equivalent. Change detection was found to be not equivalent as the accuracy of change detection for static stimuli is greater than that for dynamic stimuli at all ISIs. (Q1) Further, there was a main effect of ISI, suggesting that the ASTM underlying performance in the change detection task may be subject to decay. This addressed the secondary goal of the thesis, and answered Q2.

Experiment 1 also provides evidence that the effects of delay are not equivalent for static and dynamic stimuli, addressing both the primary and secondary goals of this study, and answering Q3 Here it was found that change detection for dynamic, but not static, stimuli is lower at the longest ISI (1800 ms), so that the change-detection accuracy difference between static and dynamic stimuli is exacerbated at this longest ISI. This finding suggests that degradation of information (presumably decay) occurs in ASTM at the longest delay, for dynamic but not for static stimuli. It also appears that change detection for static stimuli can be disrupted with a medial delay of 225 ms. Although this effect was found for 3 out of 5 stimulus sets, it was not found for 2 of the stimulus sets, so it cannot be ruled out entirely that this is an idiosyncratic effect of the exact stimuli presented here.

The analyses of false alarms on “same” trials indicates that “different” responses to identical dynamic stimuli are higher than “different” responses to identical static stimuli. This suggests that the overall lower change-detection performance for dynamic stimuli is unlikely to be simply an effect of perceptual similarity, rather it indicates poorer ASTM for static dynamic than for static stimuli.

Lastly, Experiment 1 provides evidence that the effects of stimulus and delay are not the same for all listeners, answering Q4. The individual differences in change detection indicate that some listeners (the high-sensitivity group) have good change-detection ability (and presumably ASTM) for both static and dynamic stimuli that stays intact over increasing ISIs. The low-sensitivity group of this experiment showed poorer discrimination at all four of the other ISIs, which had not been part of the selection criteria for the low-sensitivity group. Further, there was a greater discrepancy between the discrimination of static versus dynamic stimuli. These results suggest that the low-sensitivity listeners had a poorer ASTM for both static and dynamic stimuli, which was most evident in the discrimination of dynamic stimuli and (based on their inclusion in the low-sensitivity group) after a long delay.

However, these findings rest on data from just one type of dynamic stimulus. For speech stimuli, static and dynamic sounds take on a variety of durations, magnitudes, and slopes. Thus, dynamic stimuli can be defined in a number of different ways. In Experiment 2, I consider whether change detection performance is better for static stimuli than for a number of different types of dynamic stimuli.

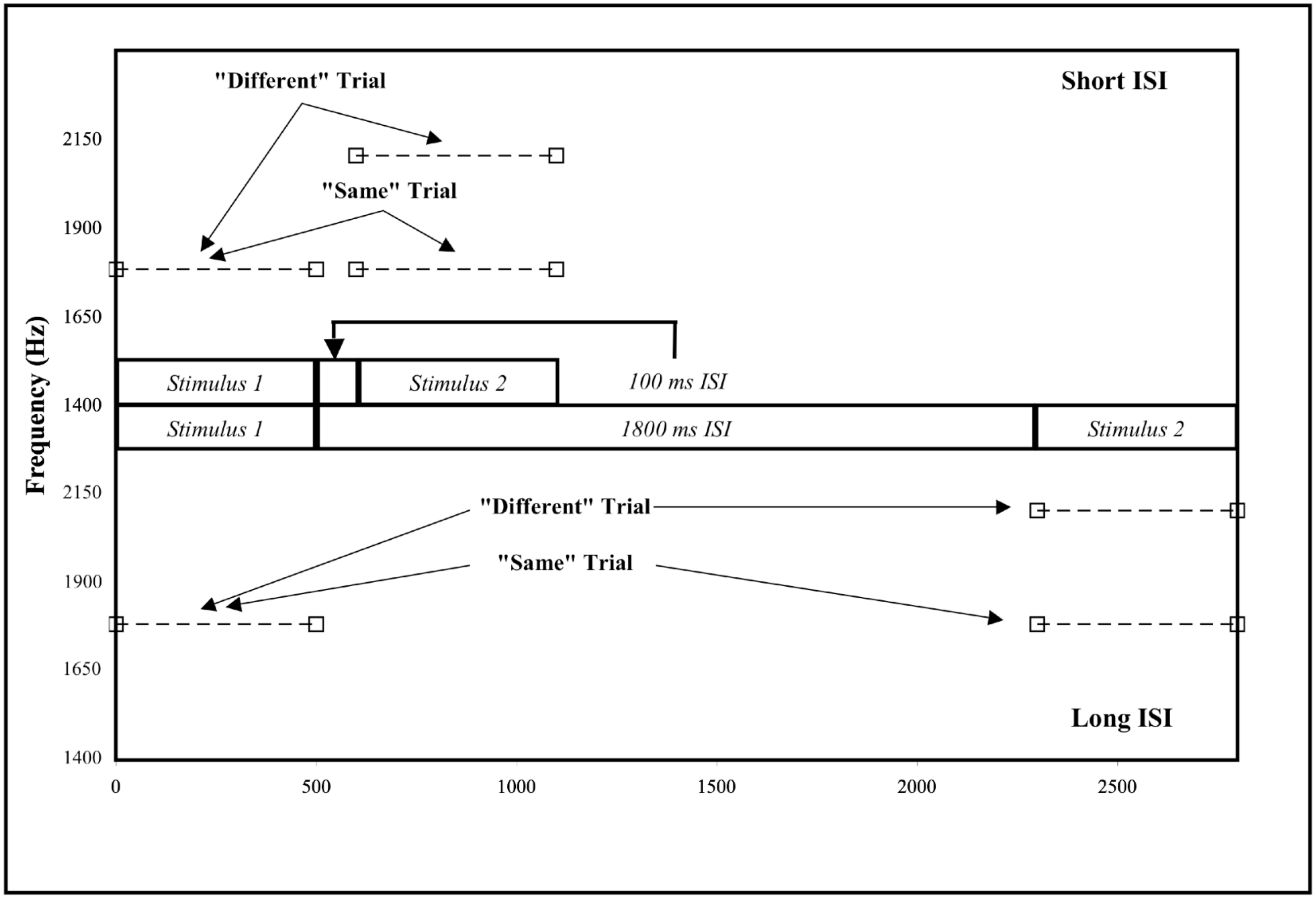

Figure 2.1. A schematic depiction of the to-be-compared stimuli on a hypothetical change trial of static stimuli in Experiment 1.

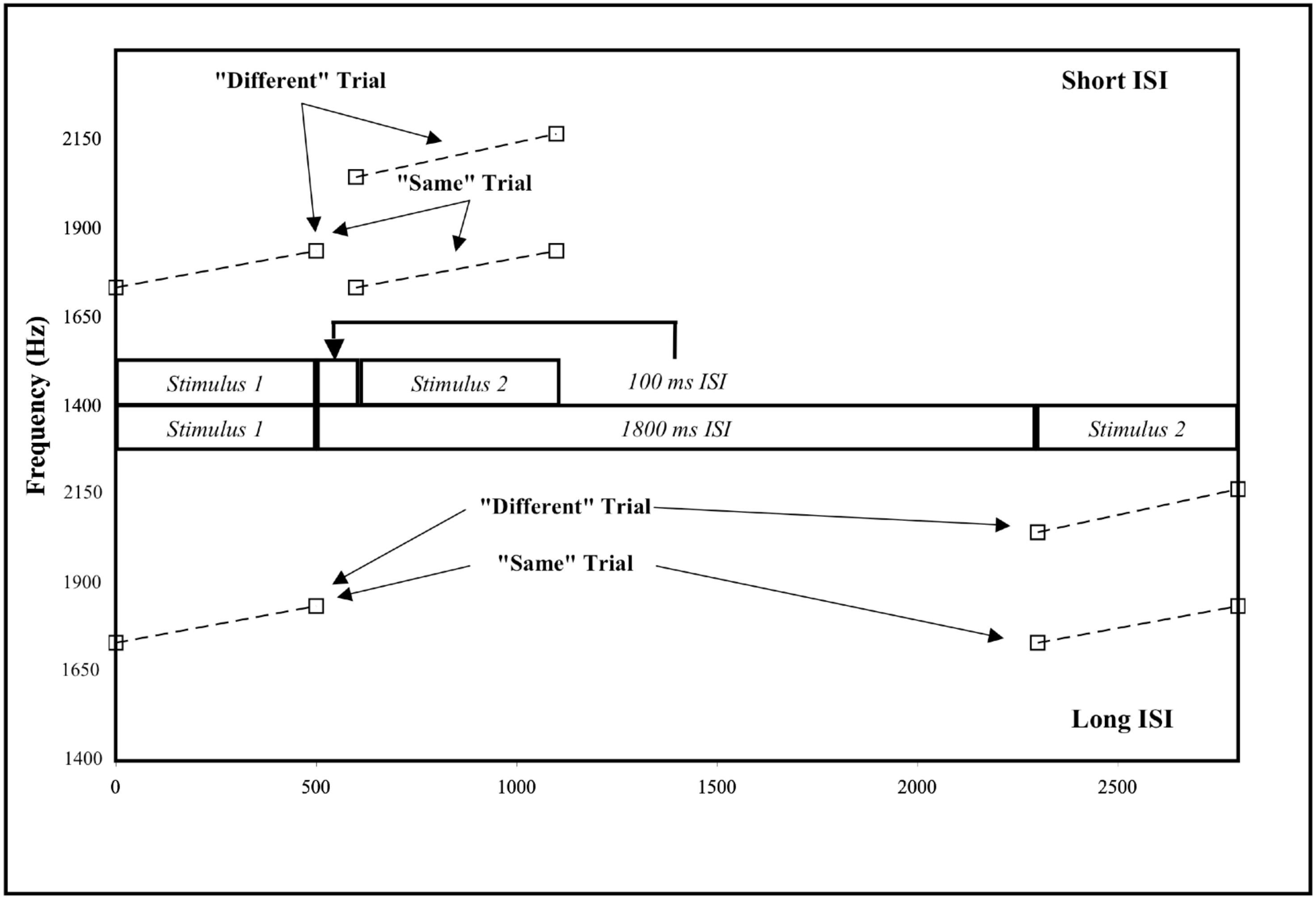

Figure 2.2. A schematic depiction of the to-be-compared stimuli on a hypothetical change trial of dynamic stimuli in Experiment 1.

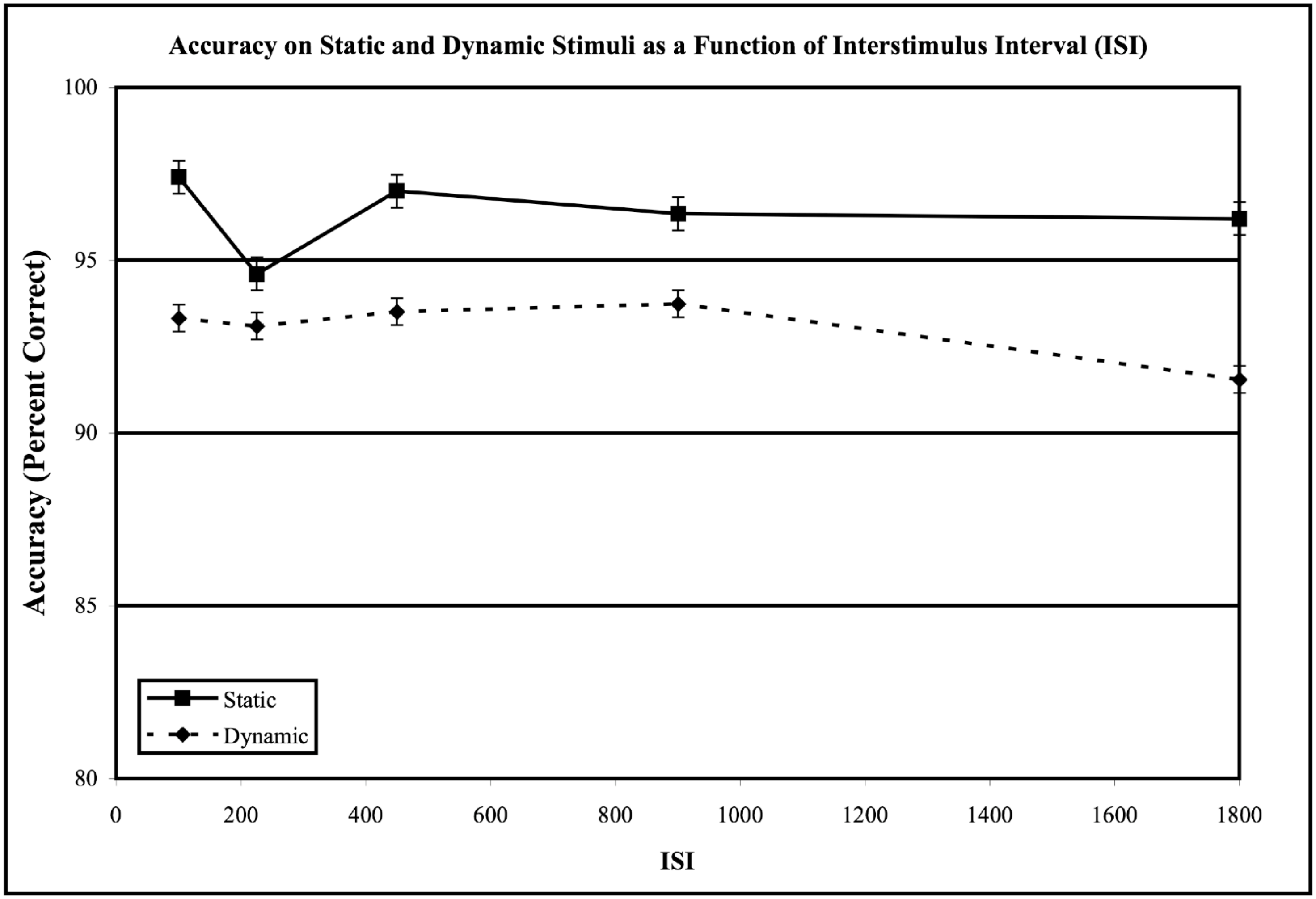

Figure 2.3. Overall percent correct for static and dynamic stimuli as a function of ISI. Bars are SEM for the parameter (static or dynamic) of the graph.

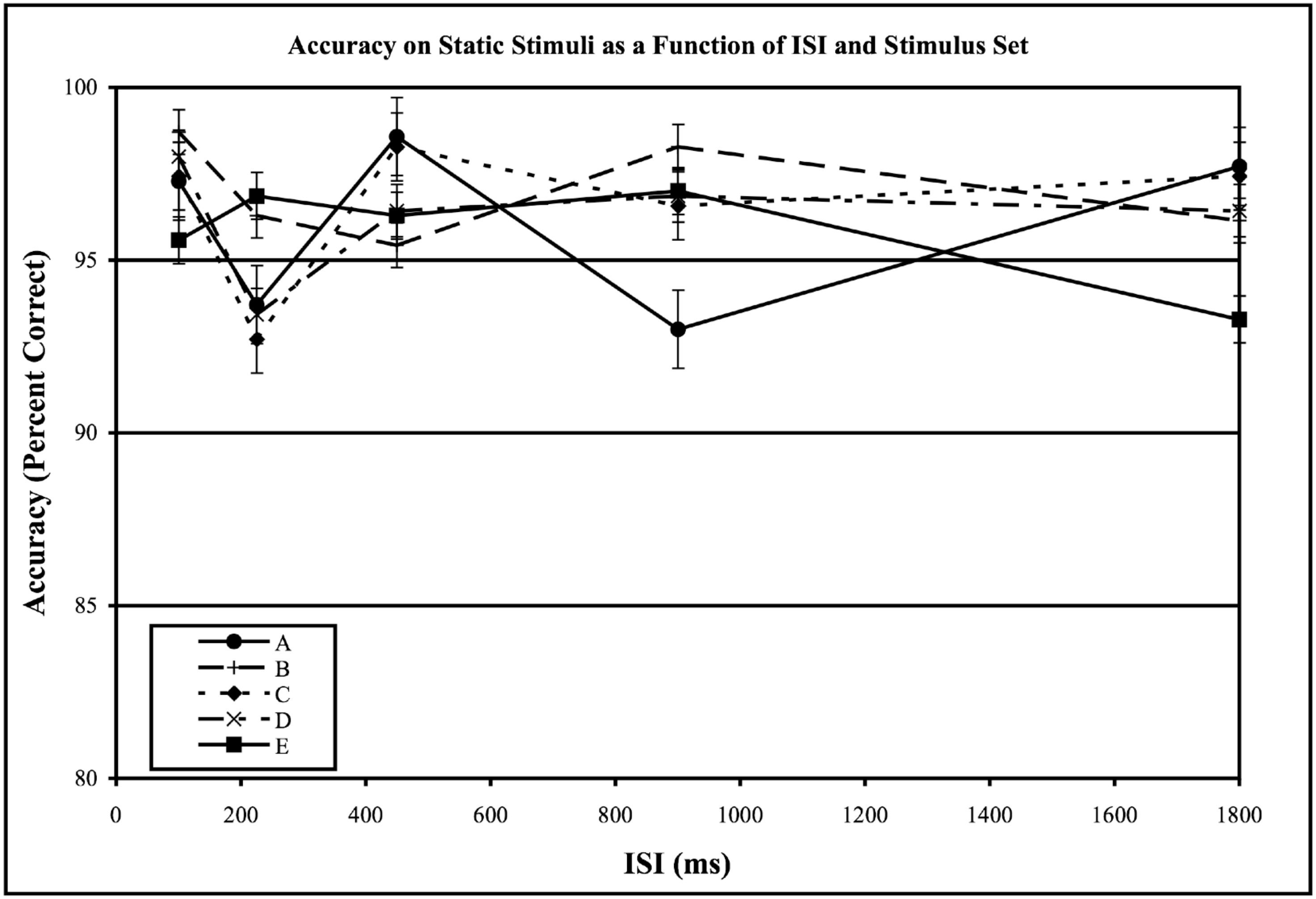

Figure 2.4. Percent correct for static stimuli as a function of stimulus set. Bars are SEM for the parameter.

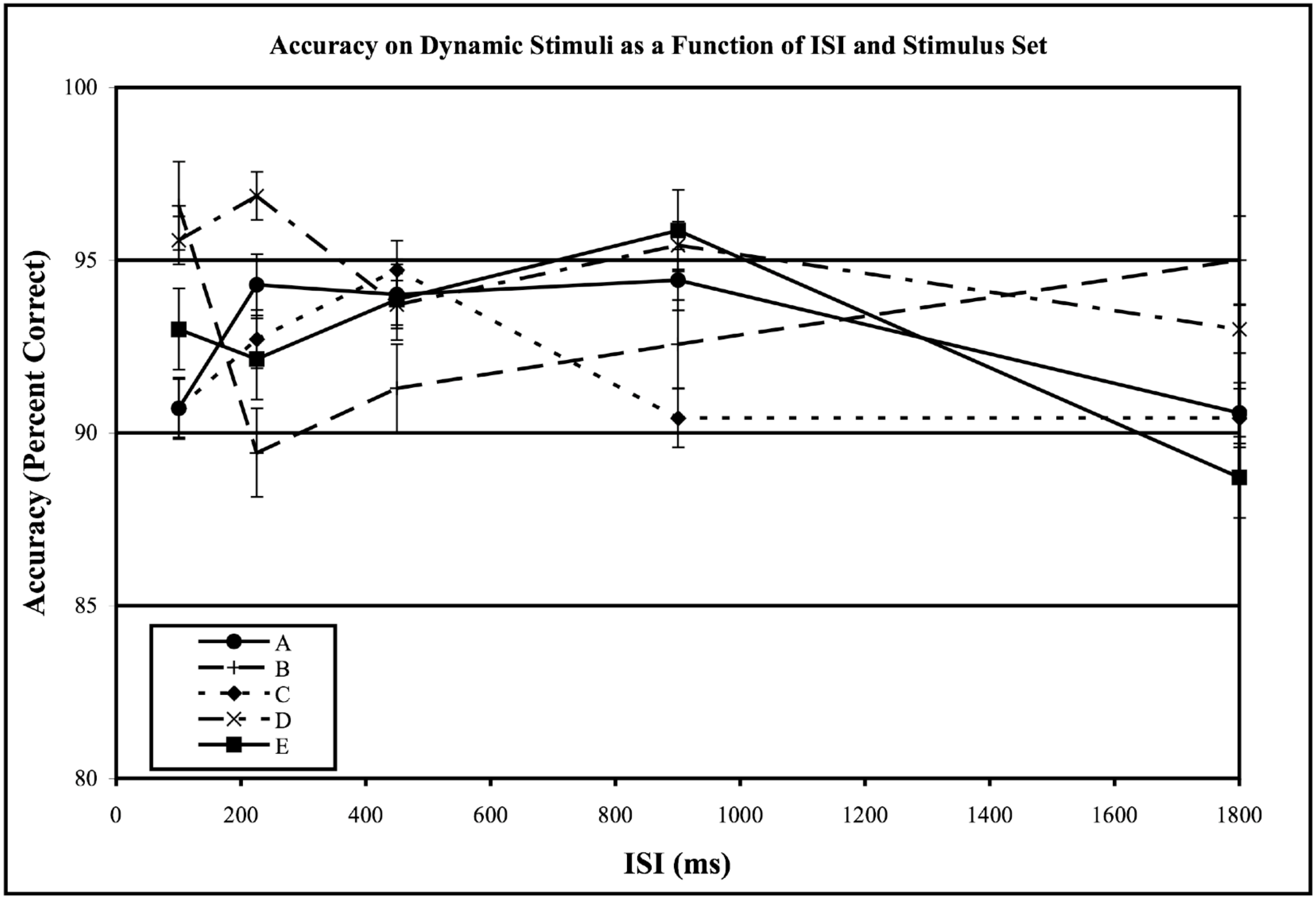

Figure 2.5. Percent correct for dynamic stimuli as a function of stimulus set. Bars are SEM for the parameter.

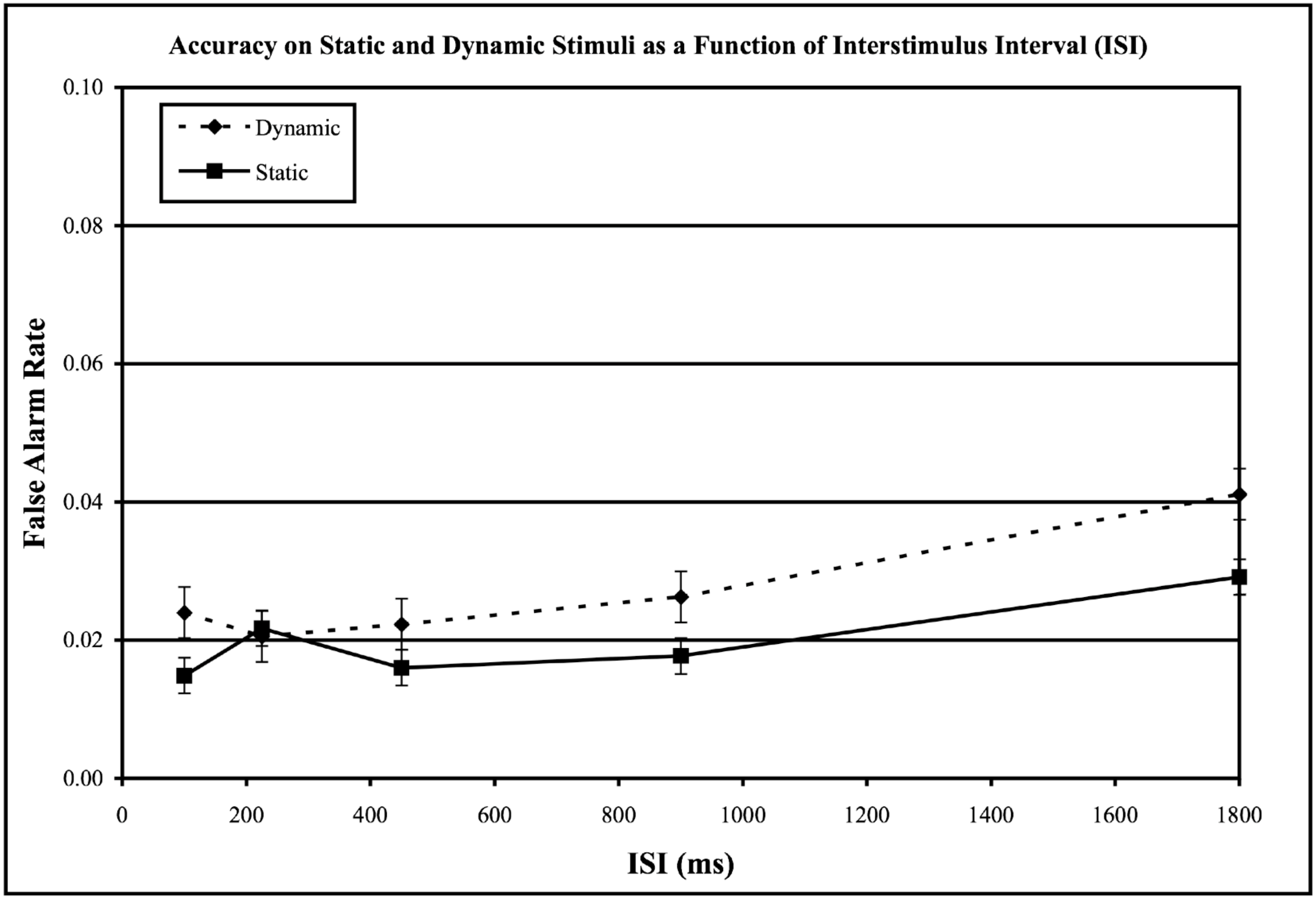

Figure 2.6. False alarm rate (proportion of "same" responses on "same" trials) for static and dynamic stimuli as a function of ISI. Bars are SEM for the parameter.

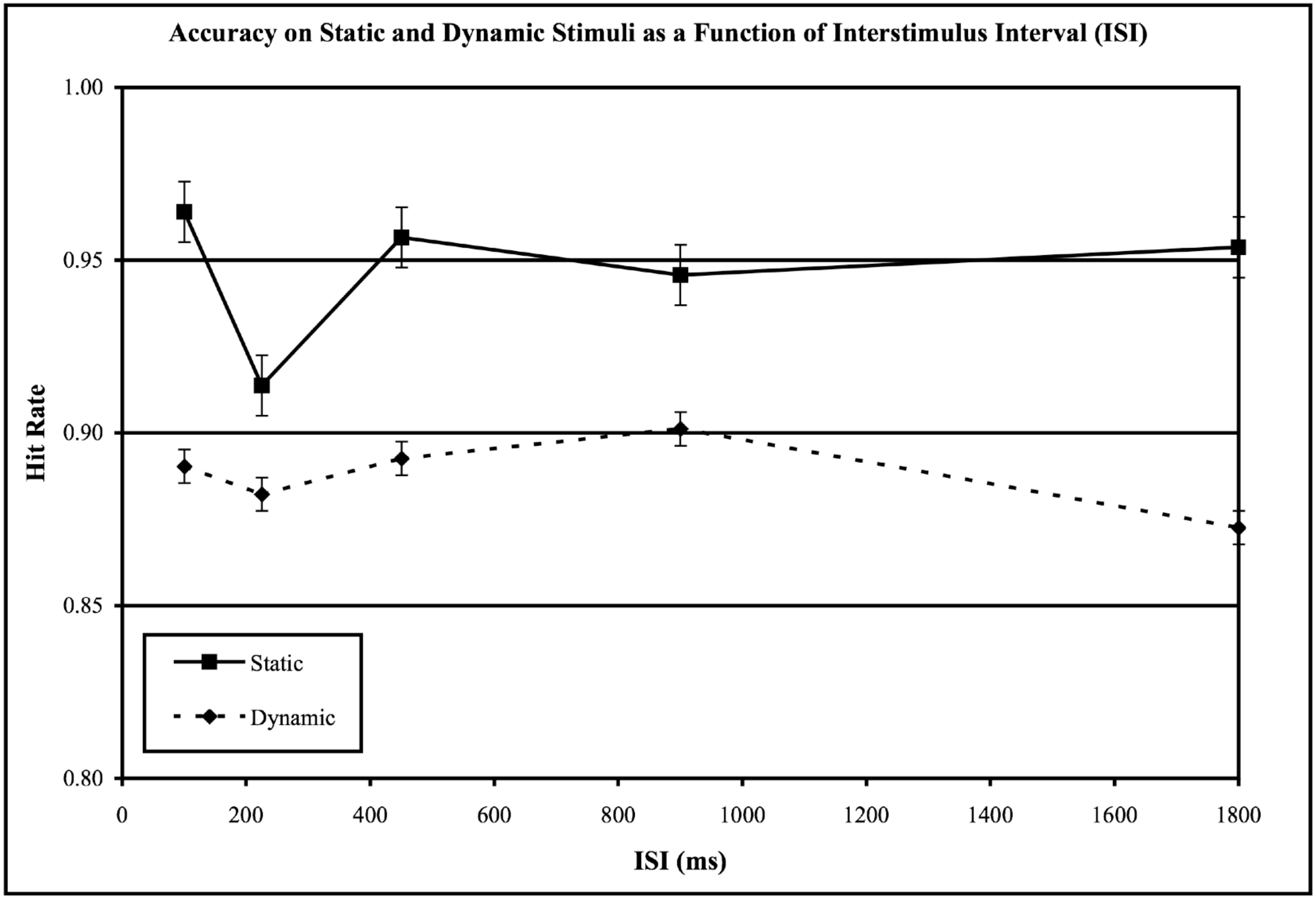

Figure 2.7. Hit rate (proportion of "different" responses on "different" trials) for static and dynamic stimuli as a function of ISI. Bars are SEM for the parameter.

