2010-08-11

Brain Mapping of the Mismatch Negativity and the P300 Response in Speech and Nonspeech Stimulus Processing

Skylee Simmons Neff
Brigham Young University - Provo

BRAIN MAPPING OF THE MISMATCH NEGATIVITY AND THE P300

RESPONSE IN SPEECH AND NONSPEECH STIMULUS PROCESSING

Skylee S. Neff

A thesis submitted to the faculty
of Brigham Young University
in partial fulfillment of the requirements for the degree of

Masters of Science

David McPherson, Chair
Ron Channell
Christopher Dromey

Department of Communication Disorders

Brigham Young University

December 2010

ABSTRACT

Brain Mapping of the Mismatch Negativity and the P300

Response in Speech and Nonspeech Stimulus Processing

Skylee S. Neff

Department of Communication Disorders

Master of Science

Previous studies have found that behavioral and P300 responses to speech are influenced by linguistic cues in the stimuli. Research has found conflicting data regarding the influence of phonemic characteristics of stimuli in the mismatch negativity (MMN) response. The current investigation is a replication of the study designed by Tampas et al. (2005), which studied the effects of linguistic cues on the MMN response. This current study was designed to determine whether the MMN response is influenced by phonetic or purely acoustic stimuli, and to expand our knowledge of the scalp distribution of processing responses to within- and across-category speech and nonspeech stimuli. The stimuli used in this study consisted of within-category synthetic speech stimuli and corresponding nonspeech frequency glides. Participants consisted of 21 (11 male and 10 female) adults between the ages of 18 and 30 years. A same/different discrimination task was administered to all participants. Data from behavioral responses and event-related potentials (MMN and P300) were recorded. Results provided additional evidence that the MMN response is influenced by linguistic information. MMN responses elicited by the nonspeech contrasts had more negative peak amplitudes and longer latencies than MMN responses elicited by speech contrasts. Brain maps of *t* scores for speech vs. nonspeech contrasts showed significant differences in areas of cognitive processing for all contrast pairs over the left hemisphere near the temporal and parietal areas. The present investigation confirms that there are significant differences in the cortical processing of speech sounds vs. nonspeech sounds.

Keywords: evoked response potentials, mismatch negativity, P300, brain mapping, scalp distribution, speech perception, categorical boundaries

ACKNOWLEDGMENTS

The past two years have been a period of incredible growth and unbelievable challenge. I would have been completely overwhelmed without the loving support of several people. First, I would like to thank my Heavenly Father for all of the blessings I have received throughout my graduate studies. I am grateful to my wonderful husband, Eric, for his confidence in me and for his ability to bolster me up “out of the dumpster”. I am indebted to my parents for their constant encouragement and for always telling me I could accomplish whatever I set my heart on. Thanks to Kalene for making sure my family ate more than frozen vegetables for dinner, and to Ken for being willing to correct my numerous grammatical errors. Without Dr. McPherson, my thesis chair, I never could have overcome the many roadblocks I ran into while completing this daunting project. I am very grateful for his confidence in me. Thanks to Amy Ragsdale and Amanda Fujiki for spending countless hours helping me collect and analyze my data. Finally, I want to say how grateful I am to my baby, Cyrus, for reminding me how precious life is.

Table of Contents

List of Tables

Table Page

List of Figures

Figure Page

List of Appendixes

Introduction

In the study of speech perception it is important to understand how listeners divide running speech into distinct categories of sounds, or phonemes. Authors from diverse fields have investigated the phenomenon of "categorical perception" across multiple sensory modalities. A well-known example of categorical perception is found in how humans perceive color categories. Colors physically differ only in their wavelengths; however, qualitative changes in color are perceived only at certain places along the electromagnetic spectrum: red to orange, yellow to green, and so forth.

Categorical Perception of Speech

The auditory system is very sensitive to changes in acoustic stimuli. For phenomena such as loudness and pitch, changes at any point along the continuum are perceptible to the listener. This is known as continuous perception, in that no categories or qualitative changes are perceived. In contrast to the perception of loudness and pitch differences from low to high in format transition frequencies or voice-onset-times (VOT) are perceived as qualitative or categorical changes only at certain points along the physical acoustic continuum (Harnard, 1987). Previous studies examining categorical perception have observed that speech perception is not dependent on the detection of gradual changes in the acoustic signal, but rather on the detection of changes perceived as discrete categories (Cutting, 1978; Harnad, 1987; Liberman, Harris, Hoffman, & Griffith, 1957). Categorical perception divides a continuum of acoustic stimuli into discrete regions of perceptual phonemic categories. While the stimuli found within categories are acoustically different but perceptually similar, the acoustic differences that exist among stimuli within the same category can be imperceptible to adult listeners. Thus, the boundaries between phonemic categories are important to the perception of speech (Cairns, 1999; Harnad, 1987).

Acoustic signals in running speech have a great deal of variation between and within speakers. Different speakers each produce speech with different fundamental frequencies and the speakers have unique vocal tracts. This means that formant frequencies and transitions vary from speaker to speaker. Likewise, there can be large variations within the same speaker. Rate of speech, intensity and coarticulation can all cause changes in the acoustic characteristics of running speech (Cairns, 1999). The theory of categorical perception of speech implies that listeners respond only to acoustic differences that cross the boundary from one category to another, thus changing the semantics of an incoming signal. Listeners perceptually ignore the acoustic differences that do not affect the overall meaning of the signal. Those changes that span a phonemic boundary yield across-category information, whereas changes that do not affect the semantics represent within-category information (Horev, Most, & Pratt, 2007). Research has supported this theory by suggesting that when listeners process speech signals, they mostly disregard any within-category acoustic differences. Instead, listeners focus on the meaningful across-category information (Fitch, Halwes, Erickson, & Liberman, 1980; Harnad, 1987; Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Liberman et al., 1957). If a listener is presented with two speech-like stimuli which vary only in VOT, one with a VOT of 20 ms and one with a VOT of 40 ms, the listener will perceive a difference. The first stimulus will be perceived as /da/ and the second stimulus will be perceived as /ta/. This is an example of an across-category contrast pair. However, if the listener is presented with two similar stimuli varying only in VOT, one with a VOT of 0 ms and one with a VOT of 20 ms, the listener will not perceive a difference. Both stimuli will be perceived as /da/ (Cairns, 1999). This is an example of a within-category contrast pair. The listener responds to the acoustic difference between the across-category contrast pair because the difference denotes a semantic change in the stimuli. The same degree of acoustic difference between two within-category stimuli is not perceived behaviorally because no meaningful information was conveyed.

Several authors have proposed that the perception of speech is different from the perception of other acoustic input. These authors hypothesize the existence of a phonemic decoder which aids in the perception of speech. This decoder facilitates the way in which we perceive discrete acoustic information by organizing the continuous acoustic stimuli into meaningful phonetic categories of speech. (Bentin & Mann, 1990; Harnad, 1987; Whalen & Liberman, 1987; Whalen & Liberman, 1996). Research has suggested two possible epochs in cognitive processing for phonemic decoding to occur. The first epoch proposed for phonemic decoding is during the acoustic perceptual epoch at approximately 100-300 ms (Whalen & Liberman, 1996). The second epoch that has been suggested as to when phonemic decoding occurs is later in the cognitive perceptual processing epoch at approximately 300-400 ms (Schouten & Hessen, 1992).

Measures of Speech Perception

Behavioral measures of speech processing only provide information related to the end result of the speech perception process. To better understand the process of speech perception in the central auditory nervous system, categorical perception has been examined through event-related potentials (ERPs). ERPs reflect electrical brain activity to sensory, motor, or cognitive events. ERP recordings are time-locked and have a high temporal resolution, making it possible to follow the time course of speech perception. Obtaining behavioral measures in conjunction with ERP recordings allows for the assessment of early versus late cortical processes associated with the temporal discrimination of speech sounds (Horev et al., 2007). Two common ERPs used in categorical perception research are the mismatch negativity (MMN) and the P300 responses (Tampas, Harkrider, & Hedrick, 2005).

MMN. The MMN refers to the central nervous system's ability to compare a deviant stimulus to a previous set of identical stimuli stored in short term memory. When the central nervous system detects a deviant stimulus, an MMN response is elicited (Lang et al., 1995; McPherson & Ballachanda, 2000). MMN responses are considered to be automatic and are therefore not confounded by attention and cognitive factors. The mismatch-generating process does not require attentive discrimination of speech sounds by the participant (Aaltonen, Tuomainen, Laine, and Niemi, 1993). This response occurs whether or not the speech sounds are relevant to the participant's task (Näätänen, 1992). Because the MMN is present regardless of task attention, it can be used to study auditory function in participants who are unable or unwilling to cooperate (Näätänen, 1995). The MMN is independent of conscious perception and tightly links the perception of sound and memory processes (Näätänen, 2001). Several studies have demonstrated that the underlying neural processes contributing to the generation of MMN originate in the auditory cortex (Kraus et al., 1994; Näätänen & Picton, 1987). Studies have also found evidence of a contralateral and biphasic MMN response elicited by deviations in sound location (Tata & Ward, 2005). The MMN waveform appears as a negative deflection in the ERP beginning at about 100 ms and peaking between 200 and 300 ms after the presentation of the stimulus. Good recordings of the MMN can be obtained by placing electrodes over the vertex, frontal area, parietal area, and in certain situations, over the occipital and prefrontal areas (McPherson & Ballachanda, 2000; Näätänen, 1992). Measuring MMN responses can provide useful information regarding speech and language processing at the cortical level of the central nervous system. Therefore, the MMN is often used as a means of investigating the mechanisms of barely perceptible speech contrasts (Sharma & Dorman, 2000).

P300. The P300 response appears when a subject consciously recognizes the occurrence of a change in an acoustic stimulus. The P300 is an endogenous ERP type that is associated with perceptual and cognitive processes which can be affected by attention state (McPherson & Ballachanda, 2000). The same oddball stimulus model that is used to evoke the MMN response can be used to elicit the P300 response. The P300 is seen as a positive potential generated between 200 and 400 ms after stimulus onset and is observed over the centro-parietal areas. It has been shown to reflect the conscious discrimination of a deviant stimulus in a string of identical sounds (McPherson, 1996; Roth, 1983). Discrimination is indicated by having participants press a button or count the presence of deviant stimuli. The cognitive response that is represented by the P300 is based on the relevance assigned to a specific stimulus by the listener (Roth, 1983; Ruchkin & Sutton, 1983).

In contrast, the MMN response represents a neural recognition and discrimination that does not require participants to actively attend to a task and consciously discriminate differences (Dalebout & Stack, 1999). It has been found that the amplitude of the P300 becomes smaller as it becomes more difficult for participants to make a conscious discrimination between stimuli (Picton, Campbell, Baribeau-Braun, & Proulx, 1978; Ruchkin & Sutton, 1983). The amplitude of the P300 becomes more robust when the predictability of a deviant stimulus is low (McPherson, 1996), when value, significance, and task relevance are increased (Ruchkin & Sutton, 1983), and when participant confidence in a threshold signal detection task response is high (Picton et al., 1978). The P300 has been found to exist in response to speech sounds regardless of the probability of the response; this is perhaps due to the high communicative value of speech sounds (Horev et al., 2007). The P300 is thought to be generated from the frontal lobe, the hippocampus, and the auditory cortex (McPherson, 1996; McPherson & Ballachanda, 2000).

ERP and Behavioral Responses

ERP responses have been used to investigate the sequential nature of the processing involved early in the auditory system (Horev et al., 2007; Näätänen, 1995). It has been suggested that the MMN, the P300, and behavioral responses could be used to test the theory of different processing levels in the perception of speech (Dalebout & Stack, 1999; Tampas et al., 2005). Previous studies have suggested that these three responses represent a hierarchy of speech perceptual levels (Pegg & Werker, 1997) for which three levels of speech processing have been proposed.

The first level of speech processing occurs when acoustic stimuli are received and converted into a neural code; the second, the phonetic/phonemic level, is when the neural code is categorized into phonemic contrasts that exist in the phonetic inventory of the native language of adults, and in the broader sensitivity of infants to contrasts in any language. The third level of speech processing involves a higher order linguistic level of processing, and is when syntactic and semantic knowledge facilitate speech perception.

It is thought that the phonetic/phonemic level of speech processing dominates the first acoustic level of processing. This hypothesis of hierarchical levels of speech processing may help explain the difficulty that listeners encounter when attempting to detect acoustic differences within phonemic categories. The MMN has been proposed to represent the acoustic level of processing (Aaltonen, Niemi, Nyrke, & Tuhkanen, 1987; Aaltonen et al., 1993), while the P300 has been suggested to represent the phonetic/phonemic level of processing (Aaltonen et al., 1987). Both the phonetic/phonemic and the linguistic levels of processing may be detected by behavioral responses to deviant stimuli (Dalebout & Stack, 1999).

Dalebout and Stack (1999) suggested that the MMN and P300 may reflect acoustic versus phonetic processing. The authors measured MMN, P300, and behavioral responses to three consonant-vowel (CV) stimulus contrasts on a continuum from /da/ to /ga/. The CV stimulus contrasts consisted of stimulus endpoints, a two-step contrast that straddled each listener's categorical boundary, and a within-category contrast. Their behavioral data were based upon same/different and oddball discrimination tasks. Dalebout and Stack hypothesized that the MMN could be recorded for contrasts that could not be perceived by participants behaviorally, thus suggesting that the MMN reflects the detection of acoustic differences and is not reliant on conscious phonemic discrimination. For consonant contrasts that were on the endpoints of the continuum, listeners demonstrated better discrimination behaviorally. In 9 of 12 participants the MMN responses were present during the stimulus contrast, and in all 12 participants the P300 responses were present. This provides support for the link between categorical perception and the P300. For the second stimulus contrast (two-step contrast), in 10 of 12 participants the MMN responses were present; only 4 participants had P300 responses. Finally, for the most difficult phoneme perception condition (within-category), an MMN from a stimulus that could not be detected as dissimilar behaviorally was exhibited in 10 of the 12 participants. No participants exhibited a P300 response from a stimulus that could not be discriminated behaviorally. These results support the theory that the MMN represents detection of unconscious acoustic differences in stimuli. Dalebout and Stack concluded that their results were consistent with the lower level detection of acoustic signal differences by MMN and higher level cognitive categorical perception of phonemic differences by the P300 and behavioral responses.

Several other studies have used across- versus within-category speech stimuli to compare MMN and behavioral labeling responses (Ylinen, Shestakova, Houtilainen, Alku, & Näätänen, 2006), MMN and P300 responses (Aaltonen et al., 1987), and MMN, P300 and behavioral labeling responses (Maiste, Wiens, Hunt, Scherg, & Picton, 1995; Sams, Aulanko, Aaltonen, &

Näätänen, 1990). These studies have found similar results to Dalebout and Stack's (1999) research. In these studies, generally the MMN responses were present during all CV stimulus contrasts with no significant differences between phonetic categories. The P300 responses were different for within- and across-category stimulus contrasts. More robust P300 responses were found in the perceptually different across-category contrasts. Behavioral labeling responses were more accurate for endpoint CV stimuli. These results support the notion of a speech discrimination model based on different levels of perception.

