2012-06-01

Brain Mapping of the Mismatch Negativity Response in Vowel Formant Processing

Elizabeth Anne Perry
Brigham Young University - Provo

Follow this and additional works at: https://scholarsarchive.byu.edu/etd

Part of the Communication Sciences and Disorders Commons

BYU ScholarsArchive Citation
Perry, Elizabeth Anne, "Brain Mapping of the Mismatch Negativity Response in Vowel Formant Processing" (2012). All Theses and Dissertations. 3226.
https://scholarsarchive.byu.edu/etd/3226

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in All Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.
Brain Mapping of the Mismatch Negativity Response in Vowel Formant Processing

Elizabeth A. Perry

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

David L. McPherson, Chair
Shawn L. Nissen
Christopher Dromey

Department of Communication Disorders
Brigham Young University
December 2012

Copyright © 2012 Elizabeth A. Perry
All Rights Reserved
ABSTRACT

Brain Mapping of the Mismatch Negativity Response in Vowel Formant Processing

Elizabeth A. Perry
Department of Communication Disorders, BYU
Master of Science

The mismatch negativity (MMN) response, a passively-elicited component of the auditory event-related potential (ERP), reflects preattentive identification of infrequent changes in acoustic stimuli. In the current study, the MMN response was examined closely to determine what extent natural speech sounds evoke the MMN. It was hypothesized that a significant MMN response results during the presentation of deviant stimuli from which spectral energy within formant bands critical to vowel identification has been removed. Localizations of dipoles within the cortex were hypothesized to yield information pertaining to the processing of formant-specific linguistic information. A same/different discrimination task was administered to 20 adult participants (10 female and 10 male) between the ages of 18 and 26 years. Data from behavioral responses and ERPs were recorded. Results demonstrated that the MMN may be evoked by natural speech sounds. Grand-averaged brain maps of ERPs created for all stimulus pairs showed a large preattentive negativity. Additionally, amplitudes of the MMN were greatest for pairs of auditory stimuli in which spectral energy not corresponding to formant frequencies was digitally eliminated. Dipoles reconstructed from temporal ERP data were located in cortical areas known to support language and auditory processing. Significant differences between stimulus type and reaction time were also noted. The current investigation confirms that the MMN response is evoked by natural speech sounds and provides evidence for a theory of preattentive formant-based processing of speech sounds.

Keywords: electroencephalography, event-related potentials, mismatch negativity, brain mapping, scalp distribution, formants, formant processing, dipole localization
ACKNOWLEDGMENTS

I would like to thank my thesis chair, Dr. McPherson, for his guidance throughout the course of this project. I would also like to thank my husband, Daniel, for his unwavering support.
# Table of Contents

List of Tables ................................................................................................................................. vi

List of Figures ............................................................................................................................... vii

List of Appendices ....................................................................................................................... viii

Introduction ..................................................................................................................................... 1

The Mismatch Negativity ............................................................................................................ 1

Formants ...................................................................................................................................... 2

Early Speech Processing ............................................................................................................. 4

Method ............................................................................................................................................ 6

Participants .................................................................................................................................. 6

Instrumentation ............................................................................................................................ 6

Pre-testing measures and preparation. ..................................................................................... 6

EEG data collection and analysis. ........................................................................................... 7

Stimuli ......................................................................................................................................... 7

Procedures ................................................................................................................................... 9

Initial screening. ....................................................................................................................... 9

Behavioral data acquisition. ................................................................................................... 11

Event related potential acquisition. ....................................................................................... 12

Data Analysis ............................................................................................................................ 13

Behavioral data. ..................................................................................................................... 13

Event related potential data. .................................................................................................. 13

Results ........................................................................................................................................... 14

Behavioral Data ...................................................................................................................... 14
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event Related Potential Data</td>
<td>15</td>
</tr>
<tr>
<td>Temporal data analysis.</td>
<td>15</td>
</tr>
<tr>
<td>Dipole localization and analysis.</td>
<td>25</td>
</tr>
<tr>
<td>Repeated Measures</td>
<td>25</td>
</tr>
<tr>
<td>Discussion</td>
<td>25</td>
</tr>
<tr>
<td>Summary and Evaluation of Results</td>
<td>26</td>
</tr>
<tr>
<td>Recommendations for Future Research</td>
<td>29</td>
</tr>
<tr>
<td>References</td>
<td>31</td>
</tr>
<tr>
<td>Annotated Bibliography</td>
<td>36</td>
</tr>
</tbody>
</table>
List of Tables