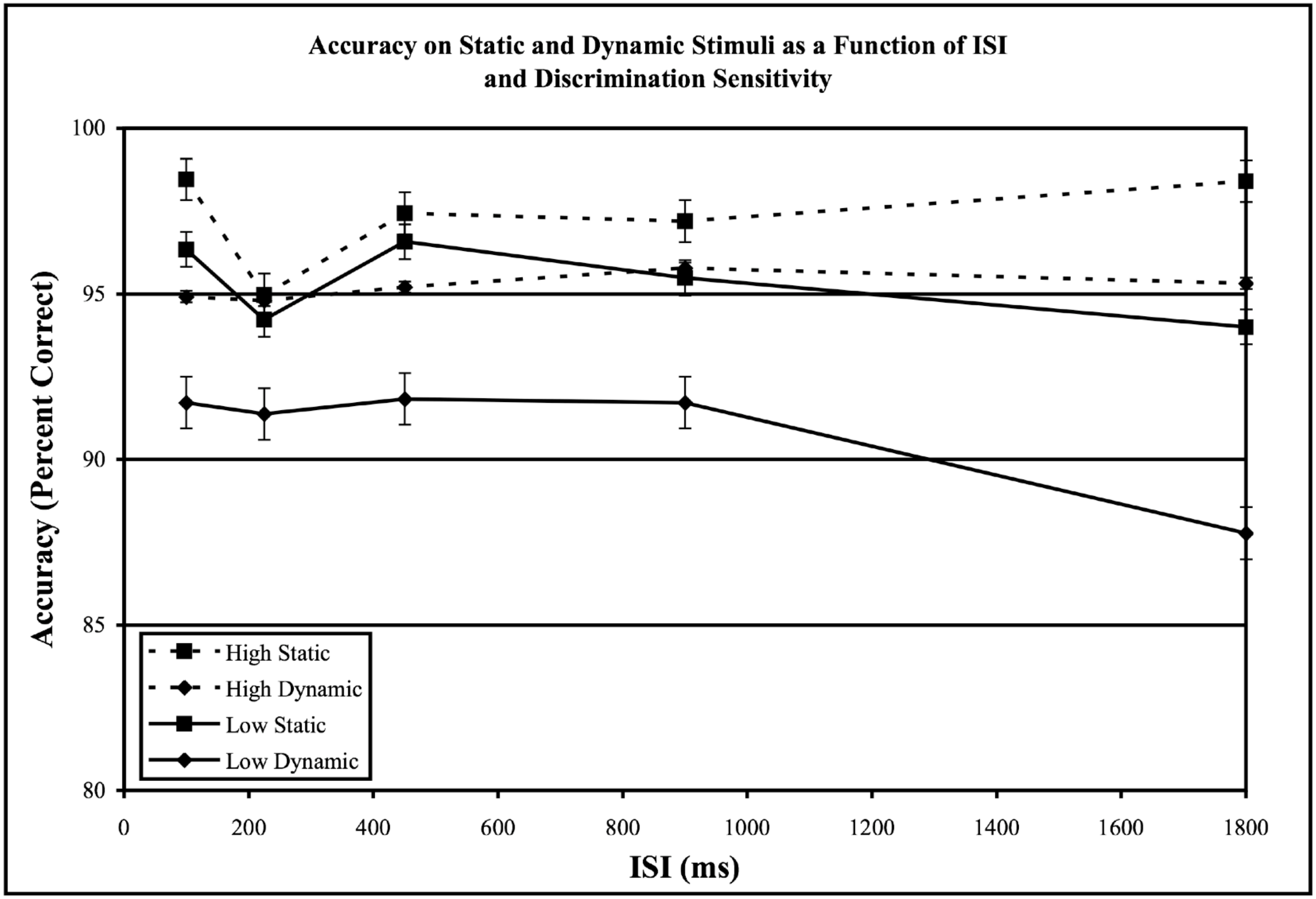

Figure 2.8. Percent correct for static and dynamic stimuli as a function of ISI and stimulus type for high- (solid line) and low-sensitivity (dashed line) groups. Bars are SEM for the parameter.

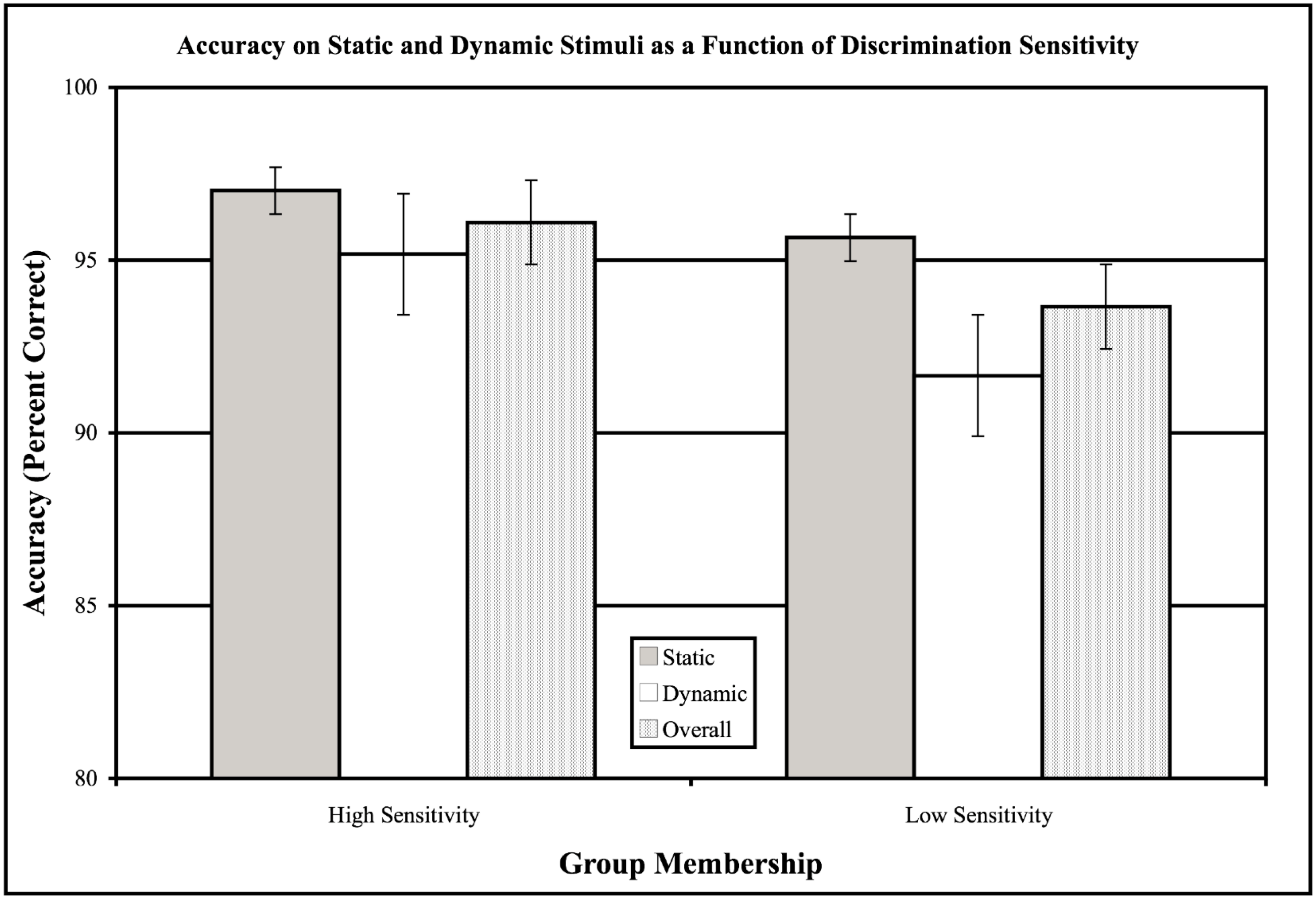

Figure 2.9. Percent correct for Static and Dynamic stimuli for both high and low sensitivity (collapsed across 100-900 ms ISI; gray and white bars) and Overall (collapsed across stimulus type and 100-900 ms ISI; striped bars). Error bars are SEM for the parameter.

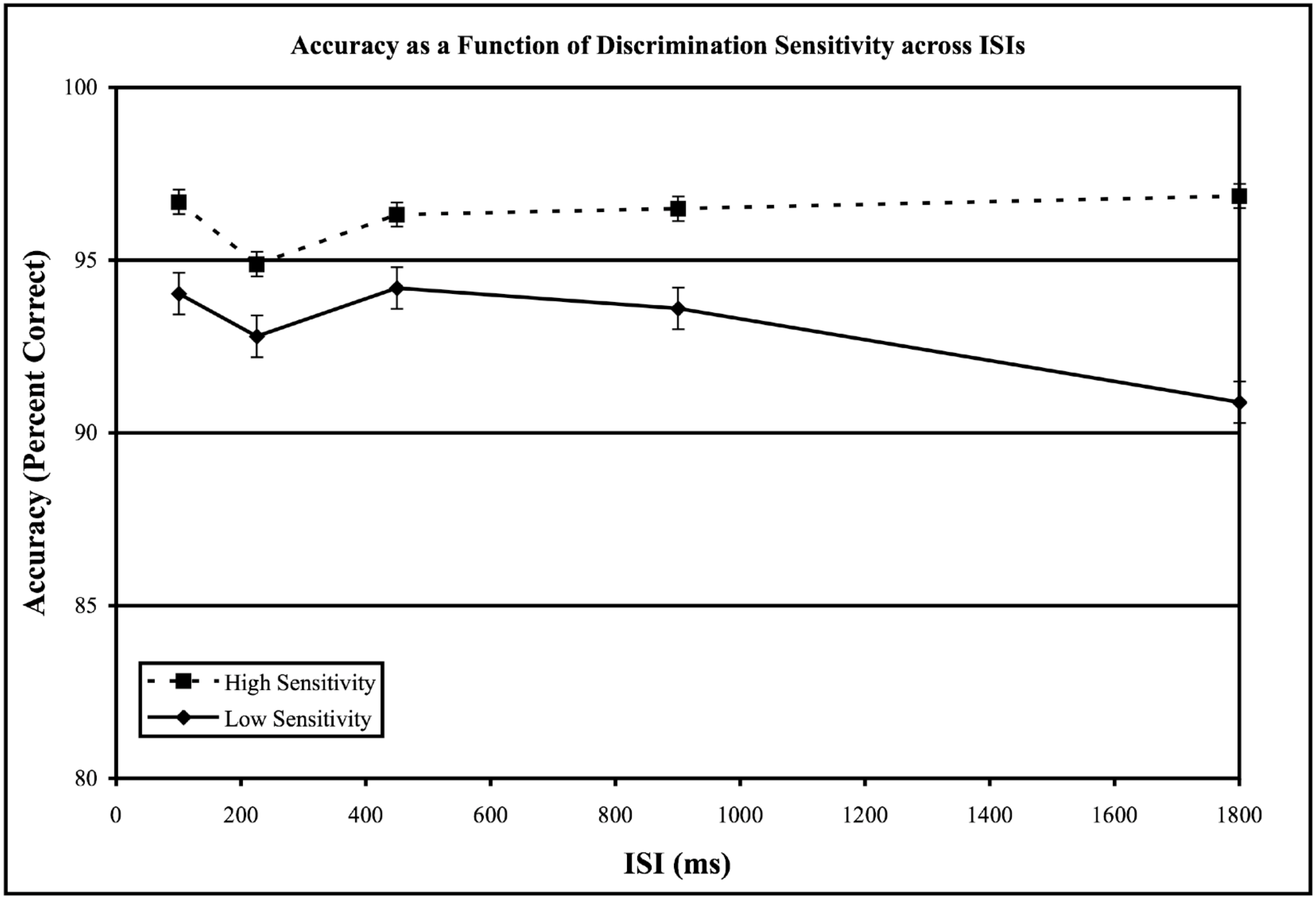

Figure 2.10. Percent correct for each ISI collapsed across stimulus type for both high and low sensitivity. Bars are SEM for the parameter.

CHAPTER 3

DIFFERENCES IN THE DISCRIMINABILITY OF STATIC AND MULTIPLE TYPES OF DYNAMIC AUDITORY STIMULI AT A SINGLE MEMORY DELAY

Experiment 2

Experiment 1 showed that there is a main effect of stimulus type on change detection accuracy, wherein changes to dynamic tones are detected less reliably than changes to static stimuli. Further, a main effect of ISI suggested that the underlying memory mechanism supporting change detection (ASTM) is subject to decay. This effect interacts with stimulus type such that dynamic stimuli show more evidence of decay than do static stimuli. Both the effects of stimulus and delay are attenuated in some listeners, suggesting that there are individual differences in the function of ASTM. The present experiment extends these findings to incorporate a number of different types of dynamic stimuli. While the precise nature of the stimuli will change, the core question will remain: 1.) **Is there a main effect of stimulus type on change detection performance?** That is, as we vary the type of static and dynamic stimuli presented, is there still a main effect of static vs. dynamic stimuli (the main question of this thesis) and is there an additional effect of stimulus type even within static and dynamic categories? The rationale and nature of these various stimuli will be presented below.

The static stimuli in Experiment 1 all incorporated a single frequency over a constant duration of 500 ms. The dynamic stimuli in Experiment 1 all incorporated the same *magnitude of the frequency change* from start frequency to end frequency, over the same constant duration of 500 ms. They also all incorporated the same *slope of the frequency change*. In Experiment 2, classes of stimuli were created that differed in their slope and magnitude of frequency change, as a result of manipulation of stimulus duration. The effect of manipulating stimulus duration is as follows. For static stimuli, different durations at a particular frequency simply increase the duration of that single frequency. For dynamic stimuli that incorporate a monotonically increasing frequency, however, there are two possibilities. First, if the *slope of the frequency change* is held constant, then two different dynamic stimuli with the same starting frequency, but different durations, will differ in the *magnitude of the frequency change*. Second, if the *magnitude of the frequency change* is held constant, then two different dynamic stimuli with the same starting frequency, but different durations, will differ in the *slope of the frequency change.* Thus, varying the stimulus duration pulls apart these two properties of dynamic stimuli, which were kept constant in Experiment 1 by virtue of stimulus duration being kept constant.

Variation of stimulus duration in Experiment 2 thus led to creation of two sets of dynamic stimuli. One set consisted of dynamic stimuli in which, for any given starting frequency *S*, the *slope of the frequency* change remained constant even though different stimuli had different durations. (And the *magnitude of the frequency change* therefore necessarily varied, even with the same start frequency). This set of dynamic stimuli will be referred to as *slope* stimuli. A second set of dynamic stimuli incorporated, for any given start frequency *S*, a constant *magnitude of the frequency change* even though different stimuli had different durations. (And the *slope of the frequency change* therefore necessarily varied, even with the same start frequency). This set of dynamic stimuli will be termed the *magnitude* stimuli.

In addition, in Experiment 2, as noted above, the *magnitude of the frequency change* was also manipulated *independent* of stimulus duration. This was instantiated as follows. For the set of *magnitude* stimuli described above, the constant *magnitude of the frequency change* was held at either 100 cents or 300 cents, creating two further classes of stimuli: *dynamic-100-magnitude* stimuli, and *dynamic-300-magnitude* stimuli. For the set of *slope* stimuli described above, two further sub-classes were created, depending on whether the magnitude of frequency change at a 1000 ms duration would have been 100 cents or 300 cents. This created two subclasses of the slope stimuli, which will be termed the *dynamic-100-slope* stimuli and the *dynamic-300-slope* stimuli.

Thus, in Experiment 2 participants were presented with static stimuli (as in the static condition from Experiment 1) and four classes of dynamic stimuli. (1) The first type of dynamic stimulus incorporated a magnitude of frequency change of 100 cents, like all dynamic stimuli in Experiment 1. However, such stimuli now had one of four durations: 500 ms, as in Experiment 1, along with two shorter durations(50 ms and 100 ms) and one longer (1000 ms) duration. The slope of the frequency change was therefore different for two such stimuli with different durations, even if their starting frequency is the same, as shown in Figure 3.1. This class of dynamic stimuli was termed the *dynamic-100-magnitude* stimuli. (2) The second type of dynamic stimulus incorporated a magnitude of frequency change of 300 cents. As in the case of the 100-cent stimuli, stimulus durations could be 50, 100, 500, or 1000 ms, so the slope of frequency change differed for stimuli of different durations, even if their starting frequency is the same; this is also shown in Figure 3.1. These stimuli were termed *dynamic-300-magnitude* stimuli. (3) The third type of dynamic stimulus was one in which the magnitude of frequency change from starting frequency to ending frequency was 100 cents *at the 1000 ms duration*. At the other durations, however, (50, 100, and 500 ms), the stimulus was truncated so that the slope of frequency change was the same as for the 1000 ms stimuli, but the magnitude of frequency change was less than 100 cents (see Figure 3.2). Such stimuli were referred to as *dynamic-100-slope* stimuli. (4) The fourth type of stimulus was one in which the magnitude of frequency change from starting frequency to ending frequency was 300 cents *at the 1000 ms duration*. As with the *dynamic-100-slope* stimuli, these *dynamic-300-slope* stimuli had constant slope across the various durations, but differed in the magnitude of the frequency change, at the various durations.

On a dynamic stimulus trial, the two stimuli were always from the same dynamic stimulus class, and had the same duration. Thus they were always identical to each other in their duration, the slope of their frequency change, and the magnitude of their frequency change. They differed from each (on a "different" trial) other only in their starting frequency. Thus "slope" and "magnitude" in the names of the stimulus classes do not indicate what differs between two dynamic stimuli on a "different" trial. Rather, it indicates the aspect of the two stimuli that is *controlled* on any trial involving dynamic stimuli. In the slope condition, as the duration of the stimuli is manipulated the slope stays the same, whereas in the magnitude condition, as the duration is manipulated the magnitude of the frequency change stays the same. On a dynamic stimulus trial, the two dynamic stimuli both incorporated either: (i) a specific magnitude of frequency change, with slope of frequency change being determined by duration (for dynamic-100-magnitude and dynamic-300-magnitude stimuli), or (ii) a specific slope of frequency change, with the magnitude of frequency change being determined by duration (for dynamic-100-slope and dynamic-300-slope stimuli).

Note that the condition in which static stimuli are presented with a 500 ms duration is the same as the static 900 ms ISI condition of Experiment 1. Similarly, the condition in which dynamic-100-magnitude stimuli are presented with a 500 ms duration is the same as the dynamic 900 ms ISI condition of Experiment 1. *Thus, these two conditions are partial replications of Experiment 1 with different stimulus sets.* Further, the dynamic-100-slope condition with a 1000 ms duration is the same as the dynamic-100-magnitude condition with a 1000 ms duration (as can be seen from comparison of Figures 3.2 and 3.1). Similarly, the dynamic-300-slope condition with a 1000 ms duration is the same as the dynamic-300-magnitude condition with a 1000 ms duration (see Figures 3.2 and 3.1). *Thus, these two conditions are identical with regard to the main aspects of the task, but not with regard to the exact stimuli that are presented.*

While Experiment 1 varied ISI keeping stimulus presentation duration constant at 500 ms, in Experiment 2, ISI was held constant at 900 ms, while duration and magnitude of frequency change were varied. Again, Experiment 2 maintained continuity with Experiment 1, however, as the 900 ms ISI condition of Experiment 1 was replicated with at the 500 ms duration of Experiment 2, using the same two stimulus classes as in Experiment 1 (static stimuli with a 500 ms duration, and dynamic-100-magnitude stimuli) as well as three additional classes of stimuli).

Again, the main question of Experiment 2 is identical to Experiment 1: 1.) ***Is there a main effect of stimulus type (static vs. dynamic) on change detection performance?*** All that differs from Experiment 1 is the precise definition of static and dynamic stimuli, in that they are much more broadly defined, and thus we would predict that change-detection performance would be higher for static stimuli overall than for dynamic stimuli collapsed across sub-type.

With regard to our secondary question of: 2.) ***Is there a main effect of duration on change detection performance?****,* two factors are likely to affect change detection for both static and dynamic stimuli at different presentation durations. On the one hand, for static and all types of dynamic stimuli, change detection may be more accurate at longer stimulus durations if a longer stimulus duration provides more opportunity for encoding the stimulus into ASTM. On the other hand, change detection may be less accurate for stimuli with a greater magnitude of frequency change, and therefore may be less accurate at longer stimulus durations for dynamic-slope stimuli, whose longer durations increases their magnitude of frequency change. This leads to the prediction that: collapsed across all stimulus types, the effect of increases in duration should be an increase in change detection accuracy, since the negative effect of duration should only affect the dynamic-slope conditions. These predictions are broadly consistent with previous research. For instance Payne (2003) suggests that, at least for what here are termed static stimuli, additional time may elevate the memory of low-sensitivity listeners to that of high-sensitivity listeners.

However, given the increase in frequency-change magnitude that does occur with increasing durations in the dynamic-slope condition we must also ask: 3.) ***Is there an interaction between stimulus type and duration*?** Even though change detection for static stimuli should benefit from a longer stimulus duration and change detection for dynamic-magnitude stimuli should also benefit from a longer stimulus duration, change detection for dynamic-slope stimuli may see a decrement as a function of longer stimulus duration owing to the increase in the magnitude of the frequency change that occurs as a result of increasing the duration of the stimulus. Predictions based on these anticipated effects for duration are presented in Figure 3.3. It should be noted that the increases and decreases depicted in this figure are categorical increments and decrements (arbitrarily set at +/- 2%) as a function of the increase in the duration being implemented. These increases and decreases are meant to show the general benefit or penalty as a function of duration and their precise numbers are not empirically-based predictions. What appear to be non-linear effects of duration emerge simply as a function of the non-linear distribution of duration conditions employed in this experiment and, again, are not empirically-based predictions.

Main effects of stimulus type are also predicted in Figure 3.3. Overall the static stimuli have the smallest frequency change magnitudes (0%) and thus performance should be best for static stimuli, starting at the shortest duration. The next smallest frequency change magnitude occurs in the *dynamic-100-slope* stimuli, followed by the *dynamic-300-slope*, *dynamic-100-magnitude,* and the *dynamic-300-magnitude* (see Methods). If the magnitude of the frequency change affects change detection as predicted above, then there should also be a main effect of stimulus type as depicted in Figure 3.3.

Change detection ability will be measured in a change-detection task like that of Experiment 1. As the contribution of hit rates and false alarms to the measure of proportion correct was established for stimuli at this ISI in Experiment 1 only the measure of proportion correct will be used to measure discrimination. With this measure we will ask the 3 questions posed above as well as the final question of: 4.) ***Are the effects of duration and stimulus type equivalent for all listeners?*** Investigating differences in change detection for stimuli in high- and low-sensitivity individuals at all stimulus durations will again aid our understanding of whether any differences in change-detection ability for static and dynamic stimuli that emerge are present for all listeners, which provides information about the natural variation in change detection ability, which in turn provides information about the function of ASTM as it is used in this task. Assessing individual differences will also provide an understanding of whether performance with a particular stimulus duration is predictive of performance at other durations.

Method

Participants

Forty undergraduates (31 = Female; 9 = Male) at the University of Iowa participated in this experiment for partial course credit.

Materials

The same beginning frequency values were used as in Experiment 1. For all stimuli, the duration for which the stimulus was presented was either 50, 100, 500, or 1000 ms. For the static stimuli, this simply meant that a particular frequency value was present in the environment for this period of time. For the dynamic stimuli, the different durations began at the same frequencies as the dynamic stimuli in Experiment 1, but could change according to two different frequency amounts, or magnitudes. The first magnitude is the same as Experiment 1, 100 cents. The second possible magnitude is 3 times larger than Experiment 1, 300 cents. Thus, two different types of dynamic stimuli are possible based on the magnitude of the frequency rise over time: dynamic-100 and dynamic-300. In addition to this difference between stimuli, stimuli with the same starting frequency could either have the same slope of the frequency change at different durations, in which case the magnitude of the frequency change increased with increases in duration (the ‘slope’ conditions; see Figure 3.2) or stimuli of different durations could have the same magnitude of the frequency change, in which case the slope of the frequency rise varied with duration (the ‘magnitude’ conditions; see Figure 3.1). For ‘slope’ stimuli, the slope was defined as the slope of the frequency change of the stimulus at the 1000 ms duration. (This was just a notational convenience, as the slope was, of course the same for all dynamic stimuli of each of the two ‘slope’ classes dynamic-100-slope and dynamic-300-slope). The end frequency of any stimulus at a particular duration (*D*) (e_D) in the slope condition would be:

(1)
$$e_D = \left(slope \bullet D \right) + S$$

S is the ‘starting frequency’ of the stimulus, and *slope* is defined by:

(2)
$$slope = \left(\frac{e_{1000} - S}{1000} \right)$$

Here ($e_{1000} - S$) is the same as the magnitude of the frequency change. When I use Equation 2 to solve Equation 1 for any stimulus, I find that the *e* frequency for any stimulus varies as a function of duration. For the ‘magnitude’ stimuli, the slope of frequency change was of course not the same for all stimuli. However, the magnitude of frequency change was the same for all stimuli within each of the two ‘magnitude’ classes (dynamic-100-magnitude and dynamic-300-magnitude). *Slope* was defined by:

(3) $$slope = \left(\frac{magnitude}{D} \right)$$

When I use Equation 3 to solve for Equation 1, I find that the duration terms cancel one another and thus the ending frequency is the same for all durations. However, the slope of the stimulus varies by duration. Altogether there are five levels of the stimulus factor of this experiment: static, dynamic-100-magnitude, dynamic-100-slope, dynamic-300-magnitude, and dynamic-300-slope. Again, the magnitude and slope distinctions refer to the variable that is being controlled and thus is held CONSTANT across a particular dynamic stimulus for each level of the duration factor.

Again, the amount of change in frequency can be expressed as both a raw frequency value and as a percentage of the starting frequency. For the dynamic-100-magnitude condition, the mean amount of change over the stimulus presentation was 130.73 Hz. (SD = 60.70) or 5.95% and the mean slopes were 2.61, 1.31, .26, and .13 Hz/ms in the 50, 100, 500, and 1000 ms condition (SDs = 1.21, .61, .12, and .06). Expressed as a percentage of the starting frequency, the slopes are .12, .06, .012, and .006 %/ms. The exact slopes for each stimulus are listed in Table B.1 of Appendix B. For the dynamic-100-slope condition, the mean amount of change over the stimulus presentation was 6.54, 13.07, 65.37, and 130.73 Hz (SDs = 3.04, 6.07, 30.35, and 60.70) for the 50, 100, 500, and 1000 ms conditions respectively. Expressed as a percentage, the magnitudes are .297, .595, 2.97, and 5.95% The mean slope was .13 (SD =.06) Hz/ms or .006 %/ms. The exact frequency changes are listed in Table B.2.

For the dynamic-300-magnitude condition, the average amount of change over the stimulus presentation was 415.99 Hz (SD = 193.14) or 18.92% and the average slopes were 8.32, 4.16, .83, and .42 Hz/ms in the 50, 100, 500, and 1000 ms condition (SDs = 3.86, 1.93, .39, and .19). Expressed as percentages of the starting frequency, the slopes are .378, .189, .0378, and .0189%/ms. The exact slopes for each stimulus are listed in

Table B.3. For the dynamic-300-slope condition, the average amount of change over the stimulus presentation was 20.80, 41.60, 207.99, and 415.99 Hz (SDs = 9.66, 19.31, 96.57, and 193.14) for the 50, 100, 500, and 1000 ms conditions respectively. Expressed as a percentage, the magnitudes are .946, 1.89, 9.46, and 18.92%. The average slope was .42 (SD =.19) Hz/ms or .0189%/ms. The exact frequency changes are listed in Table B.4.

Noise stimuli were the same as in Experiment 1, however the level of the tone was reduced to 50 units (as defined by PsyScope), or approximately 20% of the noise volume. Participants were again free to set the level of the presentation to match their comfort levels.

Design and Procedure

Presentation of a particular type of stimulus (static, dynamic-100-magnitude, dynamic-100-slope, dynamic-300-magnitude, dynamic-300-slope) was blocked, and the order of the five stimulus types was partially counterbalanced such that each block appeared in each sequence location the same number of times and came before and after each of the other conditions once.

The difference in frequency between two “different” stimuli was again randomized for each "different" trial and so I again conducted a 5 (stimulus type) X 4 (duration) analysis of variance (ANOVA) on the size of the frequency differences on "different" trials of the 10 stimulus sets to ensure that they were equal across the variables of interest. This ANOVA reported a significant interaction between stimulus type and duration, F (12, 108) = 2.05, $p < .05$, $\eta^2 = .19$, which will account for a small portion of the variance in performance discussed below, but these effects did not systematically load on any of the factors of interest as indicated by the lack of an effect of stimulus F (4, 36) < 1, and no effect of duration, F (3, 27) = 1.34, $p > .05$, $\eta^2 = .13$. Overall, the size of the difference between two different stimuli rises as a function of duration for the static condition and falls for the dynamic-100-magnitude condition. Both the dynamic-100-slope and the dynamic-300-slope conditions remained relatively unchanged as a function of duration, but show a single increase at the 500 ms duration. Lastly, the dynamic-300-magnitude condition shows a U-shaped function with the size of the differences showing elevation at the 50 and 1000 ms. These differences will be discussed in the context of overall performance when they might be impacting performance. However, it is important to again point out that this effect does not affect the main variables of interest, stimulus type and duration.

Within each block, participants heard 160 trials, 40 of each of the duration values, with 20 of these trials being change trials and 20 being no change. The duration and same/different component of each trial was randomized within a block. Participants received a 30 second break every 20 trials. Ten different sets of trials were used to eliminate set effects within this experiment. Each trial set contained a randomized combination of starting frequencies and change magnitudes for the change trials of each block.

As in Experiment 1, the analysis of Experiment 2 begins with an assessment of overall performance, directed at the main question of whether change detection is more or less accurate for static vs. dynamic stimuli and the corollary questions regarding the effects of duration. After this analysis, the frequency overlap of the dynamic stimuli will be considered followed by the individual differences between listeners in Experiment 2. In Experiment 2 ceiling effects are not an issue until the individual differences findings, thus the arcsin analysis of Experiment 1 will be limited to the individual differences of Experiment 2. Also, since the analyses of the hit rate and false alarm rate provided redundant information in Experiment 1, these analyses will not be completed in Experiment 2.

Results and Discussion

Overall Performance

For the first analysis, change detection was averaged across type of dynamic stimulus to examine change detection for static vs. dynamic stimuli. A 2 (Stimulus Type) X 4 (Duration) Repeated-Measures ANOVA on the proportion correct measure revealed a main effect of stimulus type F (1, 39) = 17.39, $p < .05$, $\eta^2 = .31$, which extends the results of Experiment 1 by showing that **the change-detection advantage for static stimuli is maintained even when multiple types of dynamic stimulus are present.** Additionally, a main effect of duration, F (3, 117) = 25.43, $p < .05$, $\eta^2 = .40$, was found. This was driven by a linear trend, F (1, 39) = 41.72, $p < .05$, $\eta^2 = .52$, indicating that the overall net effect of duration is an increase in accuracy and this increase accounts for most of the variance. There is also a quadratic trend, F (1, 39) = 5.28, $p < .05$, $\eta^2 = .12$, indicating that the slope is steeper with shorter durations than with longer durations. No cubic trend, F (1, 39) = 3.35, $p > .05$, $\eta^2 = .05$, is present. Lastly, there is a significant interaction between duration and stimulus type, F (3, 117) = 4.82, $p < .05$, $\eta^2 = .18$ indicating that the increase in accuracy with increasing duration is not as great in the case of dynamic stimuli, presumably because increases in duration for some stimuli also increase the frequency change magnitude (See Figure 3.4).

Next, change detection was examined for all five stimulus types considered separately. A 5 (Stimulus Type) X 4 (Duration) Repeated-Measures ANOVA on the proportion correct measure revealed a main effect of stimulus type, F (4, 156) = 6.11, $p < .05$, $\eta^2 = .14$ (see Figure 3.5). Here, the static condition showed the best performance (M = 93.6%, SD = 4.26), and four post-hoc paired samples *t*-tests comparing static stimuli with each of the other conditions (with an adjusted alpha level of .01) showed that performance in the static condition was not significantly better than in the dynamic-100-slope condition, t (39) = .60, $p = .55$, d = .10, but was better than all other conditions, *t*s ≥

3.29, *ps* ≤ .002, ds ≥ .52. This finding qualifies the overall change detection for static over dynamic stimuli found in the previous analysis, but overall it is clear that such an advantage exists. In addition, the difference in performance between static stimuli and the dynamic-100-magnitude condition at the 500 ms duration replicates the findings of Experiment 1 at the 900 ms ISI.

This analysis also showed a linear trend, F (1, 39) = 33.49, $p < .05$, $\eta^2 = .46$, main effect of duration, F (3, 117) = 22.30, $p < .05$, $\eta^2 = .36$, however there is also a cubic component, F (1, 39) = 9.05, $p < .05$, $\eta^2 = .19$ that indicates an inflection downward at longer durations (see Figure 3.6). These effects are in addition to a quadratic component, F (1, 39) = 6.85, $p < .05$, $\eta^2 = .15$, that indicates an overall decrease of the slope of the performance profile at longer durations. Generally, no gain in performance occurs beyond 500 ms. Thus change detection collapsed across all stimulus types is more accurate with increases in duration and asymptotes at the 500 ms duration.