However, studies have found results that contradict the suggestion that the MMN response is based only on acoustic stimuli (Cheour et al., 1998; Näätänen et al., 1997; Peltola et al., 2003; Sharma & Dorman, 2000; Winkler et al., 1999). Aaltonen, Eerola, Lang, Uusipaikka, and Tuomainen (1994) found significant differences in the amplitude of MMN responses for vowel vs. pure tone stimuli, but found no significant differences in the latency of MMN responses for vowel vs. pure tone stimuli. Recent research in speech perception across languages has suggested that the MMN reflects deviation in both acoustic and phonetic speech stimuli. Näätänen et al. (1997) investigated the MMN responses of Finnish and Estonian speaking participants when presented with a control standard stimulus sound, perceived as /e/ by speakers of both languages, and a deviant stimulus, perceived as /õ/ by the Estonian speakers but not perceived as a unique phoneme by Finnish speakers. To control for the effect of deviant stimuli, an additional deviant stimulus was presented to the participants which was perceived as /ö/ by both speakers. Näätänen et al. found that the unfamiliar /õ/ resulted in a smaller MMN response than the familiar deviant /ö/ in Finnish speakers. There was no significant difference reported in the MMN responses in Estonian speakers. These results suggest that a phonetic representation of the deviant sound was involved in creating the MMN response. In addition, when investigating

MMN responses in patients with aphasia, Becker and Reinvang (2007) found significant left hemisphere prevalence for the speech sound MMN responses of normal participants; this prevalence was not observed in the aphasic participants of Becker and Reinvang's study. The difference in MMN responses between the control group and patients with aphasia suggests that the MMN response is influenced by a subject's current language abilities.

Additional studies have produced similar findings to those of Näätänen et al. (1997) when evaluating consonant contrasts across languages (Peltola et al., 2003; Sharma & Dorman, 2000; Winkler et al., 1999). Sharma and Dorman (2000) examined MMN and behavioral responses in native and nonnative phonetic categories for speakers of English and Hindi. The participants were presented with CV pairs that were saliently different in Hindi but not in English. The Hindi participants had better discrimination and more robust MMN responses than the English participants. These results present further evidence that the MMN may signify processing beyond a purely acoustic level. The findings from cross-linguistic studies indicate that MMN responses may be partially affected by linguistic experience. The idea that language-specific categorical perception may influence speech perception at the MMN level of processing conflicts with the hypothesis that the MMN reflects a purely acoustic level of discrimination.

A review of literature suggests that the MMN has the ability to index participant discrimination of across-category consonant contrasts, as MMN responses are consistently present in across-category discrimination tasks. Reports regarding the ability of MMN to index discrimination of within-category consonant contrasts have often had conflicting results. These discrepancies may be due to differences in the way that the MMN responses were acquired. Because there are no set criteria on the most effective MMN data collection methods, different distraction techniques have been used while recording the MMN such as reading a book, watching movies with sound, or watching movies with subtitles (Pettigrew, et al., 2004). In addition, previous studies used both synthetic and natural speech stimuli to elicit MMN responses. It is more difficult to control for acoustic differences between stimuli when using spoken speech. Recording differences may account for some of the variation in the MMN responses found in previous studies, making it difficult to draw definitive conclusions about the role of the MMN as a purely acoustic response. To address this topic, Tampas et al. (2005) designed a study aimed at reconciling the conflicting results reported in the literature and finding more conclusive evidence regarding the influence of linguistic information on the MMN response. To investigate speech vs. nonspeech processing in the same group of participants, Tampas et al. compared MMN responses to within-category consonant vowel speech contrasts with frequency glides whose frequencies matched the formant transitions of the CV stimuli. Tampas et al. hypothesized that if the MMN responses to within-category CV speech pairs were absent or smaller than the MMN responses to acoustically matched frequency glide pairs, then the MMN response could be said to be influenced by categorical perception. If MMN responses to within-category CV speech contrasts were similar to the MMN responses generated by the acoustically analogous frequency glides, then MMN responses may reflect purely acoustic and not phonemic differences. The second aim of the study by Tampas et al. was to expand and integrate the information currently available regarding the neurophysiology underlying categorical perception.

Tampas et al. (2005) used the MMN, the P300, and behavioral responses to further study the hypothesis of hierarchical processing levels (acoustic, phonetic/phonemic, linguistic). They hypothesized that if the MMN responses to speech vs. nonspeech stimuli did not differ and the P300 responses did differ, the results would support the hypothesis that the MMN and the P300 represent different levels of speech processing. If, however, the MMN and the P300 responses both differ for speech vs. nonspeech stimuli, then it may be that parallel processing for speech occurs at the levels of the MMN and the P300 neural generators.

In the study by Tampas et al. (2005) it was found that both the MMN and the P300 responses for speech stimuli were smaller than the responses for nonspeech stimuli. The authors concluded that the MMN is influenced by linguistic information as well as acoustic information, and that speech processing can be represented using a parallel model at the levels of the MMN and the P300 neural generators. These findings contradict the hypothesis of hierarchical levels of speech processing proposed by Dalebout and Stack (1999).

Current Investigation

The current investigation is a replication of the study designed by Tampas et al. (2005), with some technical modifications. Participants were presented with a continuum of within- and across- category speech CV stimuli (/ba/ vs. /da/) and frequency glide stimuli. The pure tone transitions of the frequency glide stimuli matched the F2 transitions of the CV stimuli. In the current study 32 silver-silver chloride electrodes were used to record ERP data, as recommended by McPherson & Ballachanda (2000). When recording same/different behavioral responses, both hits and false alarms were measured. Hit and false alarm ratios were created to summarize the performance of each participant on the discrimination task. The purpose of this current investigation was based on the aims developed by Tampas et al.

The current investigation studied whether the scalp distribution of the MMN is influenced by linguistic information (phonetic categories based on place of articulation) or whether the scalp distribution of the MMN reflects discrimination of purely acoustic differences. This investigation also provides a scalp distribution of any significant differences found between MMN responses to synthetically generated speech-like and nonspeech stimuli. The scalp distribution provides data regarding the location of any differences in the MMN response at designated response latencies. In addition, the current investigation expands the existing literature regarding the neurophysiology underlying categorical perception.

Method

The current methodology was based on that used by Tampas et al., (2005). However, in the current investigation 32 silver-silver chloride electrodes were used in the recording of ERP, whereas Tampas et al. used only two electrodes to record ERP data. Also, in the current study both hits and false alarms were recorded while Tampas et al. only recorded hits.

Participants

Participants consisted of 21 (11 male and 10 female) adults between the ages of 18 and 30 years. Each participant passed a hearing screening test at 15 dB HL at octave frequencies from 250 Hz to 4000 Hz in both ears. Each participant had static acoustic admittance between 0.3 and 1.4 mmhos with peak pressure between –100 and +50 daPa in the test ear (ASHA, 1990; Roup, Wiley, Safady, & Stoppenbach, 1998). All participants spoke English as their first language. Each participant was right handed (Knecht et al., 2000) and had no reported neurological, cognitive, or learning impairments. In addition, all participants were screened for the ability to discriminate both CV stimuli when presented with a clear /ba/ and a clear /da/ and matching frequency glide stimuli. Those who were unable to pass the initial screening were excluded from the study. Each participant read and signed an informed consent document approved by the Institutional Review Board at Brigham Young University (see Appendix A) before participating in the study. Three participants selected for this study were tested a second time in order to establish test-retest reliability.

Instrumentation

A Grason-Stadler model GSI-10 audiometer was used to perform hearing screenings and a Grason-Stadler model GSI-33 impedance meter was used to perform tympanograms. The ERPs were collected with a NeuroScan computer using Scan 4.0 software. The raw electrical potentials were band pass filtered between DC and 500 Hz. A GSI-10 audiometer was used to present stimuli through insert earphones. During testing, participants were seated comfortably in a double-walled, sound treated test booth. The ambient noise in the booth did not exceed ANSI S3.1-1991 maximum permissible levels for air conduction testing with ears uncovered when all electronic equipment was operating.

Both the speech and nonspeech stimuli consisted of monaural files presented binaurally. Stimuli were presented at an RMS amplitude of 80 dB SPL through Etymotic EA-3 insert earphones in the four computerized same/different discrimination tasks. During the electrophysiological testing, silver-silver chloride electrodes were placed over the scalp at 32 electrode positions according to the 10-20 International System (Jasper, 1958) using an electrode cap (Electrocap International). Electrode impedances were kept below 3000 ohms. Eye movements were observed by placing electrodes above the supra-orbital foramen of one eye and on the outer cantha of the opposite eye. During computer averaging of the responses, epochs occurring within the epoch of eye movement were rejected from the resulting average.

Nine speech-like stimuli (CV) and nine nonspeech stimuli, consisting of frequency glides with matching pure tone transitions, were synthetically generated using software programs (HLsyn Version 2.2-Build 6 and Adobe Audition 3 Version 3-Build 7283.0). The stimuli were randomized before presentation to participants using NeuroScan Stim 4.2 software. The stimuli were presented binaurally at 75 dB SPL. Results were analyzed using Microsoft Excel and Sigma Plot 5.0 (SPSS) software.

Stimuli

Stimuli were selected from a nine-item, synthetically generated continuum of CVs or frequency glides. The synthetically generation of stimuli ensured they were as acoustically similar as possible by permitting control over the standard and deviant stimuli. Due to this similarity, reliable comparisons could be made between waveforms resulting from the speech stimuli and waveforms resulting from the nonspeech stimuli for the first 100 ms of the stimuli (where the differences in the stimuli are located). In previous studies, strong MMN and P300 responses to synthetically generated stop-consonant contrasts were reported (Kraus, McGee, Sharma, Carrell, & Nicol, 1992; Sharma & Dorman, 1999).

All current CV stimuli were synthetically generated using the /a/ vowel for the steady state. Stimulus one was perceptually the most /ba/ like and stimulus nine was perceptually the most /da/ like. The total duration of each individual stimulus was 600 ms, with the first 100 ms consisting of formants F1 and F2 transitioning linearly and the last 500 ms consisting of a vocalic steady-state. The nine different stimuli differed in the onset frequency of F2 which varied from 900 Hz to 1700 Hz (Figure 1). All nine stimuli reached a F2 offset frequency of 1250 Hz after 100 ms. The F1 for all stimuli had an onset frequency of 150 Hz and rose to a steady state value of 750 Hz in 100 ms. The vocalic steady-state values for /a/ during the last 500 ms of the stimuli were F1 = 750 Hz, F2 = 1250 Hz, F3 = 2400 Hz, and F4 = 3300 Hz. Speech stimuli were output from a digital-to-analog converter with a low-pass filter of 4.8 kHz, sent to an attenuator and a headphone buffer, and then routed to insert earphones.

Nonspeech frequency glide stimuli were generated using a 44,100 Hz sampling rate with a 16-bit resolution. Each stimulus consisted of two overlapping pure tone segments: the first segment consisted of a 100 ms linear frequency transition glide and the second segment consisted of a 500 ms steady state tone of 1250 Hz. The two segments were seamlessly integrated into one frequency glide by overlapping the segments by 1 ms. The onset of each glide was equal to the F2 onset frequency for the formant transitions in the speech-like CV stimuli (Figure 1). The frequency glide stimuli were output from a digital-to-analog converter, sent to an attenuator and headphone buffer, and then routed to insert headphones located inside the sound booth.

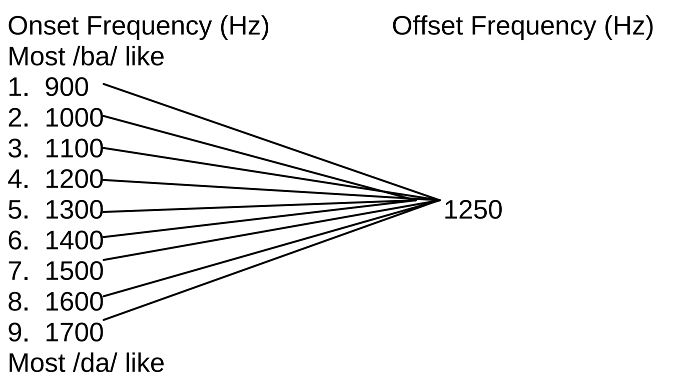

Figure 1. F2 transitions of the nine stimuli. Stimuli 1, 2 and 3 are most /ba/ like, stimuli 7, 8, and 9 are most /da/ like. This figure illustrates the onset and offset frequencies for speech and nonspeech (frequency glide) stimuli.

The stimuli were grouped into seven 1610 ms deviant trials consisting of two-step stimulus contrast pairs from the nine synthetically generated CV stimuli and the nine corresponding frequency glides (i.e., 1 vs. 3, 2 vs. 4, 3 vs. 5, 4 vs. 6, 5 vs. 7, 6 vs. 8, and 7 vs. 9). Each trial began with 105 ms of silence and had an interstimulus interval of 305 ms. Tampas et al. (2005) presented only the 2 vs. 4 speech and nonspeech stimulus contrast pairs to participants. In the current investigation behavioral and ERP data were collected on all seven speech and nonspeech stimulus contrast pairs.

Nine corresponding common stimulus pairs were also generated (i.e., 1 vs. 1, 2 vs. 2, 3 vs. 3…8 vs. 8, 9 vs. 9). Each common trial began with 105 ms of silence and had an interstimulus interval of 305 ms, consistent with the previously mentioned deviant trial contrast pairs.

Procedures

Initial screening. All participants initially went through a screening process to be sure that each participant had the ability to discriminate between stimulus one and stimulus nine. This process included a computerized labeling task in which participants were first introduced to two common trial pairs and one deviant trail pair (1 vs. 1, 9 vs. 9, and 1 vs. 9). The participants were then told to label the contrast pairs as *same* or *different*, as the pairs were presented in a randomized order with no cueing from the investigator. Participants who were unable to correctly label both the speech and nonspeech sample stimuli trial pairs as “same” or “different” were excluded from the study.

Behavioral data acquisition. After the initial screening process each participant completed four computerized same/different discrimination tasks (two consisting of 945 speech-like stimulus trial pairs each and two consisting of 945 nonspeech stimulus trial pairs each). The order of the tasks presented was randomized for each participant. In these tasks participants were presented with each two-step stimulus contrast pair from the nine synthetically generated CV stimuli and the nine corresponding frequency glides (i.e., 1 vs. 3, 2 vs. 4, 3 vs. 5, 4 vs. 6, 5 vs. 7, 6 vs. 8, and 7 vs. 9). Per participant, 270 trials of each stimulus contrast pair consisting of 40 deviant and 230 common stimulus pairs were randomly presented with a probability of occurrence of .85 and .15 respectively. To minimize test-retest effects, no feedback was provided. Participants indicated whether the pair of stimuli was the same or different by pushing a button following which subsequent trials were immediately presented. For each sample, the P300 and MMN responses were measured concurrently with behavioral responses.