1. The two corner vowels /æ/ and /u/ with their corresponding formant frequencies (in Hz) ........ 8
2. The filtered frequency ranges (in Hz) from all stimuli presented to listeners ....................... 9
3. Stimulus blocks presented to listeners ...................................................................................... 12
4. Descriptive statistics for reaction time, in ms, across subjects for each stimulus condition .... 15
5. Average latency, in ms, and amplitude, in μV, of the MMN for each stimulus block .......... 24
List of Figures

1. Spectrograms depicting common and deviant stimuli (LPC spectra included) .................. 10

2. Part A depicts Block 1 grand-averaged brain maps of ERPs for the filtered /æ/ F1 stimulus pair. Part B depicts a reconstructed image of the brain taken from axial, coronal, and sagittal views. Locations of the dipole are shown ................................................................. 17

3. Part A depicts Block 2 grand-averaged brain maps of ERPs for the filtered /æ/ F2 stimulus pair. Part B depicts a reconstructed image of the brain taken from axial, coronal, and sagittal views. Locations of the dipole are shown ................................................................. 18

4. Part A depicts block 3 grand-averaged brain maps of ERPs for the filtered random /æ/ stimulus pair. Part B depicts a reconstructed image of the brain taken from axial, coronal, and sagittal views. Locations of the dipole are shown ................................................................. 19

5. Part A depicts Block 4 grand-averaged brain maps of ERPs for the filtered /u/ F1 stimulus pair. Part B depicts a reconstructed image of the brain taken from axial, coronal, and sagittal views. Locations of the dipole for Block 4 are shown ................................................................. 20

6. Part A depicts Block 5 grand-averaged brain maps of ERPs for the filtered /u/ F2 stimulus pair. Part B depicts a reconstructed image of the brain taken from axial, coronal, and sagittal views. Locations of the dipole are shown ................................................................. 22

7. Part A depicts Block 6 grand-averaged brain maps of ERPs for the filtered random /u/ stimulus pair. Part B depicts a reconstructed image of the brain taken from axial, coronal, and sagittal views. Locations of the dipole are shown ................................................................. 23
List of Appendices

A. Informed Consent to Act as a Human Research Subject .......................................................... 71
Introduction

Considerable volumes of electroencephalographic research have been published throughout the course of the last century. Electroencephalography (EEG) is used to record the electrical activity of billions of polarized neurons within the cerebral cortex at specific locations across the scalp (Martin, Tremblay, & Korczak, 2008; Näätänen, 1995). One specific brain response measured with EEG is the event-related potential (ERP). ERPs are electrophysiological responses to sensory, motor, or cognitive stimuli and reflect perceptual processes in the cortex (Hahne & Friederici, 1999; Horev, Most, & Pratt, 2007). Many researchers have used the auditory ERP to measure the neurophysiological correlates of perceptual responses to auditory stimuli. The auditory ERP provides insight into the neural mechanisms underlying speech processing (Näätänen et al., 1997).

The Mismatch Negativity

The mismatch negativity (MMN) response is a passively-elicited component of the auditory ERP. It reflects preattentive identification of infrequent changes in the acoustic stimuli (Dalebout, 1999). The MMN is elicited using a same/different paradigm in which an auditory standard stimulus is presented repeatedly to the listener. A deviant stimulus is occasionally presented in a random fashion, replacing the standard stimulus (Aaltonen, Eerola, Lang, Uusipaikka, & Tuomainen, 1994; Dalebout, 1999; Jacobsen, 2004; Näätänen et al., 1997; Nyman et al., 1990). The MMN response is enhanced or attenuated depending on the extent of the difference between the standard and the deviant stimuli. While several studies have shown that the MMN response is evoked by both speech and nonspeech stimuli, some researchers have asserted that the MMN is only evoked by nonspeech stimuli (Tampas, Harkrider, & Hedrick,
2005). The conclusion that the MMN response detects non-phonetic acoustic changes in speech has been widely disputed (Maiste, Wiens, Hunt, Scherg, & Picton, 1995).

Becker and Reinvang (2007) and Maiste et al. (1995) demonstrated that natural speech sounds can elicit MMN responses, thereby concluding that the MMN detects both non-phonetic and phonetic changes in speech. In another study, MMN responses were elicited in infants through the use of recorded speech sounds (Cheour, Alho, et al., 1998). However, most auditory ERP studies to date have presented listeners with synthesized speech with the intent of studying the brain’s response to very specific frequencies in the speech sound (Diesch & Luce, 2000).

**Formants**

Speech sounds have been created synthetically by simulating formants, or peaks in the spectral envelope (Kent & Read, 2002). Several researchers have employed synthetic speech to determine whether synthetic formants are perceived and processed like natural speech; however, results are conflicting (Aaltonen et al., 1994; Aaltonen & Suonpaa, 1983; Ainsworth & Millar, 1972; Becker & Reinvang, 2007; Diesch & Luce, 1997; Diesch & Luce, 2000; Edmonds et al., 2010; Jacobsen, 2004; Miner & Danhauer, 1977; Peltola, Kuntola, Tamminen, Hamalainen, & Aaltonen, 2005).

