2011

Examining the Effects of Variation in Emotional Tone of Voice on Spoken Word Recognition

Maura L. Wilson
Cleveland State University

EXAMINING THE EFFECTS OF VARIATION IN EMOTIONAL TONE OF VOICE ON SPOKEN WORD RECOGNITION

MAURA L. WILSON

Bachelor of Arts in Psychology

Mercyhurst College

May, 2002

submitted in partial fulfillment of requirements for the degree

MASTER OF ARTS IN PSYCHOLOGY

at

CLEVELAND STATE UNIVERSITY

May, 2011

This thesis has been approved

for the Department of PSYCHOLOGY

and the College of Graduate Studies by

Thesis Chairperson, Conor T. M^{c}Lennan

Department & Date

Committee Member and Methodologist, Katherine S. Judge

Department & Date

Committee Member, Ernest S. Park

Department & Date

EXAMINING THE EFFECTS OF VARIATION IN EMOTIONAL TONE OF VOICE ON SPOKEN WORD RECOGNITION

MAURA L. WILSON

ABSTRACT

Despite the importance of emotional tone of voice for optimal verbal communication, how emotional speech is processed and its effects on spoken word recognition have yet to be fully understood. The current study addressed these gaps in the literature by examining the effects of intra-talker variability in emotional tone of voice on listeners' ability to recognize spoken words. Two lexical decision experiments, varying in task difficulty, were implemented to analyze participants' percent correct (PC) and reaction times (RTs). Previous research on spoken word recognition using this paradigm has found performance costs resulting from stimuli that mismatch on specific information (e.g., the identity of the talker) contained in the speech signal. Such *specificity effects* occurred only when processing was relatively slow, not when processing was relatively fast. In the current study, when processing was fast (Experiment 1), no specificity effects of emotional tone of voice emerged. When processing was slow (Experiment 2), specificity effects of emotional tone of voice emerged, but only for target words spoken in a sad emotional tone of voice and not for target words spoken in a frightened emotional tone of voice. RTs to sad target words mismatched in emotional tone of voice from prime to target blocks were longer than those that matched, but RTs to frightened target words were the same regardless of the emotional tone of voice of the word in the prime block. Separate analyses were conducted on the top and bottom performers on a

Musical Listening Test (MLT). For those who scored in the top 25%, for sad target words only (not frightened), specificity effects of emotional tone of voice emerged. For those who scored in the bottom 75% on the MLT, no specificity effects emerged, regardless of emotional tone of voice. The results of the current study have important implications for theoretical models of spoken word recognition and emotional tone of voice.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

A speaker's emotional tone of voice plays an important role in optimal verbal communication between humans. In fact, deficits in the perception of emotional stimuli are included in the diagnostic criteria for psychological abnormalities. For example, the Diagnostic and Statistical Manual of Mental Disorders-IV-TR includes "qualitative impairment in social interaction, as manifested by…lack of social or emotional reciprocity" (American Psychiatric Association, 2000) amongst other symptoms used to diagnose autistic disorder. The affective state of a speaker directly influences his or her voice, providing listeners with information about the speaker's emotional intensity and context (Bachorowski & Owren, 1995). Paralinguistic information used to draw inferences about (or index) a speaker's personal characteristics, like emotional tone of voice, has been termed *indexical* information (Abercrombie, 1967; Ladefoged & Broadbent, 1957; Laver, 1968). Indexical properties of speech supply listeners with information that can be used to speculate about the personal characteristics of a speaker, such as sex, age, physique, social status, regional origin, attitudes, occupation, and group membership (Laver, 1968). Although emotional tone of voice clearly influences human communication, the way emotional tone of voice is perceptually processed has yet to be fully understood. Moreover, the ways in which emotional tone of voice affect online spoken word recognition, including the time course of such effects, are unclear. For this reason, the following study investigated how variation in emotional tone of voice affects the time course of spoken word recognition using lexical decision tasks varying in difficulty.

CHAPTER II

REPRESENTATION OF INDEXICAL VARIABILITY

Typically, listeners have little difficulty recognizing spoken words both quickly and accurately. However, research has shown that some types of variability in the speech signal can affect spoken word recognition. One such type, *indexical variability*, refers to fluctuations in the speech signal resulting from changes in talker voice, speaking rates, emotional tones of voice, etc. from one time to another (Abercrombie, 1967; Pisoni, 1997). Importantly, indexical variability has no effect on the linguistic (e.g., semantic) content of a spoken word. For example, the word *house* still denotes a building in which people live regardless of whether it is spoken by a male or a female, in a fast or slow speaking rate, or in a happy or sad emotional tone of voice. While indexical variability has no consequences for the linguistic aspects of a spoken word, indexical and linguistic information contained in speech are nonetheless fundamentally linked. During spoken language perception, indexical properties appear to complement the processing of linguistic speech components (Nygaard & Pisoni, 1998; Remez, Fellowes, & Rubin, 1997). Furthermore, research suggests that indexical variability has long-lasting effects on the lexical representations underlying language perception, contained in the mental lexicon. For example, in a study on recall ability of word lists, Martin, Mullennix, Pisoni, and Sommers (1989) found recall performance to vary as a function of talker variability. Recall performance suffered for word lists spoken by multiple talkers compared to word lists spoken by a single talker. Other representational studies have found an attenuation in recognition accuracy when words are repeated by a different talker compared to when words are repeated by the same talker (see, e.g., Bradlow, Nygaard, & Pisoni, 1999; Craik & Kirsner, 1974; Goldinger, Pisoni, & Logan, 1991; Palmeri, Goldinger, & Pisoni, 1990, 1993). These studies suggest that indexical information is stored in memory and has consequences for perceptual processing. Perceptual performance costs observed as a result of stimuli that mismatch in talker identity are referred to as *talker effects*.

Church and Schacter (1994) investigated *talker* effects using the long-term repetition-priming paradigm much like the one used in the current study. In this paradigm, participants first respond to a block of spoken words in a prime block. Then, after a filler period, participants are asked to respond to another block of spoken words in a target block. During the target block, some of the words from the prime block are repeated. Participant responses to the repeated words are typically more rapid or accurate than responses to new words that were not presented in the prime block, known as the *long-term repetition-priming effect (or simply, priming)*. Presumably, the repeated activation of lexical representations in memory facilitates subsequent processing. Attenuation of repetition-priming effects as a result of mismatching properties in auditory stimuli (e.g., due to a change in talkers) is referred to as evidence for specificity, or *specificity effects*. Specificity effects have been observed in a number of related studies (Bradlow, Nygaard, & Pisoni, 1999; Church & Schacter, 1994; Goldinger, 1996; Houston

& Jusczyk, 2003 [in infants]; Schacter & Church, 1992; Yonan & Sommers, 2000 [in older adults]). These studies support the notion that specific indexical information contained in the speech signal is stored in memory and, consequently, influences subsequent perceptual processing.

Specificity effects appear to follow a predictable time course. This time course has been explained previously in terms of the adaptive resonance theory (ART) framework (Grossberg, 1986; Grossberg & Myers, 2000; Grossberg & Stone, 1986). According to ART, perception begins when the speech signal activates "chunks" of features of lexical and sublexical parts of one's language. A *chunk* consists of a set of learned characteristics associated with certain words and parts of words (e.g., phonemes). After chunks are activated, they resonate with the incoming speech signal; this resonance between the speech signal and the chunks forms one's perception. Some chunks occur more frequently than others in the mental lexicon. The frequency of the chunks impacts the ease with which the resonance is formed, such that more frequent chunks establish a resonance more quickly than relatively less frequent chunks. Eventually, all chunks of features, no matter their frequency, will resonate with the input over time. However, it is possible to access perceptual processing prior to the establishment of certain resonances.

As McLennan and Luce (2005) point out, abstract, general features of the speech signal tend to be more frequent than more indexical, specific features. For that reason, more abstract features should resonate with the acoustic-phonetic input relatively early during the time course of processing, and indexical features later. Thus, spoken word recognition processes depend, at least in part, on the time course of processing of these different features.

Research by M^c^Lennan and Luce (2005) suggests that specificity effects emerge relatively late during processing. Their study implemented the long-term repetition-priming paradigm to examine RTs to targets that were primed by stimuli that either matched or mismatched on the indexical properties of speaking rate (casually and carefully produced words that were spoken relatively quickly and relatively slowly, respectively) and talker identity (male and female talkers). Speed of processing was manipulated by varying the difficulty of the lexical decision task by making the nonword stimuli relatively easy or relatively hard to discriminate from the real word stimuli.[1] As a result, RTs in the more difficult discrimination experiments were significantly longer than in the easy discrimination experiments. More importantly for the time course hypothesis, when processing was fast, primes that matched and primes that mismatched on speaking rate and talker identity were equally effective, whereas when processing was slow, matched primes were more effective than mismatched primes. That is, in the hard lexical decision experiments when processing was relatively slow, talker (and rate) effects emerged, but not in the easy lexical decision experiments, when processing was relatively fast.

M^c^Lennan and Luce's (2005) results can be accounted for by the ART framework (Grossberg & Myers, 2000; Grossberg & Stone, 1986). Again, this framework posits that more frequent chunks (e.g., abstract information) resonate with the input relatively

---

[1]The current study replicated parts of the research design described in M^c^Lennan and Luce (2005). See the Methods section for a more in-depth discussion of these manipulations.

quickly, whereas less frequent chunks (e.g., specific information) do so later. In the easy discrimination tasks, the relatively fast processing only allowed the high frequency features (i.e., general, phonemic features) to establish a resonance with the input. Consequently, low frequency features (i.e., speaking rate and talker identity) took longer to establish a resonance with the input. On the other hand, in the hard discrimination tasks, relatively slow processing allowed time for both high and low frequency features of chunks to resonate with the input. Thus, these experiments provide evidence that indexical information is represented in the mental lexicon and takes time to influence spoken word processing.

The current study is relevant to both abstractionist and episodic theories of word recognition that distinguish general and specific information contained in the speech signal (Tenpenny, 1995; Tulving, 1972; 1983). Purely abstractionist theories assert that spoken word input activates abstract lexical information only and specific episodic information does not play a role in word recognition. Purely episodic approaches, on the other hand, suggest that spoken word recognition relies on prior encounters with specific words and posits that mental representations consist of episodic information, not abstract lexical entries. As Tenpenny (1995) points out, neither purely abstractionist nor purely episodic models can fully explain the process of spoken word recognition. The purely episodic approach cannot account for long-term priming effects found when priming occurs over time and with variations (e.g., in talker or emotional tone of voice) from one episode to the next. The purely abstractionist approach is unable to account for specificity effects or a reduction in priming when there is a mismatch (e.g., in talker or emotional tone of voice) from one episode to the next. A more accurate model of lexical representations involved in spoken word recognition would combine purely abstractionist and purely episodic viewpoints and take long-term priming effects of indexical specificity into account. One such possibility, described by Tenpenny (1995) as *weakly abstractionist*, suggests that abstract representations are activated first, sometimes followed by the activation of episodic representations which can facilitate access to abstract lexical entries. From this theoretical perspective, episodic and abstract input activates fundamentally different lexical representations. Another possible hybrid model, which Tenpenny (1995) calls *weakly episodic*, asserts that the same representation is activated by both abstract and episodic information, such that a prototypical representation develops for a word by abstracting over specific episodes. From this perspective, specific and general representations of a word differ only in the way they are incorporated into word representations in the mental lexicon.