Event related potential acquisition. Two types of ERPs (P300 and MMN) were measured using an electrode cap consisting of 32 electrodes. The order of presentation of each stimulus contrast pair was randomized to avoid long term habituation to the stimuli and to decrease predictability of the stimuli (McPherson, 1996). During each recording session, the participants were seated comfortably in an armchair with the neck well supported. The participants were instructed to sit quietly, but not sleep. Throughout the ERP recordings, participants were instructed to press a button in response to the deviant stimuli to promote active attention state during the recording of the ERP responses (McPherson & Ballachanda, 2000). Due to technical difficulties, behavioral and ERP data were not collected for the 5 vs. 7 speech contrast pair.

Behavioral data analysis. For the same/different discrimination task, the ratio of hits and false alarms for each of the deviant and common contrast pairs presented was measured. The ratios were used as the dependent variables for computation of a one-way analysis of variance (ANOVA). The independent variable was stimulus type (two levels: speech and nonspeech).

Event related potential data analysis. The ERP data elicited for each of the 270 trials of each common and deviant stimulus contrast pair was averaged for each participant, resulting in one individual averaged ERP file for each of the contrast pairs per participant. The individual averaged ERP responses for each participant were further averaged to obtain one grand averaged ERP file for each of the contrast pairs presented.

Individual averaged ERP responses and grand averaged ERP responses were analyzed. For the MMN responses, a difference ERP file was obtained by subtracting the deviant ERP average from the common average created for each contrast pair (Näätänen & Winkler, 1999).

The FCZ electrode was analyzed in detail to further investigate the MMN responses. For all participants, peak latency of the MMN was defined as the maximum point of negativity following the N1 component of the FCZ waveform between 80 ms and 400 ms. The P300 response peak latency was defined as the maximum point of positivity greater than 5 μV following the P2 component between 200 and 500 ms. Appropriate statistical power was expected because the sampling rates and filters were equivalent for MMN and P300 recordings and the same interval significance was used during analysis of the different types of ERP data (Kraus et al., 1995).

Results

General findings showed that there were no significant differences between the female (10) and male (11) participants for the brain mapping or the behavioral results. In addition, results for the speech contrast pair 5 vs. 7 and comparisons between the speech and nonspeech contrasts for pair 5 vs. 7 are not available due to missing data for that comparison.

Behavioral Data

Hits, misses, correct rejects, and false alarms were recorded for all speech contrast pairs presented to the participants. Discrimination ratios for the hits were obtained by dividing the total number of hits by the total number of deviant stimulus pairs presented. Discrimination ratios for the false alarms were then obtained by dividing the total number of false alarms by the total number of common stimulus pairs presented. Descriptive statistics for hit and false alarm ratios can be found in Table 1.

The participants displayed higher hit ratios for speech contrasts across categorical boundaries (4 vs. 6, and 6 vs. 8) than for speech contrasts within categorical boundaries (1 vs. 3, 2 vs. 4, and 7 vs. 9). Likewise, higher hit ratios were observed for all nonspeech contrasts than for speech contrasts within categorical boundaries. A one-way ANOVA was computed to determine if there was a difference in behavioral responses due to stimulus type for hits and false alarms. A main effect for hits was found for stimulus type, $F(12, 260) = 28.33$, $p \leq .000$, indicating a significantly higher hit ratio for nonspeech vs. speech contrasts. There was no significant difference for false alarm ratio based on stimulus type. A Bonfferoni post hoc analysis showed significant differences ($p \leq .05$) between the hit ratios for speech contrast pairs across categorical boundaries (4 vs. 6) and speech contrast pairs within categorical boundaries (1 vs. 3, 2 vs. 4, 3 vs. 5, and 7 vs. 9). Likewise, significant differences ($p \leq .05$) between hit ratios for speech contrasts within categorical boundaries,1 vs. 3, 2 vs. 4, 7 vs. 9, and all nonspeech frequency glide contrasts were observed.

Table 1

Mean Hit and False Alarm (in parenthesis) Ratios for Discrimination Task

Contrast Pair	*N*	*M*	*SD*	*SE*
Speech				
1 vs. 3	21	.11 (.04)	.13 (.06)	.03 (.01)
2 vs. 4	21	.09 (.06)	.11 (.09)	.02 (.02)
3 vs. 5	21	.21 (.07)	.17 (.09)	.04 (.02)
4 vs. 6	21	.43 (.08)	.20 (.10)	.04 (.02)
6 vs. 8	21	.37 (.16)	.17 (.13)	.04 (.03)
7 vs. 9	21	.13 (.13)	.14 (.12)	.03 (.03)
Subtotal	126	.22 (.09)	.20 (.10)	.02 (.01)
Nonspeech				
1 vs. 3	21	.65 (.07)	.13 (.09)	.03 (.02)
2 vs. 4	21	.73 (.09)	.18 (.13)	.04 (.03)
3 vs. 5	21	.67 (.11)	.19 (.13)	.04 (.03)
4 vs. 6	21	.52 (.12)	.19 (.16)	.04 (.03)
5 vs. 7	21	.46 (.11)	.18 (.15)	.04 (.03)
6 vs. 8	21	.37 (.10)	.25 (.16)	.05 (.03)
7 vs. 9	21	.39 (.11)	.27 (.16)	.06 (.04)
Subtotal	147	.54 (.10)	.24 (.14)	.02 (.01)
Total	273	.39 (.09)	.27 (.12)	.02 (.01)

Event Related Potential Data

Figures 2 through 7 display waveforms for each of the speech and nonspeech standard and deviant contrast pairs. Waveforms were created by averaging all of the individual averaged waveforms into one grand averaged waveform using the CPZ electrode (McPherson, 1996). Below the standard and deviant stimulus waveforms for each contrast pair are the MMN responses generated from the difference waveforms for each of the contrast pairs. MMN waveforms were generated using the FCZ electrode (McPherson, 1996). MMN responses were defined as the maximum point of negativity between 80 ms and 400 ms post stimulus (Tampas et al., 2005). The P300 responses were defined as present if the waveform displayed a maximum positive peak over 5 μV following N2 between 200-500 ms (Tampas et al., 2005).

MMN responses were present for all speech and nonspeech contrast pairs for individual and grand average waveforms and are highlighted in gray boxes in Figures 2 through 6. Descriptive statistics for the MMN peak latencies and amplitudes are included in Table 2 and Table 3, respectively.

Significant differences were found in the peak amplitude and latency of the MMN responses for individual averaged ERP waveforms between speech and nonspeech contrasts. A one-way ANOVA was computed for both peak amplitude and latency measures to determine if there was a difference due to stimulus type for the individual averaged ERP waveforms. For peak amplitudes, a main effect was found for stimulus type, $F(1, 513) = 32.43$, $p \leq .000$, indicating significantly more negative peak amplitudes for the nonspeech contrast responses than for the speech contrast responses. For peak latencies, a main effect of stimulus type was found, $F(1, 513) = 8.64$, $p = .003$, indicating that peak latencies were significantly longer for responses elicited by nonspeech than speech contrasts.

Table 2

Mean MMN Latency (ms) for Stimulus Types and Contrast Pairs

Contrast Pair	*N*	*M*	*SD*	*SE*
Speech				
1 vs. 3	40	215.23	81.41	12.87
2 vs. 4	37	242.14	72.58	11.93
3 vs. 5	38	232.55	81.73	13.26
4 vs. 6	38	233.71	87.03	14.12
6 vs. 8	35	226.46	80.21	13.59
7 vs. 9	38	260.12	88.05	14.28
Subtotal	226	234.94	82.34	5.48
Nonspeech				
1 vs. 3	42	271.98	98.02	15.12
2 vs. 4	41	270.48	81.86	12.79
3 vs. 5	41	260.17	97.97	15.30
4 vs. 6	44	270.64	80.69	12.16
5 vs. 7	38	234.63	72.04	11.69
6 vs. 8	43	234.26	87.83	13.40
7 vs. 9	40	255.63	81.88	12.95
Subtotal	289	257.38	86.81	5.11
Total	515	247.38	85.51	3.77

Table 3

Mean MMN Amplitude (μV) for Stimulus Types and Contrast Pairs

Contrast Pair	*N*	*M*	*SD*	*SE*
Speech				
1 vs. 3	40	-3.18	5.08	.80
1 vs. 4	37	-1.87	2.33	.38
3 vs. 5	38	-3.20	1.90	.31
4 vs. 6	38	-3.83	4.09	.66
6 vs. 8	35	-2.82	2.08	.35
7 vs. 9	38	-3.37	4.23	.69
Subtotal	226	-3.22	3.53	.23
Nonspeech				
1 vs. 3	42	-6.34	8.75	1.35
2 vs. 4	41	-6.94	9.70	1.52
3 vs. 5	41	-6.44	7.59	1.19
4 vs. 6	44	-6.21	6.78	1.02
5 vs. 7	38	-5.37	4.87	.79
6 vs. 8	43	-8.16	9.43	1.44
7 vs. 9	40	-5.03	5.64	.89
Subtotal	289	-6.38	7.75	.46
Total	515	-4.99	6.45	.28

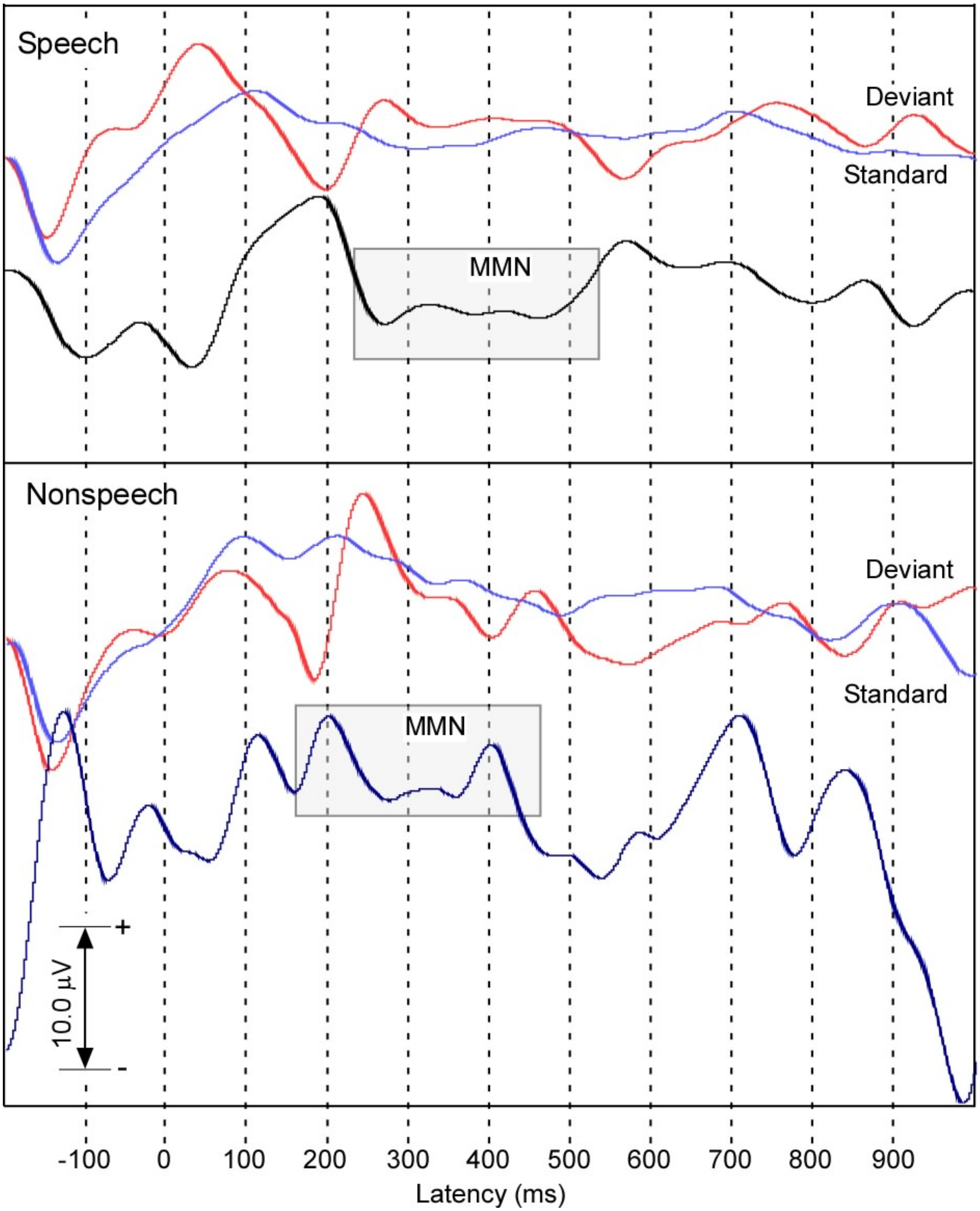

Figure 2. Contrast pair 1 vs. 3 grand averaged waveforms of speech vs. nonspeech ERP and MMN responses. Red indicates responses to deviant stimulus pairs; blue indicates responses to standard stimulus pairs.