Formants give a vowel its identity. Results from a two-formant synthesized vowel study conducted by Aaltonen & Suonpaa (1983) indicated that vowels can be identified from the presence of the first and second formants. They concluded that there may be some neural correlate linking formant frequency and vowel identification. Later, Aaltonen et al. (1994) presented participants with digitally altered speech sounds and found that alterations in formant frequency affected the MMN response. It was concluded that even small and phonetically irrelevant changes in the second formant (F2) frequency result in a heightened MMN response.
Aaltonen et al. (2008) presented participants with speech stimuli and elicited MMN responses through the intermittent violation of phonological constraints. Amplitudes of the MMN were greater for violations considered more deviant in the English language, thereby showing that the MMN response has some linguistic components. Thus, the MMN cannot be solely attributed to acoustic processing and perception.

Rather than synthesizing and adding formants together, several researchers have synthesized and attenuated formants (Ainsworth & Millar, 1972; Assmann, 1991; Diesch & Luce, 1997; Diesch & Luce, 2000; Flanagan, 1957; Horev et al, 2007; Jacobsen, 2004; Miner & Danhauer, 1977; Näätänen et al., 1997). Näätänen et al. (1997) synthesized speech by creating F1, F2, F3, and F4 components matching the Estonian vowel /ö/, altering F2 for the deviant stimulus. A large MMN response was seen following the presentation of the deviant stimuli. All spectral energy residing around F2, referred to by Näätänen et al. (1997) as the F2 formant perimeter, was removed. Large formant perimeters allowed for the attenuation or deletion of specific formants without diminishing apparent vowel indicators. Vowel indicators are the components which give a vowel its identity. Formants, specifically F1 and F2, are vowel indicators. If a specific formant frequency such as F2 is deleted, F3 may be perceived as F2. Extracting a large band of frequencies surrounding the first two formants and completely removing, rather than attenuating, all frequencies eliminates this phenomenon (Beddor & Hawkins, 1990; Dubno & Dorman, 1987). Although much is known about the MMN response to the extraction of synthesized vowel indicators, very little research pertaining to the attenuation or deletion of bands of spectral energy in nonsynthesized, or natural, speech sounds has been conducted.
Early Speech Processing

Näätänen et al. (2007) sought to determine whether the preattentive processing of natural speech sounds is in any way attributable to early speech processing within the cortex. They presented the Finnish phoneme /e/ and the Estonian phoneme /o/ to Finnish subjects with no prior knowledge of the Estonian language and found the MMN response to be greater after the presentation of a Finnish deviant stimulus, as opposed to the presentation of an Estonian deviant stimulus. The authors concluded that language-dependent memory traces exist and are activated within 200 ms after the onset of speech stimuli. The same language-dependent memory traces are not activated when acoustic non-speech stimuli are presented to the listener, even if the stimuli are acoustically equally complex.

Similar studies by Edmonds et al. (2010), Newman, Connolly, Service, and McIvor (2003), Sharma and Dorman (2000), and Ylinen, Shestakova, Huotilainen, Alku, and Näätänen (2006) yielded comparable results, although they did not attribute the early speech processing to language-dependent memory traces. They concluded that speech-specific processing may involve the primary auditory cortex, or even subcortical structures, and could influence early auditory ERPs such as the MMN response. Edmonds et al. (2010) reported not only that the MMN response reflects early speech processing, but also that formant information in steady-state vowels is the primary contributor to speech processing, thereby supporting a formant-based phonemic theory of preattentive processing of natural speech sounds.

In contrast to the aforementioned findings, Tampas et al. (2005) reported that the MMN response is not influenced by the phonetic characteristics of the stimulus. They concluded that the MMN is purely reflective of an acoustic level of preattentive processing of natural speech sounds.
sounds. Additionally, they concluded that the MMN response cannot be elicited at all by speech stimuli. The results conflict with those reported in the previously cited studies.

The current study addressed the need for additional research on acoustic and phonemic processing of nonsynthesized vowel formants. Many investigations have used ERPs to examine vowel and formant localization within the cortex and to determine whether vowel processing and discrimination are acoustic or phonemic in nature (Aaltonen, 1994; Aaltonen & Suonpaa, 1983; Alho, Lavikainen, Reinikainen, Sams, & Näätänen, 1990; Cheour et al., 1998; Diesch & Luce, 2000; Edmonds et al., 2010; Jacobsen, 2004; Miner & Danhauer, 1977; Näätänen et al., 1997; Neff, 2010; Winkler et al., 1990; Winkler et al., 1999). Many of the previous studies presented synthesized vowels as stimuli, whereas, the current study used digitally-altered natural speech.