This replication is depicted in Figure 3.7, which shows the accuracy of the *dynamic-magnitude-100* and the *static* condition. In the middle of the graph are the data from the 500 ms duration of Experiment 2, which is the condition that is a direct replication of Experiment 1. The bracket to the right of this difference is the size of the difference seen in Experiment 1 and it can be seen that the size of the difference is approximately the same, thus the stimulus effects of Experiment 1 have been replicated in Experiment 2. (The data from the 900 ms ISI condition of Experiment 1 are plotted on the right side of the graph for comparison.) On the left panel the accuracy of each stimulus condition averaged across all durations is plotted and a difference between static and dynamic-magnitude-100 stimuli is shown. Thus, the stimulus effects of Experiment 1 also generalize across all duration types. The overall decrease in performance for both stimulus types from Experiment 1 to Experiment 2 can be attributed to either participant selection or the decrease in signal-to-noise ratio.

Lastly, the ANOVA revealed a significant interaction between stimulus type and duration, *F* (12, 468) = 5.41, $p < .05$, $\eta^2 = .12$ (see Figure 3.8). In order to understand this complex interaction, ten paired samples *t*-tests were run: For each of the five stimulus types, performance was compared at (1) the 50 ms and 500 ms durations and (2) at the 500 and 1000 ms durations. These particular comparisons are of interest because they will determine which stimuli show *any* increase in accuracy as a function of duration from durations approximating the phoneme (50 ms) to the word (500 ms) level (comparison 1) and which stimuli *continue* to show an increase in accuracy moving from the word (500 ms) to longer (1000 ms) durations (comparison 2). Further At 50 ms, there is almost no visual difference between performance on the various different types of stimuli – that is, at shorter durations, change-detection accuracy for almost all stimuli is equivalent. Thus, these two comparisons assess how performance between stimulus classes diverges with increasing duration. The adjusted alpha level for this set of *t*-tests was .005.

In comparison 1, change detection for the static, dynamic-slope-100, and dynamic-magnitude-100 all increased from the 50 ms to the 500 ms duration, *t*s (39) ≥ 4.12, *p*s < .005, d ≥ .65, while the dynamic-300-slope and dynamic-300-magnitude conditions showed no difference, *t*s (39) ≤ 2.32, *p*s ≥ .03, ds ≤ .37. Thus, with stimuli containing large changes in frequency, a ten-fold increase in duration does not improve accuracy on change detection and an even greater change in duration may be required. To verify this conclusion, the dynamic-300 stimuli at 50 ms and 1000 ms were compared and only the dynamic-300-magnitude condition showed a significant improvement, *t* (39) = 3.79, $p < .001$, d = .60. The dynamic-300-slope condition showed no improvement, *t* (39) = 1.07, $p > .05$. d = .16.

In comparison 2, all comparisons were not significant, *t*s ≤ 2.30, *p*s ≥ .03, d ≤ .36, with one exception. The exception was the dynamic-300-slope condition, *t* (39) = 3.74, $p < .005$, d = .60, however the dynamic-300-slope condition at 500 ms was one of the conditions that showed elevated change magnitudes in the analysis of change magnitude across conditions and the difference between performance at the 500 and 50 ms levels was not significant. Thus, the effects of duration seem to asymptote at 500 ms for static stimuli and dynamic stimuli with 100 cent frequency changes. For dynamic stimuli with 300 cent frequency changes, there appears to be little to no effect of duration.

Overall, accuracy improves with longer durations for all classes of stimuli except the dynamic-300-slope condition. This conclusion is supported by the significant increases in discrimination at 500 ms for the static, dynamic-100-slope, and dynamic-100-magnitude stimuli, and at 1000 ms for the dynamic-300-magnitude condition. At 50 ms, there is almost no difference between performance on the various different types of stimuli suggesting that memory at this duration is the same for all stimuli.

Overall, the answer to the question: ***Is there a main effect of duration on change detection performance?*** is "Yes" given that change-detection accuracy increased with longer stimulus duration for the static and dynamic stimuli combined. For the question: ***Is there an interaction between duration and stimulus type?*** it was found that change-detection accuracy increased with longer stimulus duration for the static and dynamic-magnitude stimuli, but the pattern for dynamic-slope stimuli was less consistent. Here, change detection increased with longer durations for the dynamic-100-slope stimuli but not the dynamic-300-slope stimuli. Additionally, the difference in performance between various stimulus types increased at longer durations, and was eliminated at the 50 ms duration.

Overlapping Stimuli

Though it was ruled out as a concern for Experiment 1, overlapping stimuli is again a concern in Experiment 2 because it is again confounded, this time with both stimulus type (dynamic vs. static) and the magnitude of change for dynamic stimuli (100 vs. 300 cents). This follows from the fact that greater magnitude stimuli will involve a wider range of the frequency spectrum and thus, two different stimuli may have a greater chance of overlap in the frequency space. Thus, I again analyzed the number of times such overlap occurred in Experiment 2 in order to rule out overlap as the sole cause of differences in accuracy between static and dynamic stimuli. On average, overlap occurs on approximately 3.69% of all "different" trials with dynamic stimuli. In Experiment 2, participants correctly detected changes to dynamic stimuli with a hit rate of 87.59% and detected changes to static stimuli with a hit rate of 91.00%. Here, the overall 3.41% difference in hit rate can be wholly explained by the differential proportion of overlapping stimuli in the dynamic vs. static cases (3.69% vs. 0%). Further, the differences between false alarms for all static and all dynamic stimuli is not significant, F (1, 39) < 1. Thus, the main effect of differences between static and dynamic stimuli may possibly be explained by the overlap of stimuli in the dynamic case, however, a further examination of differences as a function of stimulus type demonstrates that the relationship between differences in overlap amount and differences in performance are not perfect.

I ruled out the effects of overlap by including only those trials on which NO overlap occurred for dynamic stimuli and re-ran the previous analysis to determine if the effects reported above remain. This new 5 (Stimulus) x 4 (Duration) Repeated Measures ANOVA revealed the same effect of stimulus, F (4, 156) = 11.47, $p < .05$, $\eta^2 = .23$, duration, F (3, 117) = 15.79, $p < .05$, $\eta^2 = .29$, and the same interaction effect, F (12, 468) = 5.22, $p < .05$, $\eta^2 = .12$. Further, these effects showed only a small decrease in their effect sizes, indicating that only a small portion of the previously reported effect was driven by stimulus overlap.

Individual Differences

In Experiment 1, participants were divided into high- and low-sensitivity groups based on performance on all stimuli at the lowest performance level, which corresponded to an ISI of 1800 ms. Similarly, in Experiment 2, participants were sorted based on performance on all stimuli at the lowest performance level, which corresponded to a duration of 50 ms. Participants who fell below the median performance of the group (91.5%) were classified as low-sensitivity (n = 20). Participants who were above the median were classified as high-sensitivity (n = 20). Performance ranged from 75.5 - 98% across the sample.

A 2 (Sensitivity) X 3 (Duration) X 5 (Stimulus) Mixed-Design ANOVA was run on proportion correct to assess the significance of differences in accuracy between groups. Again, performance on the 50 ms stimuli was left out of this ANOVA. Overall, the separation between the high- and low-sensitivity listeners was large and significant, F (1, 38) = 50.88, $p < .05$, $\eta^2 = .57$. Thus, the separation of groups on sensitivity at the 50 ms duration was predictive of their performance across durations. However, the ANOVA revealed no stimulus by sensitivity interaction, F (4, 152) < 1, but a sensitivity by duration interaction, F (2, 76) = 3.45, $p < .05$, $\eta^2 = .08$ (see Figure 3.9, indicating that performance as a function of duration, but not stimulus type was dependent on the sensitivity group that one fell in. Thus, the low-sensitivity group's difficulty was in processing stimuli of short durations, irrespective of the stimulus type. This conclusion was further supported by a lack of a stimulus by sensitivity by duration interaction, F (8, 304) < 1, $p > .05$ (see Figure 3.10).

Individual 3 (Duration) X 5 (Stimulus) Repeated-Measures ANOVAs were also run for each group to investigate the nature of the sensitivity by duration interaction. For the high-sensitivity group, there was a main effect of stimulus, F (4, 76) = 14.05, $p < .05$, $\eta^2 = .43$, a main effect of duration, F (2, 38) = 4.27, $p < .05$, $\eta^2 = .18$, and an interaction between stimulus and duration, F (8, 152) = 2.60, $p < .05$, $\eta^2 = .12$ (see Figure 3.11). This pattern of significance contrasts with the low-sensitivity group where there was a main effect of stimulus, F (4, 76) = 4.06, $p < .05$, $\eta^2 = .18$, a main effect of duration, F (2, 38) = 10.22, $p < .05$, $\eta^2 = .35$, but and ***no*** stimulus by duration interaction, F (8, 152) =

1.82, $p = .08$, $\eta^2 = .09$ (see Figure 3.12). Thus, the sensitivity by duration interaction is driven by an effect of duration that is twice as large in the low-sensitivity group than in the high-sensitivity group. The low-sensitivity group shows gains in performance between 100 and 500 ms, t (19) = 4.19, $p < .01$, d =.94, and between 100 and 1000 ms durations, t (19) = 3.50, $p < .01$, d = .78, whereas the high-sensitivity group only shows a gain between 100 and 500 ms, t (19) = 3.04, $p < .01$, d = .68, and between 100 and 1000 ms, t (19) = 1.89, $p = .07$, d = .42. Ceiling effects are a likely contributor to the lack of an effect in the case of high-sensitivity listeners, however, *t*-tests on the arcsin transform of these data do not change the results: t (19) = 3.04, $p < .01$, d = .68 and t (19) = 1.90, $p = .07$, d = .42.

In order to confirm that the gains of the low-sensitivity group did not give them equivalent performance at 1000 ms, a *t*-test on the average at this duration was completed. The groups were still significantly different at this duration, t (38) = 6.75, $p < .05$, d = 2.13. In fact, only the mean of the static condition for the low-sensitivity group made it into the range of means achieved by the high-sensitivity condition. Overall, both low- and high-sensitivity participants did show gains in performance with greater durations, but the effect for low-sensitivity participants was two times greater.

Summary

In terms of the main question of Experiment 2 (and indeed this thesis), ***Is there a main effect of stimulus type on change detection performance?***, the answer was again "Yes". Evidence for this answer comes from the fact that Experiment 2 successfully replicated the effects of Experiment 1 showing poorer discrimination for 500-ms-long dynamic stimuli as compared with 500-ms-long static stimuli. Further support for this claim comes from the change-detection advantage for static stimuli that was found in Experiment 2 for multiple types of dynamic stimuli.

For the secondary question of, ***Is there a main effect of duration on change detection performance?***, the answer was "Yes", as predicted. The evidence for this answer comes from the finding that change detection accuracy increases with longer stimulus duration for the static stimuli and dynamic stimuli overall. For the more nuanced question of, ***Is there an interaction between duration and stimulus type on change detection performance?***, the answer was also "Yes" because dynamic-slope stimuli showed an inconsistent increase in accuracy with increasing duration. Here, change detection increased with longer durations for the dynamic-100-slope stimuli but not the dynamic-300-slope stimuli. Additionally, the difference in performance between various stimulus types increased at longer durations, and were eliminated at the 50 ms duration. These differences generally patterned as predicted based on the magnitudes of the frequency changes for each stimulus type. Overall, these findings demonstrate that the effect of duration is not the same for all types of dynamic stimuli.

Experiment 2 again demonstrated individual differences in change detection, so with respect to the question, ***Are the effects of duration and stimulus type equivalent for all listeners?***, the answer is "No". Here it was found that for low-sensitivity listeners, the positive effects of duration were less (η^2 = .35) than the effects of duration for high-sensitivity listeners (η^2 = .43). Further, high-sensitivity listeners showed significant gains between durations of 500 ms and 1000 ms, whereas the low-sensitivity listeners did not. Lastly, with respect to static stimuli, performance on even these stimuli were never equivalent for high- and low-sensitivity listeners suggesting a fundamental difficulty in processing sound for the purpose of change detection in low-sensitivity listeners that that persists even with very simple stimuli presented at very long durations.

Combining the results of Experiments 1 and 2, it can be seen first of all that change-detection accuracy was generally poorer for dynamic than for static stimuli in both experiments. This is consistent with the greater complexity of dynamic stimuli, i.e., the fact that dynamic stimuli incorporate more information than static stimuli.

Change-detection accuracy for static stimuli is highly accurate in Experiment 1 across all ISIs. At the relatively long 900 ms ISI in Experiment 2, this accuracy improves as a function of stimulus duration, although there was little change at durations beyond 500 ms. This indicates that change detection for static stimuli is determined primarily by stimulus duration. This in turn suggests that ASTM for static stimuli is determined primarily by stimulus duration: the longer the duration, the better the ASTM performance for the stimulus.

In Experiment 1, change-detection accuracy for dynamic stimuli was affected by ISI, being lower at the longest ISI of 1800 ms. In Experiment 2, at the 900 ms ISI, change-detection accuracy for dynamic stimuli increased as a function of stimulus duration in the same manner as for static stimuli, with accuracy increasing from the 50 ms duration to the 500 ms duration, but not increasing further for the 1000 ms. duration. This indicates that change-detection accuracy for dynamic stimuli is a function of both stimulus duration and ISI. This in turn suggests that ASTM for dynamic stimuli is a function of both stimulus duration and ISI. This finding is consistent with the fact that dynamic stimuli incorporate more information, so that change-detection accuracy (and presumably ASTM) is likely to be more affected by various manipulations. Finally, the fact that both manipulations examined (ISI and stimulus duration) affected the time available for encoding and/or maintenance of the first stimulus suggests that ASTM is time-bound, and subject to trace decay.

With respect to individual differences, Experiment 1 showed that participants who differ in change-detection accuracy at long ISIs also differ at short ISIs. Experiment 2 showed that participants who differ (at the 900 ms ISI) in change-detection accuracy at short durations also differ at long durations. Further, in both experiments, while ASTM was better functioning overall, a difference in the processing of static and dynamic stimuli was still found. Thus, relative to low-sensitivity listener the functioning of ASTM was elevated for both static and dynamic stimuli, but within the high-sensitivity group the same (though perhaps attenuated) effect of dynamic stimuli was found on ASTM function as compared with its functioning in response to static stimuli. These results suggest that while high-sensitivity listeners may have overall greater ASTM function than their low-sensitivity peers, effects of dynamic stimuli may persist under some, but not necessarily all conditions. The greater functioning ASTM thus appears to be a very general benefit whose implications will need to be further investigated in the future.

In terms of our present objectives, Experiment 3 aims to generalize the finding of poorer memory for dynamic compared with static stimuli to one last paradigm. This is done by testing the discrimination of *sequences* of dynamic and static stimuli, thus introducing changes in the *temporal order* between stimuli, another common measure of auditory perception (Surprenant & Watson, 2001). In addition to being a common measure of auditory perception, this task assesses the memory for sequences of cues, an important ability for recognizing and remembering consonants, as they consist of sequences of short-duration sounds occurring in rapid succession.

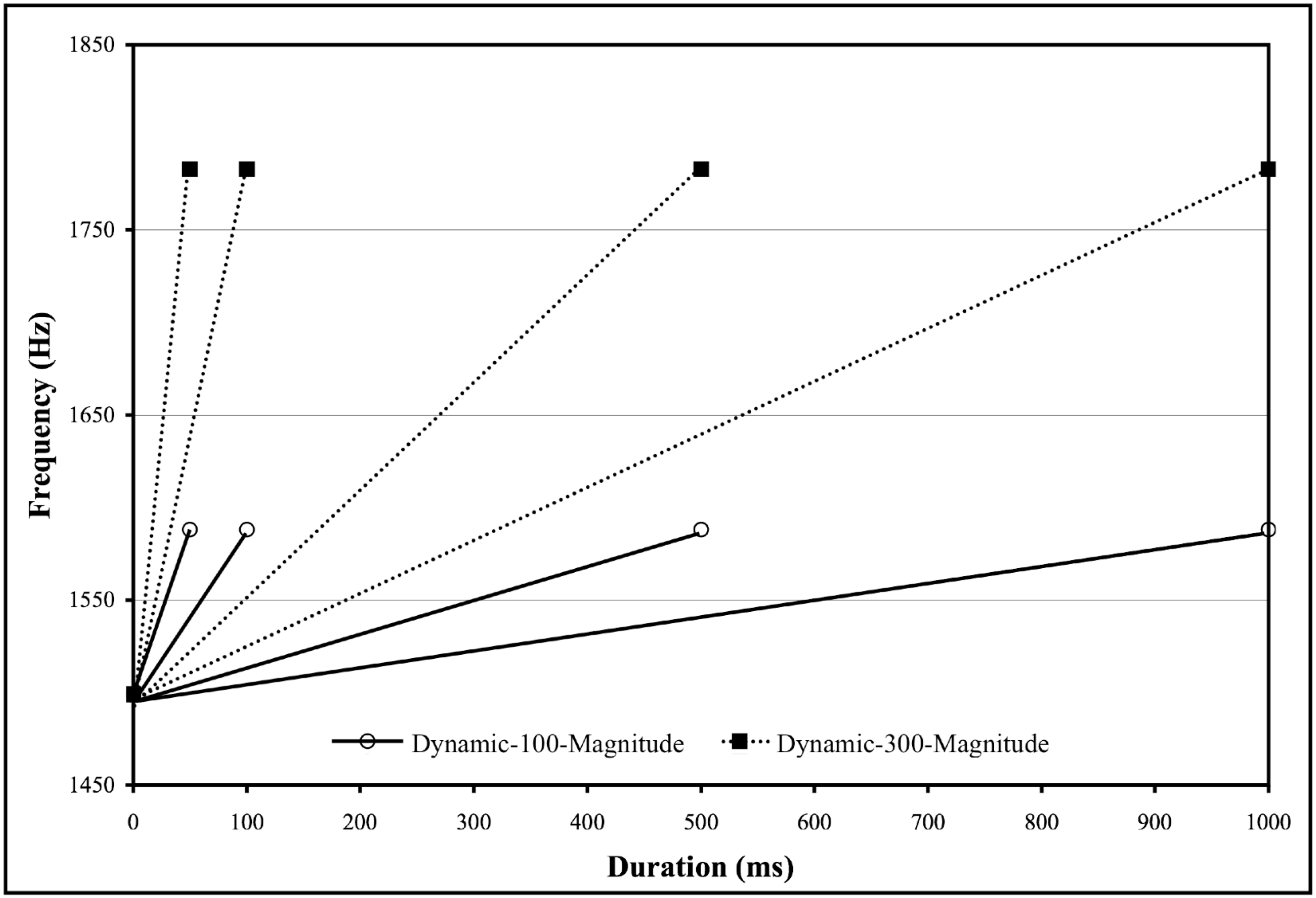

Figure 3.1. Frequency magnitude/slope profiles of a single dynamic stimulus with a starting frequency of approximately 1500 Hz. Separate plots are shown of the frequency values a 100-cent and a 300-cent magnitude stimulus passes through as a function of duration. Each stimulus at each duration covers the same number of frequencies, but in an increasingly longer period of time.

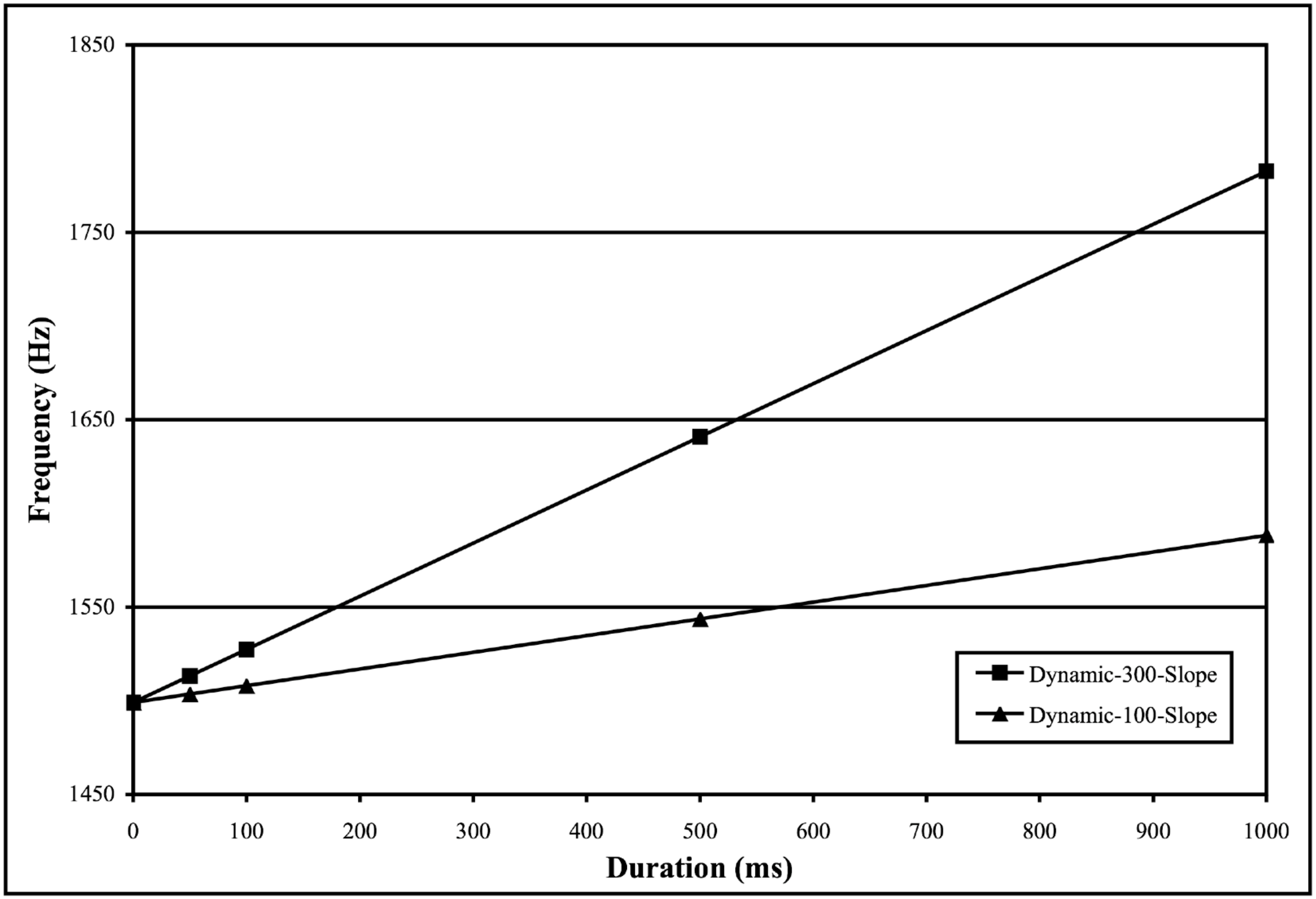

Figure 3.2. Frequency magnitude/slope profiles of a single dynamic stimulus with a starting frequency of approximately 1500 Hz. Separate plots are shown of the frequency values a 100-cent and a 300-cent magnitude stimulus passes through as a function of duration. Each stimulus at each duration covers a larger number of frequency values than the previous duration.

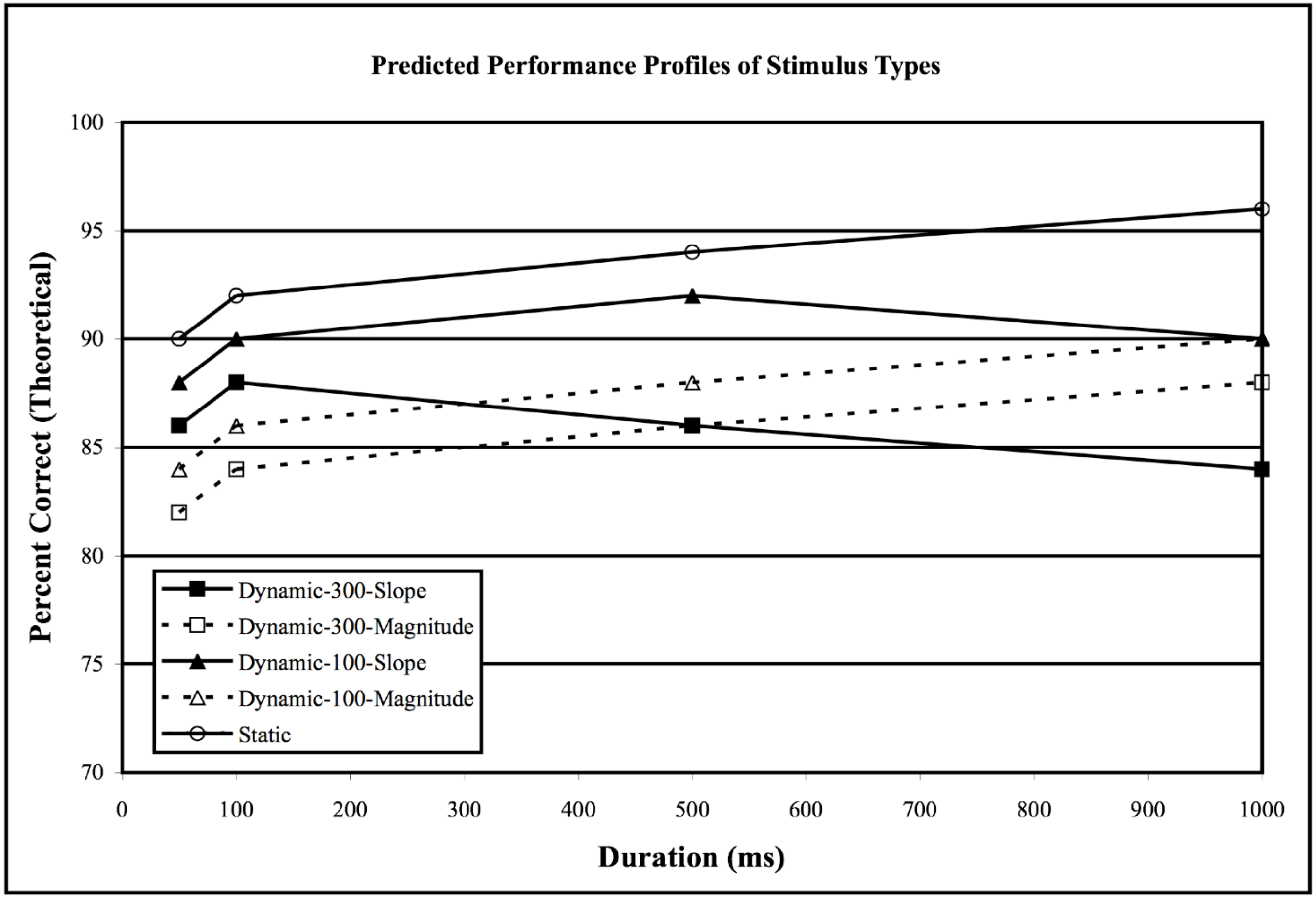

Figure 3.3. Predicted change-detection accuracy for all five stimulus types as a function of duration. The data depict qualitative differences as a function of stimulus type and duration, rather than quantitative predictions. The predicted percent correct scores simply hypothesize a main effect of stimulus driven by the frequency change magnitude at the 50 ms duration for each stimulus condition. At longer durations, an arbitrary increase or decrease of 2% was applied based on whether change detection is predicted to be easier or harder with an increase in duration.

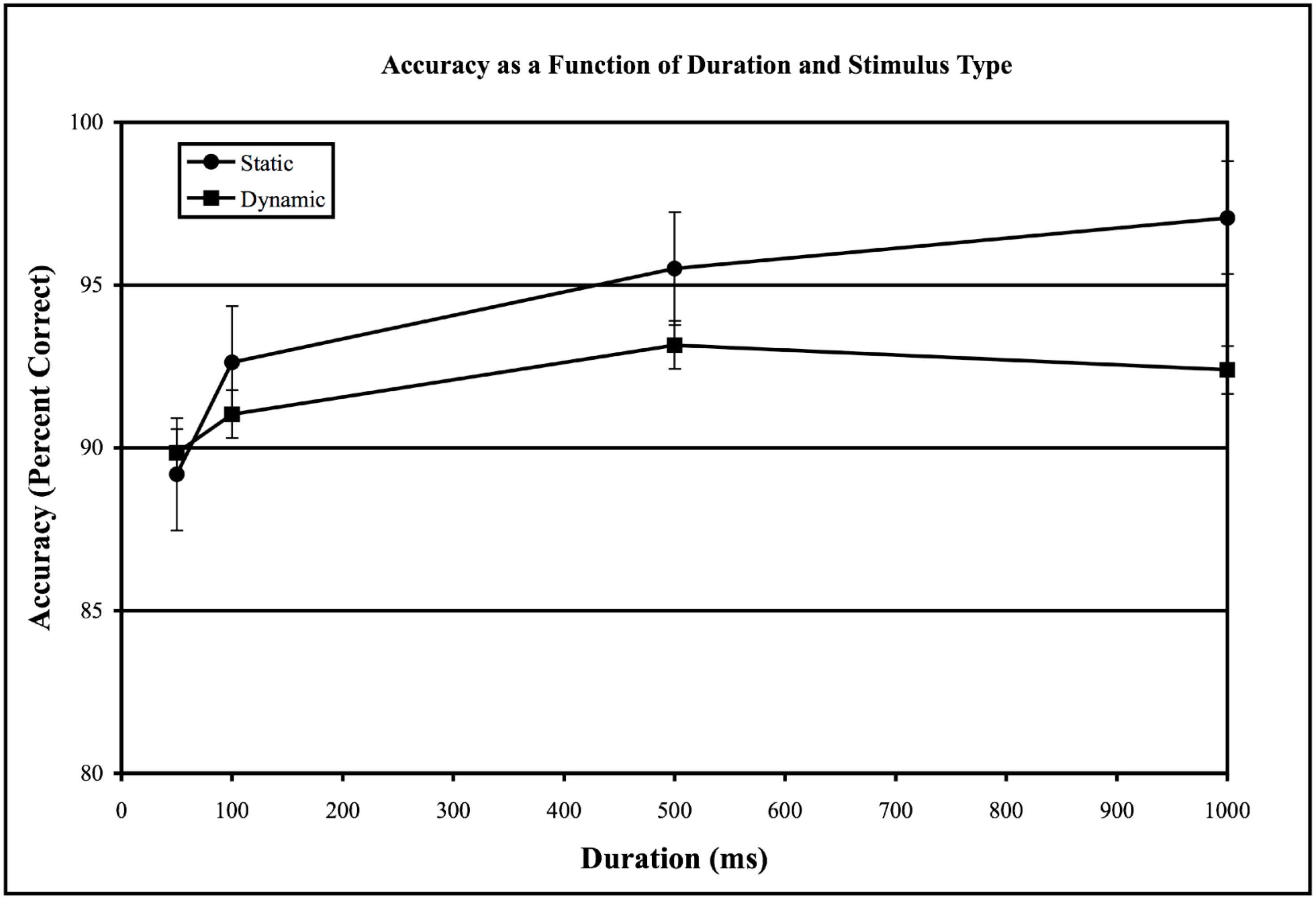

Figure 3.4. Accuracy as a function of duration and stimulus type. Bars are SEM.