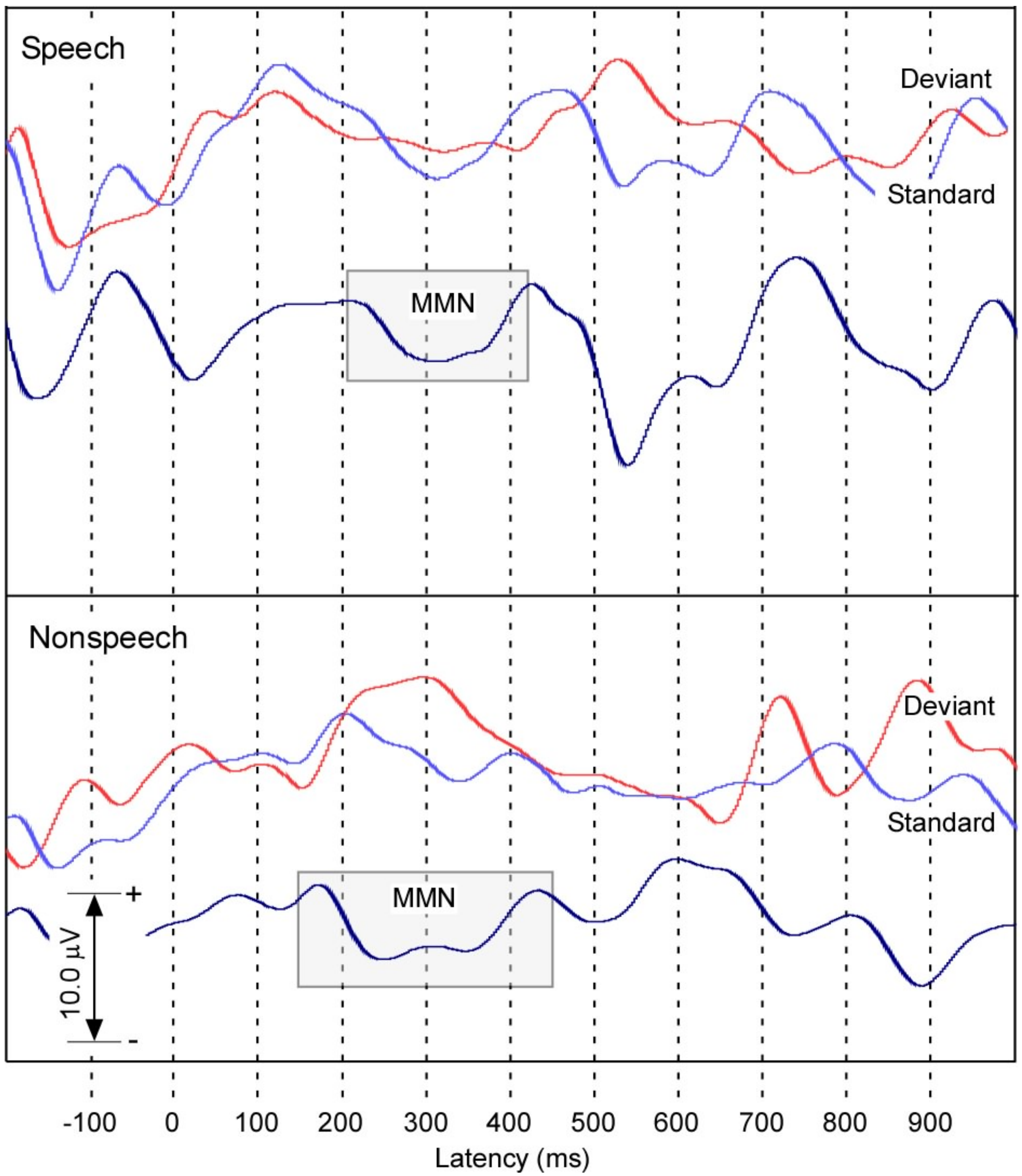

Figure 3. Contrast pair 2 vs. 4 grand averaged waveforms of speech vs. nonspeech ERP and MMN responses. Red indicates responses to deviant stimulus pairs; blue indicates responses to standard stimulus pairs.

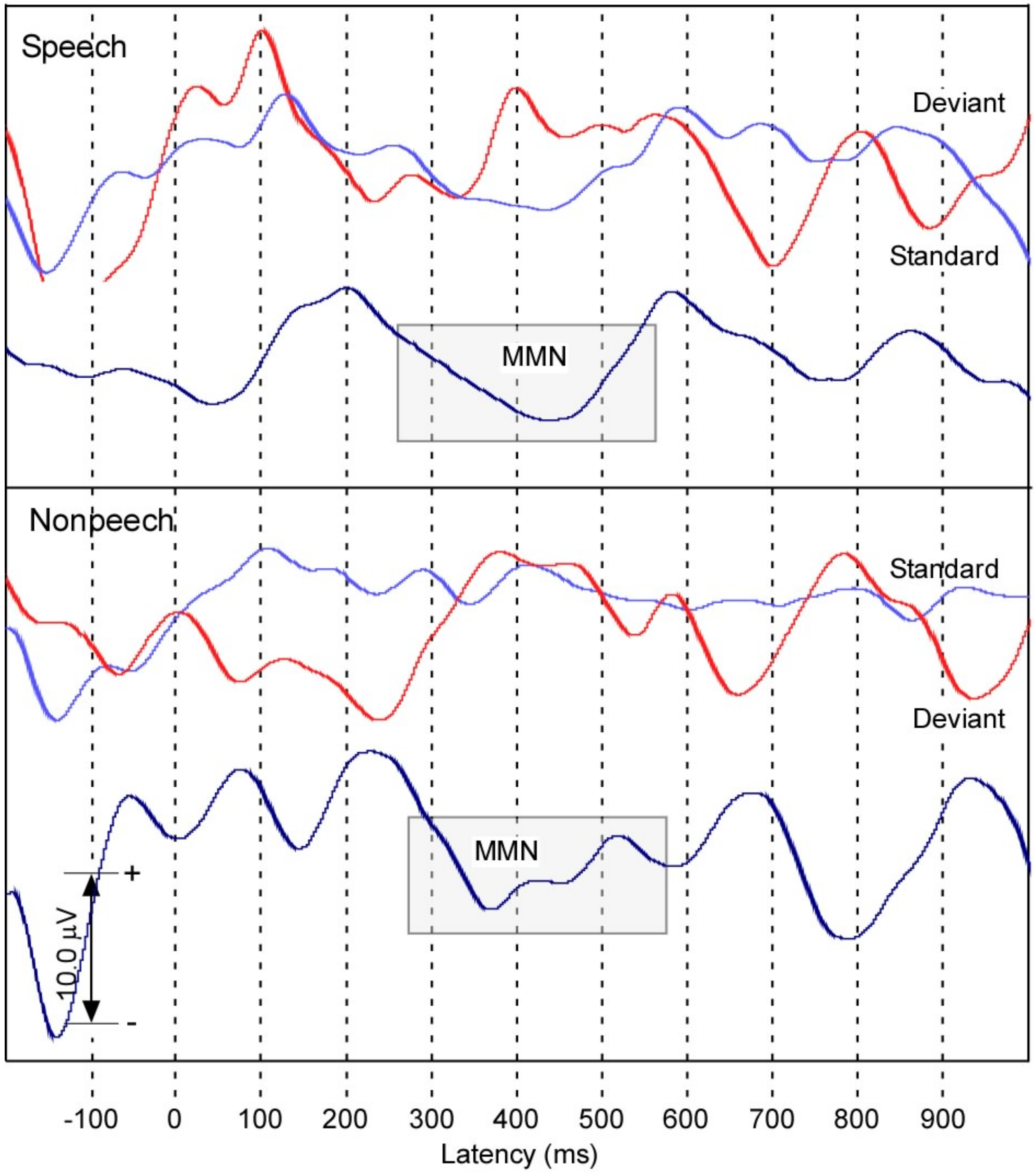

Figure 4. Contrast pair 3 vs. 5 grand averaged waveforms of speech vs. nonspeech ERP and MMN responses. Red indicates responses to deviant stimulus pairs; blue indicates responses to standard stimulus pairs.

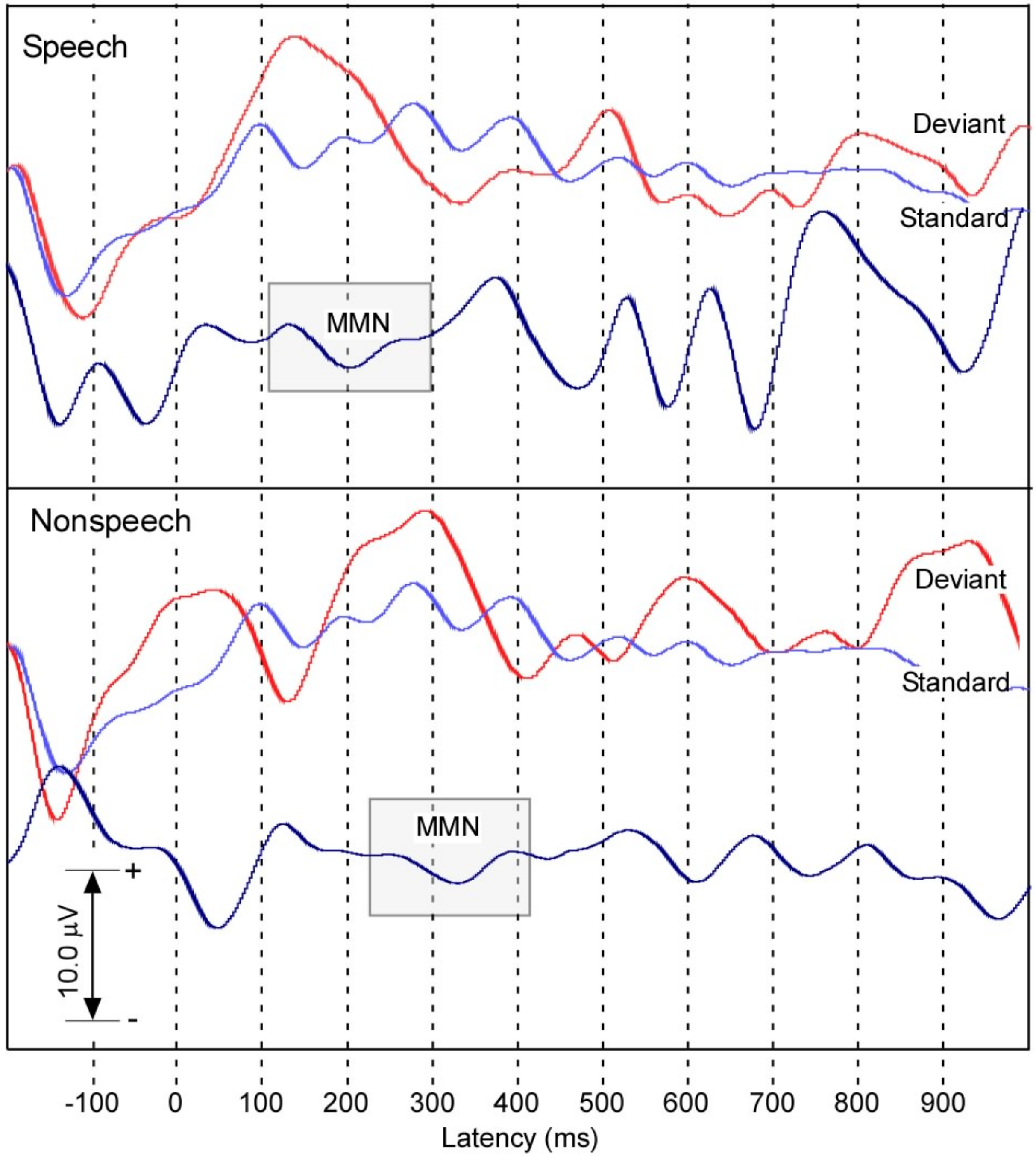

Figure 5. Contrast pair 4 vs. 6 grand averaged waveforms of speech vs. nonspeech ERP and MMN responses. Red indicates responses to deviant stimulus pairs; blue indicates responses to standard stimulus pairs.

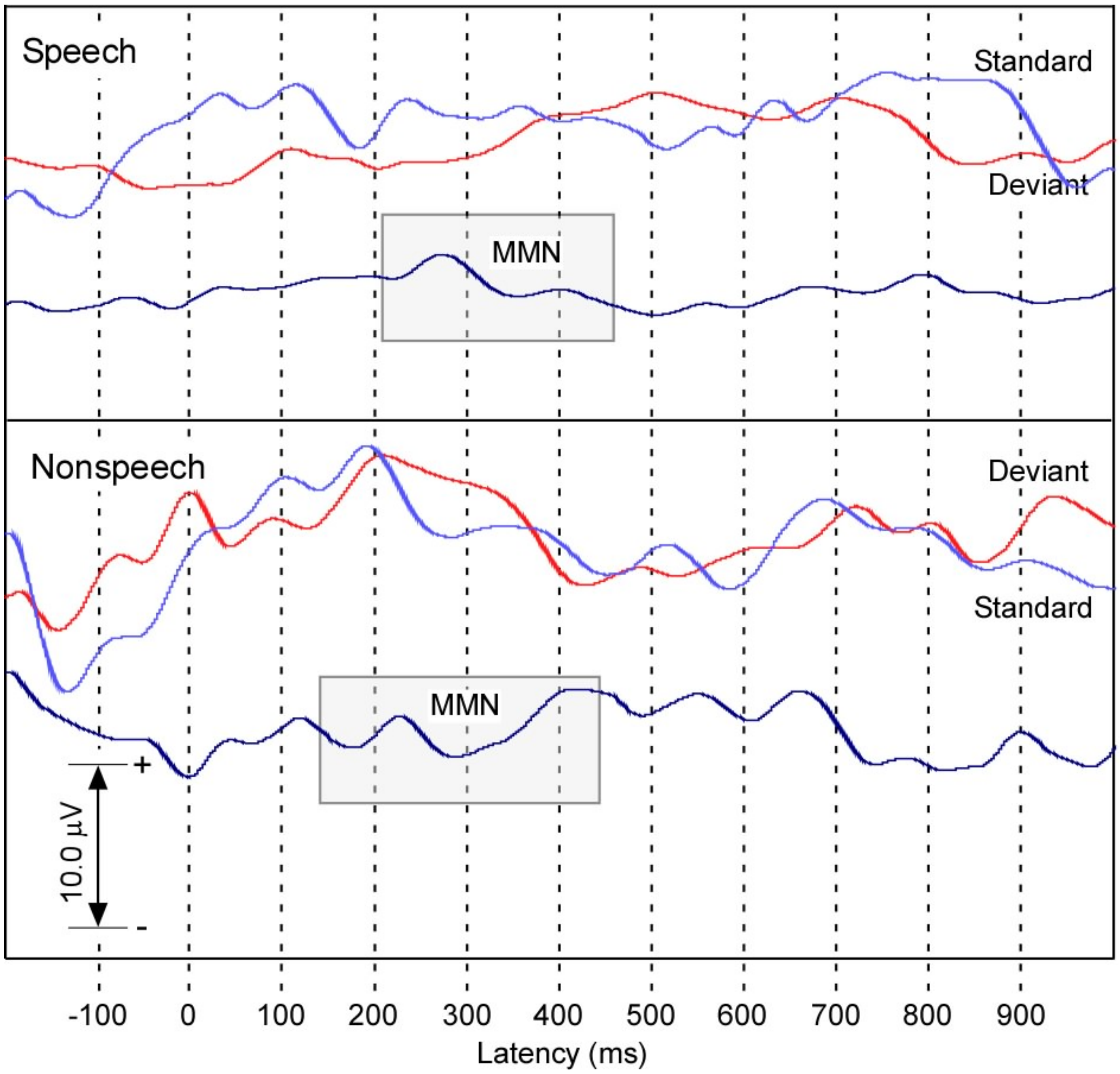

Figure 6. Contrast pair 6 vs. 8 grand averaged waveforms of speech vs. nonspeech ERP and MMN responses. Red indicates responses to deviant stimulus pairs; blue indicates responses to standard stimulus pairs.