Also addressed in the current investigation was the scalp distribution and cortical localization of formant-based linguistic information. Previous research indicates that different components of auditory information and speech processing take place in different locations within the cortex (Becker & Reinvang, 2007; Diesch & Luce, 1997; Edmonds, 2010; Knecht et al., 2000). Additionally, specific aspects of speech processing occur at different latencies after the onset of the auditory stimulus (Edmonds, 2010; Jacobsen, 2004; Kraus, McGee, Sharma, Carrell, & Nicol, 1992; Näätänen, 1995). Becker and Reinvang (2007) reported that the topographical distribution of the MMN elicited by tones and speech sounds changes depending upon the linguistic information presented to listeners.

In the current study it was hypothesized that the MMN response can be evoked by natural speech sounds. It was also expected that a significant MMN will occur during the presentation of deviant stimuli from which spectral energy within formant bands critical to vowel identification has been removed. Localization of dipoles within the cortex was hypothesized to yield
information pertaining to the processing of formant-specific linguistic information. This hypothesis supports a formant-based phonemic theory of preattentive processing of natural speech sounds.

**Method**

There were two main aims of the present study. The first was to identify the MMN response as it appears in response to natural speech stimuli. The second was to evaluate which deviant stimuli, if any, evoked an increase in the amplitude of the MMN based upon the spectral content of the stimuli.

**Participants**

Twenty native English speaking individuals (10 females and 10 males) between the ages of 18 and 26 years participated in this study. The participants were all right handed (Knecht et al., 2000) and reported no history of neurological, cognitive, or learning impairments. Prior to participating in the study, each participant read and signed an informed consent document approved by the Institutional Review Board at Brigham Young University (see Appendix A). In addition, the study was conducted in accordance with the Declaration of Helsinki (World Medical Association, 2008).

Participants received one $5 gift certificate per visit regardless of inclusion or exclusion from further EEG investigation. Test-retest reliability was established by selecting three participants from the study at random and testing them a second time, comparing results using a repeated measures design.

**Instrumentation**

**Pre-testing measures and preparation.** All stimuli were recorded in a sound-isolated chamber using a low impedance dynamic microphone (DPA 4011). The recordings were made at
44.1 kHz with the microphone placed approximately fifteen centimeters from the speaker’s mouth. Stimuli were digitized with an A/D converter by Apogee Systems. Adobe Audition 3.0 software was used to segment the recording into individual vowels and to interface with NeuroScan 4.5 software. In order to be used on the NeuroStim system, the recordings were down sampled to 22.1 kHz, 16 bits.

**EEG data collection and analysis.** Collectively, the programs used in data acquisition and analysis revealed information pertaining to the temporal and cortical localization of acoustic and phonemic processing. A standard computer equipped with NeuroScan 4.5 software (EEG data collection) and Stim 2 (stimulus presentation) software (Compumedics Neuroscan, 2008), designed to collect and record electroencephalographic data, was used for EEG data collection and stimulus presentation. The participants were fitted comfortably with a 32 channel Electrocap (Electro-Cap International, Inc., 2003). Electrode impedances did not exceed 3000 ohms. The Electrocap was equipped with 32 silver-silver chloride electrodes which rested against the scalp and distributed according to the 10-20 International System (Jasper & Radmussen, 1958). Six additional silver-silver chloride electrodes were placed with removable tape on the right and left mastoid process (linked-mastoid references), the outer cantha of each eye, and at least one inch above and below the supraorbital foramen of the left eye (to monitor eye movement and facial muscle activity). Cortical localization of speech stimuli was conducted with the CURRY 6 software (Compumedics Neuroscan, 2008) *post hoc*.

**Stimuli**

Stimuli included two corner vowels, /æ/ and /u/, which were chosen because they are more acoustically identifiable than other vowels (Kent & Read, 2002). A female native English speaker produced the stimuli (Table 1). The speaker sustained the /æ/ vowel for 378 ms and the
/u/ vowel for 447 ms. Judges unfamiliar with the study, one male and two female undergraduate students, correctly identified the recordings as /æ/ and /u/. The loudness level of each original unfiltered recording was also perceived by the judges to be equal. The first two formants, F1 and F2, were chosen for examination because of their key role in vowel identification (Kent & Read, 2002; Peterson & Barney, 1952).