Hybrid models of spoken word recognition can be used to make more feasible hypotheses about the way indexical variability is represented and processed than purely abstractionist or purely episodic models. From the weakly abstractionist approach, when indexical information (e.g., emotional tone of voice) is repeated, as it is in priming experiments, it would be expected to enhance access to the word's abstract lexical representation and aid word recognition. Alternatively, the weakly episodic model would assert that the word recognition response is improved because repetition of words matching on indexical properties (e.g., emotional tone of voice) modifies the strength of the network's response to the input. As long as other competing input does not interfere with the modification process, repetition of episodic information contained in the stimuli should give rise to indexical specificity effects. Note that these hypotheses were only based on previous descriptions of hybrid approaches and are not direct accounts of the weakly abstractionist or weakly episodic approaches.

The possibility that general abstract and specific episodic information are represented and processed in distinct areas of the brain has been addressed from a biological approach in several neurological studies. In the vast majority of right-handed and most left-handed adults, speech and language processes primarily occur in the left hemisphere (Abou-Khalil, 1995; Caplan, Hildebrandt, & Makris, 1996; Springer et al., 1999; Vanlancker-Sidtis, 2004). Nevertheless, recent functional brain imaging research has reported right-hemisphere activation during language tasks, suggesting that both hemispheres play a role in language processes (Knecht et al., 2000; Sidtis, 2000). Hence, it is possible that the representation and processing of general (i.e., abstract) and specific (i.e., indexical) properties of speech are lateralized to distinct hemispheres.

González and McLennan (2007) explored the hemispheric differences possibility in relation to indexical specificity effects during spoken word recognition. In their long-term repetition-priming lexical decision experiment, prime words matched on talker identity (male or female) served as more effective primes than prime words mismatched on talker identity, but only when auditory stimuli were presented to the left ear (right hemisphere). When presented to the right ear only, matched and mismatched prime words were equally effective. These results suggest that indexical variability can affect the perception of spoken words differently in the right and left hemispheres.

These neurological studies are particularly relevant to the present investigation of emotional tone of voice as an indexical property that influences spoken word recognition. Recent research has endorsed right hemisphere dominance for processing emotionallydriven stimuli, especially negative emotional connotations (Alpers, 2008; Sato & Aoki, 2006; Smith & Bulman-Fleming, 2005). These hemispheric differences suggest that specific emotional tone of voice information contained in the speech signal takes time to influence language recognition processes because the two hemispheres represent and process word and voice information differently (Stevens, 2004), including at different speeds (Boemio, Fromm, Braun, & Poeppel, 2005). Although the current study was not designed to examine hemispheric differences, this area of research could have theoretical implications and is discussed in greater detail in the General Discussion.

CHAPTER III

PERCEPTION OF EMOTIONAL TONE OF VOICE

During verbal communication, a speaker's emotional state is conveyed to a listener by modulating voice intonation. Emotional tone of voice is typically characterized by variations in loudness, pitch, duration of syllables, and voice quality of an utterance (Bachorowski & Owren, 2003; Cutler, Dahan, & Donselaar, 1997). According to the affective primacy hypothesis, the human mental system processes affective information contained in a stimulus pre-consciously and automatically (Murphy & Zajonc, 1993). Consequently, listeners are extremely skilled at inferring emotion from vocal cues, and do so rather quickly and accurately (Bezooijen, 1984; Frick, 1985; Scherer, 1986). More recent research measuring functional magnetic resonance imaging (fMRI) data (Wildgruber et al., 2005) and event-related potentials (ERPs; Wambacq, Shea-Miller, & Abubakr, 2004) provides further evidence that humans recognize emotional tone of voice with relative ease.

Neuropsychological studies on the perception of emotional tone of voice have identified distinct regions of cognitive processing. Voice perception is thought to involve partially dissociated functional neural pathways for processing speech (linguistic elements), speaker identification, and emotional tone of voice. Positron emission tomography (PET) and fMRI studies identify the importance of the amygdale and interior insular in the processing of vocal emotion, with greater activation of the right temporal lobe and right inferior prefrontal cortex when attention is directed to emotional prosody (Belin, Fecteau, & Bedard, 2004). Although the current study is not a neuropsychological investigation, it is concerned with the point during listeners' perceptual processing of a spoken word at which these neural pathways for processing emotional tone of voice affect the neural pathways responsible for the processing of linguistic elements of spoken words.

Humans are able to perceive emotional tone of voice from a very young age. Infants as young as two days have the ability to discriminate at least four different emotional tones of voice (happiness, sadness, anger, and neutral emotion; Mastropieri & Turkewicz, 1999). Research on infants and spoken word recognition provides a lifespan perspective on the role that emotional tone of voice plays when listening to spoken words. Singh, Morgan, and White (2004) investigated emotional tone of voice and spoken word recognition in infants. Their study found that infants prefer to listen to happy speech, as evidenced by longer listening times to words with happy affect. More importantly, the infant listeners only recognized happy words faster than unfamiliar words when the word was presented in a happy affect during a familiarization period. Younger infants (7.5 months old) showed recognition of familiarized words only when a word matched in affect (happy or neutral) across familiarization and testing periods. Thus, infants' preference for listening to speech in a happy emotional tone of voice did not bias their language processing. If this were the case, then these infants should have always recognized the words spoken with a happy affect. Instead, the match between the affect in which a particular word had been spoken during familiarization and during latter testing was apparently more important than the word simply sounding "happy." The authors of this study explain their results in terms of early lexical representations. Their results indicated that infants' recognition of new word prototypes is hindered by the highly precise acoustic and phonetic features. In other words, this study provides evidence for indexical specificity effects of emotional tone of voice in infants' spoken word recognition. As previously discussed, these effects occur when processing deficits result from variation in specific, or indexical, aspects of speech. The authors also discussed their results in relation to spoken word recognition research on adults. Young adults have demonstrated perceptual interference in linguistic processing of speech resulting from stimulus variation in specific surface details (Mullennix, Pisoni, & Martin, 1989; Newman, Clouse, & Burnham, 2001; Sommers, Nygaard, & Pisoni, 1994). In a similar manner, although 7.5 and 10.5 month-old infants did not demonstrate perceptual interference with variation in emotional tone of voice, the infants did demonstrate facilitated spoken word recognition when indexical aspects of the stimuli did not vary. While variation in emotional tone of voice did not impede infants' perception, when this variation was not present, their perception was facilitated.

Cognitive processes and emotion are closely related. This relationship is evident in research investigating the effects of emotion on memory consolidation (Dudai, 2004; Hamman, 2001). Emotion also has effects on the cognitive processes involved in decision making (Bechara, 2004; Bechara, Damasio, & Damasio, 2000). Language perception is another process that incorporates emotion (in the form of emotional tone of voice) with cognitive processing of spoken words. For this reason, studying emotional tone of voice and spoken word recognition provides important information about the connection between cognitive and affective systems. Although the current study investigated a very specific issue in spoken word recognition, the results could have more widespread implications for theories of spoken word recognition and processing of emotional tone of voice.

CHAPTER IV

EFFECTS OF EMOTION ON PERCEPTION AND MEMORY

There is a strong relationship between emotion and perception. Several studies have addressed this relationship by examining the effects of emotional priming on perception. It is important to distinguish the differences between emotional priming and the priming used in the current study (i.e., priming effects of emotional tone of voice on the perception of spoken words), as well as to recognize the implications that emotional priming studies have on the constructs currently under investigation.

In an emotional priming experiment by Halberstadt and Niedenthal (1997), individuals in emotional states (happy or sad) gave more weight to the emotional dimension and less weight to the non-emotional dimensions (gender and orientation of the head) of face stimuli in a judgment of similarity compared to emotion-neutral individuals. While this study suggests that a person's emotional state biases them to attend to emotional information in a similarity judgment, it does not explain how that state will affect the processes of online spoken word recognition. Other studies (Niedenthal & Halberstadt, 1995; Niedenthal & Setterlund, 1994) have found decreased spoken word recognition RTs (facilitative effects) for words related to the emotion induced (happy or sad). For example, when participants were primed to be happy, they responded faster to the word "smile" than they did to the word "frown". Although the current study did not prime (or induce) participants to feel a particular emotion, it is possible that a participant's emotional state at the time of participation could influence his or her RTs to words related to that current emotional state. The current study minimized the chances of these effects by implementing emotion-neutral words in the spoken recognition task. Moreover, by randomly assigning participants to groups, any effects of current emotional state should have been minimized in the current study. Future work should explore the extent to which the congruence (or lack thereof) between the listeners' emotion and the emotional tone of voice (as well as the emotional connotation) of the stimuli modulates the effects.

Previous research has shown that variation in emotional tone of voice affects the mechanisms through which humans process linguistic aspects of the speech signal. One study by Mullenix, Bihon, Bricklemyer, Gaston, and Keener (2002) investigated this relationship using a same/different speeded classification task. Participants in the experiment heard the names "Todd" and "Tom" presented in a male or female voice with either a "surprised" or "angry" emotional tone of voice and were then instructed to indicate whether the two name stimuli on each trial were the same or different. Participants' RTs were slower in this task when the talkers (male or female) or the emotional tone of voice (surprised or angry) mismatched. This study suggests that participants will take longer to respond when the emotional tone of voice of a target word differs from the prime (mismatch) compared to when the emotional tone of voice of a word matches from the prime to target blocks.

While the current proposal is partially motivated by this Mullennix et al. (2002) study, it will differ in several important ways. First, although their study provides important information regarding how adjusting for variability in emotional tone of voice affects the *processing* of spoken words, it does not address the long-lasting effects of that variability has on the *representations* underlying spoken language. By using a long-term priming paradigm, the current study provides information regarding the nature of the representations in the mental lexicon, as opposed to short-term paradigms like the one used by Mullennix, et al. (2002). Second, in addition to utilizing an alternative paradigm to investigate the effects of variation in emotional tone of voice on spoken word recognition, the current study also examines the time course of these effects. Doing so will demonstrate how indexical variability in emotional tone of voice affects listeners' perception of spoken words at different points during perceptual processing. Third, Mullennix et al. (2002) utilized inter-talker variability in emotional tone of voice, whereas the current study will utilize intra-talker variability[2].

---

[2]Indeed, the vast majority of studies examining indexical variability have focused on inter-talker variability.

CHAPTER V

PROCESSING OF EMOTIONAL TONE OF VOICE IN MUSICIANS AND NONMUSICIANS

Recent research suggests that children and adults with musical experience exhibit heightened perceptual sensitivity to emotion in speech, allowing them to more accurately identify emotional tones of voice in samples of speech (Dmitrieva, Gel'man, Zaitseva, & Orlov, 2006; Thompson, Schellenberg, & Husain, 2004). Support for this claim is provided by a study of auditory-evoked potential responses to complex emotional vocal stimuli (Strait, Kraus, Skoe, & Ashley, 2009). In this study, neural responses to emotional cues in auditory speech were much faster in musicians compared to non-musicians. More specifically, musicians showed enhanced neural responses to more acoustically complex and relatively minimized neural responses to less acoustically complex auditory stimuli. Because musicians differed in speed of processing of emotional speech, musical experience could also play a role in the time course of spoken word recognition and the role of emotional tone of voice.