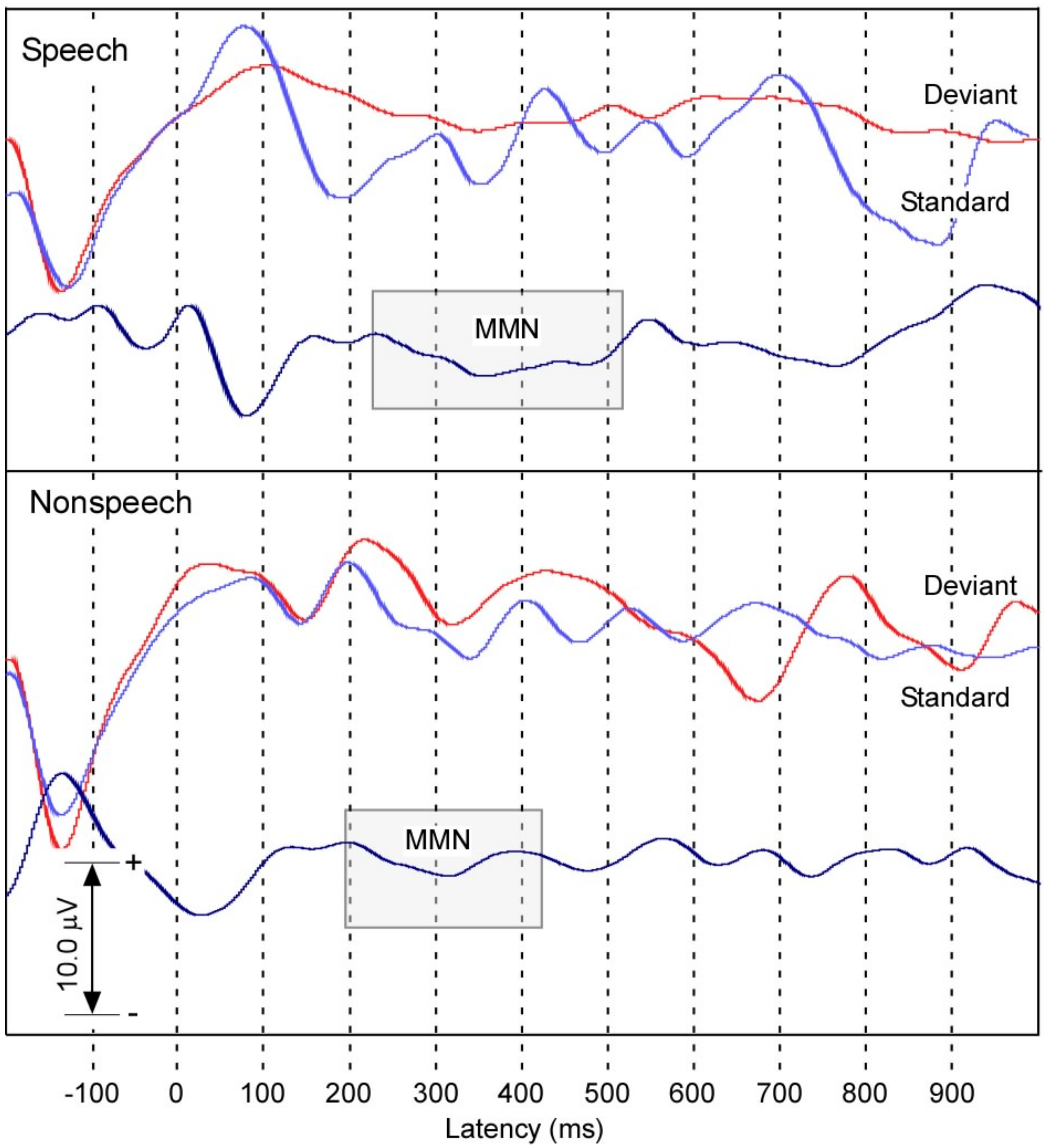

Figure 7. Contrast pair 7 vs. 9 grand averaged waveforms of speech vs. nonspeech ERP and MMN responses. Red indicates responses to deviant stimulus pairs; blue indicates responses to standard stimulus pairs.

A Bonfferoni post hoc analysis showed significant differences ($p \leq .05$) in MMN peak amplitude for the individual averaged ERP waveforms between the following contrast pairs: speech contrast pairs 1 vs. 3, 2 vs. 4, 3 vs. 5, and 6 vs. 8, and nonspeech contrast pair 6 vs. 8 (Figure 6). A Bonfferoni post hoc analysis revealed no significant differences between contrast pairs for MMN peak latency for the individual averaged ERP waveforms.

A *t* test ($df = 20$) comparing each speech contrast grand averaged ERP with its accompanying nonspeech grand averaged ERP was completed. Brain maps of these comparisons were created showing areas of significant differences in the scalp distribution of the participants' processing of speech vs. nonspeech stimuli during the first 1000 ms of stimulus processing (Figures 8 through 13).

Significant differences in the scalp distribution were noted over the anterior parietal and occipital areas during the late perceptual processing epochs (200-300 ms) in contrast pair 1 vs. 3 (Figure 8). Significant differences in cognitive processing were also noted during the linguistic processing epochs (500-600 ms) over the language centers in the left hemisphere. Differences were noted over the occipital and parietal areas during the late linguistic processing epochs (700-800 ms). The differences observed in the late linguistic epochs were accompanied by significant differences in processing over the frontal area indicating executive processing of the stimuli (700-800 ms).

Contrast pair 2 vs. 4 (Figure 9) showed greater differences between speech and nonspeech stimulus processing than contrast pair 1 vs. 3. Significant differences were noted over the frontal and temporal areas of the left hemisphere as well as the parietal and occipital areas of the right hemisphere during the perceptual processing epochs (100-300 ms). Significant differences over the temporal area of the left hemisphere were noted during the early cognitive processing epochs (300-400 ms). Differences in scalp distribution over the left frontal and temporal areas were also noted later in the cognitive process, during the linguistic processing epochs (500-600 ms).

Significant differences in the scalp distribution were noted for the contrast pair 3 vs. 5 (Figure 10) over the frontal, anterior parietal and occipital areas of the right and left hemispheres during the early cognitive processing stages (100-200 ms). Differences over the frontal area occurred early (200-300 ms) and then ebbed during the linguistic processing epoch (400-600 ms). Differences over the parietal and occipital areas continued throughout the P300 response window (300-400 ms) and were seen again during the linguistic processing time frames (500-600 ms).

Processing differences over the temporal, parietal, and occipital areas during the acoustic stages of processing (0-100 ms) were seen in contrast pair 4 vs. 6 (Figure 11). Likewise, significant differences in the scalp distribution over the temporal areas of both the right and the left hemispheres were seen in the linguistic processing epoch (400-600 ms). Contrast pair 4 vs. 6 had less significant differences in processing during the early cognitive processing epoch (200-400 ms) than other contrast pairs.

Contrast pair 6 vs. 8 (Figure 12) shows significant processing differences in the scalp distribution in the processing of speech and nonspeech contrasts during both early and late cognitive processing. Significant differences were found over the frontal, temporal, parietal, and occipital areas before and during the MMN response window (100-400 ms). These differences were also noted during the linguistic processing epoch (500-600 ms). Also, differences in cognitive processing persisted over the temporal and occipital areas throughout the linguistic processing epoch (500-900 ms).

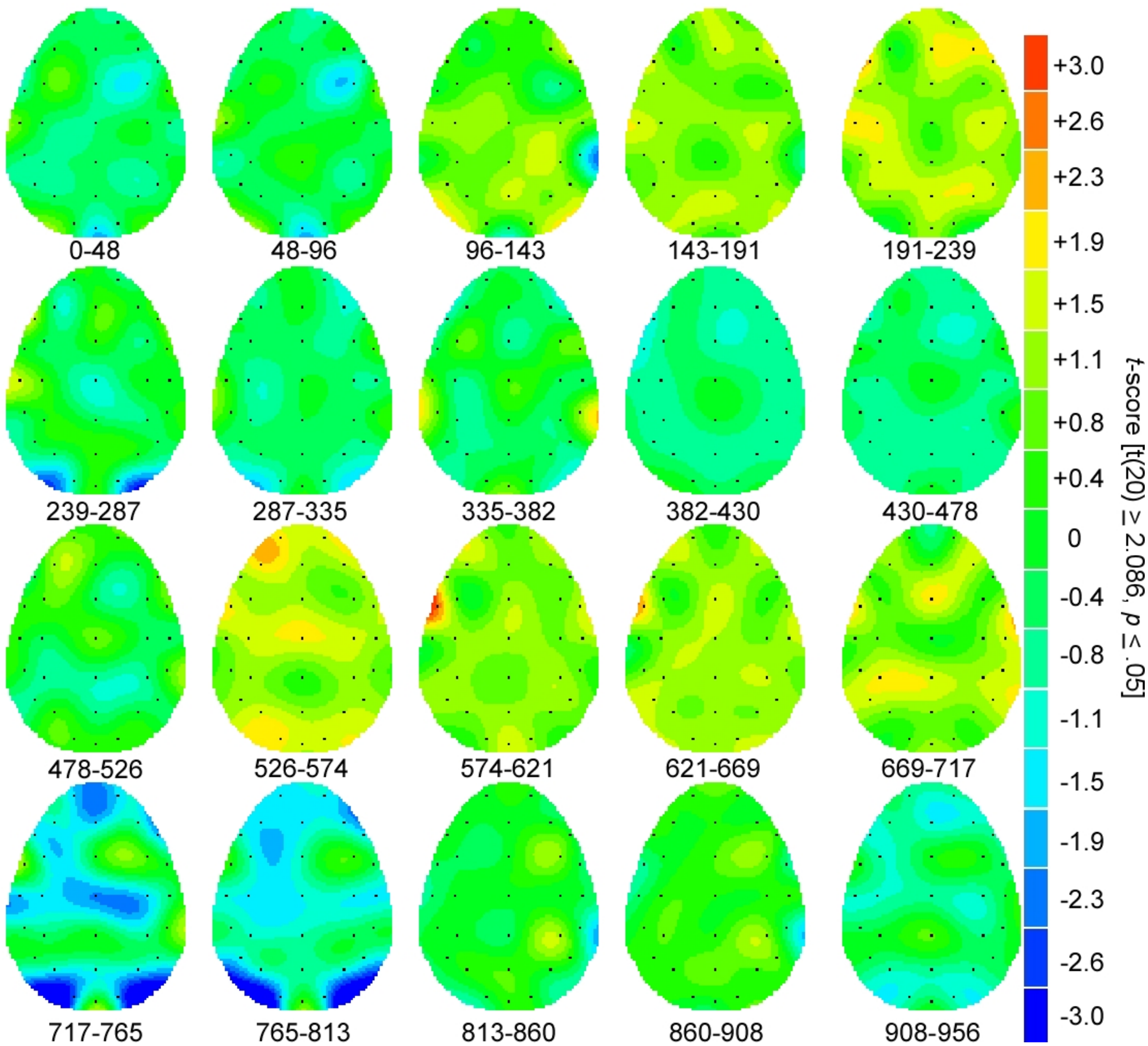

Figure 8. Contrast pair 1 vs. 3 grand averaged brain maps of speech vs. nonspeech ERPs. Areas of significant difference ($p \leq .05$) for are color coded orange/red for $t(20) \geq 2.3$ and blue for $t(20) \leq -2.3$.

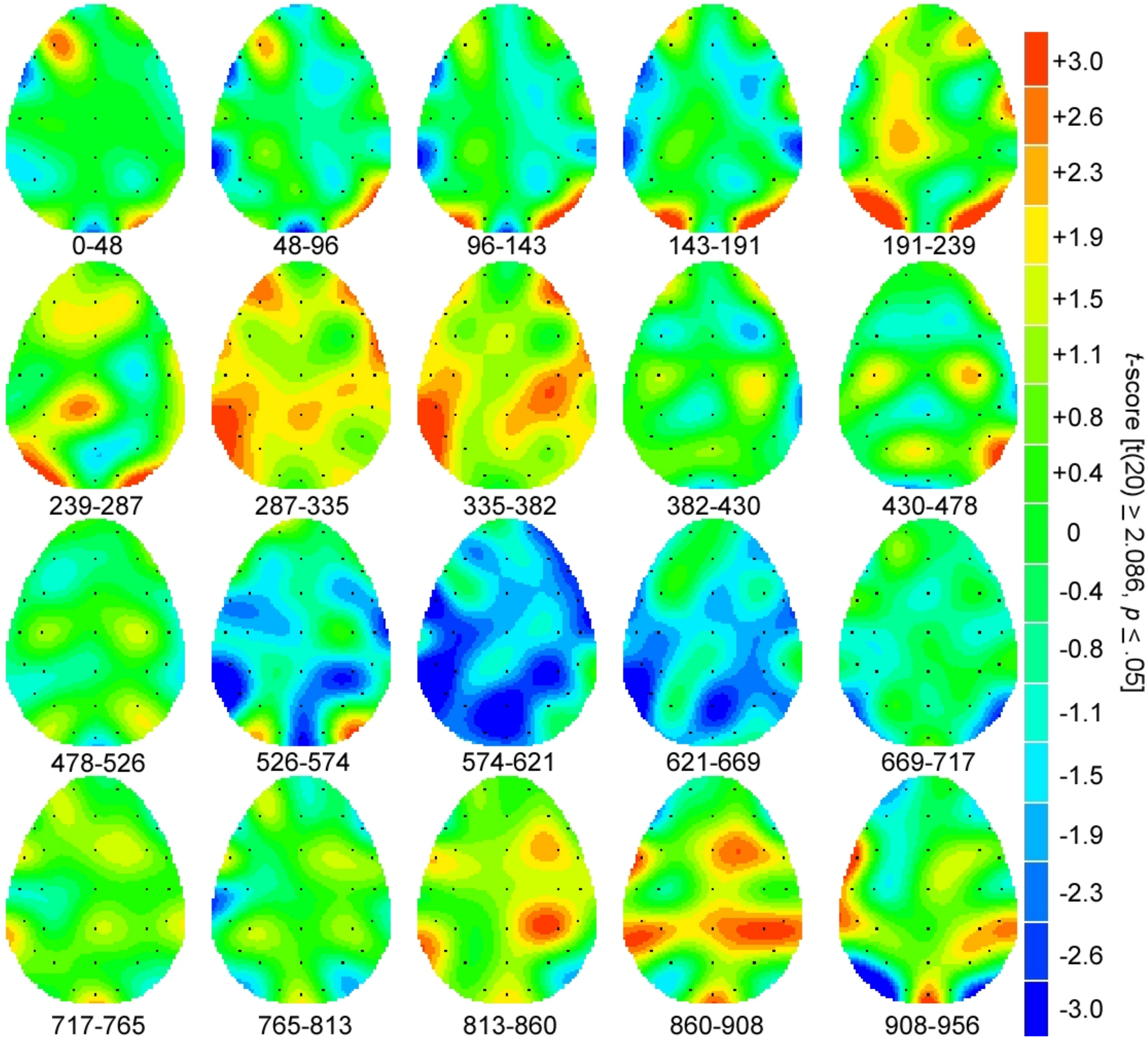

Figure 9. Contrast pair 2 vs. 4 grand averaged brain maps of speech vs. nonspeech ERPs. Areas of significant difference ($p \leq .05$) for are color coded orange/red for $t(20) \geq 2.3$ and blue for $t(20) \leq -2.3$.