Table 1

*The two corner vowels /æ/ and /u/ with their corresponding formant frequencies (in Hz)*

<table>
<thead>
<tr>
<th>Corner Vowel</th>
<th>F1</th>
<th>F2</th>
</tr>
</thead>
<tbody>
<tr>
<td>/æ/</td>
<td>1089</td>
<td>1658</td>
</tr>
<tr>
<td>/u/</td>
<td>417</td>
<td>1374</td>
</tr>
</tbody>
</table>

Once the recording was segmented into separate vowels, the initial and final 50 ms of each vowel were deleted. Consequently, 278 ms of the /æ/ recording and 347 ms of the /u/ recording were presented to participants. Since the two vowels were not compared against each other, durational differences were not a consideration in this study. The duration of each vowel effectively preserved the middle of each vowel, maximizing identifiability (Nootboom & Doodeman, 1980; Pols, Tromp, & Plomp, 1973) and minimizing shifts in fundamental frequency. Praat 5.2.26 software (Boersma & Weenink, 2004) was used to approximate perimeters of each formant.

The formant ranges were then set numerically as band reject filter parameters. All energy within the band of frequencies was eliminated, rather than attenuated, from the recording (Beddor and Hawkins, 1990; Dubno and Dorman, 1987). Random frequencies not corresponding to any formant range were also selected as filter parameters to measure whether MMN responses were evoked solely by auditory information or whether they were influenced by removal of
spectral energy residing within the range of specific formants. The amount of energy removed from the random band of frequencies, 500-550 Hz for /æ/ and 400 Hz for /u/, was the same as that removed from other filtered recordings (Table 2). A total of eight final recordings were made (Figure 1). All recordings were high-pass filtered to eliminate noise below 65 Hz. The first and last 10 ms of each recording were smoothed in the filtered stimuli to eliminate auditory clicks and increase perceived naturalness.

Table 2

<table>
<thead>
<tr>
<th>Corner Vowel</th>
<th>F1</th>
<th>F2</th>
<th>Random</th>
</tr>
</thead>
<tbody>
<tr>
<td>/æ/</td>
<td>850-1400</td>
<td>1400-1900</td>
<td>1200-1700</td>
</tr>
<tr>
<td>/u/</td>
<td>350-750</td>
<td>1250-1650</td>
<td>750-1150</td>
</tr>
</tbody>
</table>

Procedures

**Initial screening.** Participants were presented with the two unfiltered vowel recordings (/u/ and /æ/) and were asked to identify the vowels as a means of determining whether they could recognize and discriminate between the two vowel recordings. All participants were able to discriminate between the vowel recordings. In order to be included in the listening portion of the study, participants were required to pass an initial hearing screening by demonstrating normal hearing bilaterally with pure tone thresholds ≤ 15 dB HL for octave intervals between 250-8000 Hz and threshold differences between ears ≤ 5 dB. Pure tone thresholds were established using a modified Hughson-Westlake technique (American National Standards Institute, 1997; American Speech-Language-Hearing Association, 1978; Jerger, Carhart, Tillman, & Peterson, 1959). Participants also demonstrated bilateral type A tympanograms with static acoustic admittance
Figure 1. Spectrograms A and E depict common stimuli. Deviant stimuli are shown beneath the common stimuli with bands of spectral energy deleted from each recording. The insets in the upper right corner of each spectrogram are linear predictive coding (LPC) spectra in which peaks show the formants. Arrows pointing to dips in the LPC spectrum correspond with deleted bands of spectral energy.
measures between 0.3 and 1.4 mmhos, peak pressure between -100 and +50 daPa in the test ear
(American Speech-Language-Hearing Association, 1990), clear ear canals, and healthy, intact
tympanic membranes. Speech recognition thresholds (SRT) did not exceed the limits of ≤ 15 dB
HL and 5 dB between ears, and Word Recognition Scores (WRS) were ≥ 98% bilaterally.

Otoscopy was conducted with a Welch Allyn otoscope. A handheld Grason-Stadler
model GSI-33 impedance meter was used to perform the tympanograms. Participants were
seated in a double-walled, sound treated test booth. Noise levels were within the limits as
specified by ANSI S3.1-1999 R2008 Maximum Permissible Ambient Noise Levels for
Audiometric Test Rooms. The participants remained seated throughout the duration of the
session. Etymotic EA-3 insert earphones, through which all sound stimuli were presented, and a
Grason-Stadler model GSI-10 audiometer were used for stimulus presentation.

**Behavioral data acquisition.** Following the initial screening, participants were asked to
listen to six different stimulus blocks presented in a two-interval forced choice paradigm. Each
block contained a total of 200 stimuli with 160 unfiltered (standard) and 40 filtered (deviant)
stimuli interspersed randomly throughout the block. Six blocks were created by pairing
unfiltered steady-state vowels with filtered steady-state vowels (Table 3). Unfiltered vowels
served as the first pair member of each block and comprised the 160 standard stimuli per block.
Filtered vowels served as the second pair member of each block and comprised the 40 deviant
stimuli per block. Filtered stimuli included the /æ/ and /u/ recordings in which the first and
second formants were removed. Additionally, “random” deviant stimuli were created by filtering
a band of spectral energy which did not correspond to the first or second formants. As part of the
randomization, one deviant stimulus could not directly follow the presentation of another deviant
stimulus. To control for test order effects, the order of blocks presented to each participant was
randomized. The sound stimuli were presented bilaterally at 65 dB HL over a series of 1200 trials. Participants responded manually to all stimuli with a button push.