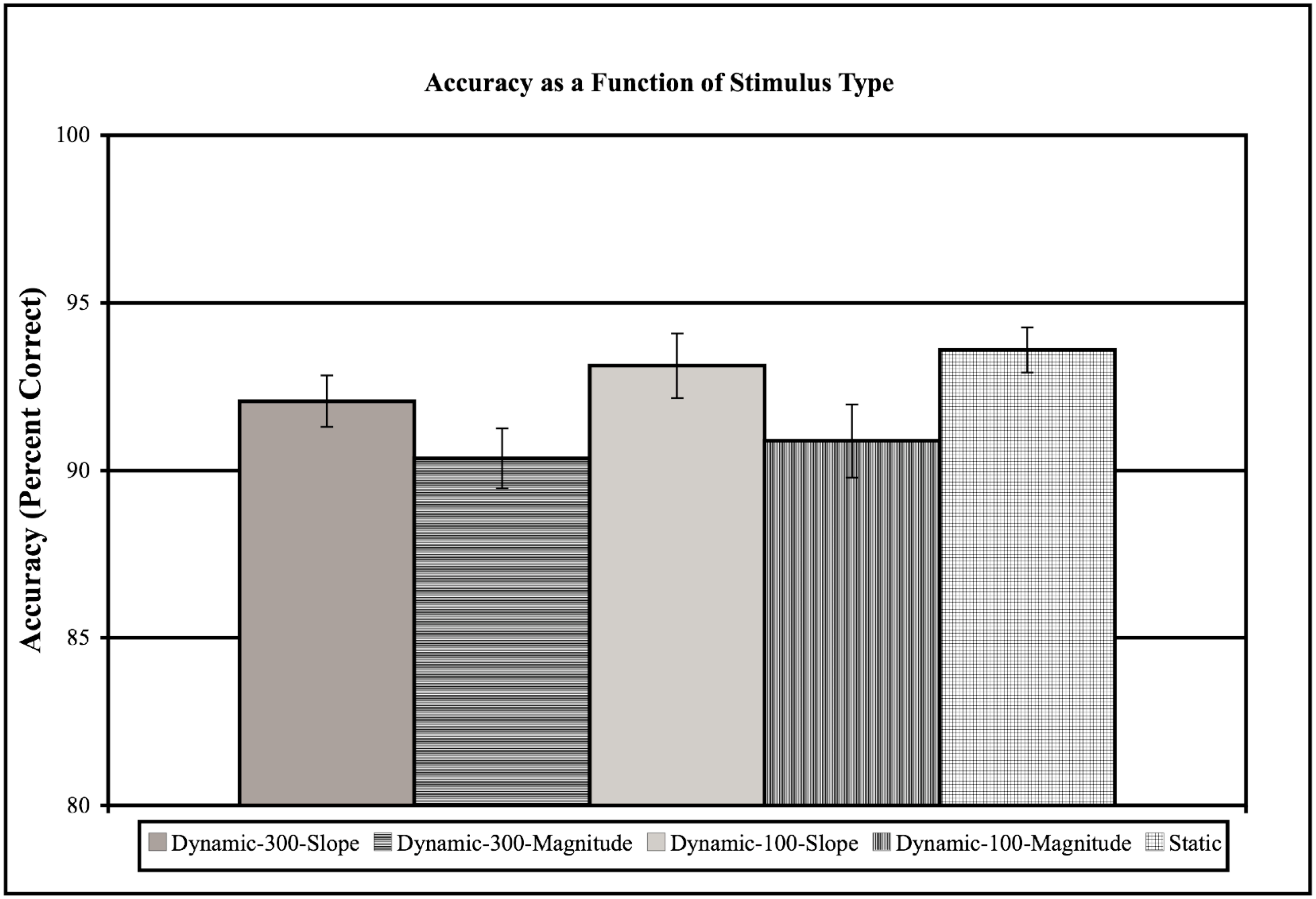

Figure 3.5. Accuracy as a function of stimulus type. Bars and SEM for the parameter.

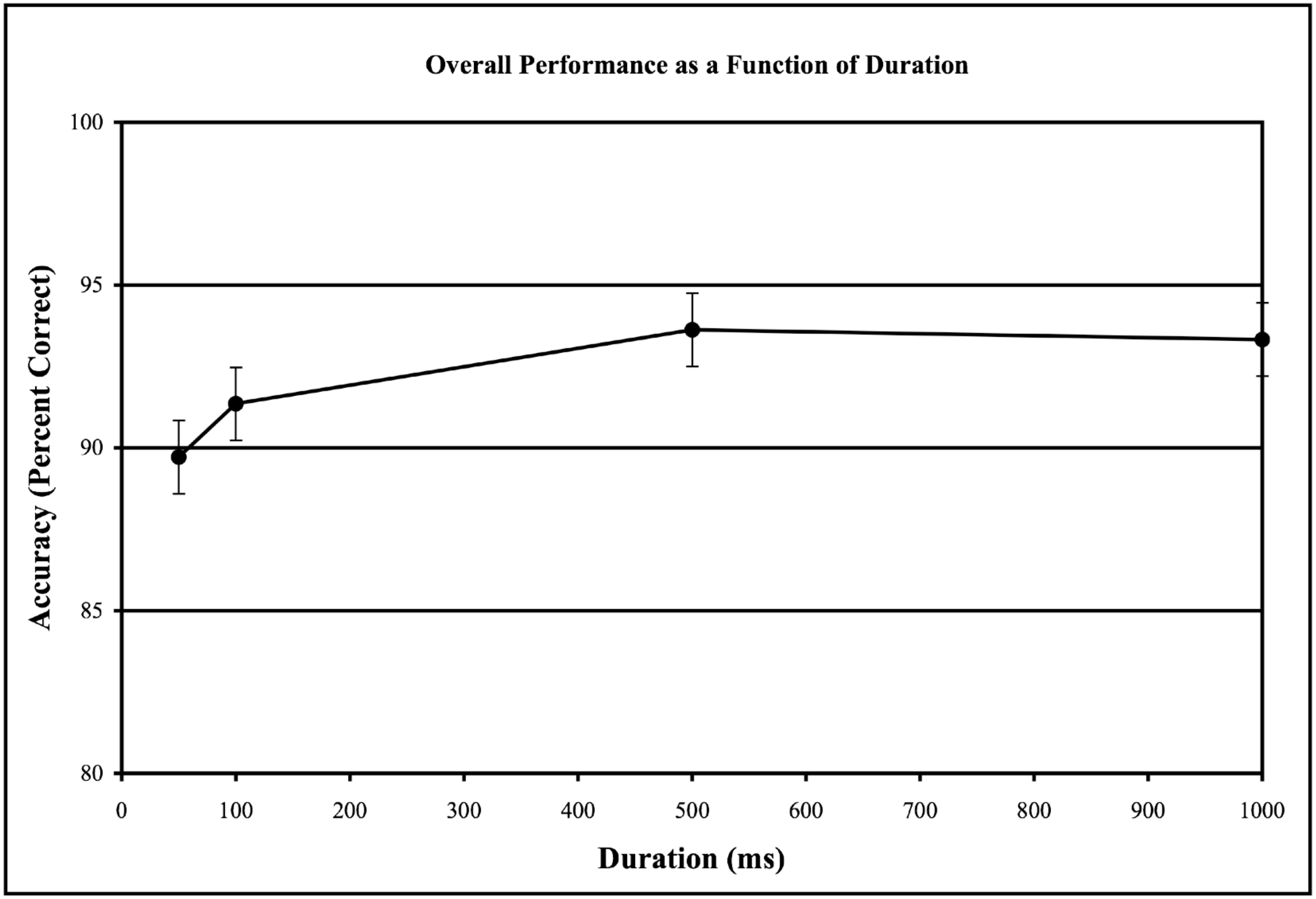

Figure 3.6. Percent correct as a function of duration. Bars are SEM.

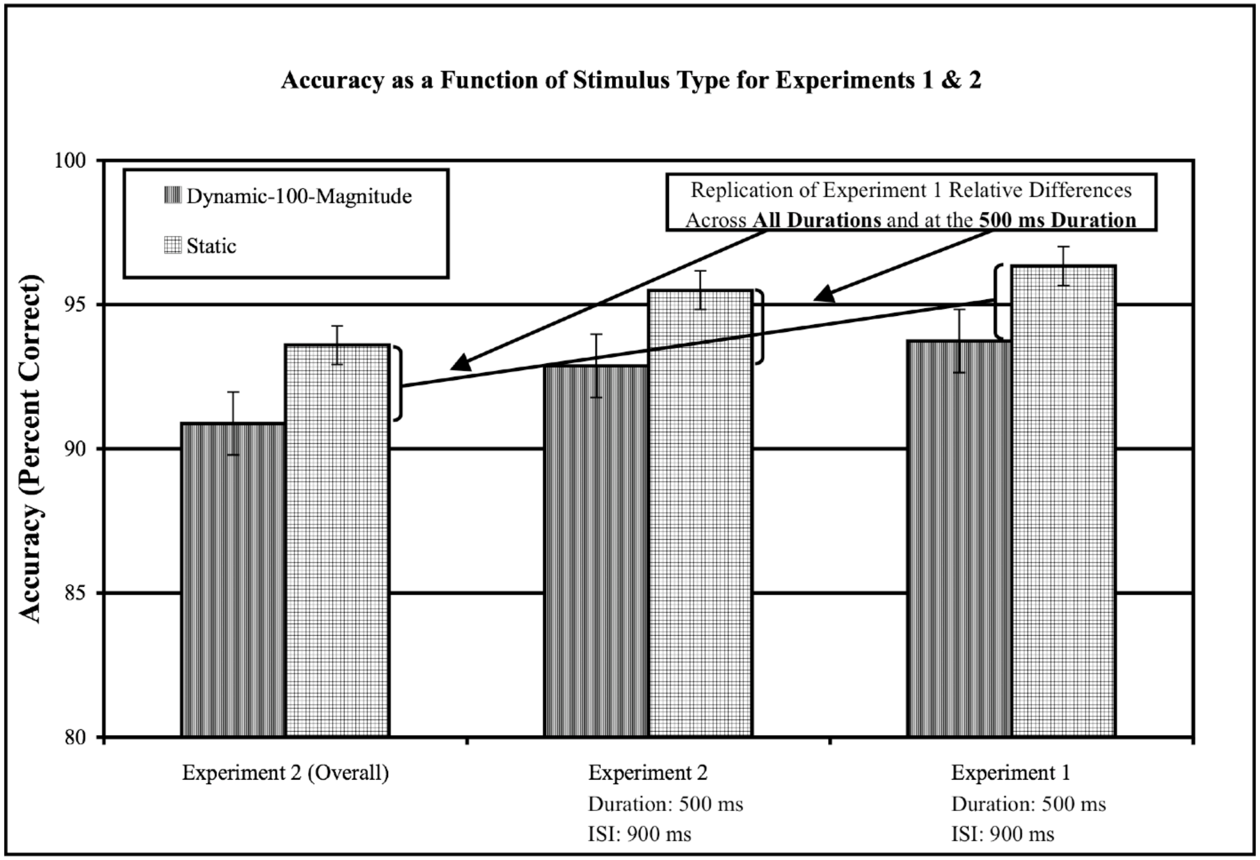

Figure 3.7. Percent correct as a function of stimulus type across Experiments 1 & 2. The left set of bars are show performance for dynamic-100-magnitude and static stimuli averaged across all durations (an extension of the effect seen in Experiment 1). The center set of bars show performance at just the 500 ms duration (with ISI set to 900 ms; a direct replication of Experiment 1 at the 900 ms ISI). The right set of bars show accuracy from the dynamic and static stimuli of from Experiment 1. The brackets show that the relative differences between static and dynamic-100-magnitude are the same in all three comparisons. Bars are SEM for the stimulus type.

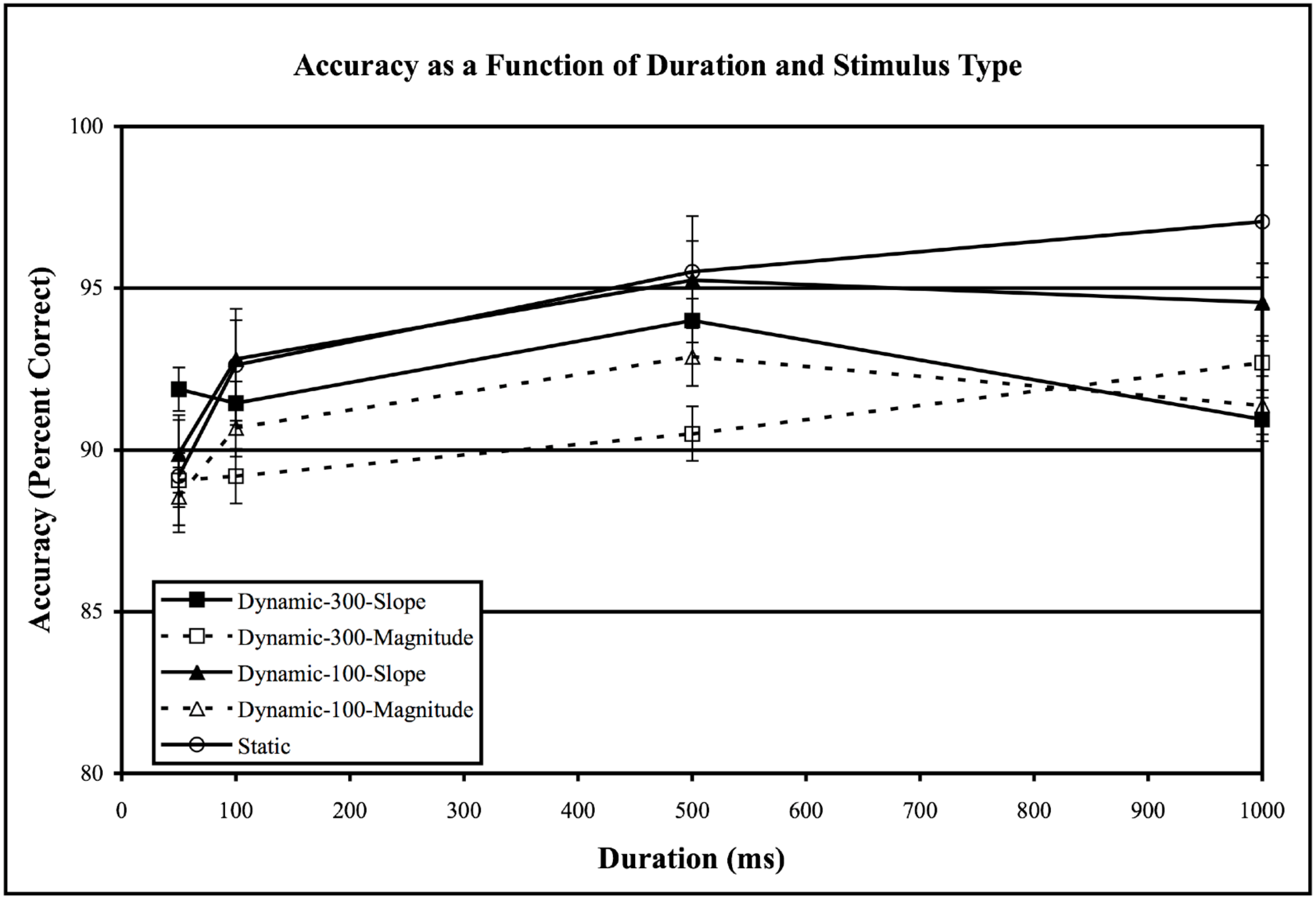

Figure 3.8. Percent correct as a function of duration and stimulus type. Bars are SEM for the parameter.

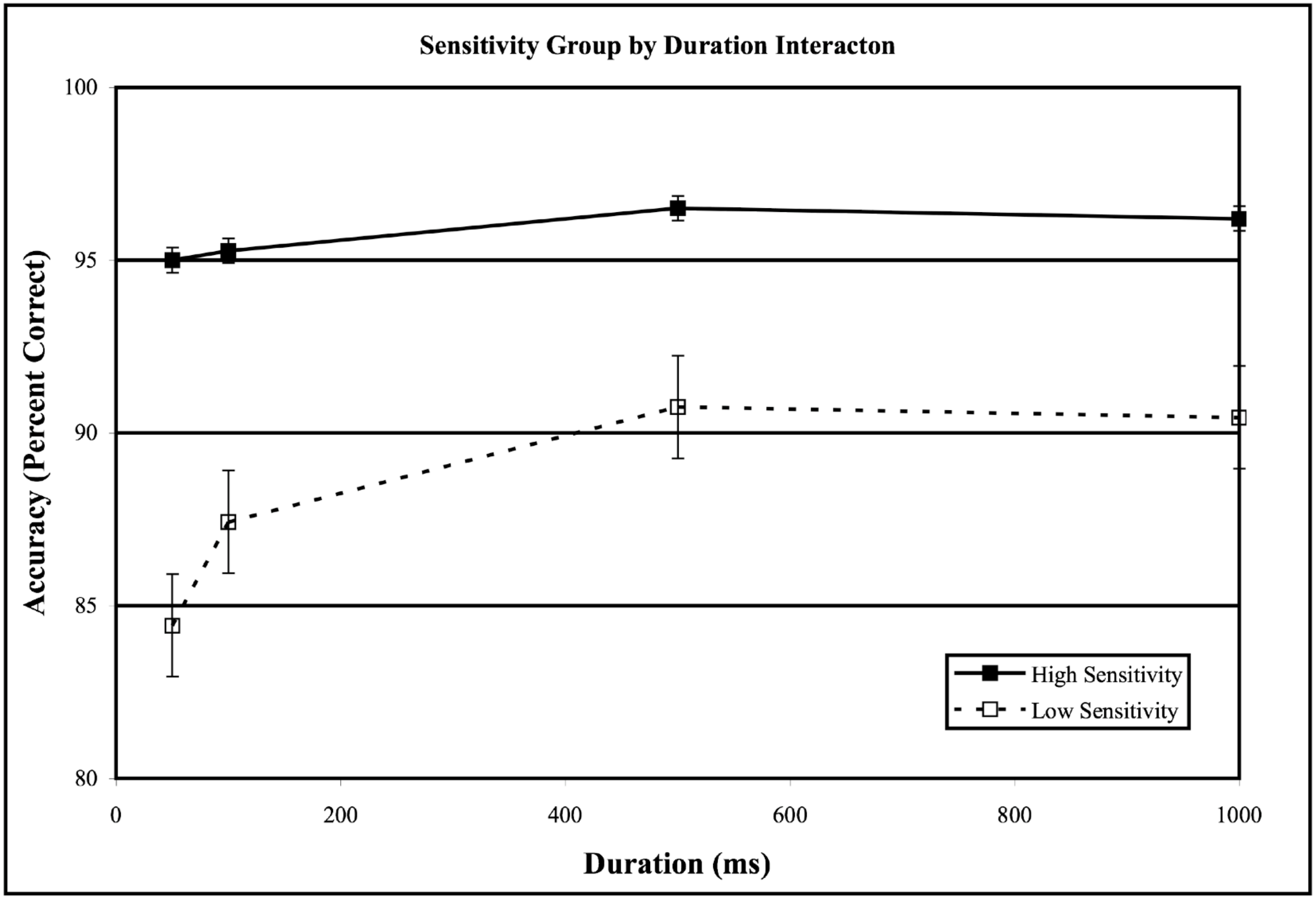

Figure 3.9. Accuracy for both sensitivity groups as a function of duration. Bars are SEM for the parameter.

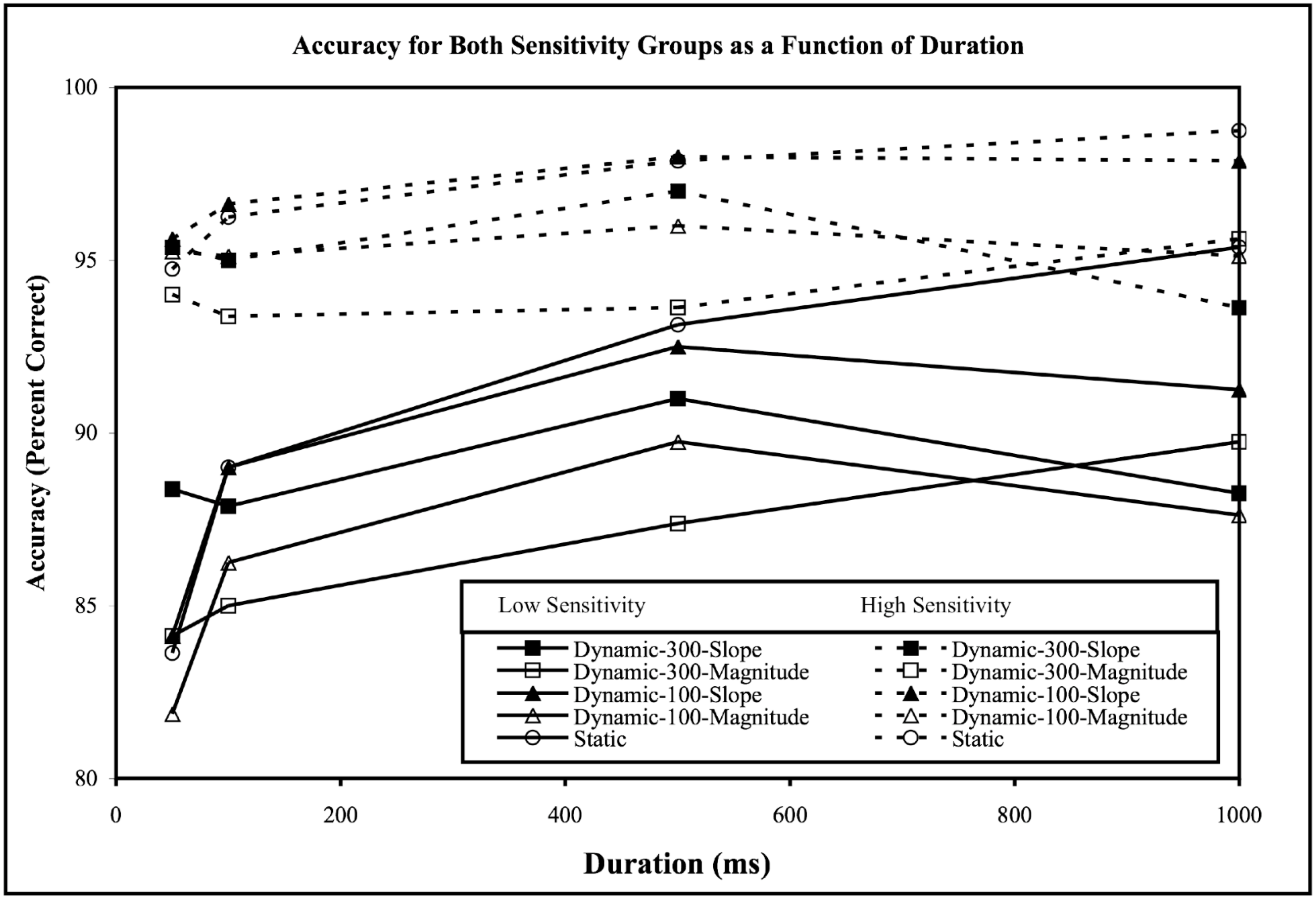

Figure 3.10. Accuracy for both the high- and low-sensitivity group as a function of stimulus and duration. Error bars have been removed for clarity, but can be found on the Figures 3.11 and 3.12.

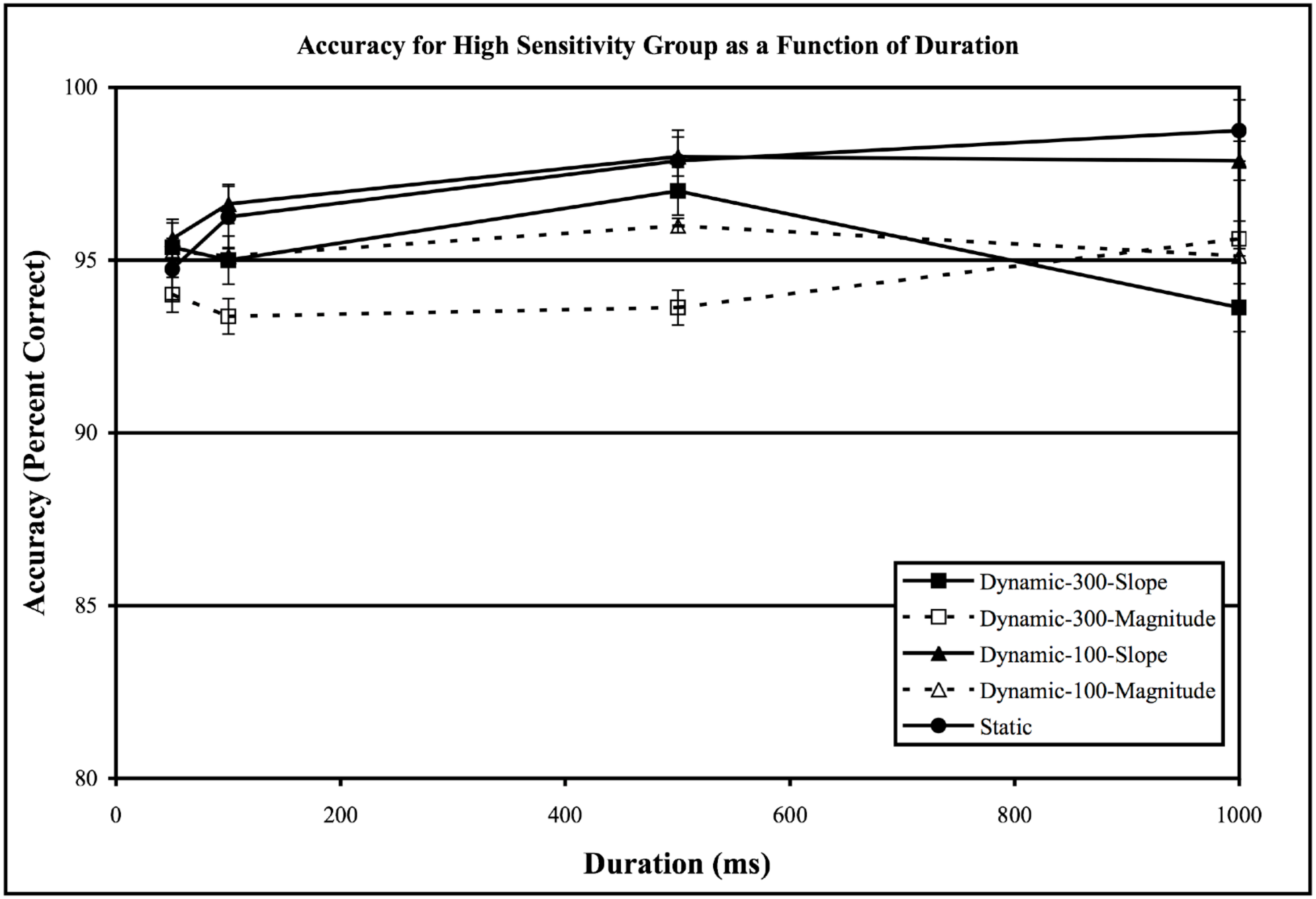

Figure 3.11. Accuracy for the high-sensitivity group as a function of duration and stimulus type. Bars are SEM for the parameter.

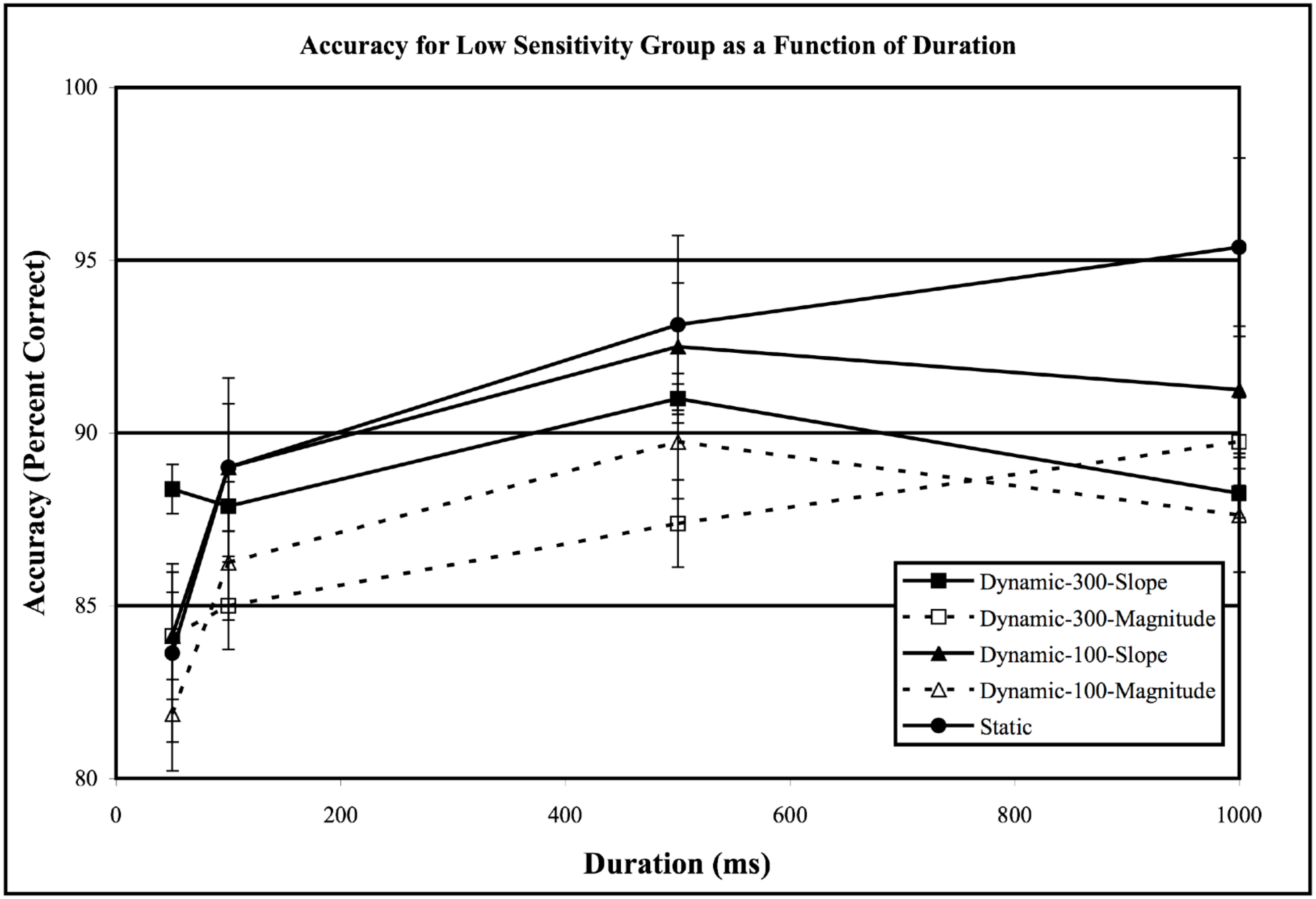

Figure 3.12. Accuracy for the low-sensitivity group as a function of duration and stimulus type. Bars are SEM for the parameter.

CHAPTER 4

CHANGE DETECTION IN SEQUENCES OF STATIC AND DYNAMIC STIMULI

Experiment 3

Experiment 3 shifts its focus to *sequences* of static and dynamic sounds in a change-detection paradigm, but asks the same fundamental question of: ***Is change detection more accurate for static than dynamic stimuli?*** Experiments 1 and 2 have shown differences between change-detection accuracy (and presumably ASTM) for static and dynamic stimuli as a function of ISI and stimulus duration, with the task requiring participants to detect changes to the frequency of a *single* stimulus presented in isolation. This task involves the comparison of only two stimuli. Experiment 3 asks whether the fundamental differences between change detection for static and dynamic stimuli exist in a change detection task involving *sequences* of stimuli. This task was chosen because it has previously been related to speech perception and language processing (Surprenant & Watson, 2001). If the differences between static and dynamic stimuli persist in this sequential task, it would leave open the possibility that both the sequencing demands of the task and the type of stimulus used in the task could potentially relate to language processing as it would again suggest that static and dynamic stimuli are not equivalently processed by ASTM. Such a finding, combined with findings from previous studies showing no relationship between static stimuli and speech processing, would predict that in future studies of the relationship between sequential change detection and language processing change detection for dynamic stimuli will have a stronger relationship with language processing than static stimuli.