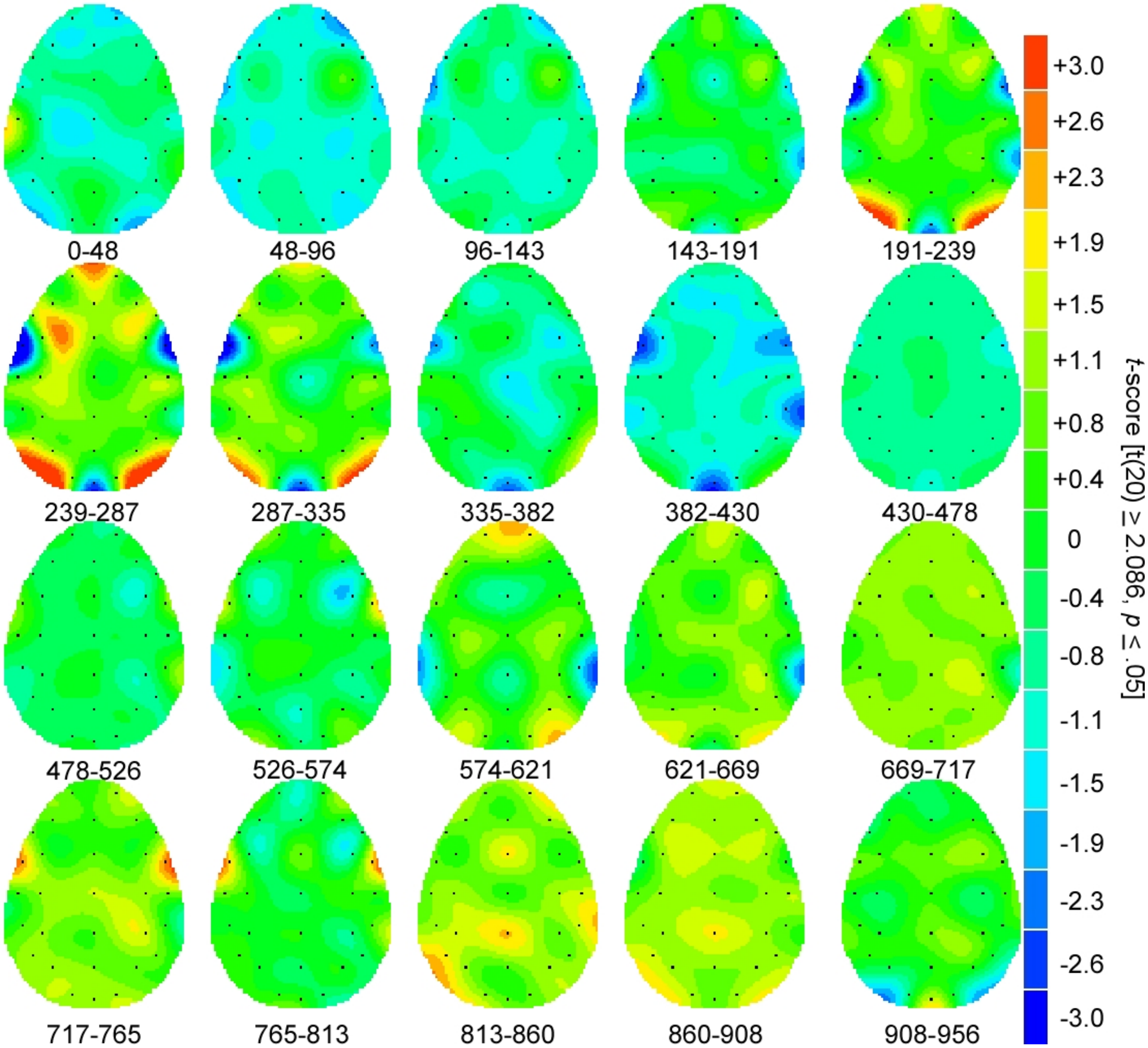

Figure 10. Contrast pair 3 vs. 5 grand averaged brain maps of speech vs. nonspeech ERPs. Areas of significant difference ($p \leq .05$) for are color coded orange/red for $t(20) \geq 2.3$ and blue for $t(20) \leq -2.3$.

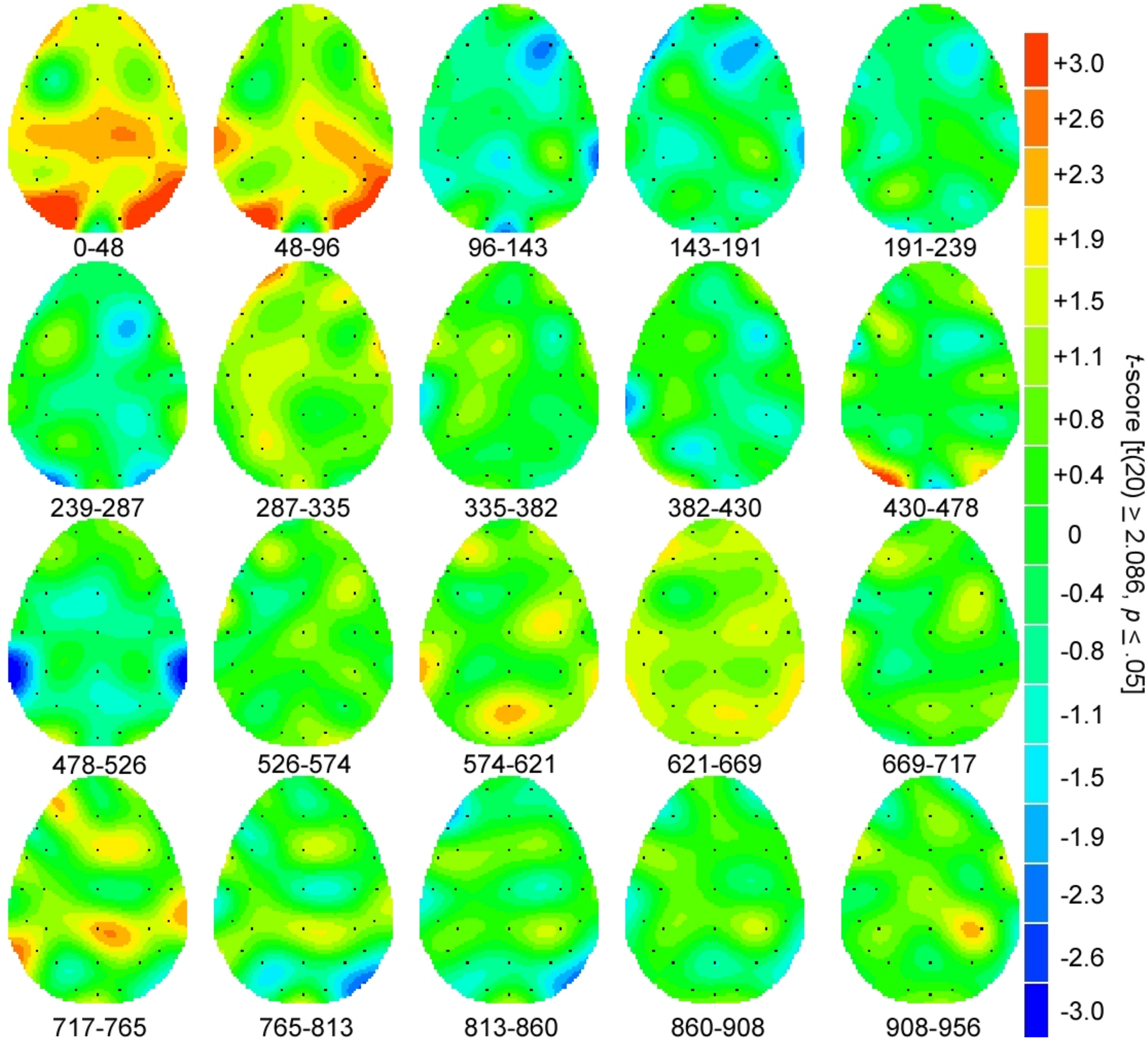

Figure 11. Contrast pair 4 vs. 6 grand averaged brain maps of speech vs. nonspeech ERPs. Areas of significant difference ($p \leq .05$) for are color coded orange/red for $t(20) \geq 2.3$ and blue for $t(20) \leq -2.3$.

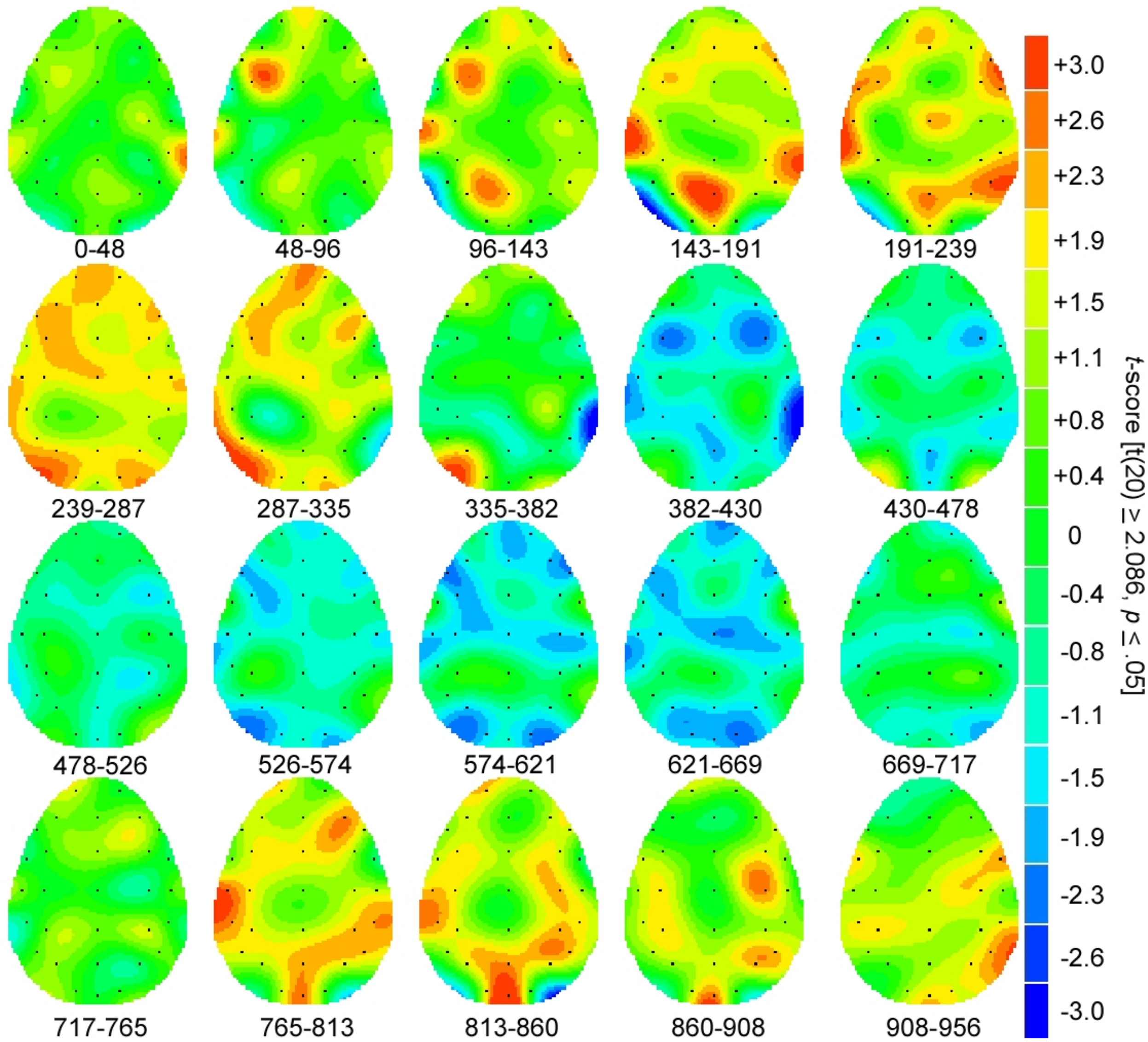

Figure 12. Contrast pair 6 vs. 8 grand averaged brain maps of speech vs. nonspeech ERPs. Areas of significant difference ($p \leq .05$) for are color coded orange/red for $t(20) \geq 2.3$ and blue for $t(20) \leq -2.3$.

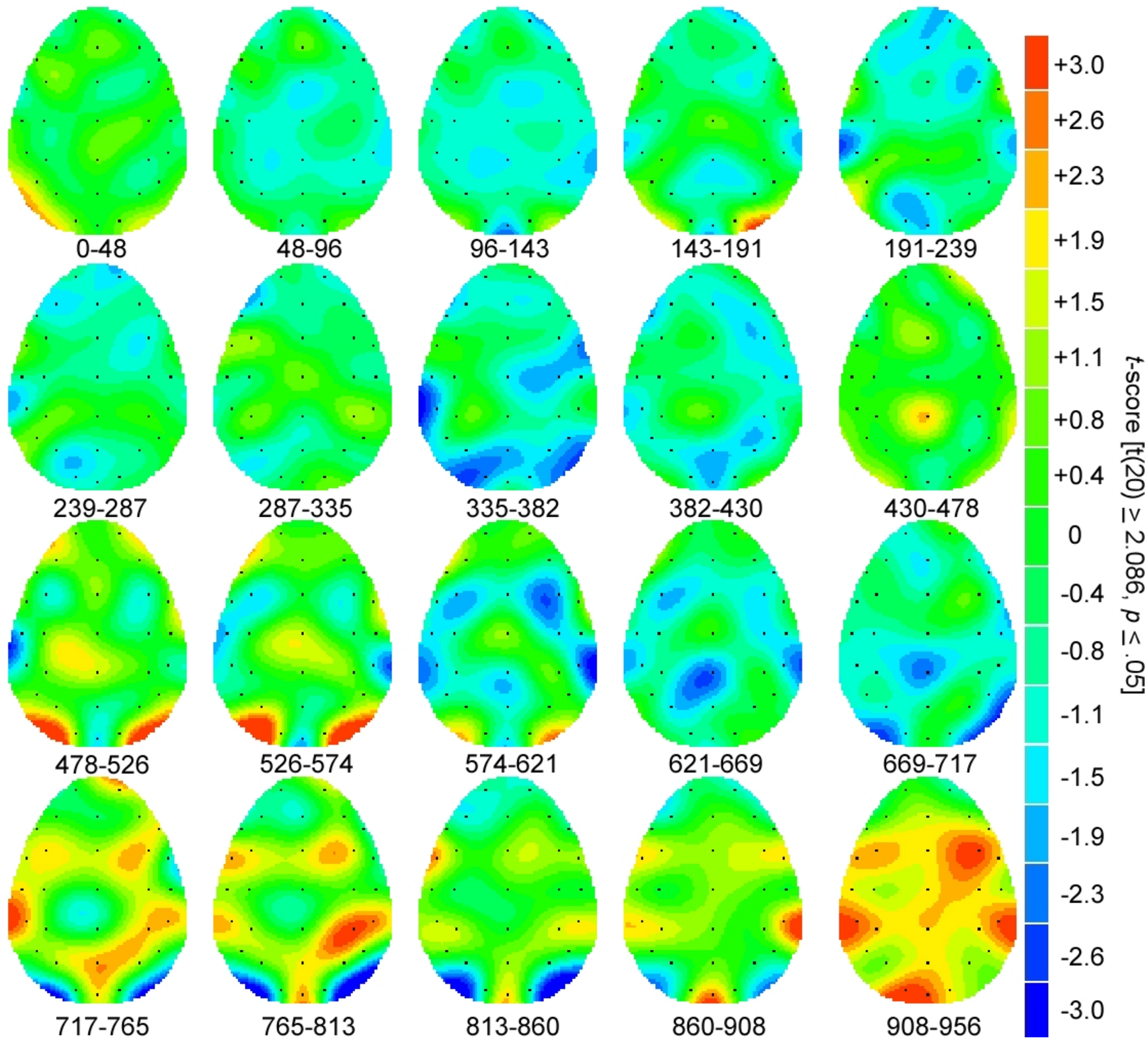

Figure 13. Contrast pair 7 vs. 9 grand averaged brain maps of speech vs. nonspeech ERPs. Areas of significant difference ($p \leq .05$) for are color coded orange/red for $t(20) \geq 2.3$ and blue for $t(20) \leq -2.3$.