According to Strait et al. (2009), by devoting neural resources according to the complexity of the speech signal, musicians demonstrate perceptual capabilities of emotional speech that are both economical and efficient. The specific types of enhancements found depend on whether an individual is classified by the age of onset of musical training or by years of consistent musical training. Musicians who had begun musical training at or before age 7 years exhibited perceptual enhancements related to the representation of pitch and timbre in emotional vocal stimuli. Musicians who had received more than 10 years of consistent musical training affected timing-related perceptual facilitation (Strait et al., 2009).

Because of musicians' enhanced sensitivity to emotion in speech, they may perceive indexical variability in spoken words differently than those without musical training. It is possible that non-musicians who are less cognizant of emotion in speech will be less susceptible to specificity effects of emotional tone of voice. Additionally, musician participants in the current study could have longer RTs if the heightened sensitivity to mismatched emotional tones interferes with their task of making lexical decisions. In this case, musicians may show specificity effects, even in the easy lexical decision task. Alternatively, the musicians' sensitivity could be associated with shorter RTs, indicating that variation in emotional tone of voice does not interfere with the lexical decision task. If this occurs, then unlike non-musicians, musicians may show reduced specificity effects, even in the hard lexical decision task. In either of these possible scenarios, a different pattern of results is predicted for musicians and non-musicians. Due to this possibility, qualitative and quantitative measurements of musical experience were taken (see Appendix G) in order to test for effects on spoken word recognition. Consequently, the current study adds to a growing body of literature demonstrating the relationship between musical training and the perception of emotional tone of voice.

CHAPTER VI

PILOT STUDIES

Method

Prior to carrying out Experiments 1 and 2, two pilot studies were conducted. Ten participants were recruited for each pilot study, resulting in a total of 20 participants.

The first pilot study assessed the intelligibility of all of the auditory target word stimuli intended for use in Experiments 1 and 2. Participants were instructed to complete a shadowing task during which they heard every target word over headphones and repeated each individual stimulus word out loud. Responses were recorded and scored for accuracy. Spoken word stimuli were included in the main experiments if at least 8 out of 10 participants correctly shadowed the spoken word stimuli. All of the screened stimuli met this criterion.

The second pilot study was conducted to determine the extent to which the two emotional tones of voice were distinguishable. Participants were instructed to indicate the emotional tone of voice of each auditory stimulus word presented over headphones by pushing one of two buttons labeled "frightened" and "sad." Responses were scored for accuracy. In order to meet the screening criteria and be suitable for use in the main experiments, mean accuracy for both sad and frightened stimuli needed to be at least

80%, which is in line with previous research (e.g., Castro & Lima, 2010). All of the screened stimuli met this criterion. See Appendix B for a complete list of the stimuli used in both experiments.

CHAPTER VII

EXPERIMENT 1: EASY DISCRIMINATION

The long-term repetition-priming paradigm and a lexical decision task were used in two experiments[3] to investigate indexical specificity effects associated with emotional tone of voice. Different sets of nonwords were used in Experiments 1 and 2. This manipulation was designed to affect the speed with which participants processed the spoken word stimuli. In Experiment 1, the nonwords were unwordlike (e.g., *yeeshgeesh*), making the discrimination between words and nonwords relatively easy. Accordingly,

---

[3]Alternatively, this study could have been designed as one experiment, manipulating difficulty of the lexical decision between participants. Such a design would have allowed us to conduct a mixed analysis with lexical decision difficulty as a between participants factor. The main advantage for doing so is that this design would allow us to directly compare specificity effects as a function of task condition within the same experiment. However, we decided to keep the design consistent with M[c]Lennan and Luce's (2005) study by conducting two separate experiments, and then subsequently using an independent t-test to directly compare the differences between the match and mismatch conditions, or magnitude of specificity (MOS), in Experiments 1 and 2.

processing of all items in Experiment 1, including the target word stimuli, was expected to be relatively fast. Therefore, based on the time-course hypothesis, indexical specificity effects associated with emotional tone of voice were expected to be attenuated (or absent) in Experiment 1. Alternatively, in Experiment 2, the words were wordlike (e.g., *bacov*), making the discrimination between words and nonwords relatively difficult. Thus, processing of all items in Experiment 2, including the target stimuli, was expected to be relatively slow. Therefore, based on the time-course hypothesis, larger indexical specificity effects associated with emotional tone of voice were predicted in Experiment 2.

Method

Participants. Thirty-six participants were recruited from the Cleveland State University community and received either partial credit for a course requirement or extra credit. Participants were right-handed native speakers of American English with no reported history of speech or hearing disorders.

Materials. The auditory stimuli consisted of (a) 12 bisyllabic target words spoken in a frightened emotional tone of voice and 12 bisyllabic target words spoken in a sad emotional tone of voice; (b) 12 bisyllabic nonwords spoken in a frightened emotional tone of voice and 12 bisyllabic nonwords spoken in a sad emotional tone of voice; and (c) 8 bisyllabic control items.

All words and nonwords were exactly the same as M^c^Lennan and Luce (2005). New auditory stimuli were recorded with a different speaker and in two different emotional tones of voice. Accordingly, to make word-nonword discrimination easy, the nonwords were unwordlike (e.g., *thushthudge*). The nonwords were created by using sequences with low phonotactic probability, determined by both positional segment frequency (i.e., how often a particular segment occurs in a position in a word) and biphone frequency (i.e., segment-to-segment co-occurrence probability). See Appendix B for a complete list of the stimuli used in both experiments.

The mean log frequency of occurrence for the target stimuli was .79 (Kučera & Francis, 1967). As expected, words spoken in a frightened emotional tone of voice had a shorter duration (M = 733 ms) than words spoken in a sad emotional tone (M = 955 ms), t (22) = -4.551, $p < .001$. The difference in average durations replicates the findings of previous research (Sobin & Alpert, 1999); thus, no attempts were made to equate the durations of the stimuli.

Recording of auditory stimuli. Auditory stimuli were recorded in a sound-attenuated room using Praat software (Boersma & Weenink, 2006). A non-actor female speaker of a Midwestern dialect was paid $25 to record all auditory stimuli[4]. The speaker was instructed to simulate, or portray, the word and nonwords in both a frightened and sad emotional tone of voice. Although such portrayals may generate more intense,

---

[4]In a spectral analysis of induced vocal affect expressions, Dmitrieva et al., (2008) found that recordings of neutral, happy, and angry expressions did not differ between actors and non-actors. Their measurements included power, fundamental frequency (F0), formant frequencies (F1, F2), utterance duration, and articulation rate. According to these findings, it is unnecessary to recruit professional actors to record the auditory stimuli because the acoustic properties of actors' and non-actors' vocal expressions should not differ.

prototypical vocal expressions when compared to "normal" emotional expression patterns, Scherer (2003) argues that all observable vocal expressions are in part simulated portrayals. Because portrayals of emotions in vocal expressions are easily recognized by listeners, they must reflect reality to some extent. Moreover, at worst, such an argument would only suggest that our predicted results would be stronger in situations in which talkers' emotional tone is rather intense. Nevertheless, the main purpose of this study was to examine whether variation in emotion *can* have the predicted effects, not to predict their likelihood of occurrence. In order to help the talker produce the appropriate emotional tone of voice, she was instructed to imitate the emotional tone of voice of a fictional character in an emotional situation described in a written passage. The passages were taken from Leinonen, et al. (1997) and modified to facilitate appropriate production of "frightened" and "sad" emotional tones of voice (Mullennix et al., 2009; See Appendix A). All words were edited into individual files and stored on computer disk for later playback.

Selection of emotional tones of voice. All of the auditory stimuli were presented in either a sad or frightened emotional tone of voice. Because of the subjective properties of emotional tone of voice, it was important that the emotions used in the current study were easily distinguishable. In order to establish conditions in which emotional tones of voice either matched or mismatched, the emotional tones needed to be qualitatively different. Using two distinct emotional tones should minimize any overlapping of acoustic properties, thereby creating a clear-cut distinction between the emotions. This should make the matching words more similar to, and the mismatching words more different from, each other. Sobin and Alpert (1999) analyzed the acoustic properties of speech spoken in four emotional tones of voice: fear, anger, sadness, and joy. Of the measurements taken, no acoustic characteristics were shared by fear and sadness, which was the greatest acoustic differentiation observed. Others (Johnstone & Scherer, 2000; pp. 226-227; Scherer, 1982; pp. 136-198) also consistently report evidence for high acoustic differentiation between the vocal expression of "fear/panic" and "sadness." For these reasons, frightened and sad were the emotional tones of voice used in the current study[5].

Design. Auditory stimuli were presented in two blocks: a prime block followed by a target block. Stimuli spoken in both emotional tones of voice served as both primes and targets. For each block, half of the stimuli were spoken in a frightened emotional tone of voice and half were spoken in a sad emotional tone of voice. Primes were

---

[5] In addition, frightened and sad emotional tones of voice were chosen because they both have relatively negative connotations, which could produce differential effects in younger and older adults. Compared to younger adults, older adult listeners tend to be less likely to attend to negative information (Mather & Carstensen, 2003) and remember negative stimuli (Charles, Mather, & Carstensen, 2003; Fung & Carstensen, 2003; Mather et al., 2004; Mather & Johnson, 2000; Spaniol, Voss, & Grady, 2008). Using negative stimuli will allow us to extend the current study to older adults in the future, providing important information about the aging lexicon. One additional point worth mentioning is that the use of highly distinctive emotional tones of voice is consistent with studies of talker variability. The majority of studies examining talker variability have used talkers of different genders in order maximize the distinctiveness of the talkers' voices.

matched, mismatched, or unrelated to the targets. Emotional tone of voice of matched primes and targets were identical (e.g., $circus_{\text{frightened}}$, $circus_{\text{frightened}}$; $circus_{\text{sad}}$, $circus_{\text{sad}}$). Emotional tone of voice of mismatched primes and targets differed (e.g., $circus_{\text{frightened}}$, $circus_{\text{sad}}$; $circus_{\text{frightened}}$, $circus_{\text{sad}}$). Both the prime and target blocks consisted of 24 stimuli, 12 words and 12 nonwords. The composition of the prime block was as follows: 8 target words, 8 nonwords, and 8 control stimuli (4 of the control stimuli were words, 4 were nonwords). The target block consisted of 12 target words and 12 nonwords. In the target block, 8 stimuli matched, 8 mismatched, and 8 were controls. Although the preparation of the nonwords and their rotation through the various conditions paralleled the real word target stimuli, the nonwords and the unrelated control stimuli (words and nonwords appearing in the prime block only in place of non-repeated stimuli) were simply fillers. That is, the focus of the experimental manipulations and later statistical analyses was limited to the target words.

Orthogonal combination of three levels of prime (match, mismatch, and control) and two levels of target (frightened and sad) resulted in six conditions, which are shown in Table 1. Across participants, each frightened and sad item participated in every possible condition. However, no single participant heard more than one version of a given word within a block. For example, if a participant heard the word *circus* in one of the blocks, he or she did not hear another version of that word again in the same block.