A sequential change detection task involves the comparison of item and order information for multiple stimuli. Accordingly, I will examine both whether participants can detect a change to the frequency of a single sound in a sequence of sounds and whether participants can detect changes to the temporal order of a sequence. Again, interest will center on whether the capability to detect such changes interacts with the type of stimulus presented in each sequence (static or dynamic). Thus, a secondary question of the present experiment can be framed as: ***Is there a difference between detecting changes in the frequency of a sound and changes in the temporal order of sounds?***

The interest in sequencing lies in the fact that auditory stimuli that are extended in time consist of both frequency and temporal order information. Specifically, this temporal order information consists of the serial order of the frequencies present in a dynamic sound. There is now considerable evidence that the maintenance of serial order information for the sequences of sounds that comprise words and nonwords is much the same as the maintenance of serial order for the sequences of words or digits within a *list* (Abbs, Gupta, Jochem, & Armogan, 2006; Gathercole & Baddeley, 1989; Gathercole, Frankish, Pickering, & Peaker, 1999; Gathercole, Willis, Emslie, & Baddeley, 1992; Gupta, 2003; Gupta, Lipinski, Abbs, & Lin, 2005; Gupta, MacWhinney, Feldman, & Sacco, 2003). On the assumption that the same relationship holds for the maintenance of serial order of frequencies within dynamic stimuli vs. sequences of dynamic or static stimuli, Experiment 3 examines change detection for sequences (i.e., lists) of dynamic and static stimuli, where maintenance of serial order information is critical for the task. The type of sequence element is manipulated (within a list, all elements are either static or dynamic stimuli), and the type of change present on "different" trials is manipulated (either the frequency of one list element changes, or two list elements are exchanged with each other). The results are intended to provide new information about factors that affect maintenance of serial order information (again presumably in ASTM), for lists, which should have implications for the maintenance of serial order *within* dynamic stimuli.

A number of specific predictions can be made. First, I hypothesize that memory for a sequence of dynamic stimuli will be more accurate than memory for a *sequence* of static stimuli, reflecting the greater informational complexity (specifically, the inherent ordering demand) of dynamic stimuli, as in Experiments 1 and 2. While I have just presented two experiments in which change detection for dynamic stimuli was poorer than for static stimuli, neither of these experiments required the maintenance of a serial order of stimuli. Here I hypothesize that because there may be a greater temporal processing demand for dynamic stimuli, this greater demand may translate in more accurate change detection for sequences of dynamic stimuli. Further, if there is a stronger relationship between dynamic stimuli and language processing, the inherent ordering demands of language may make the system more attune to the temporal properties of dynamic stimuli and performance may be elevated above that of static stimuli.

Relatedly, a “change type” by “stimulus type“ interaction is also predicted such that change-detection accuracy for static stimuli would be lower with frequency change than with temporal change, whereas change-detection accuracy for dynamic stimuli would be relatively lower with frequency change than with temporal change. Again, this hypothesis comes from the increased temporal demand that may exist for dynamic stimuli and the better availability of temporal information for dynamic stimuli in ASTM as compared with the availability of frequency information and as compared with the availability of temporal information for static stimuli.

Lastly, based on what is known about recognition memory in the case of sequences of static tones (Watson et al., 1975), I may hypothesize that accuracy will change as a function of serial position. Changes to primacy (i.e., beginning) portions of a sequence should be detected less often than changes to medial portions, which in turn should be detected less often than recency (i.e., end) portions (primacy < medial < recency).

One difficulty with this new paradigm is that *two* stimuli will necessarily change when a change to the temporal order is made, whereas only one stimulus will change when a change to the frequency of a stimulus is made. However, the primary interest of this experiment is in the prediction of an interaction with stimulus type (rather than in comparing change types themselves) and this interest circumvents the need to control for this type of difference magnitude. Difference magnitude is another variable that may be examined later.

Another key aspect of this paradigm is that only hit rate will be considered. "same" trials were included as part of the design to keep listeners unbiased, but a false alarm rate for these "same" trials cannot be calculated *as a function of serial position*: if a stimulus is "different" it may be different in a particular serial position, however by definition a "same" trial has no such specificity. Thus, because accuracy as a function of serial position is of interest, only the hit rate of the present task will be analyzed and it will be the hit rate for differences as a function of serial position. A hit rate will be used even for analyses that average across serial position, to maintain continuity between measures.

Method

Participants

Thirty-six participants (20 = Female; 16 = Male) from the Elementary Psychology pool at the University of Iowa participated in this experiment.

Materials

Static stimuli were the same as the static stimuli of 100 ms duration in Experiment 2. Dynamic stimuli were the same as in the dynamic-100-magnitude condition of Experiment 2 at the 100 ms duration. As this task was considerably more difficult than the tasks of Experiments 1 & 2, no noise was presented behind the stimulus.

Design and Procedure

On each trial, participants first heard a sequence of five tones selected from the 44 potential stimuli. A sequence was generated by randomly selecting from this set without replacement. Thus, each stimulus in the sequence was of a different frequency.

The type of tone presented in a sequence was a within-subjects factor that was manipulated between blocks of 100 trials. Thus, for every block participants went through 10 practice trials of the stimulus for that block (static or dynamic) followed by 100 experimental trials (participants received a fifteen second break every twenty trials during the experimental session). Each tone in the sequence was 100 ms and was separated by 250 ms from adjacent tones. This inter-tone interval was selected to provide perceptual boundaries between stimuli and avoid temporal masking, but maintain the need for the type of precise temporal information thought to be represented in the dynamic stimuli. Following this first sequence, there was a 900 ms pause and a second sequence presentation that followed the same stimulus duration and inter-tone parameters. On half of the trials of a block, this sequence was the same as the first ("same" trials), on the other half of trials the sequence was different ("different" trials).

On "different" trials, the to-be-changed tone was randomly selected, without replacement, from one of the five serial positions of the sequence. Two types of changes could be made to this stimulus. The type of change presented to participants (temporal versus frequency information change) was the second within-subjects factor of this experiment and was manipulated in a blocked fashion.

For blocks in which the difference between stimuli on a "different" trial was along the temporal dimension (i.e., the sequence of tones was changed), the randomly selected tone exchanged its temporal location with one of the other serial positions of the sequence. The choice of position for the exchange was randomly selected, without replacement. This random selection of changed stimuli and the position with which it is exchanged continued without replacement for ten change trials, the minimum number needed to exhaust all possible change types. For a set of 10 trials, the combinations of exchanges needed to exhaust all possible changes was: 1st-2nd, 1st-3rd, 1st-4th, 1st-5th, 2nd-3rd, 2nd-4th, 2nd-5th, 3rd-4th, 3rd-5th, and 4th-5th. This resulted in 4 changes of the 1st magnitude, 3 changes of the 2nd magnitude, 2 changes of the 3rd magnitude, and 1 change of the 4th magnitude.

For blocks in which the difference between stimuli on a "different" trial was along the frequency dimension (i.e., the frequency of a single tone was changed), a randomly selected tone was chosen for the change and it was exchanged with a stimulus from the stimulus set that was either higher or lower and of a pre-selected magnitude difference. Stimuli were changed according to a schedule that attempted to equate the change magnitude for frequency and sequence position. Thus, within a block of 10 change trials, the frequency changed by +/- 1 step (880 Hz), 4 times, 2 steps (1160 Hz), 3 times, 3 steps (1240 Hz), 2 times, and 4 steps (1320 Hz), 1 time[4]. Whether the frequency value of a stimulus was raised or lowered by a specific value was randomly determined unless one type of change would have lead to a frequency value outside of the 500 – 4000 Hz range of the stimulus set. In the case that such a change would lead to a frequency outside of this range, the change was made in the opposite direction.

In summary, this experiment manipulated the type of stimulus (dynamic and static) and the type of change (frequency and temporal) as two within-subjects factors in a 2 x 2 factorial design that yields four conditions (dynamic-frequency, dynamic-temporal, static-frequency, and static-temporal). Further, within each of these factors, the magnitude of the change (1, 2, 3, or 4 steps for frequency and magnitude) and the serial position involved in the change (1st, 2nd, 3rd, 4th, and 5th) was controlled and can be examined as a factor in performance on this task.

---

[4] Pilot studies showed that the originally proposed step sizes of 80, 160, 240, and 320 Hz made the task too difficult.

Results and Discussion

Effects of Change Magnitude on Change Detection

A 2 (Change Type) X 2 (Stimulus Type) X 4 (Magnitude) Repeated-Measures ANOVA was run on the hit rate. This ANOVA revealed no 3-way interaction between change type, stimulus, and magnitude, F (3, 105) < 1 (see Figure 4.1), but significant simple effects. Specifically, there was a main effect of stimulus type, F (1, 35) = 5.46, $p < .05$, $\eta^2 = .14$ (see Figure 4.2). The main effect of stimulus type favors static stimuli over dynamic stimuli (see Table 4.1 for means, standard errors, and η^2s related to the main effect of stimulus type in Experiment 3). Therefore, I once again have better discrimination for static as compared with dynamic stimuli, even when there is a difference in the temporal order of stimuli. Thus, the elevated change detection for static stimuli found in Experiments 1 & 2 extends to a task involving sequences of stimuli. Additionally, there is a marginal effect of change type, F (1, 35) = 3.44, $p = .07$, $\eta^2 = .09$ (see Figure 4.3) favoring temporal changes over frequency changes (see Table 4.1 for summary data related to the main effect of change type in Experiment 3). The elevated performance in the sequence change is likely due to the fact that two stimuli change their value in this condition, whereas only a single stimulus changes its value in the frequency change condition. Lastly, there is a main effect of magnitude, F (3, 105) = 6.28, $p < .05$, $\eta^2 = .15$ (see Figure 4.4). The effect of magnitude is as one would expect, with larger magnitudes leading to increasing sensitivity to change.

The main interaction of interest, between stimulus and change type, was not significant, F (1, 35) < 1 (see Figure 4.5). Thus, there is no difference in change-detection performance that depends on the exact stimulus and change type condition being presented. Additionally, the interaction between change type and magnitude was non-significant, F (3, 105) = 1.66, $p > .05$, $\eta^2 = .05$ (see Figure 4.6) suggesting that the magnitudes of changes in the temporal and frequency conditions were successfully equated. Lastly, the interaction between stimulus type and magnitude was marginally significant, *F* (3, 105) = 2.51, $p = .06$, $\eta^2 = .07$ (see Figure 4.7) suggesting that dynamic stimuli do not benefit as much from increased magnitudes of change.

To further examine the stimulus type by magnitude interaction shown in Figure 4.7, separate Repeated Measures One-Way ANOVAs with magnitude as the factor were run on hit rates for static and dynamic stimuli. An adjusted alpha level of .025 was used for these ANOVAS and it was found that the effect of magnitude was significant for static stimuli, *F* (3, 105) = 6.57, $p < .01$, $\eta^2 = .16$, but the effect for dynamic stimuli was half as large and non-significant, *F* (3, 105) = .04, $p = .04$, $\eta^2 = .07$. Further, a trend analysis showed that the effect was linear in the case of static stimuli, *F* (1, 35) = 13.36, $p < .05$, $\eta^2 = .28$, but not linear in the case of dynamic stimuli, *F* (1, 35) = 1.88, $p = .18$, $\eta^2 = .05$. To further assess the effect of magnitude and its interaction with stimulus type, four post-hoc paired-samples *t*-tests were run with an alpha level of .01. These tested the difference between static stimuli at magnitude one and magnitude four, dynamic stimuli at magnitude one and magnitude four, static stimuli at magnitude one and dynamic stimuli at magnitude one, and static stimuli at magnitude four and dynamic stimuli at magnitude four. These tests further tested the effect of magnitude on static and dynamic stimuli and tested the differences between change detection for static and dynamic stimuli at the extreme ends of the magnitude scale. It was found that the difference between magnitude one and four was significant for static stimuli, *t* (35) = 3.78, $p < .01$, d = .63, but non-significant for dynamic stimuli, *t* (35) = .283, $p > .01$, d = .05, Further, there was no significant difference between static and dynamic stimuli at magnitude one, *t* (35) = .194, $p > .01$, d = .03, but there was a significant difference between static and dynamic stimuli at magnitude four, *t* (35) = 2.33, $p < .01$, d =.39. All together, these tests indicate that magnitude has an effect on static stimuli, but not dynamic stimuli and that the difference between stimulus types is at its minimum with small magnitude changes and at its maximum with the largest magnitude changes.

The main effects of this analysis based on magnitude of changes should be given greater weight than the effects that will be considered next. This is because, in the upcoming analysis by serial position, the nature of the "different" trials for the temporal order condition (i.e., they involved changes to two stimuli at two serial positions) means that performance on a single trial will be used twice in determining the grand means for the effect, so that hit rate as a function of serial position is not a very clean measure.

Effects of Serial Position on Change Detection

To assess the effects of serial position on change detection, a 2 (Change Type) x 2 (Stimulus Type) x 5 (Serial Position) Repeated-Measures ANOVA was conducted on the hit rate for this experiment. This ANOVA again revealed a marginal main effect of change type favoring temporal changes over frequency changes, F (1, 35) = 3.82, $p = .06$, $\eta^2 = .10$, no main effect of stimulus, F (1, 35) = 2.53, $p = .12$, $\eta^2 = .07$ (see Table 4.2), and critically no interaction between change type and stimulus type F (1, 35) < 1. While the finding of no effect of stimulus is contradictory to the previous analysis that collapsed across serial positions for the magnitude factor, the present analysis includes trials more than once in the grand means that determine the significance of main effects. Thus, these means are biased and are less reliable indicators of the significance of main effects. This discrepancy can be seen by comparing the means for the stimulus factor in Table 4.1 (M = .723 for static versus M = .680 for dynamic) with the means for the stimulus factor in Table 4.2 (M = .706 versus M = .680).

Before discussing the simple effects of this ANOVA, I want to discuss an interesting finding of the present experiment that lies in a further analysis of the change type by stimulus type by serial position interaction, F (4, 140) = 3.58, $p < .05$, $\eta^2 = .09$. A visual inspection of the data (see Figure 4.8) suggest three effects, the significance of which will be tested with the appropriate *t*-test.

First, for dynamic stimuli (see Figure 4.9), the data suggest that participants did not detect frequency changes to the first position when the sequences are made up of dynamic stimuli. The mean hit rate was .54 (standard error = .04) for detecting changes to the first tone of a sequence of dynamic stimuli. The appropriate hypothesis for our evaluation of this mean is that there is no difference between this hit rate and chance (.50). The *t*-test comparing performance in this condition to chance found no significant difference, t (35) = .990, p = .33, d = .16. For position two, mean performance jumps to .64 with a .04 standard error, which is above chance levels, t (35) = 3.84, $p < .001$, d = .64, and comparable to performance in the static condition at serial position one and two. Thus, listeners showed virtually no sensitivity to changes in the frequency values of dynamic stimuli in the first position of the sequence. This finding suggests that *ASTM does not reliably maintain the first item of a sequence of dynamic stimuli in the face of interference from four subsequent dynamic stimuli*. This interference could be a memory-based effect, such as proactive interference or a capacity limitation, in which stimuli or it could be that the act of sequential processing leads to accumulated interference as a result of performing a number of processing actions in a short period of time and that this interference affects dynamic stimuli more than static stimuli. This type of central interference would be similar to the central sources of temporal masking that Massaro (1975) demonstrated by finding a temporal masking effect even when the masking tone was presented to the opposite ear of the target tone. Curiously though, this effect does not appear to be the case when the temporal order of the stimulus is changed. Though it is true that changes to temporal order would implicate other serial positions, at which changes to the stimulus would be better detected, it is not clear that this explanation completely accounts for the better performance at position 1 for frequency changes compared with temporal order changes.

Second, the data suggest that at the 3rd serial position (i.e., the exact middle of the sequence), detecting differences in the temporal order of static tones is more difficult than with dynamic stimuli. It also suggests that it is more difficult than detecting changes to the frequencies of static tones, which indicates that memory for the temporal order is less accurate in medial positions (see Figure 4.10). Three paired-samples *t*-tests were run with an alpha level of .01 to evaluate these hypotheses (as well as a final hypothesis discussed below). The first *t*-test shows that it is more difficult to detect changes to the order of tones with static stimuli compared with dynamic stimuli at this 3rd position of the tone sequence, $t\ (35) = 2.66$, $p = .01$, d = .44. The second shows that it is also more difficult (and to the same extent) than detecting changes to the frequencies of static tones at this 3rd position, $t\ (35) = 2.66$, $p = .01$, d = .44. Third, the data suggest that the difficulty with detecting changes to the serial positions of dynamic tones comes in the 4th position and that detecting changes here is more difficult than detecting these changes in static stimuli. This interpretation was confirmed with a *t*-test showing a significant difference between these means, $t\ (35) = 5.24$, $p < .01$, d = .87.

More generally, there was also a main effect of serial position, $F\ (4, 140) = 11.14$, $p < .05$, $\eta^2 = .24$ (see Figure 4.11). This effect is most evident as a recency effect, with changes being detected the best when they are made to the later items in the sequence. Overall the effect increased the hit rate on changes to the last items of the sequence by .098. There was also change by serial position interaction, $F\ (4, 140) = 9.06$, $p < .05$, $\eta^2 = .21$, that showed a primacy benefit for temporal changes, but not frequency changes (see Figure 4.12). A set of five paired-samples t-tests with an adjusted alpha level of .01 showed that the differences at serial positions 1, 2, and 3 (where frequency changes are detected more accurately) are all significant, $t\ (35) = 4.58$, 3.51, and -2.59, *ps* $\leq .01$, d = .76, .59, and .43, respectively. The final two positions were not significantly different, $t\ (35) < 1$.

In addition to a change by serial position effect, there was a stimulus by serial position effect, $F\ (4, 140) = 4.76$, $p < .05$, $\eta^2 = .12$. This effect shows elevated performance for static stimuli in the first and fourth position of the sequence, but no differences at other serial positions (see Figure 4.13). A set of paired-samples t-tests confirmed the nature of this interaction, t (35) = 3.33, $p < .01$, d = .56 for the first serial position, t (35) = 2.86, p = 2.86, $p < .01$, d = .48. All other *p*s were greater than .17.

Summary

In this third experiment, static stimuli were again easier to discriminate than dynamic stimuli**.** This marks the third time a benefit for static over dynamic stimuli were found in a change-detection task utilizing ASTM. It extends the findings of Experiment 1 & 2 to a change-detection paradigm in which sequences of five sounds were presented and differences in the temporal order and frequency of individual stimuli had to be detected. In fact, there is good evidence that only four of the five dynamic stimuli were encoded and that there was no memory for the first stimulus of a dynamic sequence when listeners were discriminating frequency changes. This is evidenced by the chance change detection of frequency changes at the first position of dynamic sequences. Thus, for all three of the experiments presented here, the answer to the question of: ***Is there a main effect of stimulus type (static vs. dynamic) on change detection performance?***, the answer is “Yes”.

In the present experiment, it was additionally found that dynamic stimuli showed no benefit for increased magnitudes in the amount of change involved in a “change” trial. Further, with the smallest change in magnitude, there was no difference in the discrimination of static and dynamic stimuli. In terms of the serial position curve, there were slight nuances in the curves of each stimulus type, but the general finding of primacy and recency for the detection of temporal changes and recency for the detection of frequency changes was found for both stimulus types. Importantly, there was no primacy for dynamic stimuli when a change was made to the frequency of the stimulus in the first position. This indicates that listeners are “change deaf” for changes to the first dynamic stimulus of a sequence of dynamic stimuli.

On the question of ***Is there a difference between detecting changes in the frequency of a sound and changes in the temporal order of sounds?***, it was found that there was a marginal advantage for the detection of temporal changes over frequency changes. This difference was the same for both static and dynamic stimuli.

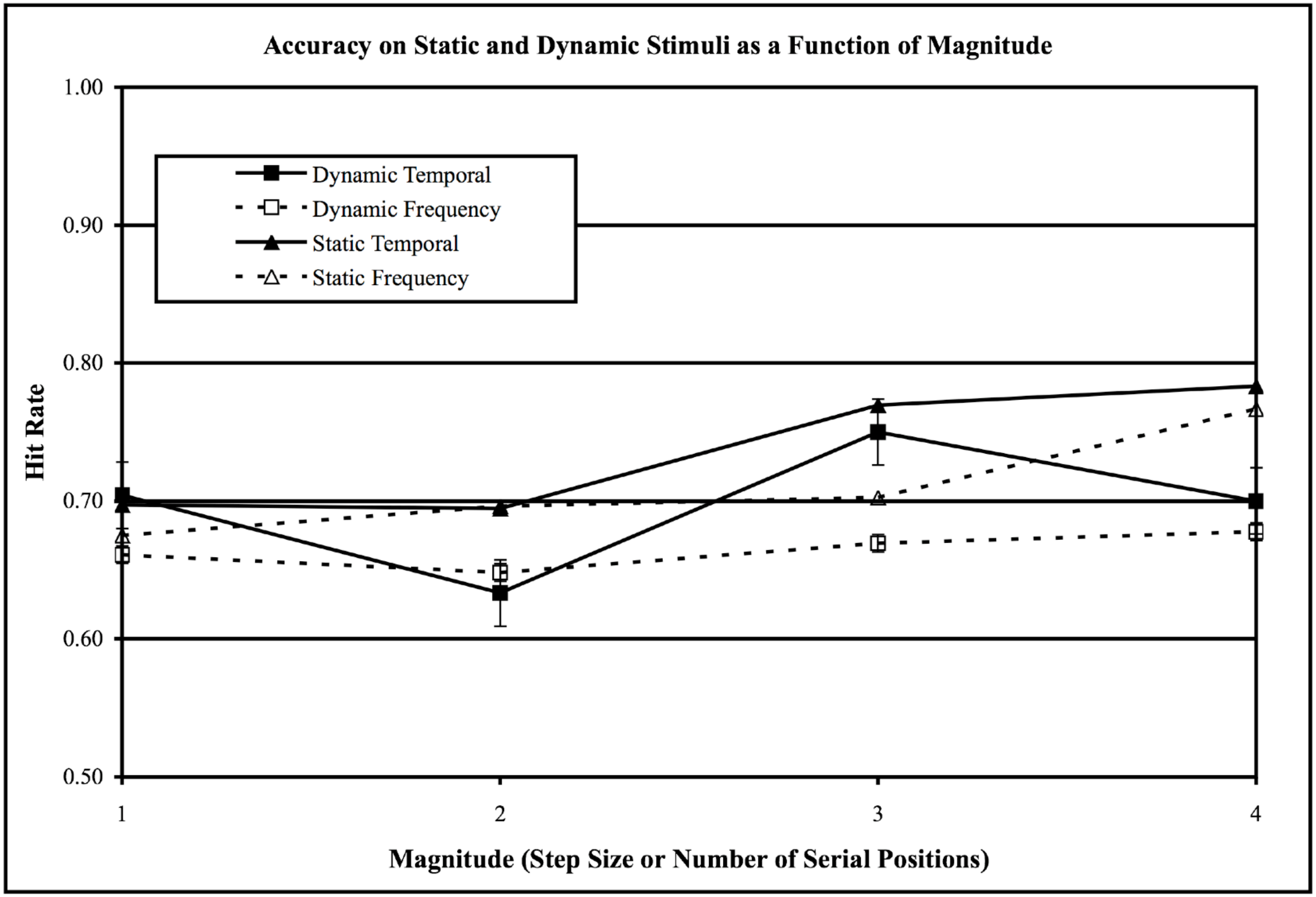

Figure 4.1. Hit rate as a function of stimulus type (squares versus triangles), stimulus type (solid line versus dashed line), and magnitude (along the x-axis). Bars are SEM for the parameter.

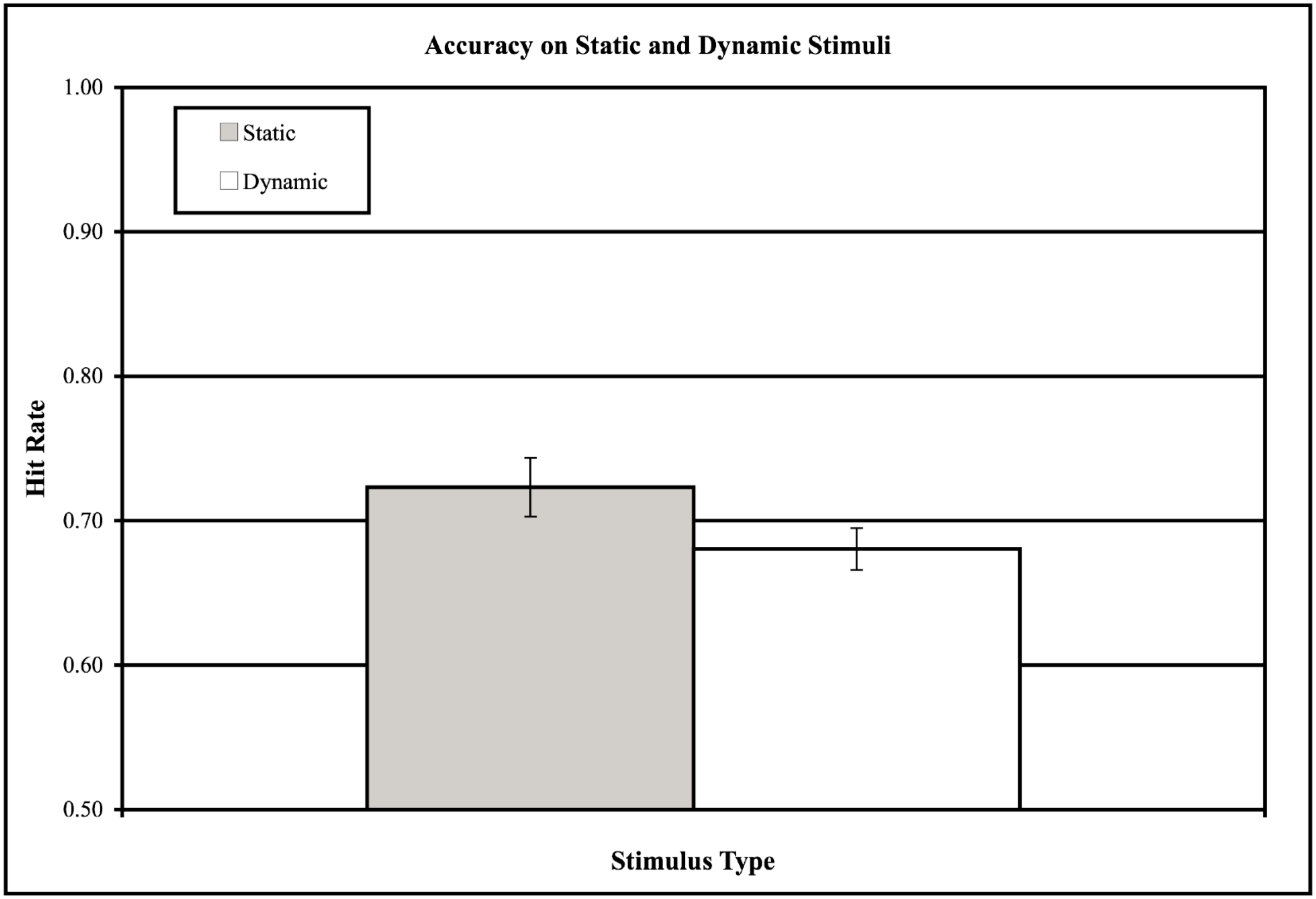

Figure 4.2. Hit rate as a function of stimulus type. Bars are SEM for the parameter.

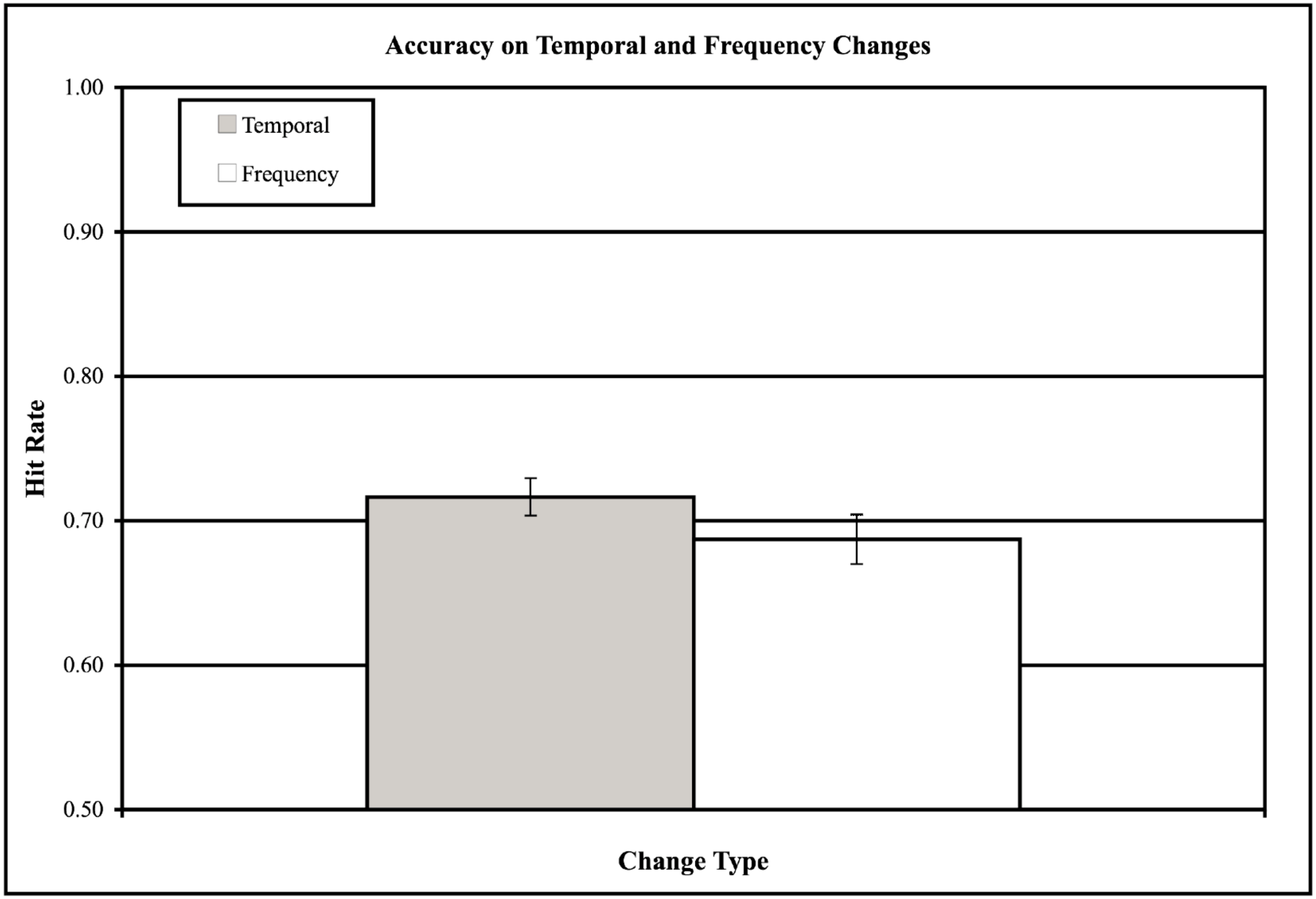

Figure 4.3. Hit rate as a function of change type (temporal order versus frequency). Bars are SEM for the parameter.