Contrast pair 7 vs. 9 (Figure 13) showed significant differences in the scalp distribution over the occipital, left temporal, and right fontal areas during the P300 response window (300-400 ms). Significant differences over the occipital and temporal areas occurred intermittently throughout the linguistic and late processing epoch (500-900 ms).

In the grand averaged ERP waveforms, the P300 responses were present for all nonspeech contrast pairs (Figures 2 through 6) suggesting that participants recognized differences between the contrasting nonspeech stimuli. The deviant waveforms for nonspeech contrast pairs 2 vs. 4 (Figure 3), 3 vs. 5 (Figure 4), and 4 vs. 6 (Figure 5) are examples of robust P300 responses. These results contrasted with the grand averaged ERP responses for speech contrast pairs where no significant P300 responses were present (Figures 2 through 6).

An analysis of the individual averaged ERP responses using a chi-square test for independent samples was completed to determine if there were significant differences between the presence and absence of the P300 responses for speech vs. nonspeech contrast type. Significant chi-square values, $\chi^2(1, N = 4) \geq 3.8, p \leq .05$, were observed for the presence of P300 responses between speech and nonspeech contrasts for the contrast pairs 1 vs. 3, 2 vs. 4, 4 vs. 6, and 7 vs. 9.

Repeated Measures

Three participants selected for this study were tested a second time to determine re-test reliability. Individual and grand averaged ERP waveforms were created for each re-test subject. A *t* test was completed on the grand averaged ERP brain maps comparing each subject’s initial data with the re-test data for the 1 vs. 3 stimulus contrast pair for speech and nonspeech. All three participants for both stimuli types failed to show significant differences between repeated measures. This indicates a reasonable level of re-test reliability for the current investigation.

Discussion

The aims of the current investigation were to study the influence of linguistic cues on the MMN response and to provide a scalp distribution of any significant differences found between MMN responses to synthetically generated speech-like and nonspeech stimuli. The results of the current investigation support and expand the existing literature regarding the neurophysiology underlying categorical perception. The results of the present investigation also provide a basis for future applications.

Comparison of Results to Current Research

The influence of categorical boundaries on speech perception was noted in the same/different discrimination task for speech contrasts. Significant differences ($p \leq .05$) were found between speech contrast pairs across categorical boundaries and speech contrasts within categorical boundaries. Similarly, significant differences ($p \leq .05$) between speech contrasts within categorical boundaries and all nonspeech frequency glide contrasts were observed. Acoustic differences within categorical boundaries were more difficult for participants to detect than acoustic differences of the same magnitude across categorical boundaries. These data are in agreement with previous research regarding categorical boundaries affecting the processing of speech sounds (Cairns, 1999; Harnad, 1987). Specifically, the research supports the theory that acoustically deviant speech sounds within a certain phonemic category are processed as the same sound and therefore are difficult to discriminate; meanwhile speech sounds across categorical boundaries are processed as different sounds and are easier to discriminate. It was also found that participants had greater difficulty in perceiving acoustic differences between within-category speech contrasts vs. those same acoustic differences between nonspeech stimulus contrasts. The behavioral response findings between speech and nonspeech contrasts support previous research regarding the difference between speech and nonspeech executive processing (Bentin, & Mann, 1990; Cairns, 1999). The findings of this study provide support to the theory that small acoustic differences between nonspeech sounds such as tones are easier to discriminate than similar acoustic differences between two speech sounds of the same phonemic category.

The differences in processing observed in the behavioral responses for speech vs. nonspeech contrasts were likewise found in the ERPs. The P300 responses were present for all nonspeech contrast pair grand averaged ERP responses. In contrast, no significant P300 responses were present for any of the speech contrast pair grand averaged ERP responses or common stimulus pair grand averaged ERP responses. This is consistent with the behavioral data where participants showed an increased ability to discriminate differences among nonspeech contrast pairs vs. speech contrast pairs. Previous research has suggested that the P300 can be used to predict behavioral responses and conscious decision making (Aaltonen et al., 1987) . Because subjects demonstrated increased hit ratios for nonspeech contrast pairs, it follows that the ERP responses for those same contrast pairs would show the presence of a P300 response. It is interesting to note that the increased hit ratios for the speech contrasts across categorical boundaries (4 vs. 6, 6 vs. 8) did not correspond with the presence of P300 responses for those same contrasts. Previous research indicates that high hit ratios for speech contrasts across categorical boundaries would be accompanied by the presence of a P300 (McPherson, 1996; Roth, 1983). It may be that higher behavioral hit ratios in the across-category speech stimulus pairs are required to consistently evoke a P300 response across a grand average of participants (Picton, et al., 1978; Ruchkin & Sutton, 1983). P300 responses were seen in 12.5% and 13.3% of the individual averaged waveforms for speech contrast pairs 4 vs. 6 and 6 vs. 8 respectively. However, when the individual responses were averaged together the P300 responses were no longer visible. Because creating a grand average requires fairly rigid time-locked responses, the individual variances due to shifts in processing time may have obscured the actual presence of the P300 in the grand average. The low percentage of P300 responses present for the across-category speech contrast pairs may have been due to individual variation in cognitive processing times.

MMN responses were present in all grand averaged ERP waveforms for all of the contrast pairs presented to participants. This suggests that the MMN responses are influenced by acoustic differences that are not consciously recognized in higher cognitive processing supporting previous research in MMN responses to deviant stimuli (Näätänen, 1995). However, MMN responses showed significant differences for speech vs. nonspeech stimuli in peak amplitude and latency. MMN responses for speech contrasts had significantly lower amplitudes than nonspeech contrasts. This suggests that participants had greater difficulty detecting differences in speech contrasts, a conclusion also supported by the P300 and behavioral responses observed for the speech contrasts. MMN responses for speech contrasts also had significantly shorter latencies than nonspeech contrasts. This phenomenon might be explained by the phonetic theory of speech perception; specifically, that the brain creates coded categorical boundaries for speech phonemes to quicken processing time. Brain maps of *t* scores for speech vs. nonspeech contrasts showed significant differences in areas of cognitive processing for all contrast pairs. Areas of greatest processing differences during the MMN response epoch were noted over the left hemisphere near the temporal and parietal areas.

Näätänen (2001) discussed the possibility of using the MMN to locate language-specific phoneme traces. These traces were located primarily in the posterior portion of the left hemisphere's auditory cortex, near Wernicke's area. The present investigation confirms that there are areas where significant differences are found in the cortical processing of speech sounds vs. nonspeech sounds. These areas are generally over the posterior portion of the auditory cortex. The results of the current study suggest that it is possible to use MMN measures to provide additional evidence that while the speech sound traces are located in and around Wernicke's area, nonspeech sound traces are not found near Wernicke's area. Tampas et al. (2005) found significant differences in the peak amplitude and latency of MMN responses for speech and nonspeech contrast stimuli. The results of the present study are in agreement with those of Tampas et. al. and support the theory that the MMN response is influenced not only by acoustic signals but by linguistic cues as well. This suggests that parallel processing takes place at the level of the MMN and P300 generators and contrasts the theory that the MMN is influenced exclusively by auditory cues. Aaltonen et al. (1987) suggested the existence of a hybrid multi-level model of speech processing. It was proposed that the MMN might represent acoustic processing of stimuli and that categorical representations were not processed until later as reflected by the P300. Dalebout & Stack (1999) found similar results concerning the influence of acoustic cues on the MMN. The present study, along with the results of Näätänen (2001) and Tampas et al. (2005), have shown that parallel processing of acoustic cues occurs early in the cognitive processing of the MMN generators, or at least during that same epoch of other generators. The results of the present investigation lend further support to the theory of parallel processing much earlier in the auditory system than was previously thought.

The results of this study support previous evidence suggesting that the MMN is not purely an acoustic response. The present study suggests that the MMN is influenced by linguistic information as well as acoustic information and that the brain imposes categorical boundaries on speech sounds as early in the auditory process as the MMN generators (100-300 ms).

Future Applications

Findings from the current investigation may have significant theoretical and clinical applications. For example, further research is required to analyze whether any significant differences exist for P300 response amplitude and latency for speech vs. nonspeech contrast pair responses. The MMN provided information regarding the perceptual processing of differences in stimuli. Significant differences were found over the temporal, parietal, and occipital areas of the scalp during the acoustic and cognitive perceptual processing epochs (100-300 ms). Although it is similar to the MMN, the P300 provides additional information regarding the cognitive processing differences between two stimuli. The P300 is reflected by perceptual and cognitive processes and is evoked by the conscious discrimination of a deviant stimulus. The P300 is influenced by the relevance assigned to a specific stimulus (McPherson & Ballachanda, 2000; Roth, 1983; Ruchkin & Sutton, 1983). It would be interesting to know whether the differences noted in the MMN are carried into the cognitive processing of speech vs. nonspeech stimuli. Aaltonen et al. (1987) had asked as to which processing changes were reflected by the P300 if categorical perception is already completed by that stage of auditory processing. The current investigation supports research suggesting that categorical boundaries influence the MMN response (Becker and Reinvang, 2007; Näätänen et al., 1997; Tampas, et al., 2005). However, the extent to which those same linguistic cues influence the P300 still remains unknown.

In addition, research regarding differences between normal and participants with different types and extents of hearing loss would be beneficial. This information would assist in expanding the existing literature on deviations in ERPs based on differences in the central nervous system. Clinically, the normative patterns involved in speech vs. nonspeech processing found in the present study could be compared with those of individuals with known neural processing disorders. Deviations in the subject's ERPs could help the clinician create a more complete differential diagnosis. Likewise, monitoring ERPs throughout progressive therapy sessions could create an objective means to assess the effectiveness of auditory training programs.

References

Aaltonen, O., Eerola, O., Lang, A. H., Uusipaikka, E., & Tuomainen, J. (1994). Automatic discrimination of phonetically relevant and irrelevant vowel parameters as reflected by mismatch negativity. *The Journal of Acoustical Society of America, 96*(3), 1489-1493. doi:10.1121/1.410291

Aaltonen, O., Niemi, P., Nyrke, T., & Tuhkanen, M. (1987). Event-related brain potentials and the perception of a phonetic continuum. *Biological Psychology, 24*(3), 197-207. doi:10.1016/0301-0511(87)90002-0

Aaltonen, O., Tuomainene, J., Laine, M., & Niemi, P. (1993). Cortical differences in tonal versus vowel processing as revealed by an ERP component called mismatch negativity (MMN). *Brain and Language 44*(2), 139-152. doi:10.1006/brln.1993.1009

ASHA. (1990). Guidelines for screening for hearing impairments and middle-ear disorders. *ASHA, 32*(2), 17–24.

Becker, F., & Reinvang, I. (2007). Mismatch negativity elicited by tones and speech sounds: Changed topographical distribution in aphasia. *Brain and Language 100*(1), 69-78. doi:10.1016/j.bandl.2006.09.004

Bentin, S., & Mann, V. (1990). Masking and stimulus intensity effects on duplex perception: A confirmation of the dissociation between speech and nonspeech modes. *The Journal of the Acoustical Society of America, 88*(1), 64-74. doi:10.1121/1.399845

Cairns, H. S. (1999). *Psycholinguistics: An Introduction*. Austin, TX: PRO-ED, Inc.

Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., et al. (1998). Development of language-specific phoneme representations in the infant brain. *Nature Neuroscience, 1*(5), 351-353. Retrieved from http://neurosci.nature.com

Cutting, J. E. (1978). There may be nothing peculiar to perceiving in a speech mode. In J. Requin (Ed.) *Attention and Performance VII* (pp. 229-243). Hillsdale, NJ: Erlbaum Associates.

Dalebout, S. D., & Stack, J. W. (1999). Mismatch negativity to acoustic differences not differentiated behaviorally. *Journal of the American Academy of Audiology, 10*, 388-399.

Fitch, H. L., Halwes, T., Erickson, D. M., & Liberman, A. M. (1980). Perceptual equivalence of two acoustic cues for stop-consonant manner. *Perception and Psychophysics, 27*(4), 343-350. Retrieved from http://www.psychonomic.org

Harnad, S. (Ed.) (1987). *Categorical perception: The groundwork of cognition*. Cambridge, England: Cambridge University Press.