Table 3

*Stimulus blocks presented to listeners: numbers are assigned to each block for convenience*

<table>
<thead>
<tr>
<th>Block</th>
<th>Pair Member 1</th>
<th>Pair Member 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Unfiltered /æ/</td>
<td>Filtered /æ/ F1</td>
</tr>
<tr>
<td>2</td>
<td>Unfiltered /æ/</td>
<td>Filtered /æ/ F2</td>
</tr>
<tr>
<td>3</td>
<td>Unfiltered /æ/</td>
<td>Random /æ/</td>
</tr>
<tr>
<td>4</td>
<td>Unfiltered /u/</td>
<td>Filtered /u/ F1</td>
</tr>
<tr>
<td>5</td>
<td>Unfiltered /u/</td>
<td>Filtered /u/ F2</td>
</tr>
<tr>
<td>6</td>
<td>Unfiltered /u/</td>
<td>Random /u/</td>
</tr>
</tbody>
</table>

All participants were asked to listen to the test administrator read the following script prior to the beginning of data collection:

You will hear many pairs of speech sounds. Each pair of sounds will be presented as one sound followed by another. You must press a button after each pair. If the two sounds are the same, press 1. Here is an example of what you might hear if the sounds are the same [/æ/-/æ/ and /u/-/u/]. If the two sounds are different, press 4. Here is an example of what you might hear if the sounds are different [/æ/-/æ/ and /u/-/ũ/]. If you wish to discontinue the test at any time, you may say, “I want to stop now.” Are there any questions? OK, we still start the test.

Following the script reading, participants performed a training task. They were familiarized with the task and all stimuli through the presentation of 30 stimuli pairs. The 1200 trials and EEG recordings began following the training task.

**Event related potential acquisition.** The ongoing EEG was streamed onto the computer using NeuroScan 4.5 software and saved for *post hoc* averaging of the MMN. The participants
sat quietly in a padded armchair which provided neck and head support for the duration of the data acquisition period. They were permitted to get up and stretch (without removing or altering the electrode cap) after the presentation of three blocks of stimuli to maintain an awake and alert state. The participants’ responses, response times, and EEG were recorded and stored on a secure digital computer for off-line processing.

**Data Analysis**

**Behavioral data.** Descriptive statistics were computed for correct responses and reaction times less than 2000 ms. An analysis of variance was conducted for reaction time versus stimulus type.

**Event related potential data.** Epochs were created from the raw EEG data and recordings were individually investigated. Artifacts such as eye and jaw movement were removed as part of the artifact removal procedure embedded in the Neuroscan 4.5 software prior to completing epoch averaging. All collected data resulting from these artifacts were removed prior to compiling averages. Individual averages were calculated by averaging the ERP data acquired for each block of stimuli per participant. Eight grand averaged ERP files were obtained by further averaging the individual averaged ERP files for each stimulus. A difference ERP file was created for MMN response analysis by subtracting the deviant ERP average from the common ERP average for each stimulus block (Aaltonen et al., 1994). The CURRY 6.0.12 neuroimaging software was used to determine dipoles, cortical sites to which the source of electrical activity have been traced (Näätänen, 2008), within the cortex for all individual averaged ERP files and again for the grand averaged ERP files. Locations for each dipole were compared between individuals and the grand average for all deviant and standard responses within each block of stimuli.
The dipole resulting from the grand averaged ERP file was used to determine the electrode site at which to compute EEG results. Peak latency of the MMN, or peak negativity occurring before the P3 component and after the N1 component of the elicited ERP at the dipole site, was measured (Lang et al., 1995; Martin et al., 2008; Näätänen, 1995, 2008). Descriptive statistics including means and standard deviations were determined for the MMN latency and amplitude in all listeners. Variability across participants was determined by analysis of variance.

Results

A t-test was conducted to determine whether significant differences were present between male and female participants. No significant gender differences, $t(20) \geq 2.00, p \geq .05$, were found.