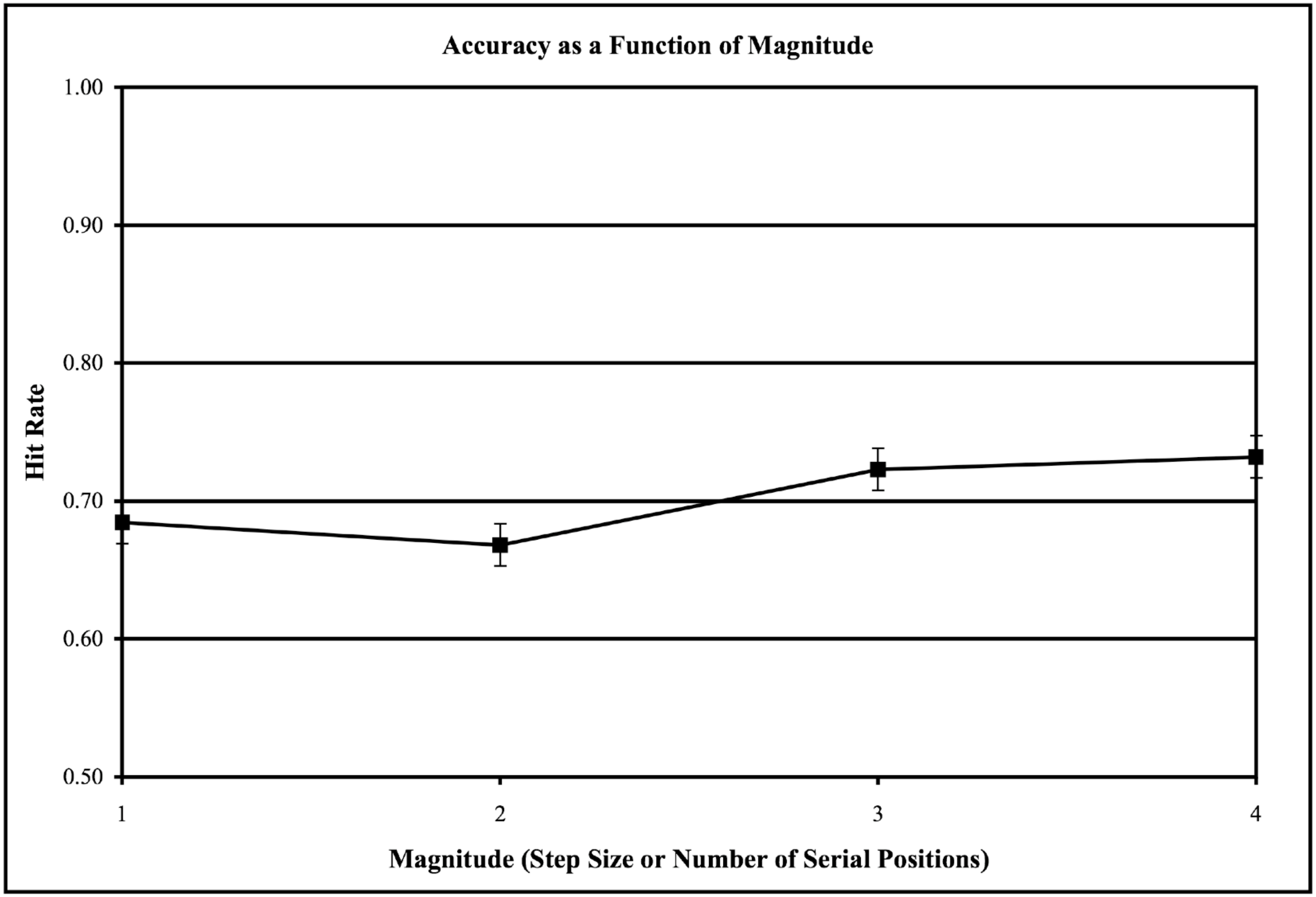

Figure 4.4. Hit rate as a function of magnitude. Bars are SEM for the parameter.

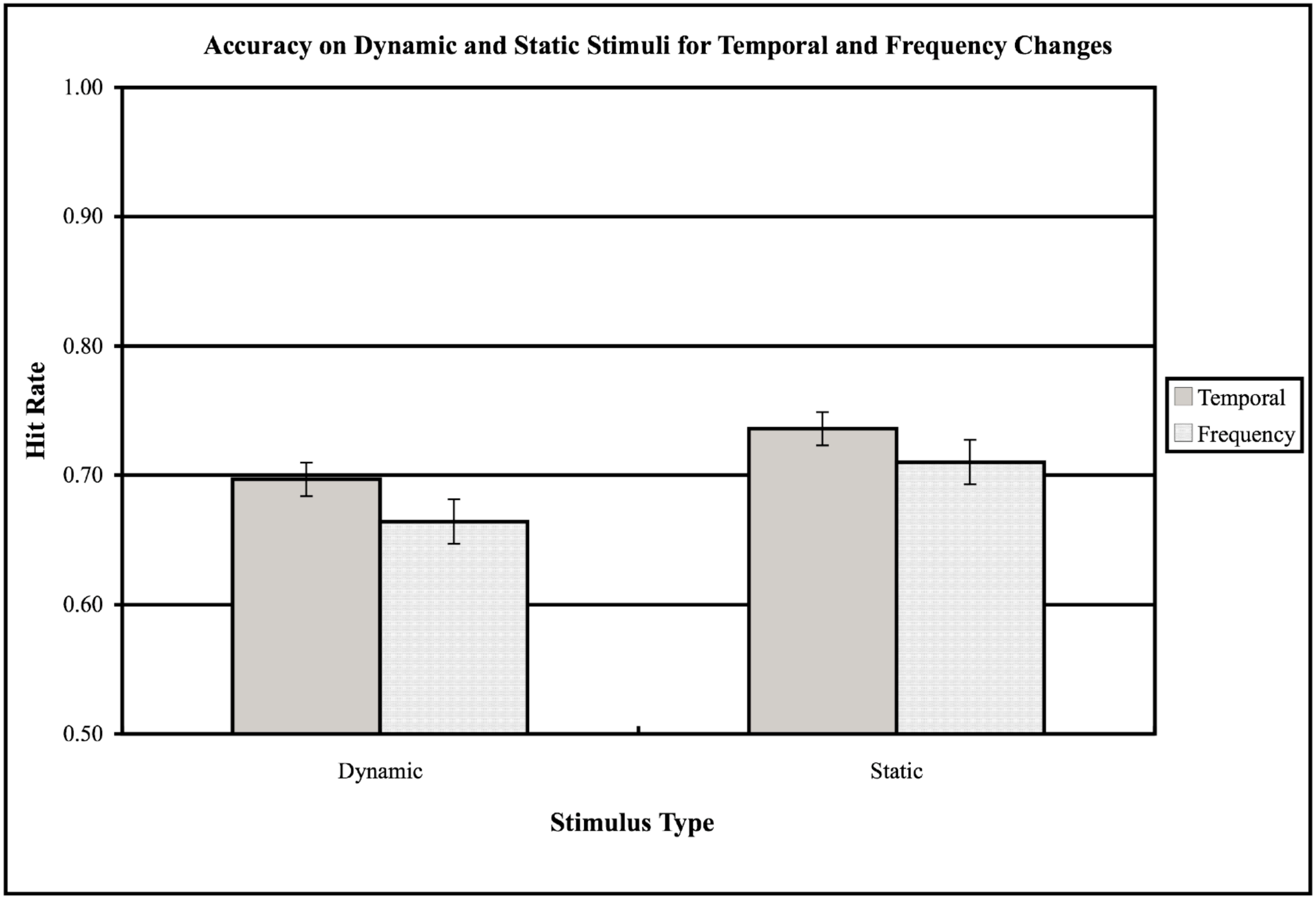

Figure 4.5. Hit rate as a function of change type and stimulus type. Bars are SEM for the parameter.

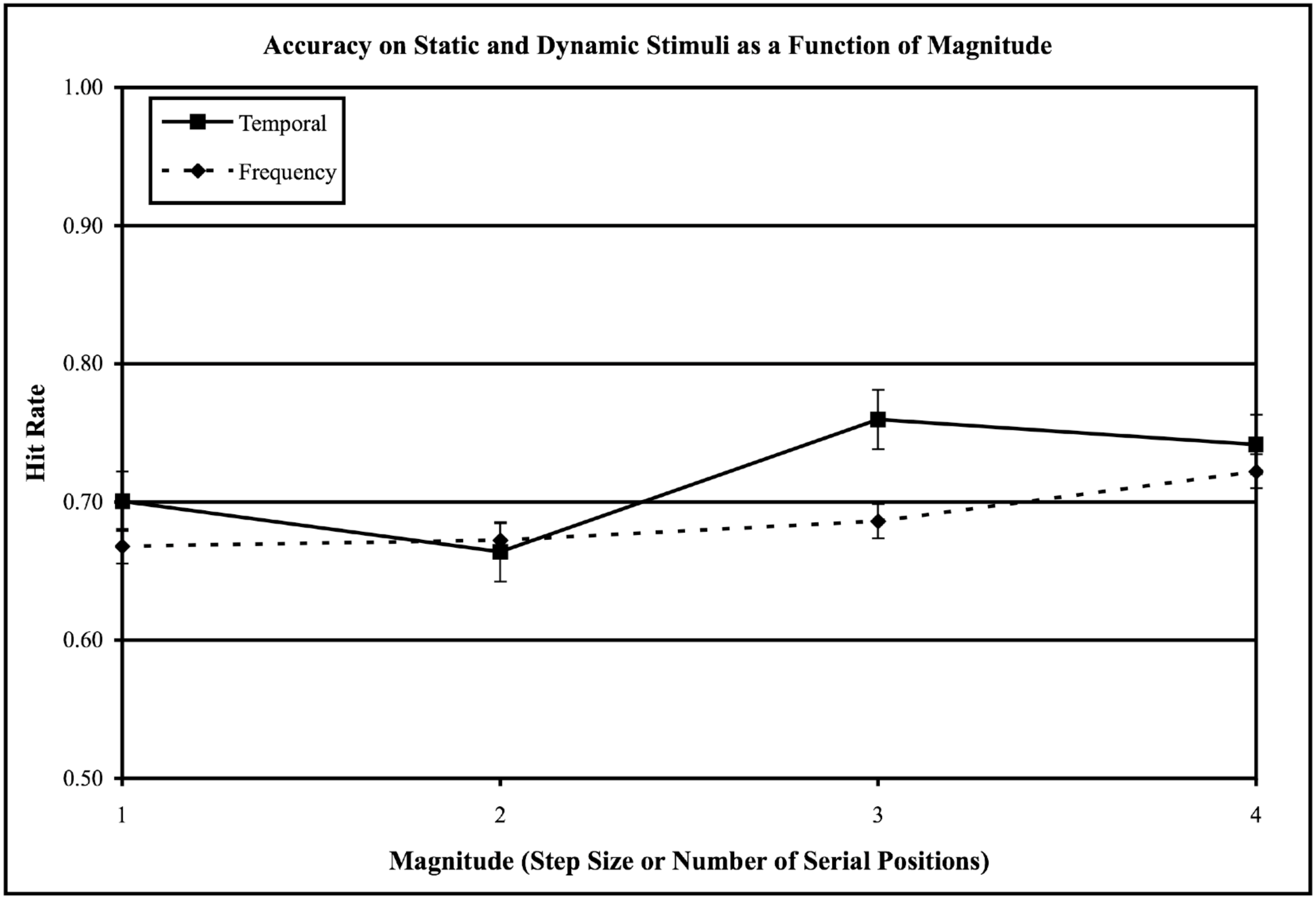

Figure 4.6. Hit rate as function of change type and magnitude. Bars are SEM for the parameter.

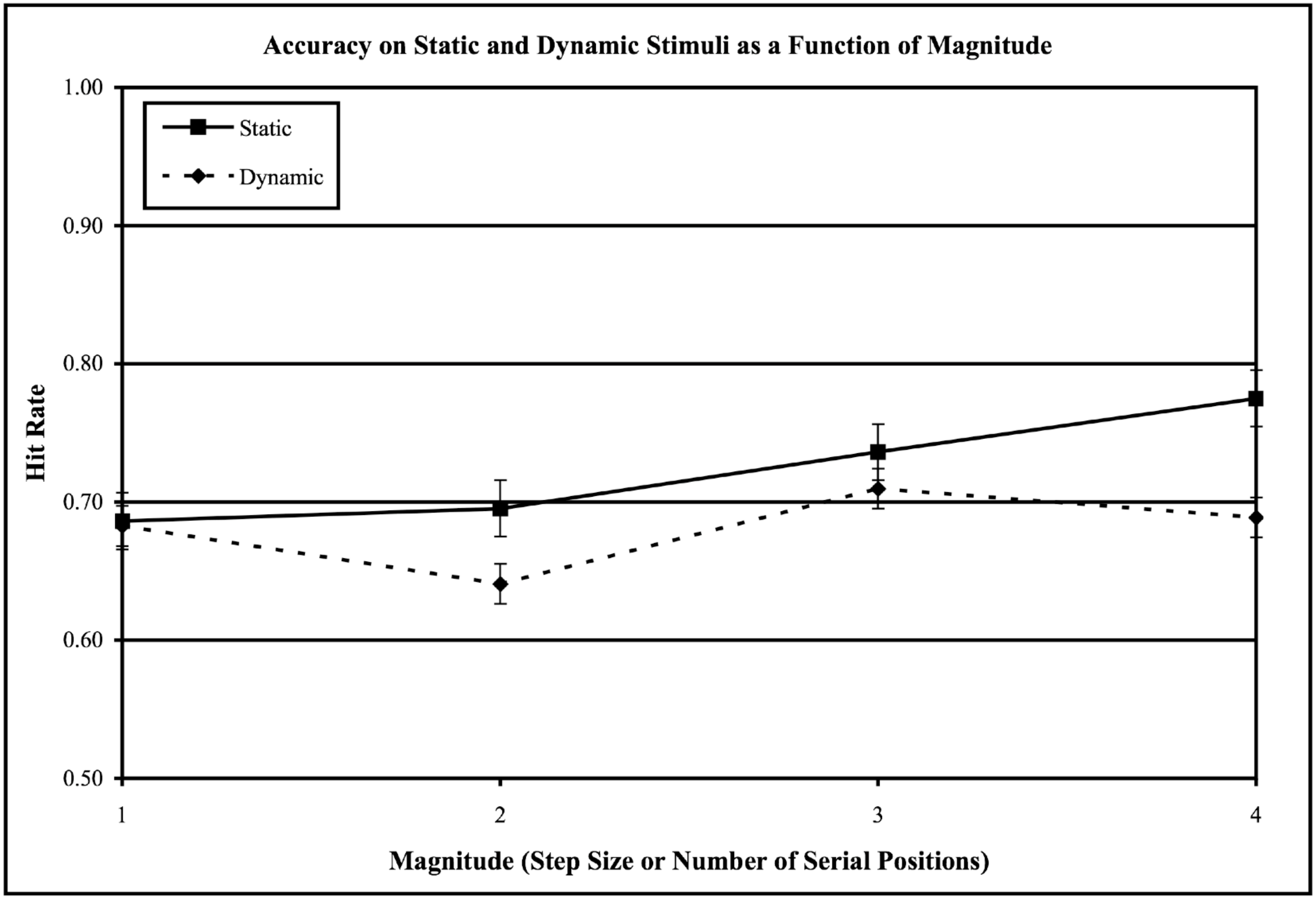

Figure 4.7. Hit rate as a function of stimulus type and magnitude. Note the linearly increasing function in the case of static, but not dynamic stimuli. Bars are SEM for the parameter.

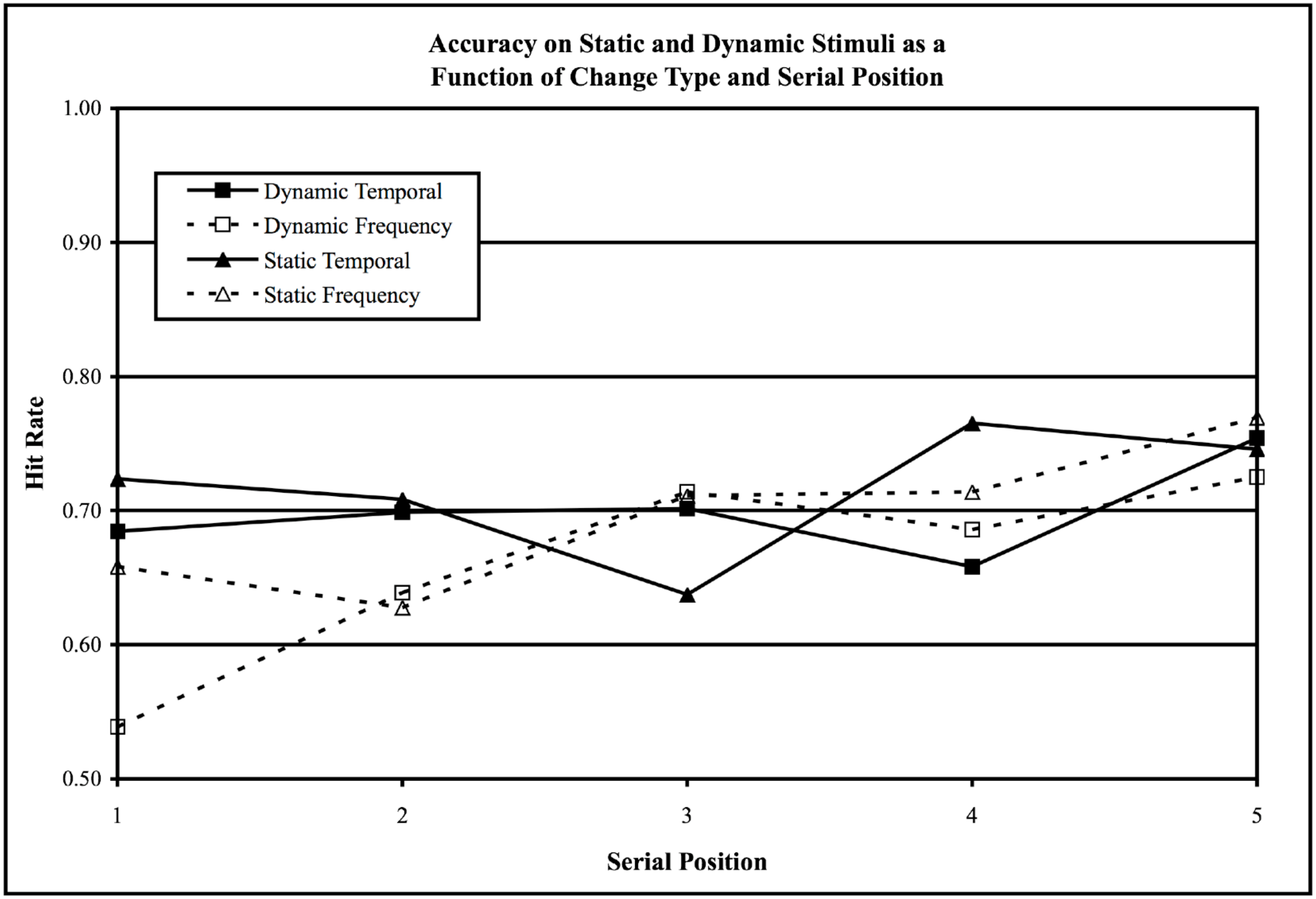

Figure 4.8. Hit rate as a function of serial position, stimulus type, and change type. Error bars are not plotted, but can be found on Figure 4.9 and 4.10.

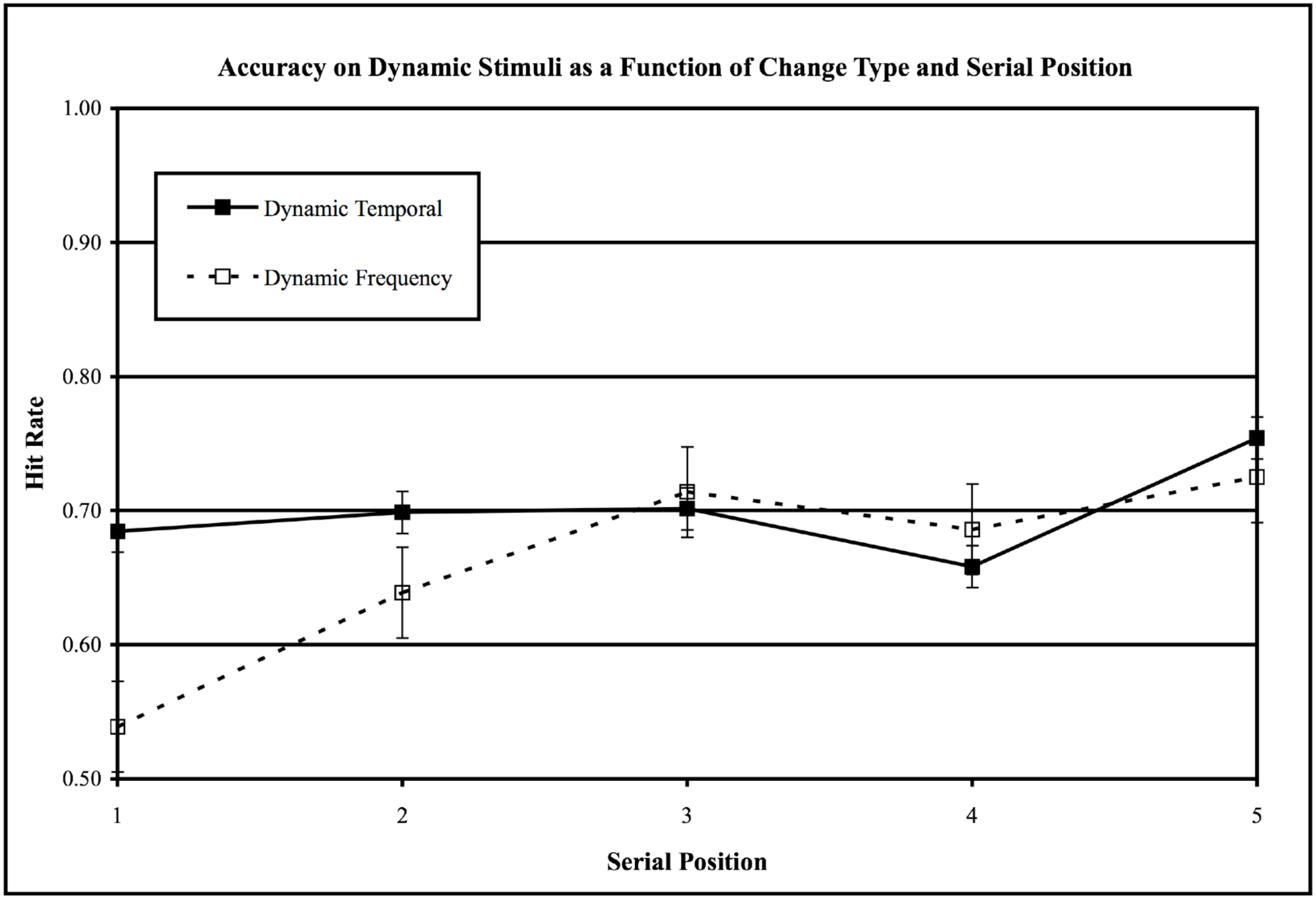

Figure 4.9. Accuracy on dynamic stimuli as a function of change type and serial position. Bars are SEM for the parameter.

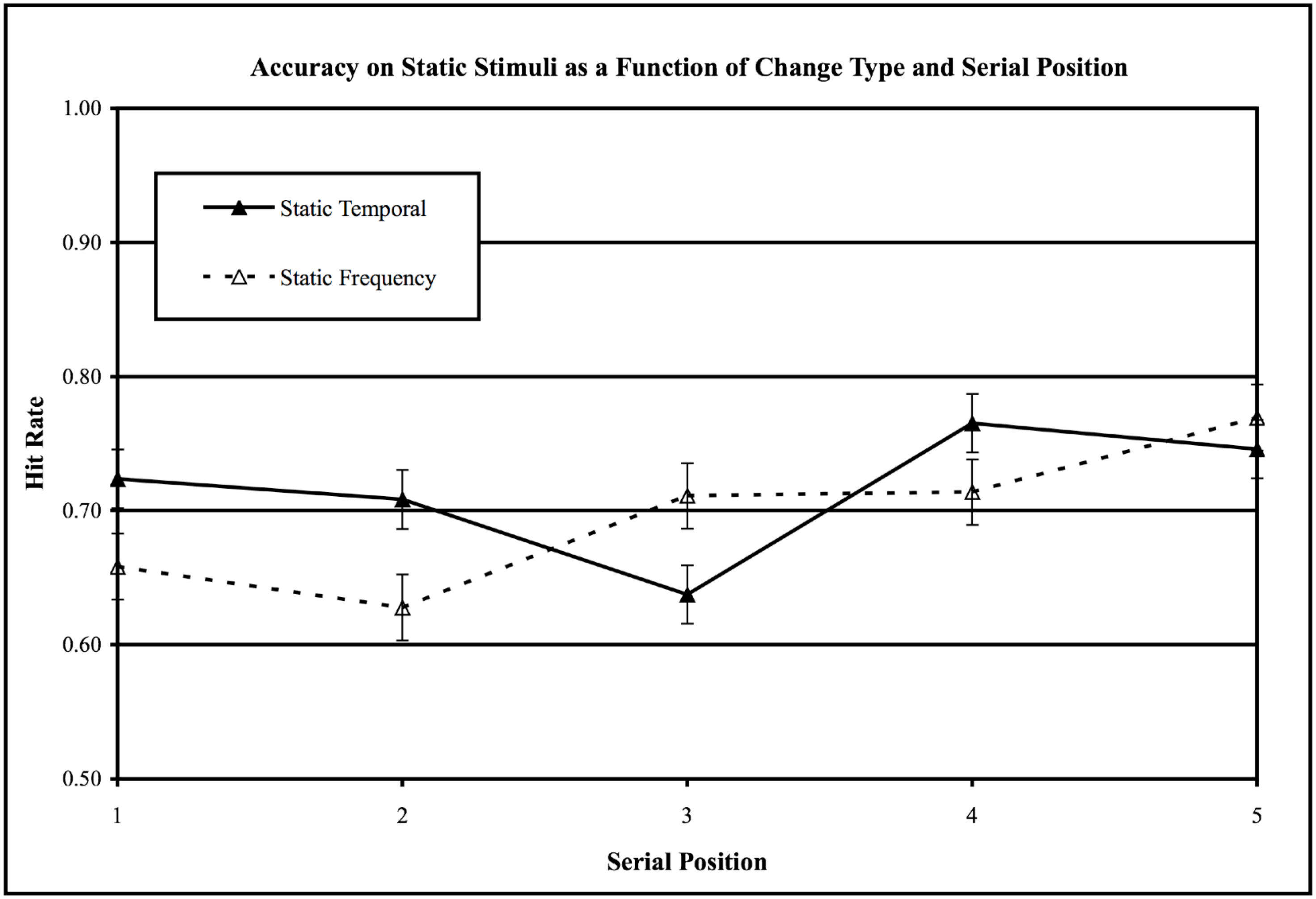

Figure 4.10. Accuracy on static stimuli as a function of change type and serial position. Bars are SEM for the parameter.

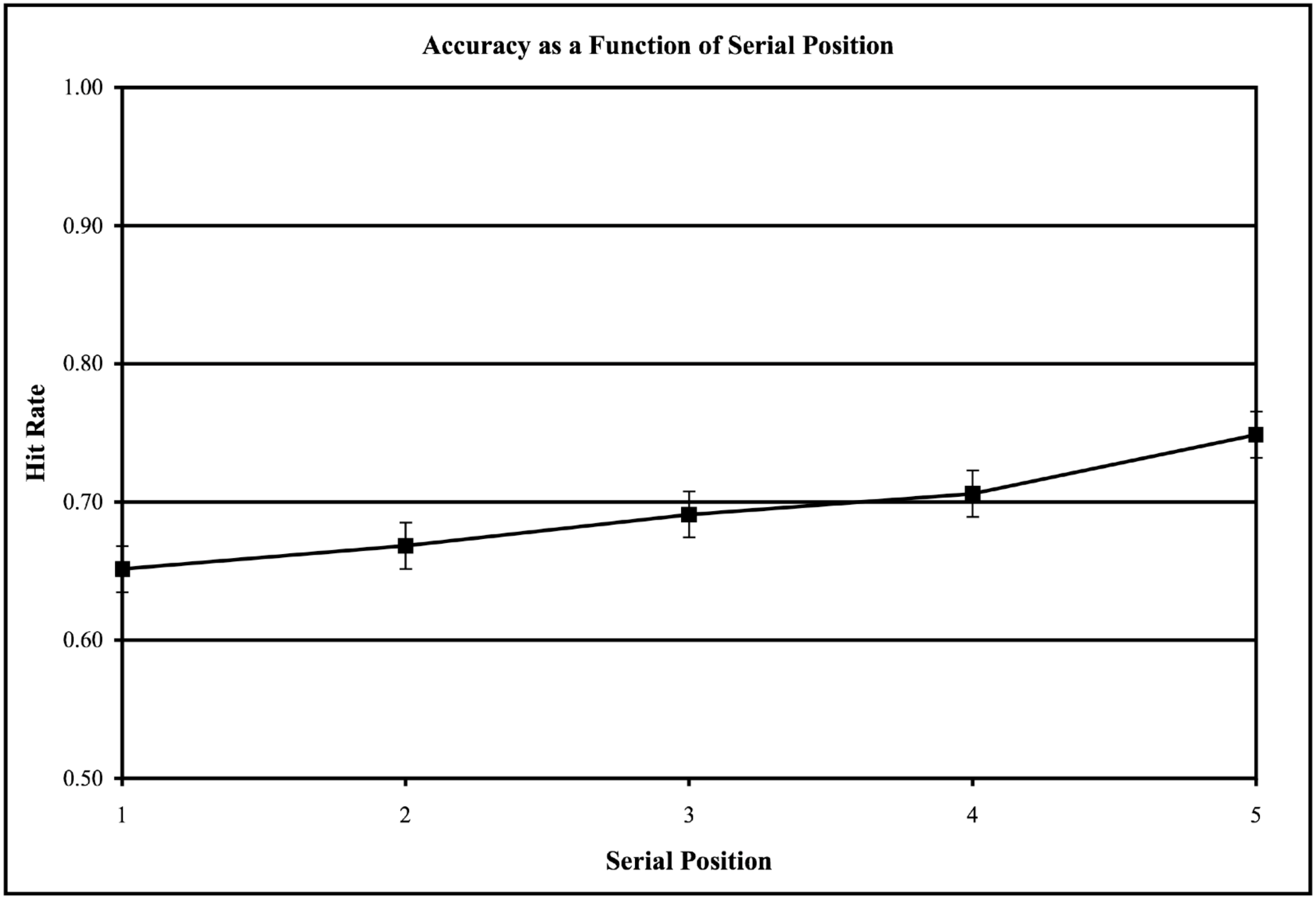

Figure 4.11. Hit rate as a function of the serial position. Bars are SEM for the parameter.