Horev, N., Most, T., & Pratt, H. (2007). Categorical perception of speech (VOT) and analogous non-speech (FOT) signals: Behavioral and electrophysiological correlates. *Ear and Hearing, 2007 28*(1), 111-128. doi:10.1097/01.aud.0000250021.69163.96

Knecht, S., Deppe, M., Dräger, B., Bobe, L., Lohmann, H., Ringelstein, E. B., et al. (2000). Language lateralization in healthy right-handers. *Brain, 123*(1), 74-81. doi:10.1093/brain/123.1.74

Kraus, N., McGee, T., Carrell, T., King, C., Littman, T., & Nicol, T. (1994). Discrimination of speech-like contrasts in the auditory thalamus and cortex. *The Journal of the Acoustical Society of America, 96*(5), 2758-2768. doi:10.1121/1.411282

Kraus, N., McGee, T., Carrell, T. D., King, C., Tremblay, K., & Nicol, T. (1995). Central auditory system plasticity associated with speech discrimination training. *Journal of Cognitive Neuroscience, 7*(1), 25-32. doi:10.1162/jocn.1995.7.1.25

Kraus, N., McGee, T., Sharma, A., Carrell, T. D., & Nicol, T. (1992). Mismatch negativity event-related potential elicited by speech stimuli. *Ear and Hearing, 13*(3), 158-164. doi:10.1097/00003446-199206000-00004

Lang, A. H., Eerola, O., Korpilahti, P., Holopainen, I., Salo, S., & Aaltonen, O. (1995). Practical issues in the clinical application of mismatch negativity. *Ear and Hearing, 16*(1), 118-130. doi:10.1097/00003446-199502000-00009

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review, 74*(6), 431-461. doi:10.1037/h0020279

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology, 54*(5), 358-368. doi:10.1037/h0044417

Maiste, A. C., Wiens, A. S., Hunt, M. J., Scherg, M., & Picton, T. W. (1995). Event-related potentials and the categorical perception of speech sounds. *Ear and Hearing, 16*(1), 68-89. doi:10.1097/00003446-199502000-00006

McPherson, D. (1996). *Late Potentials of the Auditory System*. San Diego, CA: Singular Publishing Group, Inc.

McPherson, D., & Ballachanda, B. (2000). *Middle and long latency auditory evoked potentials*. New York: Thieme Medical Publishers, Inc.

Näätänen, R. (1992). *Attention and Brain Function*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Näätänen R. (1995). The mismatch negativity: A powerful tool for cognitive neuroscience. *Ear and Hearing, 16*(1), 6-18. doi:10.1097/00003446-199502000-00002

Näätänen, R. (2001). The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology, 38*(1), 1-21. doi:10.1111/1469-8986.3810001

Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., et al. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature, 385*(6615), 432-434. doi:10.1038/385432a0

Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology, 24*(4), 375-425. doi:10.1111/j.1469-8986.1987.tb00311.x

Näätänen, R., & Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin, 125*(6), 826-859. doi:10.1037/0033-2909.125.6.826

Pegg, J. E., & Werker, J. F. (1997). Adult and infant perception of two English phones. *The Journal of the Acoustical Society of America, 102*(6), 3742-3753. doi:10.1121/1.420137

Peltola, M. S., Kujala, T., Tuomainen, J., Ek, M., Aaltonen, O., & Näätänen, R. (2003). Native and foreign vowel discrimination as indexed by the mismatch negativity (MMN) response. *Neuroscience Letters, 352*(1), 25-28. doi:10.1016/j.neulet.2003.08.013

Pettigrew, C., Murdoch, B., Ponton, C., Kei, J., Chenery, H., & Alku, P. (2004). Subtitled videos and mismatch negativity (MMN) investigations of spoken word processing. *Journal of the American Academy of Audiology*, *15*(7), 469-485. doi:10.3766/jaaa.15.7.2

Picton, T., Campbell, K. B., Baribeau-Braun, J., & Proulx, G. B. (1978). *The neurophysiology of human attention: A tutorial review* (Vol. 7). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Roth, W. T. (1983). A comparison of P300 and skin conductance response. In A. W. K. Gaillard & W. Ritter (Eds.), *Tutorials in event related potential research: Endogenous components* (pp. 177–199). Amsterdam: North-Holland Publishing.

Roup, C. M., Wiley, T. L., Safady, S. H., & Stoppenbach, D. T. (1998). Tympanometric screening norms for adults. *American Journal of Audiology, 7*, 55–60. doi:10.1044/1059-0889(1998/014)

Ruchkin, D. S., & Sutton, S. (1983). Positive slow wave and P300: Association and disassociation. In A. W. K. Gaillard & W. Ritter (Eds.), *Tutorials in event related potential research: Endogenous components* (pp. 233–250). Amsterdam: North-Holland Publishing.

Sams, M., Aulanko, R., Aaltonen, O., & Näätänen, R. (1990). Event-related potentials to infrequent changes in synthesized phonetic stimuli. *Journal of Cognitive Neuroscience, 2*(4), 344-357. doi:10.1162/jocn.1990.2.4.344

Schouten, M. E. H., & Hessen, A. J. V. (1992). Modeling phoneme perception. I: Categorical perception. *The Journal of the Acoustical Society of America, 92*(4), 1841-1855. doi:10.1121/1.403841

Sharma, A., & Dorman, M. F. (1999). Cortical auditory evoked potential correlates of categorical perception of voice-onset time. *The Journal of the Acoustical Society of America, 106*(2), 1078-1083. doi:10.1121/1.428048

Sharma, A., & Dorman, M. F. (2000). Neurophysiologic correlates of cross-language phonetic perception. *The Journal of the Acoustical Society of America, 107*(5), 2697-2703. doi:10.1121/1.428655

Tampas, J. W., Harkrider, A. W., & Hedrick, M. S. (2005). Neurophysiological indices of speech and nonspeech stimulus processing. *Journal of Speech, Language and Hearing Research 48*(5), 1147-1164. doi:10.1044/1092-4388(2005/081)

Tata, M. S., & Ward, L. M., (2005). Early phase of spatial mismatch negativity is localized to a posterior "where" auditory pathway. *Experimental Brain Research, 167*(3), 481-486. doi:10.1007/s00221-005-0183-y

Whalen, D. H., & Liberman, A. M. (1987). Speech perception takes precedence over nonspeech perception. *Science, New Series, 237*(4811), 169-171. doi:10.1126/science.3603014

Whalen, D. H., & Liberman, A. M. (1996). Limits on phonetic integration in duplex perception. *Perception and Psychophysics, 58*(6), 857-870. Retrieved from http://www.psychonomic.org

Winkler, I., Lehtokoski, A., Alku, P., Vainio, M., Czigler, I., Csépe, V., et al. (1999). Pre-attentive detection of vowel contrasts utilizes both phonetic and auditory memory representations. *Cognitive Brain Research, 7*(3), 357-369. doi:10.1016/S0926-6410(98)00039-1

Ylinen, S., Shestakova, A., Houtilainen, M., Alku, P., & Näätänen, R. (2006). Mismatch negativity (MMN) elicited by changes in phoneme length: A cross-linguistic study. *Brain Research 1072*(1), 175-185. doi:10.1016/j.brainres.2005.12.004

Appendix A

Informed Consent to Act as a Human Research Subject
Brain Mapping of the MMN and P300 in Speech and Nonspeech Stimulus Processing
David L. McPherson, Ph.D.
Communication Science and Disorders
Brigham Young University
(801) 422-6458

Name of Participant: ________________________

Purpose of Study
The purpose of the proposed research project is to study whether specific aspects of brain activity are influenced by the speech sounds associated with certain stimuli or whether these measurements are influenced by purely acoustic differences.

Procedures
You have been asked to participate in this study by Skylee Neff, BS, a graduate student conducting research under the direction of Dr. David L. McPherson.
The study will be conducted in room 111 of the John Taylor Building on the campus of Brigham Young University. The testing will consist of one session including orientation and testing and will last for 2-3 hours. You may ask for a break at any time during testing. Basic hearing tests will be administered during the first half-hour of the session.
Surface electrodes (metal discs about the size of a dime) will be used to record electrical activity of your brain. These discs will be applied to the surface of the skin with a cream or gel and are easily removed with water. Blunt needles will be used as a part of this study to help apply the electrode gel. They will *never* be used to puncture the skin.
Acoustic and linguistic processing will be measured using an electrode cap, which simply measures the electrical activity of your brain and *does not* emit electricity; no electrical impulses will be applied to the brain. These measurements of the electrical activity are similar to what is known as an "EEG" or brain wave test. These measurements are of normal, continuous electrical activity in the brain.
You will wear the electrode cap while you listen to different consonant vowel combinations, during which time the electrical activity of your brain will be recorded on a computer. The sounds will be presented through insert earphones at a comfortable but not loud level. You will be seated comfortably in a sound treated testing room. You will be asked to give responses during the hearing test and portions of the electrophysiological recording by pressing a button.
The procedures used to record the electrophysiological responses of the brain are standardized and have been used without incident in many previous investigations. The combination of sounds presented is experimental, but the recording procedure is not.

Risks/Discomforts
There are very few potential risks from this procedure, and these risks are minimal. The risks of this study include possible allergic reactions to the gel used in applying the electrodes. Allergic reactions to the gel are extremely rare. There is also a possibility for an allergic reaction to the electrodes. If any of these reactions occur, a rash would appear. Treatment would include

removing the electrodes and gel and exposing the site to air, resulting in removal of the irritation. If there is an allergic reaction, testing procedures would be discontinued. Another unlikely risk is a small abrasion on the scalp when the blunt needle is used to place electrode gel. Treatment would also include removing the electrode and gel, exposing the site to air and testing procedures would be discontinued.

Benefits
You will receive a copy of your hearing assessment at no charge. You will be notified if any indications of hearing loss are found in this area. The information obtained may help to further the understanding of language processing, which will be beneficial to professionals involved in treating speech and hearing disorders.

Confidentiality
All information obtained from testing is confidential and is protected under the laws governing privacy. All identifying references will be removed and replaced by control numbers. Data collected in this study will be stored in a secured area accessible only to personnel associated with the study. Data will be reported in aggregate form without individual identifying information.

Compensation
You will be given a voucher for a free pizza at each session you attend for this study; you will receive a voucher whether or not you complete the study.

Participation
Participation in this research study is voluntary. You have the right to withdraw at anytime or refuse to participate entirely without affecting your standing with the University.

Questions about the Research
If there are any further questions or concerns regarding this study, you may ask any of the investigators or contact David McPherson, Ph.D., Communication Science and Disorders, room 129, Taylor Building, Provo, Utah 84602; phone (801) 422-6458; email: david_mcpherson@byu.edu.

Questions about your Rights as Research Participants
If you have questions regarding your rights as a research participant, you may contact the BYU IRB Administrator at (801) 422-1461, A-285 ASB, Brigham Young University, Provo, UT 84602, irb@byu.edu.

Other Considerations
There are no charges incurred by you for participation in this study. There is no treatment or intervention involved in this study.

The procedures listed above have been explained to me by: ________________________ in a satisfactory manner and any questions relating to such risks have been answered.

I understand what is involved in participating in this research study. My questions have been answered and I have been offered a copy of this form for my records. I understand that I may withdraw from participating at any time. I agree to participate in this study.

Printed Name:___________________________________

Signature: _____________________________________

Date:____________________

Appendix B

Informed Consent to Act as a Human Research Subject
Brain Mapping of the MMN and P300 in Speech and Nonspeech Stimulus Processing
David L. McPherson, Ph.D.
Communication Science and Disorders
Brigham Young University
(801) 422-6458

Name of Participant: ________________________

Purpose of Study
The purpose of the proposed research project is to study whether specific aspects of brain activity are influenced by the speech sounds associated with certain stimuli or whether these measurements are influenced by purely acoustic differences.

Procedures
You have been asked to participate in a research study conducted by Skylee Neff, BS, a graduate student conducting research under the direction of Dr. David L. McPherson.
The study will be conducted in room 111 of the John Taylor Building on the campus of Brigham Young University. The testing will consist of two identical sessions including orientation and testing. Each session will last for 2-3 hours. Total participation time will be approximately 4-6 hours. You may ask for a break at any time during testing. Basic hearing tests will be administered during the first half-hour of the first session.
Surface electrodes (metal discs about the size of a dime) will be used to record electrical activity of your brain. These discs will be applied to the surface of the skin with a cream or gel and are easily removed with water. Blunt needles will be used as a part of this study to help apply the electrode gel. They will *never* be used to puncture the skin.
Acoustic and linguistic processing will be measured using an electrode cap, which simply measures the electrical activity of your brain and *does not* emit electricity; no electrical impulses will be applied to the brain. These measurements of the electrical activity are similar to what is known as an “EEG” or brain wave test. These measurements are of normal, continuous electrical activity in the brain.
You will wear the electrode cap while you listen to different consonant vowel combinations, during which time the electrical activity of your brain will be recorded on a computer. The sounds will be presented through insert earphones at a comfortable but not loud level. You will be seated comfortably in a sound treated testing room. You will be asked to give responses during the hearing test and portions of the electrophysiological recording by pressing a button.
The procedures used to record the electrophysiological responses of the brain are standardized and have been used without incident in many previous investigations. The combination of sounds presented is experimental, but the recording procedure is not.

Risks/Discomforts
There are very few potential risks from this procedure, and these risks are minimal. The risks of this study include possible allergic reactions to the gel used in applying the electrodes. Allergic reactions to the gel are extremely rare. There is also a possibility for an allergic reaction to the

electrodes. If any of these reactions occur, a rash would appear. Treatment would include removing the electrodes and gel and exposing the site to air, resulting in removal of the irritation. If there is an allergic reaction, testing procedures would be discontinued. Another unlikely risk is a small abrasion on the scalp when the blunt needle is used to place electrode gel. Treatment would also include removing the electrode and gel, exposing the site to air and testing procedures would be discontinued.

Benefits
You will receive a copy of your hearing assessment at no charge. You will be notified if any indications of hearing loss are found in this area. The information obtained may help to further the understanding of language processing, which will be beneficial to professionals involved in treating speech and hearing disorders.

Confidentiality
All information obtained from testing is confidential and is protected under the laws governing privacy. All identifying references will be removed and replaced by control numbers. Data collected in this study will be stored in a secured area accessible only to personnel associated with the study. Data will be reported in aggregate form without individual identifying information.

Compensation
You will be given a voucher for a free pizza at each session you attend for this study; you will receive a voucher whether or not you complete the study.

Participation
Participation in this research study is voluntary. You have the right to withdraw at anytime or refuse to participate entirely without affecting your standing with the University.

Questions about the Research
If there are any further questions or concerns regarding this study, you may ask any of the investigators or contact David McPherson, Ph.D., Communication Science and Disorders, room 129, Taylor Building, Provo, Utah 84602; phone (801) 422-6458; email: david_mcpherson@byu.edu.

Questions about your Rights as Research Participants
If you have questions regarding your rights as a research participant, you may contact the BYU IRB Administrator at (801) 422-1461, A-285 ASB, Brigham Young University, Provo, UT 84602, irb@byu.edu.

Other Considerations
There are no charges incurred by you for participation in this study. There is no treatment or intervention involved in this study.

The procedures listed above have been explained to me by: ________________________ in a satisfactory manner and any questions relating to such risks have been answered.

I understand what is involved in participating in this research study. My questions have been answered and I have been offered a copy of this form for my records. I understand that I may withdraw from participating at any time. I agree to participate in this study.

Printed Name:_____________________________________

Signature: ____

Date:_____________________