Table 1

Experimental Conditions and Examples of Primes and Targets in Experiments 1 and 2

Condition	Prime	Target
Match		
Emotion 1 prime → Emotion 1 target	*circus*$_{frightened}$	*circus*$_{frightened}$
Emotion 2 prime → Emotion 2 target	*circus*$_{sad}$	*circus*$_{sad}$
Mismatch		
Emotion 2 prime → Emotion 1 target	*circus*$_{sad}$	*circus*$_{frightened}$
Emotion 1 prime → Emotion 2 target	*circus*$_{frightened}$	*circus*$_{sad}$
Control		
Unrelated prime → Emotion 1 target	*jagged*	*circus*$_{frightened}$
Unrelated prime → Emotion 2 target	*jagged*	*circus*$_{sad}$

Procedure. Upon arrival to the laboratory, each participant provided informed consent and completed a demographics questionnaire (See Appendix E). Each participant also completed a handedness inventory (See Appendix D; Cohen, 2008), which was adapted from the Edinburgh inventory (Oldfield, 1971), an objective measure of the extent of right- or left-handedness of the individual. Finally, each participant completed the Musical Listening Test (MLT; Delosis Limited, 2009), which is an abbreviated version of The Montreal Battery for the Evaluation of Amusia (MBEA; Peretz, 2003). The test involves two phases of listening to pairs of musical tunes and deciding if they are the same or different. The first phase includes tunes with same or different scales and the second includes tunes with same or different rhythms. This assessment provided a quantitative measure (two numbers designating number correct out of 30) of musical ability in addition to qualitative descriptions that were obtained during the post-experiment questionnaire described below.

Participants were tested individually in a quiet room and were not told at the beginning of the experiment that there would be two blocks of trials. Participants performed a lexical decision task in which they were instructed to decide as quickly and accurately as possible whether the item they heard was a real English word or a nonword. They indicated their decision by pressing one of two appropriately labeled buttons (a green button for *word* on the right and a red button for *nonword* on the left) on a response box positioned directly in front of them. In both the prime and target blocks, the stimuli were presented binaurally over headphones. A Macintosh computer controlled stimulus presentation and recorded participants' RTs to make correct lexical decisions and PC. Stimulus presentation within each block was randomized for each participant.

After the participant read instructions on the computer screen, a given trial proceeded as follows: The participant was presented with a stimulus word or nonword binaurally over the headphones and was instructed to make a lexical decision as quickly and accurately as possible. RTs were measured from the onset of the presentation of the stimulus word or nonword to the onset of the participant's button press response. After the participant responded, the next trial was initiated. If the maximum RT (5 s) expired, the computer automatically recorded an incorrect response and presented the next trial.

Upon completion of the target block, the participant was given three minutes to recall as many of the *words* presented in the prime and/or target block as possible by typing responses directly into an Excel spreadsheet. Then, the participant was given an additional two minutes to indicate whether each word he or she recalled was spoken in a frightened or sad emotional tone of voice by typing either "frightened" or "sad" next to each word recalled.

Following the recall task, the participant was instructed to complete a post-experiment questionnaire by typing open-ended answers to questions (unless otherwise noted) displayed on a computer screen (See Appendix G). First, the participant was asked what the purpose of the experiment was to determine whether or not RTs may have been affected by knowledge of the experiment's purpose. Next, a series of questions relating to musical training and experience were presented. These questions probed for information regarding the type of musical training, the age of onset of musical training, and the amount of consistent musical training. Finally, the post-experimental questionnaire asked if the participant had any difficulty hearing or understanding the auditory stimuli. [6]

Results

None of the RTs in Experiment 1 met the exclusion criteria set by M[c]Lennan and Luce (2005) of less than 500 ms or greater than 2,500 ms; thus, no RTs were excluded from the analyses. Moreover, any participant whose overall mean RT fell two standard

---

[6]In addition, alpha-memory span and speeded classification tasks were also given to provide objective measures of working memory and speed of processing, respectively. These tasks were performed in order to determine whether the MOS would vary as a function of working memory or processing speed. Similarly, a neuroticism questionnaire (See Appendix H; Goldberg, 1999; Goldberg et al., 2006) was given because trait neuroticism is related to a predisposal for negative emotional states (e.g., Costa & McCrae, 1980; Larsen & Ketelaar, 1989) and both emotional tones of voice used in the current study (frightened and sad) had negative connotations. No relationship was found between memory, processing speed, or narcissism and MOS; thus, these data are not discussed further.

deviations beyond the grand mean was eliminated from analyses, resulting in the elimination of 1 participant. There was no missing data in Experiment 1.

RT data are rarely identically and independently distributed, due to practice effects, fatigue, and other influences that are usually ignored and considered minor (Wheelan, 2008). However, because RT distributions are not normal (they are positively skewed), RT data violate assumptions for statistical analysis. This violation can lead to a substantial reduction in the ability to detect differences in RT using ANOVA. For that reason, in the following statistical analyses, RTs were inverse-transformed (i.e., speed), according to suggestions from Whelan (2008). Such a transformation is commonly used and accepted in analyses of RT data (e.g., Orfanidou, Davis, Ford, & Marslen-Wilson, 2011). Consequently, all of the following statistical analyses reported for RTs are presented as inverse-transformed speed data, but the means and standard deviations reported were calculated using raw RTs and are used in the tables to facilitate interpretation of the results.

Prime (match, mismatch, control) × Target (sad, frightened) participant analyses of variance (ANOVAs) were performed on mean RTs for correct responses and PCs for the target stimuli. Accuracy was 96% overall and did not yield significant effects. Mean PCs as a function of prime type are reported in Table 2. Although responses to nonwords were not the focus of the current study, the overall mean RT and mean PC for the nonword stimuli were 1,205 ms and 96%, respectively (*SE*s = 218 ms and 9%), indicating that participants were both highly accurate and relatively fast in responding to the nonwords.

Traditional item analyses are not appropriate for the current experiments. The stimuli were carefully selected on the basis of many variables known to affect the dependent variables under investigation. Thus, performing traditional ANOVAs with items as random factors was not justified (see Raaijmakers, Schrijnemakers, & Gremmen, 1999). Furthermore, the design of the current experiments used counterbalanced lists, such that each item appeared in every condition. It has been argued (see Raaijmakers, 2003; Raaijmakers et al., 1999) that conducting separate item analyses in such counterbalanced lists is inappropriate. Furthermore, as noted by M^{c}Lennan and Luce (2005), the nature of the long-term repetition-priming paradigm limits the number of items that can be used in a within-participants manipulation because increasing the number of items tends to decrease the likelihood of attaining long-term repetition-priming effects. Thus, statistical power also would have been necessarily weak as a consequence of the small number of items used. Nevertheless, given that the design of our experiments included counterbalanced lists, such that each of the test items appeared in every condition, two dummy variables representing allocation of participants to experimental lists were included in the ANOVA. Because these dummy variables were included solely to reduce the estimate of random variation (see Pollatsek & Well, 1995), effects involving the dummy variables are not reported.

Table 2

Mean Percentage Correct as a Function of Prime Type for Experiments 1 and 2

Experiment	Match	Mismatch	Control
1: Easy Discrimination	97	98	93
2: Hard Discrimination	90	92	86

Mean RTs as a function of condition, MOS, and magnitudes of priming (MOP) are shown in Table 3. Recall that MOS is the difference in RT between the match and mismatch conditions. MOP is the difference in RT between the match and control conditions.

Frightened items (*M* = 878 ms) were responded to significantly more quickly than were sad items (*M* = 1,045 ms), F (1, 31) = 99.456, *MSE* = 1.826E-6, p < .001, as a function of durational differences. However, this main effect target was not of theoretical interest in the current study.

Table 3

Reaction Times, Standard Errors, and Magnitudes of Specificity (MOS) and Priming (MOP) for Experiments 1 and 2

	Match		Mismatch		Control			
Experiment	RT	*SE*	RT	*SE*	RT	*SE*	MOS	MOP
1: Easy	924	16	931	16	1029	23	-7	-105
2: Hard	1060	25	1080	26	1164	31	-20	-104

Of primary interest was the main effect of prime, which was significant, F (2, 62) = 15.667, *MSE* = 2.321E-7, p < .001. As expected, prime and target did not interact (F < 1.0). Planned comparisons based on the main effect of prime revealed significant differences between the match and control conditions (MOP; p < .001) and between the mismatch and control conditions (p < .001), indicating that both the match and mismatch conditions served as effective primes. As expected, there was no difference between the match and mismatch conditions (no MOS). Mean RTs as a function of prime condition and emotional tone of voice are shown in Table 4.

Table 4

Mean RTs as a Function of Prime Condition and Emotional Tone of Voice for Experiments 1 and 2

Prime-Target	S-S		F-S		C-S		F-F		S-F		C-F	
Experiment	RT	*SE*	RT	*SE*	RT	*SE*	RT	*SE*	RT	*SE*	RT	*SE*
1: Easy	992	124	1013	127	1137	234	858	160	853	118	927	108
2: Hard	1137	180	1174	152	1216	179	979	165	975	146	1099	215

Note. S: Sad emotional tone of voice, F: Frightened emotional tone of voice, C: Control word that is completely different from the target word

Discussion

Both matched and mismatched primes significantly facilitated lexical decision responses. Moreover, mismatched primes in the easy lexical decision experiment facilitated responses to targets as much as matched primes. These results are consistent with the time-course hypothesis discussed previously: When processing was fast (as a result of the easy discrimination allowed by the unwordlike nonwords), indexical specificity effects of emotional tone of voice (i.e., effects related to differences in emotional tone of voice) did not emerge.

Experiment 2 was conducted to test the hypothesis that when processing is slowed by the use of unwordlike nonwords, indexical specificity effects of emotional tone of voice should emerge with the same experimental target word stimuli used in Experiment 1.

CHAPTER VIII

EXPERIMENT 2: HARD DISCRIMINATION

Method

Participants. Thirty-seven different participants were recruited from the same population and met the same criteria as Experiment 1.

Materials. The stimuli consisted of (a) the same 12 bisyllabic spoken target words used in Experiment 1, (b) 12 new spoken bisyllabic nonwords, and (c) 8 bisyllabic spoken control items.

To increase the difficulty of the word-nonword discrimination task, the nonwords were wordlike. Wordlike nonwords were taken from McLennan and Luce (2005)[7], which were created by changing the endings of real words so that they became nonwords (e.g., *bygone, bygups*). Nonwords were spoken by one speaker in both frightened and sad emotional tones of voice.

---

[7]Their research found the same pattern of results in lexical decision RTs between words and nonwords that overlap phonologically and words and nonwords that did not overlap.

The stimuli were recorded in a sound-attenuated room by the same speaker, as previously described. All words were edited into individual files and stored on computer disk for later playback.

Design and Procedure. The design and procedure were identical to those described in Experiment 1.

Results

Two RTs and 1 participant were excluded from the analyses based on the same criteria used in Experiment 1. Moreover, missing cells in the RT data (as a result of errors in both of the trials in a given condition) were not included in the analysis for their respective prime conditions. There were 6 empty cells overall in the target block (3 RTs for match-sad, 0 RTs for mismatch-sad, 3 for control-sad, 0 for match-frightened, 0 for mismatch-frightened, and 0 for control-frightened) in Experiment 2.

Mean PCs as a function of prime type are shown in Table 2 and mean RTs as a function of condition and MOS and MOP are shown in Table 3. Accuracy was 96% overall and did not yield significant effects. Participants in Experiment 2 were significantly more accurate when responding to frightened target words (93%) compared to sad (86%), F (1, 31) = 5.812, MSE = 28.6, p = .022. The overall mean RT and mean PC for the nonword stimuli were 1,299 ms and 92%, respectively.