Behavioral Data

Descriptive statistics were computed for correct responses and for reaction times less than 2000 ms (Table 4). The longest mean reaction time for the /u/ stimulus pairs was indicated in the /u/ Random stimulus condition (1585.43 ms). The longest mean reaction time for the /æ/stimulus pairs, however, was observed in the /æ/ F2 stimulus condition (1523.98 ms); the /æ/ Random mean reaction time was nearly 100 ms shorter (1432.61 ms). The shortest mean reaction times for the /u/ stimulus pairs and /æ/ stimulus pairs were seen in the unfiltered stimulus conditions /u/ (1407.46 ms) and /æ/ (1320.57 ms), respectively. Second shortest mean reaction times were seen in the F1 stimulus conditions for both /u/ (1430.74 ms) and /æ/ (1361.66 ms) stimulus pairs.

An analysis of variance was conducted for reaction time versus stimulus type, $F(7, 260) = 292.94, p \leq .00$. Significant differences were found for reaction time between groups for stimulus type. An LSD post-hoc comparison was conducted for reaction time versus stimulus
type and significant differences were identified between each stimulus type within a block for \( p < .01 \).

Table 4

*Descriptive statistics for reaction time, in ms, across subjects for each stimulus condition*

<table>
<thead>
<tr>
<th>Stimulus condition</th>
<th>( M )</th>
<th>( SD )</th>
<th>Minimum response time</th>
<th>Maximum response time</th>
</tr>
</thead>
<tbody>
<tr>
<td>/u/</td>
<td>1407.46</td>
<td>164.87</td>
<td>907</td>
<td>2000</td>
</tr>
<tr>
<td>/u/ F1</td>
<td>1430.74</td>
<td>145.71</td>
<td>1115</td>
<td>1995</td>
</tr>
<tr>
<td>/u/ F2</td>
<td>1496.48</td>
<td>162.77</td>
<td>1178</td>
<td>1995</td>
</tr>
<tr>
<td>/u/ Random</td>
<td>1585.43</td>
<td>198.79</td>
<td>1251</td>
<td>1991</td>
</tr>
<tr>
<td>/æ/</td>
<td>1320.57</td>
<td>174.96</td>
<td>817</td>
<td>1999</td>
</tr>
<tr>
<td>/æ/ F1</td>
<td>1361.66</td>
<td>156.21</td>
<td>1063</td>
<td>1994</td>
</tr>
<tr>
<td>/æ/ F2</td>
<td>1523.98</td>
<td>213.83</td>
<td>1086</td>
<td>1999</td>
</tr>
<tr>
<td>/æ/ Random</td>
<td>1432.61</td>
<td>185.37</td>
<td>1113</td>
<td>2000</td>
</tr>
</tbody>
</table>

**Event Related Potential Data**

Figures 2 through 7, included below, depict grand-averaged waveforms of Blocks 1 through 6. Table 5, also included below, contains information regarding latency and amplitude of the MMN response for Blocks 1 through 6 and the grand-averaged block.

**Temporal data analysis.** Once difference ERP files were created for each grand-averaged stimulus block, two-dimensional brain maps were created (Figures 2 through 7). Each brain map depicts areas of greatest positivity and negativity in the scalp distribution of the participants’ processing of deviant stimuli during the first 400 ms of stimulus processing.

The Block 1 (Figure 2) scalp distribution yielded greatest positivity in the left frontal area from 74 ms to 149 ms, 124 ms to 149 ms, and 248 to 273 ms. Further positivity was noted in the left temporo-occipital area from 223 ms to 248 ms and in the right posterior frontal area at 372 ms. Areas of greatest negativity were noted in the left temporo-occipital region at 99 ms and 347 ms. Other areas with negativity included the left frontal region at 198 ms and again from 297 ms
to 347 ms, the occipital area from 223 ms to 298 ms, the right frontal area from 248 ms to 273 ms, and the left parietofrontal area from 248 ms to 273 ms.

Block 2 (Figure 3) demonstrated areas of greatest negativity in the scalp distribution in the right posterior frontal region beginning at 124 ms and ending at 223 ms. Negativity was also noted in the left frontal area from 174 ms to 248 ms, again at 372 ms, and in the right posterior parietotemporal area at 248 ms. Large positivity was observed in the left frontal area from 75 ms to 99 ms, at 149 ms, and again from 273 ms to 322 ms. Positivity was also noted in the left parieto-occipital area from 149 ms to 174 ms and in the left posterior frontal area at 198 ms.

Areas of greatest negativity were observed in the Block 3 (Figure 4) scalp distribution from 124 ms to 198 ms and again from 322 ms to 372 ms in the left frontal region. Further negativity was noted in the left temporo-occipital area from 124 ms to 198 ms, the left parietofrontal area from 223 ms to 248 ms, the right frontal area at 248 ms, and in the left temporo-occipital area from 347 ms to 372 ms. Cortical regions marked by large areas of positivity included the left frontal area from 50 ms to 99 ms and again from 223 ms to 273 ms. Further positivity was noted in the right frontal region from 149 ms to 174 ms and again from 347 ms to 382 ms.