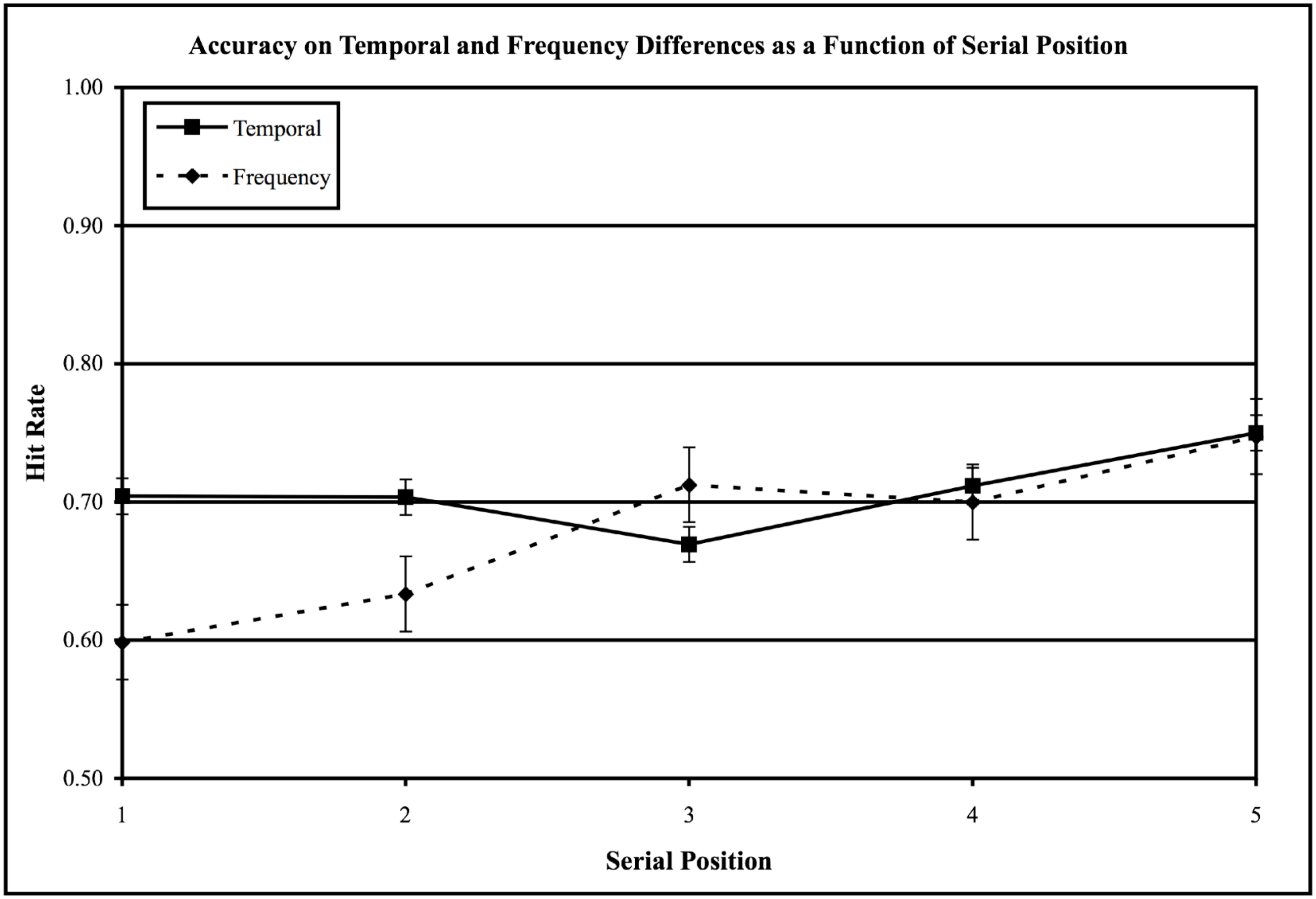

Figure 4.12. Hit rate as a function of serial position and change type. In these data, primacy is elevated for sequence trials compared with frequency trials. Bars are SEM for the parameter.

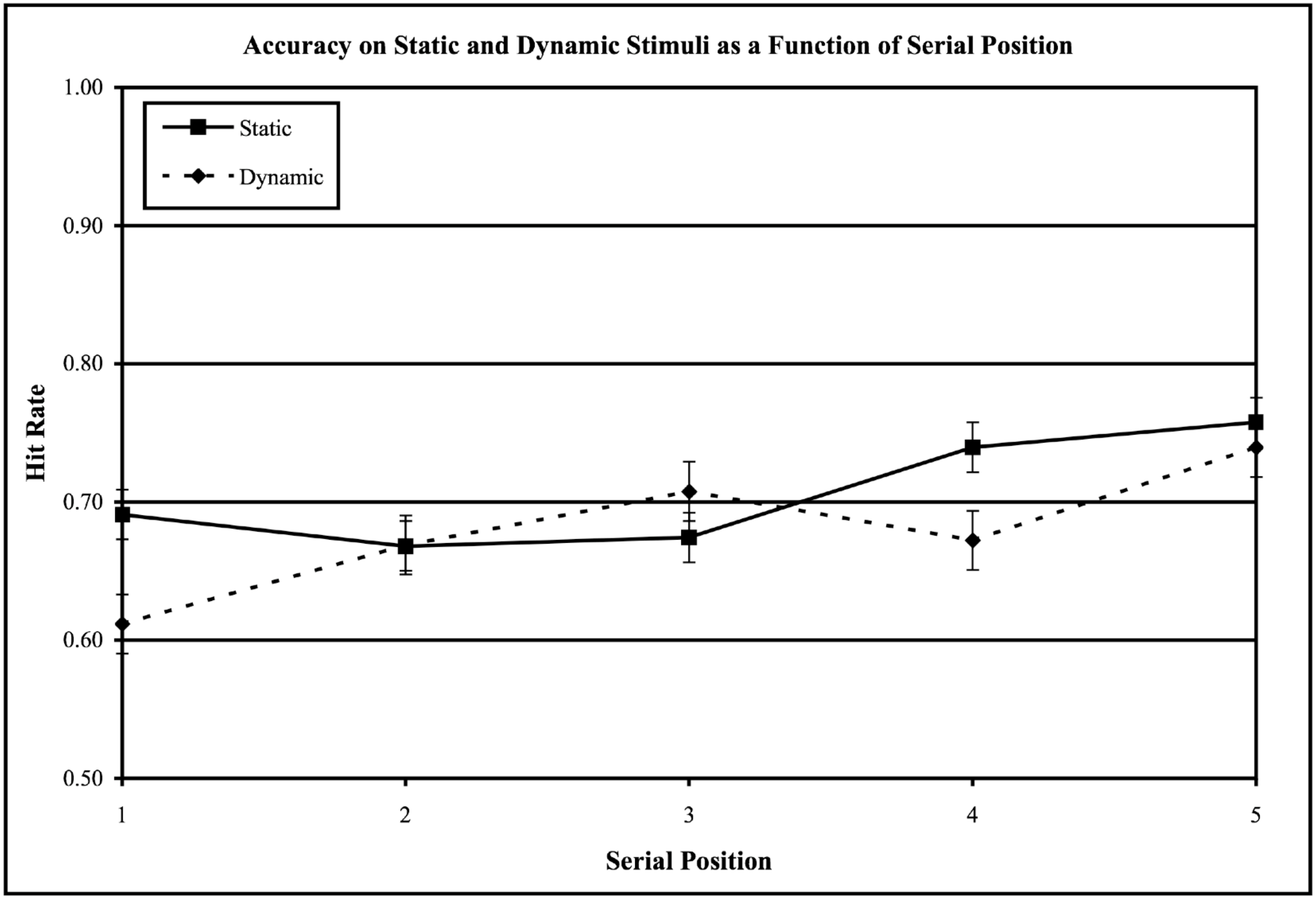

Figure 4.13. Hit rate as a function of serial position and stimulus type. Significantly better performance is found in the static condition at positions 3 and 4. Differences are not significantly different at any other position. Bars are SEM for the parameter.

Factor	Level	Mean	Standard Error	η^2
Change Type	Temporal Change	.716	.020	.09*
	Frequency Change	.687	.026	
Stimulus Type	Static Stimulus	.723	.023	.14**
	Dynamic Stimulus	.680	.025	

Table 4.1. Mean, standard, error, and η^2 values for two main factors of Experiment 3 (change type and stimulus type) collapsed across change magnitude.

Note: Differences between these factors were significant at the $p < .10$ (*) and $p < .05$ (**) level. These summary data do not have single trials included more than once and are thus a better measure of main effects of change type and stimulus.

Factor	Level	Mean	Standard Error	η^2
Change Type	Temporal Change	.716	.021	.14**
	Frequency Change	.678	.025	
Stimulus Type	Static Stimulus	.718	.022	.17**
	Dynamic Stimulus	.677	.024	

Table 4.2. Mean, standard error, and η^2 values for two factors of Experiment 3 collapsed across serial position.

Note: Differences between these factors were significant at the $p < .05$ (**) level. These data are collapsing across serial position and performance on a single trial will be included more than once in the summary data.

CHAPTER 5

OVERALL DISCUSSION

Summary of Results

During the course of these three experiments, much has been learned about listener's ability to detect changes between static and dynamic stimuli as well as individual differences in this ability. The primary, overarching, finding was that change-detection accuracy for static and dynamic auditory stimuli is not equivalent. By consistently finding a performance difference in change-detection across a range of stimulus and experimental parameters, a robust difference in memory as a function of stimulus type has been found. Further, the more nuanced performance profiles of each study as a function of other manipulated variables (the ISI, duration, and relevant features of a change-detection task) provide information as to how we might characterize the memory system responsible for performance in this task, our so-called ASTM. Lastly, the individual differences findings suggest that limitations in performance imposed by dynamic stimuli are related to its interactions with individual ASTM systems, not due to a functional incompatibility between ASTM and dynamic stimuli.

In Experiment 1, with stimulus duration held constant at 500 ms for all stimuli, the ISI between the to-be-remembered tone and the probe tone was varied from between 100 and 1800 ms and it was found that discrimination between dynamic stimuli is less accurate than discrimination between static stimuli, no matter how short the delay interval. Concurrently it was found that change detection for dynamic but not static stimuli decreases as a function of increasing ISI, at least in the range of 100-1800 ms. These findings reflect differences in *ASTM-related processes (encoding, maintenance, and/or comparison)* for dynamic vs. static stimuli, *not* greater perceived similarity of the dynamic over the static stimuli. Individual differences in the ability to detect changes as a function of stimulus type and ISI suggests that some listeners (the high-sensitivity group) have good change-detection ability (and presumably a high-functioning ASTM) for both static and dynamic stimuli that stays intact over increasing ISIs. However, even in this high-sensitivity group, the ability to detect changes for static stimuli remains a *relative* strength compared to the ability to detect changes for dynamic stimuli, though this effect *is* less consistent than for the low-sensitivity group and the overall results.

In Experiment 2, ISI was held constant at 900 ms, but the duration of the stimuli was manipulated, which effectively created a number of different types of dynamic stimuli. The data from change detection on a variety of durations suggests that the change-detection advantage for static stimuli versus dynamic stimuli overall is maintained even when multiple types of dynamic stimulus are present. The effect of duration in-and-of itself was also examined and it was found that change-detection accuracy increased with longer stimulus duration for the static and dynamic-magnitude stimuli. However, for dynamic-slope stimuli, the pattern of increase was less consistent, with change detection accuracy increasing with longer durations for the dynamic-100-slope stimuli but not the dynamic-300-slope stimuli. Additionally, the difference in performance between various stimulus types decreased at shorter durations and was eliminated at the 50 ms duration.

The individual differences data for this experiment showed that increases in duration did not lead to equal performance in low-sensitivity listeners as compared with high-sensitivity listeners, though the low-sensitivity group did show substantial gains in performance for static stimuli that put change detection for these stimuli within the range of performance shown by high-sensitivity listeners for dynamic (but not static) stimuli. The differences between change detection for static and dynamic generally held for participants in the high-sensitivity group, but were again not as consistent as for the low-sensitivity group or the overall results.

In Experiment 3, the auditory stimuli of Experiments 1 and 2 were presented in 5-tone sequences and listeners now had to detect frequency changes (in one within-subjects condition) and temporal order changes (in another within-subjects condition). Here static stimuli were again easier to discriminate than dynamic stimuli. There was also evidence that only four of the five dynamic stimuli were remembered and that there was no memory for the first stimulus of a dynamic sequence when asked to discriminate sequences based on frequency. In terms of the size of the change and the ability to detect changes for sequences, dynamic stimuli showed no benefit for increased magnitudes in the amount of change involved in a “change” trial. Further, with the smallest change in magnitude, there was no difference in the discrimination of static and dynamic stimuli. In terms of change detection as a function of serial position, there were slight nuances in the curves of each stimulus type, but the general finding of primacy and recency for the detection of temporal changes and recency for the detection of frequency changes was found for both stimulus types. Lastly, there was a marginal advantage for the detection of temporal changes over frequency changes this difference was the same for both static and dynamic stimuli.

Overall these findings unequivocally answer the main question of the present thesis by demonstrating consistently across a variety of conditions that ***change detection for static stimuli is more accurate than for dynamic stimuli.*** These data thus suggest that a fundamental shortcoming of studies that have previously attempted to relate performance on a psychoacoustic measure to performance on a language processing measure may be the failure to include dynamic stimuli in their psychoacoustic measures. Further, the key correspondence in studies that have successfully found a relationship between psychoacoustic and language processing measures may be partially attributed to the inclusion of dynamic stimuli in the psychoacoustic measure. By additionally answering many subsequent questions about change detection across a variety of conditions, these findings also contribute to our understanding of how ASTM, a fundamental memory mechanism involved in auditory change detection (just as VSTM is a fundamental memory mechanism involved in visual change detection), functions. We turn now to an elaboration of this function.

Information Processing in ASTM

A schematic of the chain of processing I hypothesize to be involved in a change-detection task is presented in Figure 5.1. Here the perceptual system separately extracts temporal and frequency information (Plack, 2005, pp. 10-14; Møller, 1983; Rose, Brugge, Anderson, & Hind, 1967) from the first stimulus and this information is encoded into a stable memory representation in ASTM. Encoding begins at the earliest stages of processing, the level of sensory memory, where the frequency information (i.e., the particular frequency that is present in the environment) of a sound is represented within the auditory nerve, the entry point of auditory information into the human cognitive system. Without further processing this information will be quickly overwritten (i.e., the memory will be lost due to interference from subsequent auditory information) as the auditory nerve represents new frequency information present in the environment. The task of representing this information must be quickly passed off to a new set of neural processes that respond to more specific patterns of firing in the auditory nerve (e.g., firing that results from sound within a single band or set of bands or firing that results from the particular timing of frequency presentations). This more specific firing can be thought of as an acoustic feature (or in speech terms, a cue). An acoustic feature (Abbs & Sussmann, 1971) is a combination of frequency and temporal information and it is a more specific form of information simpler features, such as a single frequency component. Frequency sweeps (i.e., a rise in frequency over time or a particular set of frequencies presented in a particular order within a particular time frame) also constitute an acoustic feature. Work on animals have found cells in high-level neural areas such as auditory cortex that are selective for particular frequency sweeps (Versnel & Shamma, 1997), suggesting that the encoding process abstracts away from simple frequency and temporal information as processing proceeds. Computationally, acoustic features would be less susceptible to interference because similar information is less likely to occur in the environment (i.e., they are conjunctions of simple features, thus are statistically less likely than the simple features that the auditory nerve responds to).

Acoustic features can then be cues to more abstract categories found in the environment that are defined by sets of acoustic features, or acoustic features occurring in specific sequences, much like acoustic features themselves are abstractions of simple frequency and temporal information. An example of a more abstract category is a phonological category such as that is defined by VOT. Here the presence of low-frequency voicing cues in combination with particular high-frequency linguistic cues indicate the presence of voicing and distinguish, for example, a /b/ from a /p/. I hypothesize that this chain of processing (from simple frequency to acoustic/phonological category) can be thought of as the encoding process of ASTM (though the process takes place via sensory memory) and is affected by the manipulation of duration in Experiment 2 and the extension of change detection to sequences in Experiment 3. In this way encoding does not exist at a single level of processing, but is a culmination of a multiple levels of processing that occur within sensory memory and determines the form of the memory representation in ASTM.

Once the stimulus is encoded into a memory representation, this information must be maintained across any memory delay. This operation is a critical component of any STM or working memory system. In Figure 5.1. this operation is depicted as a circular arrow that may seem to indicate a ‘refreshing’ or ‘active rehearsal’ over time, but this is not a theoretical claim of the present work. This arrow is simply a convenient way for indicating that the stimulus representation is made available at a later time, when the comparison stimulus is presented to ASTM. I hypothesize that this maintenance operation is also affected by the manipulation of ISI in Experiment 1, the manipulation of duration in Experiment 2, and the change detection of sequences in Experiment 3. Each of these different task configurations presents a different challenge to ASTM.

In a change-detection task, a second stimulus is presented and encoded into ASTM. In order to generate a response, some ability to compare the second stimulus to the first must be possible. This ability is indicated by a second circular arrow. A response (same/different) can then be generated based on a (mis)match between the two stimuli. Like the depiction of the maintenance process, the form of this arrow should not be taken as a hypothesis as to how the comparison process is implemented in ASTM. The output of the comparison process will form the basis of a response by some other system and thus can be though of as the ‘output’ of ASTM. Since this comparison process occurs after a delay and involves representations of both stimuli (i.e., the results of both encoding processes) the manipulations of all three experiments (ISI, duration, and sequencing) may also affect the comparison of stimulus 2 to stimulus 1. However, because the manipulation of magnitude, serial position, and type of change employed by Experiment 3 all impact either the feature change to-be-detected or the size of the feature change, these manipulations will most directly affect the comparison process.

The results of Experiments 1 – 3 will now be interpreted with respect to the overall processing scheme just outlined. Even though effects of duration may affect all three processes related to ASTM (including encoding via sensory memory), this is mostly the case for dynamic stimuli given the changes in the form of the stimulus that occur as a function of duration and the possibility that these changes in form (and not simply the increase in duration) result in a different representation of the stimulus, thus affecting both maintenance and comparison processes. However, given that the static stimuli do not change in form as a function of duration and performance on almost all stimulus types is the same at the 50 ms level, it could be argued that the overall increase in change detection for static stimuli across durations may be seen as a baseline increase purely as a function of increased time of encoding. Deviations from the gain achieved by an increase in duration by other stimuli may indicate that the increase in performance normally achieved through change in duration is attenuated by changes in the form of the stimulus that affect either maintenance or comparison processes.

Consistent with this interpretation, increasing the amount of time available for encoding all stimuli (through duration) generally produces better change-detection accuracy. However, whether or not there was an increase in accuracy with an increase in duration did vary for the different duration conditions, as did the precise amount of the increase. These data showed a pattern of accuracy that generally favored dynamic stimuli with smaller change magnitudes (i.e., are more like static stimuli) those with greater change magnitudes. Again, these data depended on the exact duration being examined, especially in the case of 50 ms where no difference between conditions was found.

Overall, the effect of a change in duration on ASTM via changes in encoding is constrained by the changes in the stimulus form that occur as a result (e.g., the subtraction or addition of frequency information to the stimulus). Additionally, changes in the stimulus form may impact the maintenance and comparison of dynamic stimuli. In the present set of experiments, it was found that factors that are beneficial for detecting changes in one type of stimulus (for example, extending static stimuli from 50 to 1000 ms) are not always beneficial for other types of stimuli (for example extending dynamic-300-slope from 50 to 1000 ms).

These factors in performance as a function of duration had the greatest effect in the low-sensitivity group (who had the most to gain through increases in duration), though the high-sensitivity group did show some of the same effects. Further, increases in duration only were not enough to achieve equivalent change-detection accuracy between the low- and high-sensitivity groups in the case of static stimuli, suggesting that encoding is not the only affected process in this group. However, there does seem to be a substantial difficulty in encoding given that the static stimuli did show the greatest improvement and performance did move into the range of high-sensitivity listeners on other types of stimuli.

The changes in the form of the stimulus (as a function of stimulus type and changes in duration) may impact performance by affecting the representation of the stimulus, which may in turn impact the maintenance and/or the comparison process in ASTM. In Experiment 1, maintenance of dynamic stimuli was found to be more difficult, but only at the longest delay whereas discrimination of dynamic stimuli was found to be more difficult at all delays. Thus, the impact of delay on ASTM via changes in maintenance are also constrained by the type of stimulus being processed. Further, at delays where maintenance is not contributing to differences between ASTM for static and dynamic stimuli (the 100 ms ISI), differences in performance between static and dynamic stimuli are found that can not be attributed to perceived similarity. Instead, these differences must be attributed to either encoding or comparison processes.

The independent contributions of these processes to ASTM function can not be assessed from the present set of data, however differences in their function do appear to contribute to performance in both high- and low-sensitivity listeners. High-sensitivity listeners in Experiment 1 show no impact of increases in the demands on the maintenance function, but do show an impact of the stimulus manipulation, thus their encoding and/or comparison processes are impacted by the type of stimulus.

These results are consistent with the individual differences results of Experiment 2. In this experiment high-sensitivity listeners were selected on the basis of performance at the 50 ms level where the encoding, maintenance, and comparison processes appear to be contributing to performance equally for almost all stimuli. However, these high-sensitivity listeners were still impacted by changes in stimuli as a function of duration, suggesting that while their ASTM is more adept for change detection, it is not impervious to changes in the form of the stimulus and their affect on ASTM processes. Similarly, in the case of Experiment 1, the high-sensitivity individuals were selected on the basis of the function of their maintenance process, but were still impacted by changes in stimulus type that impacted encoding and/or comparison processes.

Taking both the overall and individual differences data into account, there is no one group that shows unaffected performance for all processes. Both high-sensitivity groups show systematic changes in performance that seem to relate to non-optimal processing in one or another ASTM process, even if they were selected based on near-optimal processing involving particular ASTM processes. In terms of the overall data, there is only one condition where performance is equal for almost all stimuli. Otherwise, performance is conditionalized on the exact stimulus and the exact discrimination context (i.e., the feature to be discriminated and the task configuration), suggesting that the impact of any change in stimulus or task variable for any group of people (let alone individuals) is currently unknowable. The data presented here fill just a small part of a contingency table that describes the potential impact of the variables of consideration on ASTM and these variables have all been examined in isolation. For any higher-level ASTM task (e.g., phoneme perception) these variables may be interacting and contributing to performance in non-obvious ways.

Experiment 3 reveals information about the function of the comparison process, but in a more complex discrimination context that complicates a direct comparison to Experiment 1 and 2. Here it was found that static stimuli, but not dynamic stimuli, benefited from larger change magnitudes between stimulus 1 and stimulus 2. Further, with the smallest change magnitude, performance on change detection for static and dynamic stimuli were equal. Just as in Experiment 2, it seems that static stimuli enjoy a greater benefit from a manipulation that purportedly makes the function of a particular process of change detection easier. In Experiment 2 it was the duration aiding encoding and in Experiment 3 it is change magnitude aiding the comparison process. The impact of ISI on maintenance (in terms of aiding this component) is the same for both static and dynamic stimuli. The impact of primacy and recency, as well as change type (in which case there is a marginal impact), on the comparison process is also the same for both static and dynamic stimuli. This experiment also demonstrates that the processing of sequences of stimuli by ASTM is more difficult for sequences of dynamic than static stimuli, to the point that not all items may contribute to performance in certain discrimination contexts (i.e., ASTM may be 'agnostic' as to changes at particular locations when frequency changes are being detected). In this situation, the first item of the sequence is the item that loses its influence over performance, though whether the failure is due to an impact on encoding, maintenance, or comparison is not known.

Speech Perception and Psychoacoustics

The motivation for the present work was a number of findings of weak correlations between measures of psychoacoustics and speech perception. In both cases, the tasks can be conceptualized as change detection. In the case of speech a single sound is presented to listeners followed by a pause and a second stimulus that is either the "same" or "different". Stimuli generally vary along a particular dimension that is a reliable cue for a particular cue category (e.g., the onset time of the low-frequency energy associated with voicing). Similarly, 'basic' auditory capacities (Watson, Johnson, Lehman, Kelly, & Jensen, 1982) such as frequency resolution are measured through change detection, except that simple sine-waves are used as stimuli and the measure is the minimum amount of change along a particular dimension needed to discriminate two simple stimuli. Despite the similarities in the general measures of the tasks and the acoustic basis for both behaviors, correlations between these two abilities have been found to be relatively low (Christopherson & Humes, 1992; Espinoza-Varas & Watson, 1988; Era et al., 1986; Johnson et al., 1987; Karlin, 1942; Van Rooij & Plomp, 1990, 1992) while correlations between these tasks and tasks that measure higher-level functioning (such as IQ, verbal SAT scores, or GPA) have been higher (Hirsch & Watson, 1996; Surprenant & Watson, 2001).

The analysis of this problem presented in Chapter 1 focused on the assumptions of these researchers that static and dynamic stimuli are processed in the same manner. This interpretation was bolstered by the findings of Fu (2002) who demonstrated a correlation between speech identification and change detection for dynamic stimuli that explained a large portion of the variance for the perception of vowels (71.9%) as well as consonants (97.1%) in cochlear implant users. It was further bolstered by findings demonstrating similar performance profiles on change detection and/or phoneme identification for stimuli that mimicked key features of speech (Cutting & Rosner, 1974; Jusczyk, et al., 1989; Triesman, et al., 1995).

The results that are presented here do not directly address the relationship between psychoacoustics and speech perception as they utilized a range of time scales that have varying applicability to a number of different aspects of speech and language. However, these results do directly address the role of memory in the tasks traditionally utilized to examine the relationship between psychoacoustics and speech. In many different contexts, it was found that changes to static and dynamic stimuli were detected with different levels of accuracy and the processing of these two types of stimuli were differentially affected by changes in the configuration of the change-detection task. Thus, static and dynamic stimuli are not processed in the same manner and these differences impact memory for the stimulus, as well as the ability to respond to changes in the stimulus.

While these data support the notion that static and dynamic stimuli should not be treated equally in psychoacoustic measures, less was discovered regarding why this is the case. For example, in the Fu (2002) study, the measure of performance on the dynamic stimuli was the threshold for modulation detection, expressed as the depth of modulation. This is a measure of one's ability to discriminate rates of change. In the present set of experiments the main measure was of one's ability to discriminate the frequency of a dynamic stimulus. Further, the present set of experiments demonstrated that the ability to discriminate stimuli is dependent on the discrimination context, which includes the task configuration; a finding that has also been demonstrated in speech perception (Carney, et al. 1977).

The Carney et al. (1977) work, conditionalized a popular finding that the acoustic features of a native-language speech stimulus cannot be finely discriminated on a change-detection task. Instead, it was thought that listeners could only perform a categorical discrimination of speech stimuli. Categorical discrimination is defined by responding (same/different) in a non-linear way to linear variations in the acoustic stimulus. For example, though a voicing cue (voice-onset time [VOT]) can be varied linearly (e.g., increasing the VOT by 10 ms to create a continuum of speech stimuli) the responses of listeners discriminating these stimuli generally do not show sensitivity to these 10 ms differences unless the stimuli correspond to different *categories*. Voiced categories are defined by a VOT of between 0 and 25 ms and voiceless categories are defined by a VOT of greater than 25 ms. Unless the difference between the two stimuli cross this categorical boundary (e.g., a stimulus changes from a 20 ms VOT to a 30 ms VOT) the stimuli are heard as the "same".

What Carney et al. (1977) showed was that this categorical responding was partially a function of the task environment and could be manipulated by changing the discrimination context. First, listeners were trained with feedback to discriminate fine-grained VOT differences that did not cross category boundaries (e.g., 10 ms VOT vs. 20 ms VOT) in a change-detection task. Later, they were given a traditional change-detection task in which categorical perception had previously been found and these listeners they failed to respond categorically even with no-feedback. Essentially, listeners had learned the relevant feature dimension and learned to respond to fine-grained distinctions along this dimension. Second, it was shown that stimuli could be arbitrarily assigned to different category boundaries (i.e., the features that defined categories, or the discrimination context, could be changed) and artificial VOT categories could be created in listeners. The critical aspects of the training were (1) that listeners were given a fixed standard (i.e., the first stimulus was the same every time) and (2) that listeners were given feedback as to the correctness of the response. In terms of the ASTM functions discussed earlier, these aspects of the task ensured that the initial stimulus was *encoded* and *maintained* (through repeated presentation of the same stimulus) and ensured that feedback as to the outcome of the *comparison* process was given, which helped listeners learn to respond to the results of the comparison.

Given the context-dependent nature of speech as well as the abundance of auditory cues for speech recognition (Repp, 1982), the more important ability for speech and language learning may be the ability to adapt to a perceptual environment quickly in order to make linguistically-relevant distinctions between stimuli. That is, the important ability for speech perception and learning may be less about the stimulus itself and more about responding appropriately to the stimulus. Thus, the shared variance of a psychoacoustic and speech perception task is a function of the correspondence between the stimuli (i.e., how well the acoustic features of the artificial stimulus map onto the acoustic features of a speech stimulus and thus map onto corresponding acoustic categories), the correspondence between the contexts of the discrimination (i.e., what feature distinction is the listener asked to make), and one's ability to respond to the stimulus given the context of the task (i.e., the listeners ability to rapidly learn to respond to this distinction). All three of these are highly dependent on the nature of the stimuli, but are not wholly determined by the stimuli. Further, the more closely all three of these align with the problems faced during speech perception and learning, the better the relationship between one's ability in the artificial task and one's ability in the speech task.

In Fu (2002) this possibility is demonstrated by the status of the listeners, namely that they all were cochlear implant users. In these listeners, temporal cues generally become the main cue to speech perception as the frequency-related details of speech are processed more coarsely (and are under-utilized in the absence of noise even when processed at finer levels; Fu, Shannon, & Wang, 1998). Following implant, these listeners must relearn the mapping between the new signals they are receiving and the previously-learned responses to this information (i.e., relearn certain aspects of speech recognition). The relationship between the psychoacoustic ability and the speech recognition ability that is being measured is a function of the stimulus (i.e., the presence of temporal modulations), the discrimination context (i.e., the need to discriminate based on temporal modulations), and the listeners' ability to respond in that context (i.e., how well they have learned to respond to temporal modulations). Given that the listeners have not been trained in the lab to discriminate based on the temporal modulation of the stimuli, their ability to do so will be dependent on a general acoustic sensitivity as well as their experiences outside the lab relearning the importance of temporal modulations in everyday speech processing. Thus, the artificial situation corresponds to the speech situation not only in stimulus (the presence of temporal modulations), but also in the discrimination context (the need to discriminate temporal modulations) and the patients' need to respond to the feature being tested (ability to learn to respond to temporal modulations).

Empirical measures may reflect the correspondence between both the stimulus properties and the discrimination context. The separate contributions of these are difficult to disentangle and the performance of listeners is based on a general ability to learn to respond in these tasks, which will reflect their experience with these stimuli and situations outside of the laboratory that provide experience with this type of responding. Thus, as demonstrated by the present set of experiments, the precise ability to discriminate two stimuli will be based on the stimuli themselves as well as the listeners ability to respond to the stimuli. This ability is reflected in measures of ASTM function and can be understood by understanding how components of the stimuli and the task affect the *encoding, maintenance, and comparison* of auditory stimuli. Once I begin to understand these processes in the context of artificial tasks, I may then begin to understand how these processes function in the context of language acquisition.

Future Directions

In order to disentangle these separate contributions a number of different future lines of research are needed. (1) Continued investigation into the discrimination of static and dynamic stimuli utilizing the present paradigm while carefully controlling the participants, stimuli, and discrimination contexts that are used in these endeavors. As was mentioned earlier, the present investigation, while extensive in its parametric manipulation of duration and ISI, is far from complete in this regard. Understanding the situations in which dynamic and static stimuli are discriminated equally will lead to better theories on how these stimuli might be represented in ASTM, which may then lead to predictions as to when and how the stimuli might correspond to speech, as well as how sequential presentations correspond to paired presentations. One follow-up that has been completed is the change-detection ability for sequences of static and dynamic stimuli that are 50 ms in duration. This was done to begin to understand how changing the stimulus characteristics in the sequence change-detection paradigm impacts ASTM and whether there is a correspondence between the sequence change-detection paradigm and the single tone paradigm.