Frightened items (M = 2,332 ms) were again responded to more quickly than sad items (M = 2,745 ms), F (1, 26) = 62.100, MSE = 9.962E-7, p < .001. Again, a main effect of prime condition resulted, F (2, 52) = 11.540, MSE = 9.976-E8, p < .001. Planned comparisons based on the significant main effect of prime revealed the predicted differences between the match and control conditions (MOP), p < .001, and between the mismatch and control, $p = .009$, indicating that as in Experiment 1, both the match and mismatch conditions served as effective primes. Moreover, the predicted significant difference between the match and mismatch conditions (MOS) failed to emerge, $p = .534$. However, there was a significant interaction between prime and emotional tone of voice, F (2, 52) = 3.292, *MSE* = 2.809E-8, $p = .045$.

Based on the significant interaction, separate ANOVAs were conducted on RTs to the frightened and sad target words. For RTs to the frightened target words, the pattern of results resembled those in Experiment 1: The main effect of prime was significant, F (2, 62) = 9.708, *MSE* = 1.079E-7, $p < .001$; planned comparisons revealed significant differences between the match and control conditions (MOP), $p = .004$, and the mismatch and control conditions, $p = .002$, but not between the match and mismatch conditions (MOS), $p = 1.00$. For RTs to the sad target words, the predicted pattern of specificity effects of emotional tone of voice emerged: There was a main effect of prime, F (2, 52) = 7.100, *MSE* = 4.595E-8, $p = .002$, qualified by significant differences between the match and control conditions (MOP), $p = .008$, no difference between the mismatch and control conditions, $p = 1.00$, indicating that the mismatch condition did not serve as an effective prime, and most importantly, significant differences between the match and mismatch conditions, $p = .024$. Mean RTs by condition and emotional tone of voice are displayed in Table 4.

Discussion

In Experiment 2, both matched and mismatched prime words facilitated lexical decision responses, regardless of emotional tone of voice. However, primes that matched in emotional tone of voice were more effective than primes that mismatched in emotional tone of voice (which did not serve as effective primes, relative to the control condition), but only for sad target words. The predicted pattern did not emerge for the frightened target words. These results are only partially consistent with the time-course predictions outlined above. When processing was relatively slow and difficult, indexical specificity effects of emotional tone of voice emerged for targets in a sad, but not frightened, emotional tone of voice. On the other hand, when processing was fast, indexical specificity effects of emotional tone of voice did not emerge. Thus, it is possible that time course is an important factor in determining how indexical variability in emotional tone of voice affects spoken word recognition.

CHAPTER IX

COMBINED ANALYSIS OF EXPERIMENTS 1 AND 2

Two combined analyses were conducted. The first analysis compared overall RTs in the easy and difficult lexical decision experiments to verify that RTs were significantly longer in the difficult experiment, a necessary requisite for evaluating time course effects. The second analysis compared the MOS in the easy and difficult lexical decision experiments.

As expected, RTs to make lexical decisions in the difficult discrimination experiment were significantly longer than in the easy discrimination experiment, F (1, 68) = 13.440, $MSE < .001$, $p < .001$, indicating that the difficulty of discrimination manipulation between Experiments 1 and 2 was successful. The main effect of discrimination difficulty on MOS was not significant, $F < 1$, which suggests that the magnitudes of indexical specificity of emotional tone of voice in the easy and difficult discrimination experiments were equivalent. Because there was a significant difference between the match and mismatch conditions for the sad target word stimuli but not the frightened in Experiment 2, but not in Experiment 1, a final combined analysis directly compared the MOS for the sad target word stimuli in Experiments 1 and 2. As expected, the magnitude of the MOS for the sad target word stimuli was greater in Experiment 2

(MOS = -38 ms) than in Experiment 1 (MOS = -5 ms), though this difference did not reach significance, F (1, 59) = 1.205, MSE = 19,549.341, p = .277.

Finally, in the recall task, the number of false alarms and PCs for words and emotional tones of voice recalled were calculated for words that appeared in the prime condition only (12 possible words and emotional tones of voice), target condition only (12 possible words and emotional tones of voice), and both the prime and target conditions (8 possible words, 4 possible emotional tones of voice which matched from prime to target conditions. No significant differences were found between participants in Experiment 1 (Easy Lexical Decision) compared to Experiment 2 (Hard Lexical Decision), all ts (85) < 1, except for the correct identification of the emotional tone of voice of the words that matched from prime to target, t (85) = 1.411, p = .161. Thus, the difficulty of the lexical decision task did not have an effect on recall. However, when participants were collapsed across experiments, trends in PCs emerged. A one-way ANOVA on the 3 possible times (prime block only, target block only, prime and target blocks) words could be presented showed a main effect for word identification, F (2,172) = 100.229, MSE = 37.6, p < .001. Pairwise comparisons showed significant differences between all comparisons (prime/target, prime/prime and target, target/prime and target), all ps < .001. Emotional identification resulted in the same pattern of differences: a main effect of prime condition, F (2,172) = 11.228, MSE = 16.9, p < .001. Again, all comparisons showed significant differences, all ps < .001. Mean PCs for word recall are displayed in Table 5. Participants recalled the most words and emotions correctly when they appeared in both the prime and target blocks, followed by the words that appeared only in the target block, then the words that appeared only in the prime block. The superior performance for words and emotional tones of voice repeated in the prime and target blocks suggests priming effects also emerged in the recall data. In addition, superior recall for words in the target block only compared to those in the prime block only could be explained by recency effects.

Table 5

PCs Recalled for Words and Emotional Tones of Voice Presented in Prime Block, Target Block, or Both Prime and Target Blocks

Prime		Target		Prime and Target	
Word	Emotion	Word	Emotion	Word	Emotion
32	17	40	21	45	26

CHAPTER X

MUSICIAN AND MUSICAL LISTENING ABILITY ANALYSES

Musicians were identified according to the criteria used by Strait et al. (2009), which qualified participants who learned how to play a musical instrument prior to age 7 (Experiment 1 n = 4, Experiment 2 n = 6) and those who have had at least 10 years of consistent musical training (Experiment 1 n = 4, Experiment 2 n = 1). Due to overlap between the age and experience classifications in addition to the small sample of musicians, participants were collapsed across experiments and grouped as musicians and non-musicians, whether they met one or both of the above criteria. This resulted in a total of 11 musicians and 59 non-musicians.

Contrary to predictions, musicians (M = 1,025, SD = 148) did not respond faster than non-musicians (M = 1,030, SD = 124), t (67) = .092, p = .93. Another surprising result was that musicians (M = -31, SD = 120) and non-musicians (M = -23, SD = 181) did not have different MOS, t (69) = .247, p = .81.

To examine the RT data in relation to musical listening ability more closely, participants were collapsed across Experiments 1 and 2 and divided into quartiles according to their total score on the MLT (Delosis, 2009), as patterns of RTs across prime conditions were similar across experiments. Recall that the MLT assessed ability to determine if two sets of tunes are the same or different in rhythm and pitch. Separate 3 (prime: match, mismatch, control) X 2 (emotional tone of voice: sad, frightened) ANOVAs were performed on the top 25% (n = 21) and the remaining 75% (n = 52) of scorers on the MLT. The results of the bottom 75% mirrored those from Experiment 1: A significant main effect of prime emerged, F (2, 86) = 18.045, MSE = 2.459E-7, $p <$ .001 and prime and target did not interact ($F < 1.0$). Planned comparisons based on the main effect of prime revealed significant differences between the match and control conditions (MOP; $p < .001$) and between the mismatch and control conditions ($p < .001$), indicating that both the match and mismatch conditions served as effective primes. There was no difference between the match and mismatch conditions ($p = .985$; no MOS). The results of the top 25% showed a similar pattern to that of Experiment 2: Again, a main effect of prime condition resulted, F (2, 26) = 13.467, MSE = 1.038E-7, $p < .001$. Planned comparisons based on the significant main effect of prime revealed the predicted differences between the match and control conditions (MOP), $p = .001$, and between the mismatch and control, $p = .027$, indicating that as in Experiment 1, both the match and mismatch conditions served as effective primes. Moreover, the predicted significant difference between the match and mismatch conditions (MOS) failed to emerge, $p = .205$. However, there was a significant interaction between prime and emotional tone of voice, F (2, 26) = 4.989, MSE = 3.409E-8, $p = .015$.

Based on the significant interaction, separate analyses were conducted on RTs to the frightened and sad target words. For RTs to the frightened target words, the pattern of results resembled those in Experiment 1: The main effect of prime was significant, F (2, 30) = 8.996, MSE = 7.289E-8, $p = .001$; planned comparisons revealed significant differences between the match and control conditions (MOP), $p = .036$, and the mismatch and control conditions, $p = .010$, but not between the match and mismatch conditions (MOS), $p = .381$. For RTs to the sad target words, the predicted pattern of specificity effects of emotional tone of voice emerged: There was a main effect of prime, F (2, 26) = 11.582, *MSE* = 6.553E-8, $p < .001$, qualified by significant differences between the match and control conditions (MOP), $p = .004$, no difference between the mismatch and control conditions, $p = .745$, indicating that the mismatch condition did not serve as an effective prime, and most importantly, significant differences between the match and mismatch conditions, $p = .002$. Mean RTs by condition and emotional tone of voice are displayed in Table 4.

Thus, for participants scoring high in musical listening ability, both matched and mismatched prime words facilitated lexical decision responses, regardless of emotional tone of voice. Primes that matched in emotional tone of voice were more effective than primes that mismatched in emotional tone of voice (which did not serve as effective primes, relative to the control condition), but only for sad target words. The predicted pattern did not emerge for the frightened target words. For those with relatively good musical listening ability, indexical specificity effects of emotional tone of voice emerged for targets in a sad, but not frightened, emotional tone of voice. On the other hand, for those with relatively average or poor musical listening ability, indexical specificity effects of emotional tone of voice did not emerge. Thus, it is possible that musical listening ability plays a role in the way variation in emotional tone of voice affects spoken word recognition.

CHAPTER XI

GENERAL DISCUSSION

The current study aimed to investigate the effects of intra-talker variability in emotional tone of voice on lexical decision responses. Results of Experiments 1 and 2 were expected to follow a time-course pattern similar to those obtained by McLennan and Luce (2005). Indexical specificity effects were expected to emerge when processing was relatively slow (Experiment 2) but not when processing was relatively fast (Experiment 1).

Theoretical Implications

Experiments 1 and 2 were designed to assess processing using a lexical decision task varying in difficulty of discrimination between words and nonwords. In Experiment 1, when processing was fast (as a result of the easy discrimination allowed by the unwordlike nonwords), indexical specificity effects of emotional tone of voice were not predicted (i.e., no effects related to differing emotional tones of voice). This prediction was confirmed by the data: There was no difference between the match and mismatch conditions. Experiment 2 was conducted to test the hypothesis that when processing was slowed by the use of wordlike nonwords, specificity effects were expected to emerge with the same target stimuli that were used in Experiment 1. When processing was slowed by the wordlike nonwords in the hard lexical decision task, indexical specificity effects of emotional tone of voice were predicted. These predictions for the hard lexical decision were confirmed for the sad word target stimuli, but not the frightened target words. In other words, effects on RTs emerged due to intra-talker variation from frightened emotional tone of voice in the prime block to sad emotional tone of voice in the target block, but not intra-talker variation from sad in the prime block to frightened in the target block.

The lack of specificity effects on frightened target words in the hard lexical decision could be tied to the duration of the spoken word stimuli. Recall that in both experiments sad target words were significantly longer in duration than frightened target words; the same durational differences emerged in RTs to sad and frightened target words. In addition, participants were significantly more accurate when responding to frightened words in the target block in Experiment 2. This pattern of results could mean that due to the quicker responses to frightened target words, there was not ample time for specific aspects of the speech signal (i.e., emotional tone of voice) to affect processing. It is possible that even in the difficult lexical decision discrimination task (Experiment 2), the processing of frightened stimuli was still relatively fast. Perhaps processing of sad stimuli in the difficult lexical decision task was slowed down enough to allow time for specificity effects of emotional tone of voice to emerge, but even the slowed processing of frightened stimuli was not sufficiently slow to allow time for such effects. Future research could directly address this possibility by controlling for the duration of emotionally-charged stimuli in order to determine if it is duration, or some inherent difference between the emotional tones of voice, that is causing the different pattern of specificity.

Another possible explanation for the different pattern for the sad and frightened stimuli is that the listeners attended to the sad stimuli more than the frightened stimuli, which resulted in the sad stimuli serving as more effective primes. In Experiment 2, words spoken in a sad emotional tone of voice in the prime block produced equivalent priming for words spoken in both sad and frightened emotional tones of voice in the target block. On the other hand, words spoken in a frightened emotional tone of voice in the prime block only primed words spoken in a frightened tone of voice in the target block. Because the words spoken in a sad emotional tone of voice appeared to serve as a more effective prime, regardless of the emotional tone of voice of the word in the target block, a direct comparison was made between the two conditions in which a word in a sad emotional tone of voice served as the prime (i.e., S-S and S-F conditions in Table 4) with the two conditions in which a word in a frightened emotional tone of voice served as the prime (i.e., F-F and F-S in Table 4). This analysis provided statistical support that the words spoken in a sad emotional tone of voice did indeed serve as a more effective prime – for words spoken in either a sad or a frightened emotional tone of voice, $t\ (19) = 33.411$, $p < .001$.

Implications for Musical Experience and Musical Ability

Contrary to predictions, when classified according to musical experience, musicians did not differ from non-musicians in overall RTs. It is possible that although neuronal responses to emotional speech are faster in musicians compared to non-musicians, this difference does not emerge in RTs to make lexical decisions. More importantly for the current study, musicians and non-musicians, as defined by Strait et al. (2009), did not differ in MOS across experiments. On the other hand, when listeners were classified according to scores on the MLT, notable patterns emerged in RTs. For those scoring in the bottom 75%, indexical specificity effects of emotional tone of voice did not emerge (i.e., no effects related to differing emotional tones of voice); there was no difference between the match and mismatch conditions. For those scoring in the top 25% on the MLT, indexical specificity effects of emotional tone of voice emerged for the sad target words, but not the frightened target words. In other words, effects on RTs emerged due to intra-talker variation from frightened emotional tone of voice in the prime block to sad emotional tone of voice in the target block, but not intra-talker variation from sad in the prime block to frightened in the target block. Thus, musical listening ability could play a role in why specificity effects of emotional tone of voice emerge when switching from one emotional tone of voice from another, but not vice versa.

The results of the current study have important implications for theories of the representational aspects of spoken word recognition. More specifically, these results provide information regarding the ways in which indexical variation in emotional tone of voice is represented in the mental lexicon. The emergence of the predicted results for sad, but not frightened, stimuli provides evidence that specificity effects of some emotional tones of voice, but not others, may follow a precise time course: For sad target words, during earlier perceptual processing, more abstract, underlying features dominated and during later perceptual processing, more specific, detailed surface information dominated. For frightened target words, specific, detailed information did not affect spoken word recognition.

Applied/Practical Implications

The results of the current study could serve as the foundation for several important areas of applied research. First, the results of the proposed research could be applicable to any situation in which auditory stimuli is expressed in a particular emotional tone of voice. By providing a deeper understanding of how spoken word recognition is affected by variation in emotional tone of voice, the current research could be applied in the development of improved speech communication techniques between individuals. The findings of this study could have implications for the development of synthetic speech as well by contributing to current understandings of the mechanisms involved in recognizing words spoken with emotional intonation.

Another domain that could benefit from the results of the current research is forensic voice identification, or "earwitness" testimony (Hollien, 2002). Such implications would presumably extend far beyond the current study. Nevertheless, a deeper understanding of variability in emotional tone of voice could provide information about the validity of earwitness testimony. For example, a spoken word recognition task could be given to witnesses. Doing so would provide an objective measure of the likelihood of valid earwitness testimony provided by a witness. In addition, the results of this study can be applied to foil voices used in a line up of potential criminals to be identified by their voice alone. By extending the current study's findings on how emotional tone of voice is represented and processed during spoken word recognition, forensic investigators could maximize the likelihood of matching foil voices in a line up to a criminal suspect's voice.

Future Directions

Recall that emotion can affect decision making processes. Consequently, future work examining the time course of specificity effects in emotional tone of voice in other tasks without an explicit decision component, such as speeded and delayed shadowing, and eye-tracking, will help to determine the extent to which the current findings are (or are not) task specific. Using different methodologies will help to ensure that our results are indicative of spoken word recognition and not limited to the lexical decision task.

Future research could also explore the possibility of replicating the current study with emotion-related words (e.g., "cry", "scary") instead of emotion-neutral words. Recall that facilitative spoken word recognition effects have been found for emotion-related words heard by individuals with congruent induced emotional states. The current study could be replicated using emotion-related word stimuli recorded in both emotion-matching and emotion-mismatching contexts (e.g., recording "cry" and "scary" each in both a sad and frightened emotional tone of voice). Such an extension could provide further information regarding how emotional tone of voice affects processes involved in online spoken word recognition.

Future research could also investigate hemispheric differences in the effects of variability in emotional tone of voice on spoken word recognition. As previously discussed, such differences have been found for variability in talkers. Whether similar patterns emerge for emotional tone of voice could provide valuable information about the way emotional speech is perceived and the way it is processed in the brain.

REFERENCES

Abercrombie, D. (1967). *Elements of general phonetics.* Chicago: University of Chicago Press.

Abou-Khalil, B. (1995). Insights into language mechanisms derived from the evaluation of epilepsy. In H. S. Kirshner (Ed.), *Handbook of neurological speech and language disorders* (pp. 213–275). New York: Marcel Dekker.

Alpers, G. W. (2008). Eye-catching: Right hemisphere attentional bias for emotional pictures. *Laterality, 13(2),* 158-178.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed. text revision). Washington, D.C.: Author.

Bachorowski, J., and Owren, M. J. (1995). Vocal expression of emotion: acoustic properties of speech are associated with emotional intensity and context. *Psychological Science, 6,* 219-224.

Bechara, A., Damasio, H., and Damasio, A.R. (2000). Emotion, decision-making, and the orbitofrontal cortex. *Cerebral Cortex, 10,* 295-307.

Bechara, A. (2004). The role of emotion in decision-making: Evidence from neurological patients with orbitofrontal damage. *Brain and Cognition, 55,* 30-40.

Belin, P., Fecteau, S., & Bédard, C. (2004). Thinking the voice: neural correlates of voice perception. *TRENDS in Cognitive Sciences, 8(3),* 129-135.

Bezooijen, R. van (1984). The characteristics and recognizability of vocal expressions of emotion. Drodrecht, The Netherlands: Foris.

Boemio, A., Fromm, S., Braun, A., & Poeppel, D. (2005). Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nature Neuroscience, 8,* 389–395.

Boersma, P. and Weenink, D. (2006). Praat: doing phonetics by computer (Version 4.5.08) [Computer program]. Retrieved January 20, 2010, from http://www.praat.org/

Bradlow, A. R., Nygaard, L. C., and Pisoni, D. B. (1999). Effects of talker, rate, and amplitude variation on recognition memory for spoken words. *Perception and Psychophysics, 61 (2),* 206-219.

Caplan, D., Hildebrandt, N., and Makris, N. (1996). Location of lesions in stroke patients with deficits in syntactic processing in sentence comprehension. *Brain, 119,* 933-949.

Castro, S. L., and Lima, C. F. (2010). Recognizing emotions in spoken language: A validated set of Portuguese sentences and pseudosentences for research on emotional prosody. *Behavior Research Methods, 42,* 74-81.

Charles, S. T., Mather, M., and Carstensen, L. L. (2003). Aging and emotional memory: The forgettable nature of negative images for older adults. *Journal of Experimental Psychology: General, 132,* 310-324.

Church, B. A., and Schachter, D. L. (1994). Perceptual specificity of auditory priming: Implicit memory for voice intonation and fundamental frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20(3),* 521-533.

Cohen, M. (2008). *Handedness inventory.* Retrieved January 13, 2010, from http://www.brainmapping.org/shared/Edinburgh.php

Costa, P. T., and McCrae, R. R. (1980). Influence of extraversion and neuroticism on subjective well-being: Happy and unhappy people. *Journal of Personality and Social Psychology, 38,* 668-678.

Craik, F. M., and Kirsner, K. (1974). The effect of speaker's voice on word recognition. *Quarterly Journal of Experimental Psychology, 26,* 274-284.

Cutler, A., Dahan, D., Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech, 40,* 141-201.

Delosis Limited. (2009). *Musical Listening Test.* Retrieved January 13, 2010, from http://www.delosis.com/listening/home.html

Dmitrieva, E. S., Gelvin, V. Ya., Zaitseva, K. A., and Orlov, A. M. (2008). Peculiarities of vocal emotion expression: Analysis of acoustic parameters. *International Journal of Psychophysiology, 69,* 284.

Dudai, Y. (2004). The neurobiology of consolidations, or, how stable is the engram. *Annual Review of Psychology, 55,* 51-86.

Frick, R. W. (1985). Communicating emotion: The role of prosodic features. *Psychological Bulletin, 97,* 412-429.

Fung, H. H., and Carstensen, L. L. (2003). Sending memorable messages to the old: Age differences in preferences and memory for advertisements. *Journal of Personality and Social Psychology, 85,* 163-178.

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, and F. Ostendorf (Eds.), *Personality Psychology in Europe,* Vol. 7 (pp. 7-28). Tilburg, The Netherlands: Tulburg University Press.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., and Gough, H. C. (2006). The International Personality Item Pool and the future of public-domain personality measures. *Journal of Research in Personality, 40,* 84-96.

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22(5),* 1166-1183.

Goldinger, S. D., Pisoni, D. B., and Logan, J. S. (1991). On the nature of talker variability effects on recall of spoken word lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17(1),* 152-162.

González, J. and M^{c}Lennan, C. T. (2007). Hemispheric differences in indexical specificity effects in spoken word recognition. *Journal of Experimental Psychology: Human Percepton and Performance, 33(2),* 410-424.

Grossberg, S. (1986). The adaptive self-organization of serial order in behavior: Speech, language, and motor control. In E. C. Schwab and H.c. Nusbaum (Eds.), *Pattern recognition by humans and machines: Vol. 1. Speech perception* (pp. 187-294). New York: Academic Press.