Predictions as to how stimuli and tasks correspond may also be generated if (2) precise models of auditory perception are linked to process-based models of ASTM. Essentially, the goal of this modeling work will be to utilize the vast amount of precise research that has been done on the response of the auditory perceptual system and model the link between this level of processing and ASTM. This will be done in order to model how stimuli are represented in ASTM and how manipulations of stimuli and discrimination contexts changes the performance of ASTM (and thus the ability to respond based on the output of ASTM). The strength of psychoacoustics is the control over the stimulus and both of these arms of research utilize these strengths to gain precision on how static and dynamic stimuli are represented in ASTM.

In conjunction with a precise investigation into the correspondence between the processing of simple acoustic stimuli and speech stimuli, (3) a concurrent investigation into the link between ASTM, speech perception, and language acquisition. While the theoretical link between these areas is motivating, substantial work is required to understand how vibrations at the level of the basilar membrane turn into representations of the complex series of sounds that make up words and sentences and how cognitive processes utilize these representations to acquire language. These types of investigations will include simple correlation studies between change detection and speech perception. This work could then be used to understand correlations between speech perception, PSTM (which can now be thought of as a more language-specific version of ASTM), and vocabulary that have recently been identified (Abbs et al., 2007).

It is important to work on each of these levels of language processing not only because it is theoretically comprehensive, but also because individual differences in performance are likely to arise from combinations of differences in performance interacting with one another, especially during the time that language is first developing. In fields concerned with disordered populations, such as speech pathology and psychiatry, a full understanding of the connection between these areas may lead to not only better diagnosis and understanding of the origins of the disorder, but also better treatment of the symptoms of the disorder as well as language-related behaviors that limit the functioning of these patients. These gains will be made as I understand the complex relationship between multiple levels of auditory processing and the sets of factors that can impact each one in the learning and use of language in the act of communication.

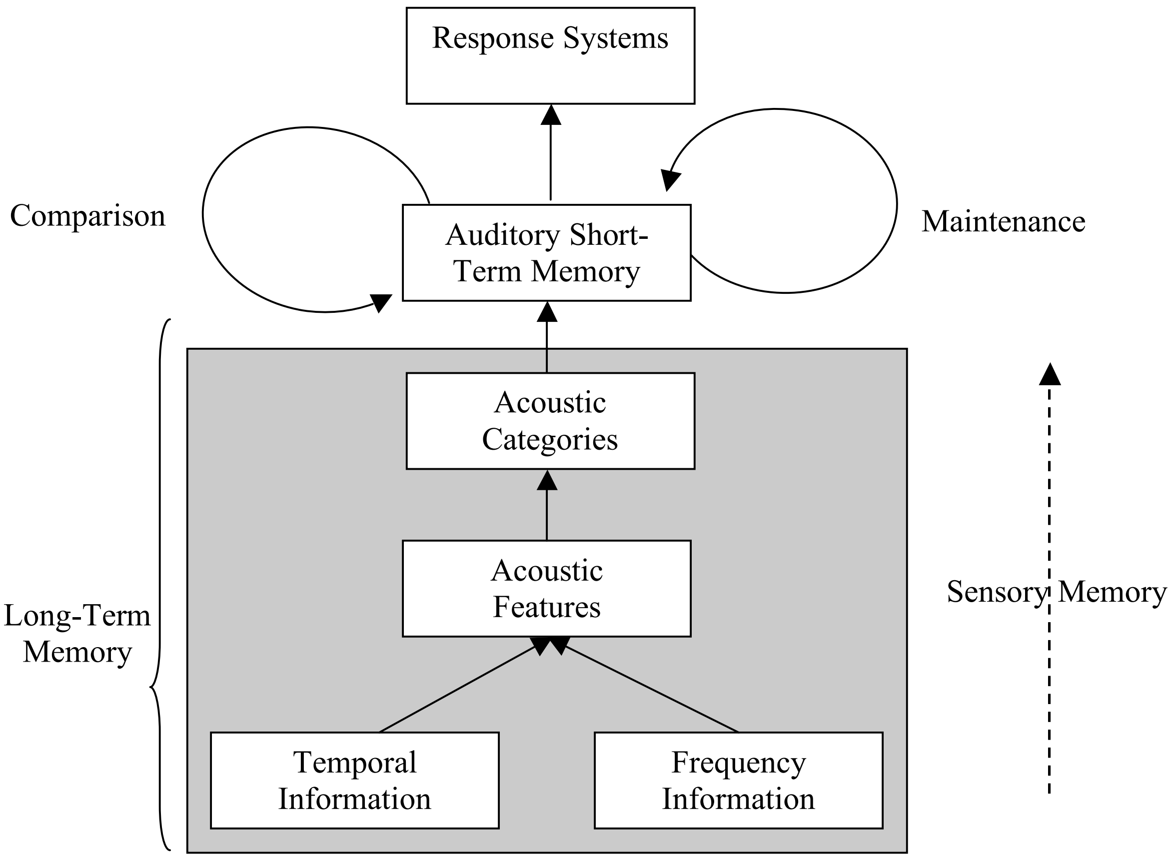

Figure 5.1. The processing of information in ASTM. When an input stimulus enters the auditory system, the specific temporal and spectral features (i.e., temporal and frequency information) are combined to create acoustic features and categories, which form the basis for the representation in ASTM. This process of abstraction away from the frequency and temporal information is the 'encoding' process. This representation can be maintained across a delay by maintenance processes. When a second input stimulus appears, it is also encoded in ASTM and these representations can be compared along a feature dimension (e.g., frequency) that is specified by the task instructions. The result of this comparison can then be used to generate a response (same/different).

APPENDIX A

MASTER SET OF EXPERIMENTAL STIMULI

Frequency (Hz)	Bark Value
503	4.76
583	5.44
663	6.09
743	6.71
823	7.31
903	7.87
983	8.40
1063	8.90
1143	9.38
1223	9.83
1303	10.25
1383	10.65
1463	11.03
1543	11.39
1623	11.72
1703	12.04
1783	12.35
1863	12.63
1943	12.91
2023	13.16
2103	13.41
2183	13.65
2263	13.87
2343	14.09
2423	14.29
2503	14.49
2583	14.68
2663	14.86
2743	15.04
2823	15.20

Table A.1. Beginning frequencies and bark values of all stimuli used in Experiment 1.

Frequency (Hz)	Bark Value
2903	15.37
2983	15.52
3063	15.67
3143	15.82
3223	15.96
3303	16.09
3383	16.23
3463	16.35
3543	16.48
3623	16.60
3703	16.72
3783	16.83
3863	16.94
3943	17.05

Table A.1. Continued

Mid-Freq. (Hz)	Beg. Freq.	End Freq.	△f	Slope (Hz/ms)
503	488.68	517.74	29.06	0.06
583	566.40	600.08	33.68	0.07
663	644.13	682.43	38.30	0.08
743	721.85	764.77	42.92	0.09
823	799.57	847.12	47.55	0.10
903	877.29	929.46	52.17	0.10
983	955.02	1011.80	56.78	0.11
1063	1032.74	1094.15	61.41	0.12
1143	1110.46	1176.49	66.03	0.13
1223	1188.18	1258.84	70.66	0.14
1303	1265.91	1341.18	75.27	0.15
1383	1343.63	1423.52	79.89	0.16
1463	1421.35	1505.87	84.52	0.17
1543	1499.07	1588.21	89.14	0.18
1623	1576.80	1670.56	93.76	0.19
1703	1654.52	1752.90	98.38	0.20
1783	1732.24	1835.25	103.01	0.21
1863	1809.96	1917.59	107.63	0.22
1943	1887.67	1999.93	112.26	0.22
2023	1965.41	2082.28	116.87	0.23
2103	2043.13	2164.62	121.49	0.24
2183	2120.85	2246.97	126.12	0.25
2263	2198.58	2329.31	130.73	0.26
2343	2276.30	2411.66	135.36	0.27
2423	2354.02	2494.00	139.98	0.28
2503	2431.74	2576.34	144.60	0.29
2583	2509.47	2658.69	149.22	0.30
2663	2587.19	2741.03	153.84	0.31
2743	2664.91	2823.38	158.47	0.32
2823	2742.63	2905.72	163.09	0.33
2903	2820.36	2988.06	167.70	0.34
2983	2898.08	3070.41	172.33	0.34
3063	2975.80	3152.75	176.95	0.35

Table A.2. Specification of dynamic stimuli for Experiment 1.

Mid-Freq. (Hz)	Beg. Freq.	End Freq.	△f	Slope (Hz/ms)
3143	3053.52	3235.10	181.58	0.36
3223	3131.25	3317.44	186.19	0.37
3303	3208.97	3399.79	190.82	0.38
3383	3286.69	3482.13	195.44	0.39
3463	3364.42	3564.47	200.05	0.40
3543	3442.14	3646.82	204.68	0.41
3623	3519.86	3729.16	209.30	0.42
3703	3597.58	3811.51	213.93	0.43
3783	3675.31	3893.85	218.54	0.44
3863	3753.03	3976.19	223.16	0.45
3943	3830.75	4058.54	227.79	0.46

Table A.2. Continued

APPENDIX B

CHARACTERISTICS OF DYNAMIC STIMULI IN EXPERIMENT 2

Beg. Freq. (Hz.)	△f 50 ms	△f 100 ms	△f 500 ms	△f 1000 ms
488.68	1.45	2.91	14.53	29.06
566.40	1.68	3.37	16.84	33.68
644.13	1.92	3.83	19.15	38.30
721.85	2.15	4.29	21.46	42.92
799.57	2.38	4.75	23.77	47.54
877.29	2.61	5.22	26.08	52.17
955.02	2.84	5.68	28.39	56.79
1032.74	3.07	6.14	30.70	61.41
1110.46	3.30	6.60	33.02	66.03
1188.18	3.53	7.07	35.33	70.65
1265.91	3.76	7.53	37.64	75.27
1343.63	3.99	7.99	39.95	79.90
1421.35	4.23	8.45	42.26	84.52
1499.07	4.46	8.91	44.57	89.14
1576.80	4.69	9.38	46.88	93.76
1654.52	4.92	9.84	49.19	98.38
1732.24	5.15	10.30	51.50	103.00
1809.96	5.38	10.76	53.81	107.63
1887.69	5.61	11.22	56.12	112.25
1965.41	5.84	11.69	58.43	116.87
2043.13	6.07	12.15	60.75	121.49
2120.85	6.31	12.61	63.06	126.11
2198.58	6.54	13.07	65.37	130.73
2276.30	6.77	13.54	67.68	135.36
2354.02	7.00	14.00	69.99	139.98
2431.74	7.23	14.46	72.30	144.60
2509.47	7.46	14.92	74.61	149.22
2587.19	7.69	15.38	76.92	153.84

Table B.1. Frequency characteristics of dynamic stimuli (100 cent magnitude) in Experiment 2 when slope is held constant.

Beg. Freq. (Hz.)	△f 50 ms	△f 100 ms	△f 500 ms	△f 1000 ms
2664.91	7.92	15.85	79.23	158.46
2742.63	8.15	16.31	81.54	163.09
2820.36	8.39	16.77	83.85	167.71
2898.08	8.62	17.23	86.16	172.33
2975.80	8.85	17.70	88.48	176.95
3053.52	9.08	18.16	90.79	181.57
3131.25	9.31	18.62	93.10	186.19
3208.97	9.54	19.08	95.41	190.82
3286.69	9.77	19.54	97.72	195.44
3364.42	10.00	20.01	100.03	200.06
3442.14	10.23	20.47	102.34	204.68
3519.86	10.47	20.93	104.65	209.30
3597.58	10.70	21.39	106.96	213.92
3675.31	10.93	21.85	109.27	218.55
3753.03	11.16	22.32	111.58	223.17
3830.75	11.39	22.78	113.89	227.79
3908.47	11.62	23.24	116.20	232.41

Table B.1. Continued

Beg. Frequency (Hz.)	Slope 50 ms	Slope 100 ms	Slope 500 ms	Slope 1000 ms
488.68	0.58	0.29	0.06	0.03
566.40	0.67	0.34	0.07	0.03
644.13	0.77	0.38	0.08	0.04
721.85	0.86	0.43	0.09	0.04
799.57	0.95	0.48	0.10	0.05
877.29	1.04	0.52	0.10	0.05
955.02	1.14	0.57	0.11	0.06
1032.74	1.23	0.61	0.12	0.06
1110.46	1.32	0.66	0.13	0.07
1188.18	1.41	0.71	0.14	0.07
1265.91	1.51	0.75	0.15	0.08
1343.63	1.60	0.80	0.16	0.08
1421.35	1.69	0.85	0.17	0.08
1499.07	1.78	0.89	0.18	0.09
1576.80	1.88	0.94	0.19	0.09
1654.52	1.97	0.98	0.20	0.10
1732.24	2.06	1.03	0.21	0.10
1809.96	2.15	1.08	0.22	0.11
1887.69	2.24	1.12	0.22	0.11
1965.41	2.34	1.17	0.23	0.12
2043.13	2.43	1.21	0.24	0.12
2120.85	2.52	1.26	0.25	0.13
2198.58	2.61	1.31	0.26	0.13
2276.30	2.71	1.35	0.27	0.14
2354.02	2.80	1.40	0.28	0.14
2431.74	2.89	1.45	0.29	0.14
2509.47	2.98	1.49	0.30	0.15
2587.19	3.08	1.54	0.31	0.15
2664.91	3.17	1.58	0.32	0.16
2742.63	3.26	1.63	0.33	0.16
2820.36	3.35	1.68	0.34	0.17

Table B.2. Slope characteristics of dynamic stimuli (100 cent magnitude) in Experiment 2 when magnitude is held constant.

Beg. Frequency (Hz.)	Slope 50 ms	Slope 100 ms	Slope 500 ms	Slope 1000 ms
2898.08	3.45	1.72	0.34	0.17
2975.80	3.54	1.77	0.35	0.18
3053.52	3.63	1.82	0.36	0.18
3131.25	3.72	1.86	0.37	0.19
3208.97	3.82	1.91	0.38	0.19
3286.69	3.91	1.95	0.39	0.20
3364.42	4.00	2.00	0.40	0.20
3442.14	4.09	2.05	0.41	0.20
3519.86	4.19	2.09	0.42	0.21
3597.58	4.28	2.14	0.43	0.21
3675.31	4.37	2.19	0.44	0.22
3753.03	4.46	2.23	0.45	0.22
3830.75	4.56	2.28	0.46	0.23
3908.47	4.65	2.32	0.46	0.23

Table B.2. Continued

Beg. Freq. (Hz.)	△f 50 ms	△f 100 ms	△f 500 ms	△f 1000 ms
488.68	4.62	9.25	46.23	92.46
566.40	5.36	10.72	53.58	107.17
644.13	6.09	12.19	60.94	121.87
721.85	6.83	13.66	68.29	136.58
799.57	7.56	15.13	75.64	151.28
877.29	8.30	16.60	83.00	165.99
955.02	9.03	18.07	90.35	180.70
1032.74	9.77	19.54	97.70	195.40
1110.46	10.51	21.01	105.05	210.11
1188.18	11.24	22.48	112.41	224.81
1265.91	11.98	23.95	119.76	239.52
1343.63	12.71	25.42	127.11	254.22
1421.35	13.45	26.89	134.46	268.93
1499.07	14.18	28.36	141.82	283.64
1576.80	14.92	29.83	149.17	298.34
1654.52	15.65	31.30	156.52	313.05
1732.24	16.39	32.78	163.88	327.75
1809.96	17.12	34.25	171.23	342.46
1887.69	17.86	35.72	178.58	357.16
1965.41	18.59	37.19	185.93	371.87
2043.13	19.33	38.66	193.29	386.58
2120.85	20.06	40.13	200.64	401.28
2198.58	20.80	41.60	207.99	415.99
2276.30	21.53	43.07	215.35	430.69
2354.02	22.27	44.54	222.70	445.40
2431.74	23.01	46.01	230.05	460.10
2509.47	23.74	47.48	237.40	474.81
2587.19	24.48	48.95	244.76	489.51
2664.91	25.21	50.42	252.11	504.22
2742.63	25.95	51.89	259.46	518.93
2820.36	26.68	53.36	266.82	533.63
2898.08	27.42	54.83	274.17	548.34

Table B.3. Frequency characteristics of dynamic stimuli (300 cent range) in Experiment 2 when slope is held constant.

Beg. Freq. (Hz.)	△f 50 ms	△f 100 ms	△f 500 ms	△f 1000 ms
2975.80	28.15	56.30	281.52	563.04
3053.52	28.89	57.77	288.87	577.75
3131.25	29.62	59.25	296.23	592.45
3208.97	30.36	60.72	303.58	607.16
3286.69	31.09	62.19	310.93	621.87
3364.42	31.83	63.66	318.29	636.57
3442.14	32.56	65.13	325.64	651.28
3519.86	33.30	66.60	332.99	665.98
3597.58	34.03	68.07	340.34	680.69
3675.31	34.77	69.54	347.70	695.39
3753.03	35.50	71.01	355.05	710.10
3830.75	36.24	72.48	362.40	724.81
3908.47	36.98	73.95	369.76	739.51

Table B.3. Continued

	Slope	Slope	Slope	Slope
Beg. Freq. (Hz.)	50 ms	100 ms	500 ms	1000 ms
488.68	1.85	0.92	0.18	0.09
566.40	2.14	1.07	0.21	0.11
644.13	2.44	1.22	0.24	0.12
721.85	2.73	1.37	0.27	0.14
799.57	3.03	1.51	0.30	0.15
877.29	3.32	1.66	0.33	0.17
955.02	3.61	1.81	0.36	0.18
1032.74	3.91	1.95	0.39	0.20
1110.46	4.20	2.10	0.42	0.21
1188.18	4.50	2.25	0.45	0.22
1265.91	4.79	2.40	0.48	0.24
1343.63	5.08	2.54	0.51	0.25
1421.35	5.38	2.69	0.54	0.27
1499.07	5.67	2.84	0.57	0.28
1576.80	5.97	2.98	0.60	0.30
1654.52	6.26	3.13	0.63	0.31
1732.24	6.56	3.28	0.66	0.33
1809.96	6.85	3.42	0.68	0.34
1887.69	7.14	3.57	0.71	0.36
1965.41	7.44	3.72	0.74	0.37
2043.13	7.73	3.87	0.77	0.39
2120.85	8.03	4.01	0.80	0.40
2198.58	8.32	4.16	0.83	0.42
2276.30	8.61	4.31	0.86	0.43
2354.02	8.91	4.45	0.89	0.45
2431.74	9.20	4.60	0.92	0.46
2509.47	9.50	4.75	0.95	0.47
2587.19	9.79	4.90	0.98	0.49
2664.91	10.08	5.04	1.01	0.50
2742.63	10.38	5.19	1.04	0.52
2820.36	10.67	5.34	1.07	0.53
2898.08	10.97	5.48	1.10	0.55

Table B.4. Frequency characteristics of dynamic stimuli (300 cent range in Experiment 2 when slop is held constant.

Beg. Freq. (Hz.)	Slope 50 ms	Slope 100 ms	Slope 500 ms	Slope 1000 ms
2975.80	11.26	5.63	1.13	0.56
3053.52	11.55	5.78	1.16	0.58
3131.25	11.85	5.92	1.18	0.59
3208.97	12.14	6.07	1.21	0.61
3286.69	12.44	6.22	1.24	0.62
3364.42	12.73	6.37	1.27	0.64
3442.14	13.03	6.51	1.30	0.65
3519.86	13.32	6.66	1.33	0.67
3597.58	13.61	6.81	1.36	0.68
3675.31	13.91	6.95	1.39	0.70
3753.03	14.20	7.10	1.42	0.71
3830.75	14.50	7.25	1.45	0.72
3908.47	14.79	7.40	1.48	0.74

Table B.4. Continued

REFERENCES

Abbs, B., Gupta, P., Tomblin, B., & Lipinski, J. (2007). A behavioral and computational integration of phonological, short-term memory, and vocabulary acquisition processes in nonword repetition. *Proceedings of the Twenty-Ninth Annual Conference of the Cognitive Science Society,* 59-64.

Abbs, B., Gupta, P., Jochem, E., & Armogan, B. (2006, November). What's in a nonword? linguistic factors that distinguish lists from nonwords. *47th Meeting of the Psychonomic Society,* Houston, TX.

Abbs, J.H., & Sussman, H.M. (1971). Neurophysiological feature detectors and speech perception: A discussion of theoretical implications. *Journal of Speech and Hearing Research, 14*, 23-36.

Alvarez, G.A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science, 15*(2), 106.

Baddeley, A.D., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*(1), 158-173.

Baddeley, A., & Lewis, V. (1984). When does rapid presentation enhance digit span? *Bulletin of the Psychonomic Society, 22*(5), 403-405.

Carney, A.E., Widin, G.P. and Viemeister, N.F. (1977) Non categorical perception of stop consonants differing in VOT. *Journal of the Acoustical Society of America, 62,* 961-970.

Christopherson, L.A., & Humes, L.E. (1992). Some psychometric properties of the test of basic auditory capabilities (TBAC). *Journal of Speech and Hearing Research, 35*(4), 929-935.

Cutting, J. E. & Rosner, B. S. (1974). Categories and boundaries in speech and music. *Perception & Psychophysics, 16*, 564-570.

Egan, J.P., & Hake, H.W. (1950). On the masking pattern of a simple auditory stimulus. *Journal of the Acoustical Society of America, 22*(5), 622-630.

Elliott, D. N., Riach, W. D., Sheposh, J. P., & Trahiotis, C. (1966). Discrimination performance of high school sophomores on a battery of auditory tests. *Acta Oto Laryngol Supplement, 216,* 1-59.

Era, P., Jokela, J., Qvarnberg, Y., & Heikkinen, E. (1986). Pure-tone thresholds, speech understanding, and their correlates in samples of men of different ages. *Audiology, 25*(6), 338-352.

Espinoza-Varas, B., & Watson, C.S. (1988). Low commonality between tests of auditory discrimination and of speech perception. *The Journal of the Acoustical Society of America, 84*(S1), S143.

Fant, C.G.M. (1960). *Acoustic theory of speech production*. The Hague: Mouton.

Fant, C.G.M. (1970). *Acoustic theory of speech production*. The Hague: Mouton.

Fu, Q. (2002). Temporal processing and speech recognition in cochlear implant users. *NeuroReport, 13*(13), 1635-1639.

Fu, Q., Shannon, R. V., & Wang, X. (1998). Effects of noise and spectral resolution on vowel and consonant recognition: Acoustic and electric hearing. *The Journal of the Acoustical Society of America, 104*(6), 3586-3596.

Gathercole, S.E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics, 27*, 513-543.

Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the rold of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language, 28*(2), 200-213.

Gathercole, S.E., Frankish, C.R., Pickering, S.J., & Peaker, S. (1999). Phonotactic influences on short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(1), 84-95.

Gathercole, S.E., Willis, C.S., Emslie, H., & Baddeley, A.D. (1992). Phonological memory and vocabulary development during the early school years: A longitudinal study. *Developmental Psychology, 28*(5), 887-898.

Goldinger, S.D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review, 105*(2), 251-279.

Gupta, P. (2003). Examining the relationship between word learning, nonword repetition, and immediate serial recall in adults. *The Quarterly Journal of Experimental Psychology: A, 56*(7), 1213-1236.

Gupta, P., Lipinski, J., Abbs, B., & Lin, P. (2005). Serial position effects in nonword repetition. *Journal of Memory and Language, 53*(1), 141-162.

Gupta, P., MacWhinney, B., Feldman, H.M., & Sacco, K. (2003). Phonological memory and vocabulary learning in children with focal lesions. *Brain and Language, 87*(2), 241-252.

Harris, D.M., & Dallos, P. (1979). Forward masking of auditory nerve fiber responses. *Journal of Neurophysiology, 42*, 1083-1107.

Harris, J. D. (1964). A factor analytic study of three signal detection abilities. *Journal of Speech and Hearing Research, 7,* 71-78.

Henry, B.A., Turner, C. ., & Behrens, A. (2005). Spectral peak resolution and speech recognition in quiet: Normal hearing, hearing impaired, and cochlear implant listeners. *The Journal of the Acoustical Society of America, 118*(2), 1111-1121.

Hirsh, I.J., & Watson, C.S. (1996). Auditory psychophysics and perception. *Annual Review of Psychology, 47*(1), 461-484.

Johnson, D.M., Watson, C.S., & Jensen, J.K. (1987). Individual differences in auditory capabilities I. *Journal of the Acoustical Society of America, 81*(2), 427-438.

Jusczyk, P.W., Rosner, B.S., Reed, M.A., & Kennedy, L.J. (1989) Could temporal order differences underlie 2-month-olds' discrimination of English voicing contrasts? *Journal of the Acoustical Society of America, 85(4)*, 1741-1749.

Karlin, J.E. (1942). A factorial study of auditory function. *Psychometrika, 7*(4), 251-279.

Kent, R.D., & Read, C. (2002). *Acoustic analysis of speech*. San Diego, CA: Singular/Thomson Learning.

Klatt, D.H. (1979). Speech perception: A model of acoustic-phonetic analysis and lexical access. *Journal of Phonetics, 7*(3), 279-312.

Kluender, K.R., Coady, J.A., & Kiefte, M. (2003). Sensitivity to change in perception of speech. *Speech Communication, 41*(1), 59-69.

Liberman, A.M., Delattre, P.C., & Cooper, F.S. (1952). The role of selected stimulus variables in the perception of unvoiced stop consonants. *American Journal of Psychology, 65*, 497-516.

Liberman, A.M., Harris, K.S., Hoffman, H.S. & Griffith, B.C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology, 54*, 358 - 368.

Luck, S.J., & Vogel, E.K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*, 279-281.

Massaro, D.W. (1975). Backward recognition masking. *Journal of the Acoustical Society of America, 58(5)*, 1059-1065.

McLeish, J. (1950). The validation of Seashore's measures of musical talent by factorial method. *British Journal of Psychology, 3,* 129-140.

McClelland, J.L., & Elman, J.L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18*, 1-86.

Middelweerd, M. J., Festen, J.M., & Plomp, R. (1990). Difficultis with speech intelligibility in noise in spite of normal pure-tone audiogram. *Audiology, 29,* 1-7.

Moore, B.C.J. (1997). *An introduction to the psychology of hearing* (4th ed.). San Diego: Academic Press.

Møller, A. R. (1983). Frequency selectivity of phase-locking of compex sounds in the auditory nerve of the rat. *Hearing Research, 11*, 267-284.

Oden, G. C., & Massaro, D. W. (1978). Integration of featural information in speech perception. *Psychological Review, 85*, 172-191.

Payne, T. (2003). *Working memory capacity and pitch discrimination.* Unpublished Ph.D Thesis, Georgia Institute of Technology.

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception and Psychophysics, 16*, 283-290.

Plack, C. J. (2005). *The sense of hearing*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Plomp, R. (1964). The ear as frequency analyzer. *The Journal of the Acoustical Society of America, 36*(9), 1628-1636.

Relkin, E.M., & Turner, C.W. (1988). A reexamination of forward masking in the auditory nerve. *The Journal of the Acoustical Society of America, 84*(2), 584-591.

Repp, B. H. (1982). Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception. *Psychological Bulletin, 92*(1), 81-110.

Rodriguez, G. P., DiSarno, N. J., and Hardiman, C. J. (1990). Central auditory processing in normal-hearing elderly adults. *Audiology, 29,* 85-92.

Rose, J. E., Brugge, J. F., Anderson, D. J., & Hind, J. E. (1967). Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. *Journal of Neurophysiology, 30*, 769-793.

Rupp, R. R., & Phillips, D. (1969). The effect of noise background on speech discrimination function in normal-hearing individuals. *Journal of Audiology Research, 9,* 60-63.

Schouten, M.E.H., Gerrits, E. & van Hessen, A. (2003). The end of categorical perception as we know it. *Speech Communication, 41*, 71-80.

Shower, E.G., & Biddulph, R. (1931). Differential pitch sensitivity of the ear. *The Journal of the Acoustical Society of America, 3*(1A), 7.

Smith, R.L. (1977). Short-term adaptation in single auditory nerve fibers: Some poststimulatory effects. *Journal of Neurophysiology, 49*, 1098-1112.

Stankov, L., & Horn, J. L. (1980). Human abilities revealed through auditory tests. *Journal of Educational Psychology,* 72, 21-44.

Surprenant, A., & Watson, C.S. (2001). Individual differences in the processing of speech and nonspeech sounds by normal-hearing listeners. *Journal of the Acoustical Society of America, 110*, 2085-2095.

Treisman, M., Faulkner, A., Naish, P. L. N. and Rosner, B. S. (1995) Voice-onset time and tone-onset time: The role of criterion-setting mechanisms in categorical perception. *Quarterly Journal of Experimental Psychology, 48A*, 334-366.

van Rooij, J.C.G.M., & Plomp,R. (1990). Auditive and cognitive factors in speech perception by elderly listeners. II: Multivariate analyses. *The Journal of the Acoustical Society of America, 88*(6), 2611-2624.

van Rooij, J.C.G.M., & Plomp, R. (1992). Auditive and cognitive factors in speech perception by elderly listeners. III. additional data and final discussion. *The Journal of the Acoustical Society of America, 91*(2), 1028-1033.

Versnel, H., & Shamma, S.A. (1998). Spectral-ripple representation of steady-state vowels in primary auditory cortex. *The Journal of the Acoustical Society of America, 103*(5), 2502-2514.

Viemeister, N.F. (1974). Intensity discrimination of noise in the presence of band-reject noise. *The Journal of the Acoustical Society of America, 56*(5), 1594-1600.

Viemeister, N.F. (1983). Auditory intensity discrimination at high frequencies in the presence of noise. *Science, 221*(4616), 1206-1208.

Vogel, E. K., Woodman, G.F., & Luck, S.J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 92-114.

Watson, C.S., Johnson, D.M., Lehman, J.R., Kelly, W.J., & Jensen, J.K. (1982). An auditory discrimination test battery. *Journal of the Acoustical Society of America Supplement 1, 71*, S73.

Watson, C. S., Kelly, W. J., & Wroton, H. W. (1976). Factors in the discrimination of tonal patterns. II. selective attention and learning under various levels of stimulus uncertainty. *Journal of the Acoustical Society of America, 60*(5), 1176-1186.

Watson, C.S., Wroton, H.W., Kelly, W.J., & Benbassat, C.A. (1975). Factors in the discrimination of tonal patterns. I. component frequency, temporal position, and silent intervals. *Journal of the Acoustical Society of America, 57*(5), 1175-1184.

Wegel, R.L., & Lane, C.E. (1924). The auditory masking of one pure tone by another and its probable relation to the dynamics of the inner ear. *Physical Review, 23*, 266-285.

Woodman, G.F., Vecera, S.P., & Luck, S.J. (2003). Perceptual organization influences visual working memory. *Psychonomic Bulletin & Review, 10*, 80-87.

Wright, B.A., Lombardino, L.J., King, W.M., Puranik, C.S., Leonard, C.M., & Merzenich, M.M. (1997). Deficits in auditory temporal and spectral resolution in language-impaired children. *Nature, 387*, 176-178.

Zhang, W., & Luck, S.J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature, 453*(7192), 233-235.