Grossberg, S., and Myers, C. W. (2000). The resonant dynamics of speech perception: Interword integration and duration-dependent backward effects. *Psychological Review, 107(4),* 735-767.

Grossberg, S., and Stone, G. (1986). Neural dynamics of word recognition and recall: Attentional priming, learning, and resonance. *Psychological Review, 93,* 46-74.

Halberstadt, J. M., & Niedenthal, P. M. (1997). Emotional state and the use of stimulus dimensions in judgment. *Journal of Personality and Social Psychology, 72(5),* 1017-1033.

Hamman, S. (2001). Cognitive and neural mechanisms of emotional memory. *Trends in Cognitive Science,s 5*, 394-400.

Hollien, H. F. (2002). *Forensic voice identification.* San Diego: Academic Press.

Houston, D. M., and Jusczyk, P. W. (2003). Infants' long-term memory for the sounds patterns of words and voices. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 1143-1154.

Johnstone, T., and Scherer, K. R. (2000). Vocal communication of emotion. In M. Lewis and J. Haviland (Eds.), *Handbook of emotion* (2nd ed., pp. 226-227). New York: Guilford.

Knecht, S., Deppe, M., Drager, B., Bobe, L., Lohmann, H., Ringelstein, E.-B. et al. (2000). Language lateralization in healthy right-handers. *Brain, 123,* 74-81.

Kučera, H., and Francis, W. (1967). *Computational analysis of present day American English.* Providence, RI: Brown University Press.

Ladefoged, P., and Broadbent, D. E. (1957). Information conveyed by vowels. *Journal of the Acoustical Society of America, 29,* 98-104.

Larsen, R. J., and Ketelaar, T. (1989). Extraversion, neuroticism, and susceptibility to positive and negative mood induction procedures. *Personality and Individual Differences, 10(12)*, 1221-1228.

Laver, J. (1968). Voice quality and indexical information. *British Journal of Disorders of Communication, 3,* 43-54.

Leinonen, L., Hiltunen, T., Linnankoski, I., and Laasko, M. (1997). Expression of emotional-motivational connotations with a one-word utterance, *Journal of the . Acoustical Society of America, 102*, 1853-1863.

Martin, C. S., Mullennix, J. W., Pisoni, D. B., and Sommers, W. V. (1989). Effects of talker variability on recall of spoken word lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15(4),* 676-684.

Mastropieri, D. & Turkewicz, G. (1999). Prenatal experience to neonatal responsiveness to vocal expression of emotion. *Developmental Psychobiology, 35,* 204-214.

Mather, M., Canli, T., English, T., Whitfield, S., Wais, P., Ochsner, K. et al. (2004). Amygdala responses to emotionally valenced stimuli in older and younger adults. *Psychological Science, 15(4),* 259-263.

Mather, M. and Carstensen, L. L. (2003). Aging and attentional biases for emotional faces. *Psychological Science, 14,* 409-415.

Mather, M. and Johnson, M. K. (2000). Choice-supportive source monitoring: Do our decisions seem better to us as we age? *Psychology and Aging, 15(4),* 596-606.

McLennan, C. T. & Luce, P. A. (2005). Examining the time course of indexical specificity effects in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 306-321.

Mullennix, J.W., Bihon, T. Bricklemyer, J., Gaston, J., and Keener, J.M. (2002). Effects of variation in emotional tone of voice on speech perception. *Language and Speech, 45,* 255-283.

Mullennix, J.W., Gaston, J., Long, J., Wilson, M., May, E., & Tessmer, B. (2009). Encoding emotional tone of voice into memory. Manuscript submitted for publication.

Mullennix, J. W., Pisoni, D. B., and Martin, C. S. (1989). Some effects of talker variability on spoken word recognition. *Journal of the Acoustical Society of America, 85*, 365-378.

Murphy, S. T., & Zajonc, R. B. (1993). Affect, cognition, and awareness: Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology, 64,* 723-739.

Newman, R. S., Clouse, S. A., and Burnham, J. L. (2001). The perceptual consequences of within-talker variability in fricative production. *Journal of the Acoustical Society of America, 109,* 1181-1196.

Niedenthal, P. M., & Halberstadt, J. B. (1995). The acquisition and structure of emotional response categories. In D. L. Medin (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 33, pp. 23-64). New York: Academic Press.

Niedenthal, P. M., & Setterlund, M. B. (1994). Emotional congruence in perception. *Personality and Social Psychology Bulletin, 20,* 401-410.

Nygaard, L.C., and Pisoni, D. B. (1998). Talker-specific learning in speech perception. *Perception and Psychophysics, 60(3),* 355-376.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia, 9(1),* 97-113.

Orfanidou, E., Davis, M. H., Ford, M. A., and Marslen-Wilson, William D. (2011). Perceptual and response components in repetition priming of spoken words and pseudowords, *The Quarterly Journal of Experimental Psyhcology, 64(1),* 96-121.

Palmeri, T. J., Goldinger, S. D., and Pisoni, D. B. (1990). Episodic encoding of voice and recognition memory for spoken words. *Research on Speech Perception Progress Report, 16* (pp. 391-412). Bloomington: Indiana University.

Palmeri, T. J., Goldinger, S. D., and Pisoni, D. B. (1993). Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19(2),* 309-328.

Peretz, I., Champod, A. S., and Hyde, K. (2003). Varieties of musical disorders: The Montreal Battery of Evaluation of Amusia. *Annals of the* New York *Academy of Sciences, 999*, 1-18.

Pisoni, D. B. (1997). Some thoughts on "normalization" in speech perception. In K. Johnson and J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 9-32). San Diego, CA: Academic Press.

Pollatsek, A. & Well, A.D. (1995) On the Use of Counterbalanced Designs in Cognitive Research: A Suggestion for a Better and More Powerful Analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 785-794.

Raaijmakers, J. G. W. (2003). A further look at the "Language-as-Fixed-Effect Fallacy." *Canadian Journal of Experimental Psychology, 57,* 141-151.

Raaijmakers, J. G. W., Schrijnemakers, J. M. C., and Gremmen, F. (1999). How to deal with "The Language-as-Fixed-Effect Fallacy": Common misconceptions and alternative solutions. *Journal of Memory and Language, 41,* 416-426.

Remez, R. E., Fellowes, J., and Rubin, P. E. (1997). Talker identification based on phonetic information. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 651-666.

Sato, W., and Aoki, S. (2006). Right hemispheric dominance in processing of unconscious negative emotion. *Brain and Cognition, 62,* 261-266.

Schacter, D. L., and Church, B. A. (1992). Auditory priming: Implicit and explicit memory for words and voices. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 915-930.

Scherer, K. R. (1982). Methods of research on vocal communication: Paradigms and parameters. In K. R. Scherer and P. Ekman (Eds.), *Handbook of methods in nonverbal behavior research* (pp. 136-198). New York: Cambridge University Press.

Scherer, K. R. (1986). Vocal affect expression: A review and a model for future research. *Psychological Bulletin, 99,* 143-165.

Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication, 40,* 227-256.

Sidtis, J. J. (2000). From chronograph to functional image: What's next? *Brain and Cognition, 42,* 75-77.

Singh, L., Morgan, J. L., & White, K. S. (2004). Preference and processing: The role of speech affect in early spoken word processing. *Journal of Memory and Language, 51,* 173-189.

Smith, S. D., and Bulman-Fleming, B. (2005). An examination of the right-hemisphere hypothesis of the lateralization of emotion. *Brain and Cognition, 57,* 210-213.

Sobin, C. and Alpert, M. (1999). Emotion in speech: The acoustic attributes of fear, anger, sadness, and joy. *Journal of Psycholinguistic Research, 28,* 347-365.

Sommers, M. S., Nygaard, L. C., and Pisoni, D. B. (1994). Stimulus variability and spoken word recognition: Effects of variability in speaking rate and overall amplitude. *Journal of the Acoustical Society of America, 96,* 1314-1324.

Spaniol, J., Voss, A., and Grady, C. L. (2008). Aging and emotional memory: Cognitive mechanisms underlying the positivity effect. *Psychology and Aging, 23(4),* 859-872.

Springer, J. A., Binder, J. R., Hammeke, T. A., Swanson, S. J., Frost, J. A., Bellgowan, P. S. F. et al. (1999). Language dominance in neurologically normal and epilepsy subjects: A functional MRI study. *Brain, 122,* 2033-2045.

Stevens, A. A. (2004). Dissociating the cortical basis of memory for voices, words, and tones. *Cognitive Brain Research, 18,* 162–171.

Strait, D. L., Kraus, N., Skoe, E., & Ashley, R. (2009). Musical experience and neural efficiency-effects of training on subcortical processing of vocal expressions of emotion. *Cognitive Neuroscience, 29,* 661-668.

Tenpenny, P. L. (1995). Abstractionist versus episodic theories of repetition priming and word identification. *Psychonomic Bulletin and Review, 2(3),* 339-363.

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2004). Decoding speech prosody: do music lessons help? *Emotion, 4,* 46-64.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving and W. Donaldson (Eds.), *Organization of memory* (pp. 381-403). New York: Academic Press.

Tulving, E. (1983). *Elements of episodic memory.* Oxford: Oxford University Press.

Vanlancker-Sidtis, D. (2004). When only the right hemisphere is left: Studies in language communication. *Brain and Language, 91,* 199-211.

Wambacq, I. J. A., Shea-Miller, K. J., & Abubakr, A. (2004). Non-voluntary and voluntary processing of emotional prosody: an event-related potentials study. *Cognitive Neuroscience and Neuropsychology, 15(3),* 555-559.

Whelan, R. (2008). Effective analysis of reaction time data. *The Psychological Record, 58 (3),* Article 9. Available at: http://opensiuc.lib.siu.edu/tpr/vol58/iss3/9

Wildgruber, D., Riecker, A., Hertrich, I., Erb, M., Grodd, W., Ethofer, T., and Ackermann, H. (2005). Identification of emotional intonation evaluated by fMRI. *NeuroImage, 24,* 1233-1241.

Yonan, C. A., and Sommers, M. S. (2000). The effects of talker familiarity on spoken word identification in younger and older adults. *Psychology and Aging, 15,* 88-99.

APPENDICES

APPENDIX A

Preparatory passages used to guide recording of different emotional tones of voice used in Experiments 1 and 2, adapted from Mullennix et al. (2009).

Sad. The word was spoken by his friend who had not passed the final examination in spite of much hard work.

Frightened. The word was spoken by his friend when they met a dangerous dog which started to charge toward them on a path in the woods.

APPENDIX B

Target, Nonword, Filler, and Control Stimuli Used in Experiments 1 and 2

Experiment 1	Experiment 2
Targets	
bacon	*bacon*
baggage	*baggage*
boycott	*boycott*
bucket	*bucket*
bygone	*bygone*
bypass	*bypass*
cabbage	*cabbage*
cabin	*cabin*
caucus	*caucus*
circuit	*circuit*
circus	*circus*
coping	*coping*
Nonwords	
jʌʃðʌtʃ	bekəv
θʌsjʌdʒ	bægənt
tʃʌʃθʌdʒ	boɪkɔf
jʌtʃtʃʌdʒ	bʌkəm
θʌtʃʃʌdʒ	baɪgəps
ðaɪðʃaɪð	baɪpæb
ʃaɪðtʃaɪð	kæbəv
gaɪððaɪz	kɔkəg
ðaɪbdʒaɪz	sɜkɚ
ðaɪvʃaɪb	kopɚg
tʃaɪzwaɪð	dʒægʌp
jiʃgiʃ	wɛpʌks
Controls	
luggage	*luggage*
jagged	*jagged*
ribbon	*ribbon*
dʒeθʃeʒ	kɪkbæp
θedʒʃeð	mædkʌs
jɜzjɜθ	bamʃɛz
ʃɜθjɜg	kalfæp

APPENDIX C

PARTICIPANT INFORMATION FORM
PAGE 1
MAURA L. WILSON, GRADUATE STUDENT
LANGUAGE RESEARCH LABORATORY
CLEVELAND STATE UNIVERSITY: DEPARTMENT OF PSYCHOLOGY
CHESTER BUILDING 32
(216) 687-3834

FOR LRL USE:
Room #____________________
Participant #________________
______ (credits) OR $__________
Experiment __________________
Date ____________________________
Experimenter ________________

Please fill in the following information:

Name: ____
*Address: ____
____
E-mail address(es): ____________________________________
____
Telephone Number: ________________ Cell Phone Number: ______________
Date of Birth: ______________________ Place of birth (City):______________
Gender: _________ Major: ____________________
Place of Longest Residence (City): ____________________________
First language spoken: _______________________________________
Are you (circle one): right-handed left-handed ambidextrous
What languages do you speak fluently? _________________________
Would you like to be added to (or remain on) our "Paid Participants Database" so that we can notify you in the future of paid experiments for which you are eligible to participate? ___________________________

* *Note: If you would prefer not to provide your full address and phone number(s), you may simply provide your zip code. Thank you.*

PARTICIPANT INFORMATION FORM
PAGE 2
MAURA L. WILSON, GRADUATE STUDENT
LANGUAGE RESEARCH LABORATORY
CLEVELAND STATE UNIVERSITY: DEPARTMENT OF PSYCHOLOGY
CHESTER BUILDING 32
(216) 687-3834

FOR LRL USE:
Room #__________________
Participant #________________
_____ (credits) OR $__________
Experiment ________________
Date ____________________
Experimenter ______________

Please note that your responses to the following questions will *not* be directly linked to your name. As with any part of your experience as a research participant in our study, please feel free to ask the experimenter if you have any questions. Thank you.

Have you ever had a hearing or speech disorder?
(circle one) YES NO
If yes, please explain: ____________________________________

Have you ever had a visual or reading disorder (other than glasses/contacts)?
(circle one) YES NO
If yes, please explain: ____________________________________

Have you ever been diagnosed with Attention Deficit Disorder (ADD) or Attention Deficit Hyperactivity Disorder (ADHD)?
(circle one) YES NO
If yes, please explain: ____________________________________

APPENDIX D

Edinburgh Handedness Inventory (modified and completed on computer)

You can further help us by providing answers to the following questions. There are no right or wrong answers. Please indicate your preferences in the use of hands in the following activities by answering L for Left hand OR R for Right hand, OR X for No preference. After answering L, R, or X, please answer whether or not you ever use the other hand for each activity by typing Y for Yes OR N for No. Please answer all of the questions. If you have any questions, please ask the experimenter. Please type in your assigned ID number.

Which hand do you write with?
L)Left R) Right X) No Preference

Writing
Do you ever use the other hand?
Y for Yes OR N for No

Which hand do you draw with?
L) Left R) Right X) No Preference

Drawing
Do you ever use the other hand?
Y for Yes OR N for No

Which hand do you throw with?
L)Left R) Right X) No Preference

Throwing
Do you ever use the other hand?
Y for Yes OR N for No

Which hand do you use when using scissors?
L)Left R) Right X) No Preference

Scissors
Do you ever use the other hand?
Y for Yes OR N for No

Which hand do you put your toothbrush in?
L)Left R) Right X) No Preference

Toothbrush
Do you ever use the other hand?
Y for Yes OR N for No

Which hand do you use when using a knife without a fork?
L)Left R) Right X) No Preference

Knife
Do you ever use the other hand?
Y for Yes OR N for No

Which hand do you use when using a spoon?
L)Left R) Right X) No Preference

Spoon
Do you ever use the other hand?
Y for Yes OR N for No

Which hand is your upper hand when using a broom?
L)Left R) Right X) No Preference

Broom
Do you ever use the other hand?
Y for Yes OR N for No

Which hand do you use when striking a match?
L)Left R) Right X) No Preference

Striking a match
Do you ever use the other hand?
Y for Yes OR N for No

Which hand do you use when opening a lid to a box?
L)Left R) Right X) No Preference

Opening a lid to a box
Do you ever use the other hand?
Y for Yes OR N for No

Thank you! Please inform the researcher that you have completed this questionnaire.

APPENDIX E

Race, Ethnicity and Gender Questionnaire (completed on computer)

Your gender is:
Male
Female
x.) Skip

Your ethnic background is:
Hispanic or Latino
Not Hispanic or Latino
x.) Skip

Your racial background is:
American Indian/Alaska Native
Native Hawaiian or Other Pacific Islander
White
Unknown
Asian
Black or African American
More than One Race
x.) Skip

Thank you! Please inform the researcher that you have completed the questionnaire.

APPENDIX F

Participant Consent Form

Maura L. Wilson, Graduate Student: m.l.wilson90@csuohio.edu
Dr. McLennan, Faculty Advisor: c.mclennan@csuohio.edu (216) 687-3750
Language Research Laboratory - Chester Building 32
LANGUAGERESEARCH@MAC.COM (216) 687-3834
http://web.mac.com/languageresearch
Cleveland State University: Department of Psychology

This research project is being conducted as part of Maura Wilson's Master's Thesis under the supervision of Dr. McLennan. If you have any questions about this project, please feel free to contact Ms. Wilson and/or Dr. McLennan at any time (contact information above).

There are two copies of this letter. After signing them, please keep one copy for your records and return the other one. Thank you in advance for your cooperation and support.

"I agree to participate in a perceptual experiment in which I will hear spoken words over headphones. I agree to respond to these sounds by pressing a response button. I also understand that I may be asked to complete a few questionnaires and to participate in a musical listening test. I further understand that confidentiality of my identity will be maintained at all times (i.e., a participant ID code will be assigned to all of my data).

I understand that the procedures to be followed in this experiment have been fully explained to me and that I may ask questions regarding the experiment at the end of the experimental session. I understand the approximate time commitment involved and that I will receive 1.0 research credit for my participation.

I understand that participation in this experiment involves minimal risk beyond those associated with daily living. I further understand that thinking about my answers to some of the questions may make me upset. However, I am also aware that I may choose not to respond to any question that makes me uncomfortable, that I may withdraw at any time without penalty, and that the location and telephone number for Cleveland State University's Counseling Center will be provided to me before I leave the lab today.

I understand that the purpose of this research is to add knowledge to the field of spoken word recognition. I understand that although there may be several indirect benefits of this study, its direct benefit is adding to the current body of knowledge on human perception.

I, the undersigned, am 18 years or older and have read and understood this consent form and hereby agree to give my consent to voluntarily participate in this experiment

I understand that if I have any questions about my rights as a research subject I can contact the Cleveland State University Institutional Review Board at (216) 687-3630.

______________________________ __________________
Signature of Participant *Date*

______________________________ __________________
Name of Participant (PLEASE PRINT) *Date*

APPENDIX G

Post-Experiment Questionnaire (completed on computer)

You can further help us by providing answers to the following questions. There are no right or wrong answers. We are simply interested in your experience in the experiment that you have just participated in and your musical experience. If you have any questions, please ask the experimenter. Please type in your assigned number.

Have you ever learned how to play a musical instrument? If yes, please specify the instrument(s).

How old were you when you began musical training? If you haven't had musical training, please respond "not applicable."

How many years of consistent musical training did you receive, on average? If known, provide the years or ages during which this training occurred (e.g., from 1995 to 2010; from age 10 to 14). If you haven't had musical training, please respond "not applicable."

Approximately how many hours per week did you spend practicing and/or attending group/private lessons for the musical instrument(s) you have trained on? If you haven't had musical training, please respond "not applicable."

Do you ***currently*** play a musical instrument? If yes, please specify the instrument(s).

How many hours per week do you ***currently*** practice and/or attend group or private lessons for the musical instrument(s) you currently play. If you are not involved in musical training/practice currently, please respond "not applicable."

Have you ever trained in any other music-related activities (e.g., voice lessons, dance classes, choir, etc.)? Please describe the activity/activities in as much detail as possible.

At what age did you begin training in the music-related activity/activities? If you haven't had music-related training, please respond "not applicable."

How many years of consistent training have you experienced with this activity? If known, provide the years or ages during which this training occurred (e.g., from 1995 to 2010, from age 10 to 14). If you haven't had music-related training, please respond "not applicable."

Approximately how many hours per week did you spend practicing and/or attending group/private lessons for the music-related activity/activities you have trained on? If you haven't had musical training, please respond "not applicable."

Are you ***currently*** training in any other music-related activities (e.g., voice lessons, dance classes, choir, etc.)? Please describe the activity/activities in as much detail as possible. How many hours per week do you ***currently*** practice and/or attend group or private lessons for the music-related activity/activities you're involved in. If you are not involved in musical training/practice currently, please respond "not applicable."

What do you think was the purpose of this experiment?

Did you have any problem hearing or understanding the words and nonwords you were presented?

Do you have any general comments or observations about the experiment?

Thank you!
Please inform the researcher that you have completed this questionnaire.

APPENDIX H

Neuroticism Questionnaire

On the following pages, there are phrases describing people's behaviors. Please use the rating scale below to describe how accurately each statement describes ***you***. Describe yourself as you generally are now, not as you wish to be in the future. Describe yourself as you honestly see yourself, in relation to other people you know of the same sex as you are, and roughly your same age. So that you can describe yourself in an honest manner, your responses will be kept in absolute confidence. Please read each statement carefully, and then circle the number that corresponds to the number on the following scale:
1: Very Inaccurate
2: Moderately Inaccurate
3: Neither Inaccurate nor Accurate
4: Moderately Accurate
5: Very Accurate

Often feel blue.	1	2	3	4	5
Dislike myself.	1	2	3	4	5
Am often down in the dumps.	1	2	3	4	5
Have frequent mood swings.	1	2	3	4	5
Panic easily.	1	2	3	4	5
Am filled with doubts about things.	1	2	3	4	5
Feel threatened easily.	1	2	3	4	5
Get stressed out easily.	1	2	3	4	5
Fear for the worst.	1	2	3	4	5
Worry about things.	1	2	3	4	5
Seldom feel blue.	1	2	3	4	5
Feel comfortable with myself.	1	2	3	4	5
Rarely get irritated.	1	2	3	4	5
Am not easily bothered by things.	1	2	3	4	5
Am very pleased with myself	1	2	3	4	5
Am relaxed most of the time.	1	2	3	4	5
Seldom get mad.	1	2	3	4	5
Am not easily frustrated.	1	2	3	4	5
Remain calm under pressure.	1	2	3	4	5
Rarely lose my composure.	1	2	3	4	5