The greatest areas of negativity in the Block 4 scalp distribution (Figure 5) were noted in the left posterior temporo-occipital region from 124 ms to 173 ms after the onset of the stimulus. Negativity was also observed in the left frontal areas from 149 ms to 198 ms. Areas of positivity were noted in the left posterior parietal region from 149 ms to 174 ms, in the left frontal region at 223 ms, and in the left posterior temporal region at 273 ms. Further negativity was observed in the left- and mid-frontal area at 298 ms and 322 ms. The greatest positivity was seen in the right posterior frontal area at 322 ms and the left frontal area from 347 ms to 372 ms.
Figure 2. Part A depicts Block 1 grand-averaged brain maps of ERPs for the filtered /æ/ F1 stimulus pair. Orange/red indicates areas of greatest positivity (≥ 2.0 μV) and blue indicates areas of greatest negativity (≤ -2.0 μV). Part B depicts a reconstructed image of the brain taken from axial, coronal, and sagittal views. Locations of the dipole for Block 1 are shown in red. The axial view depicts a source in the inferior parietal lobule. The coronal view shows a source in the postcentral gyrus. The sagittal view yields a source in the precuneus.
Figure 3. Part A depicts Block 2 grand-averaged brain maps of ERPs for the filtered /æ/ F2 stimulus pair. Orange/red indicates areas of greatest positivity ($\geq 2.0 \mu V$) and blue indicates areas of greatest negativity ($\leq -2.0 \mu V$). Part B depicts a reconstructed image of the brain taken from axial, coronal, and sagittal views. Locations of the dipole for Block 2 are shown in red. The axial view depicts a source in the postcentral gyrus. The sagittal view yields a source in the medial frontal gyrus. No dipole was identified from the coronal view.
Figure 4. Part A depicts Block 3 grand-averaged brain maps of ERPs for the filtered random /æ/ stimulus pair. Orange/red indicates areas of greatest positivity ($\geq 2.0 \mu V$) and blue indicates areas of greatest negativity ($\leq -2.0 \mu V$). Part B depicts a reconstructed image of the brain taken from axial, coronal, and sagittal views. Locations of the dipole for Block 3 are shown in red. The axial view depicts a source in the precentral gyrus. The sagittal view yields a source in the medial frontal gyrus and paracentral lobule. No dipole was identified from the coronal view.
Figure 5. Part A depicts Block 4 grand-averaged brain maps of ERPs for the filtered /u/ F1 stimulus pair. Orange/red indicates areas of greatest positivity (≥ 2.0 μV) and blue indicates areas of greatest negativity (≤ -2.0 μV). Part B depicts a reconstructed image of the brain taken from axial, coronal, and sagittal views. Locations of the dipole for Block 4 are shown in red. The axial view depicts a source in the precuneus. The coronal view shows a source in the paracentral lobule, subgyral area, and cingulate gyrus. The sagittal view yields a source in the cingulate gyrus and precuneus.
Additional positivity was noted in the left posterior temporo-occipital region from 322 ms to 372 ms. Positivity observed within the cortex from 322 ms to 372 ms most likely corresponds with the P3 component (Lang et al., 1995; Martin et al., 2008; Näätänen, 1995; Näätänen, 2008).

Pre-attentive negativity in the Block 5 scalp distribution (Figure 6) was similar to that seen in Block 4 in terms of latency and cortical location. A large area of negativity was noted in the left frontal area at 74 ms after the onset of the stimulus. Areas of positivity were found in the left frontal area from 124 ms to 149 ms and in the left posterior temporoparietal region at 149 ms. From 174 ms to 223 ms, negativity was observed in the left frontal area. Positivity was seen in the left frontal area from 248 ms to 273 ms and again from 347 ms to 372 ms. Areas of negativity were noted in left temporo-occipital areas from 273 ms to 322 ms, in the left frontal area from 298 ms to 322 ms, and again in the left posterior frontal area at 372 ms. Positivity was found in the left frontal area from 347 ms to 372 ms and in the right posterior frontal area at 347 ms.

Areas of greatest negativity in Block 6 (Figure 7) appeared in more widespread cortical regions and lasted for longer durations of time than in Blocks 4 and 5. Negativity was first noted in the left frontal region at 74 ms. It was also seen in the left temporo-occipital and left posterior parietofrontal areas at 124 ms. Greatest positivity was observed in the left frontal region from 99 ms to 149 ms and in right posterior frontal regions from 99 ms to 174 ms. Regions marked by greatest negativity were found in the left frontal area from 174 ms to 372 ms, the right parietofrontal area from 223 ms to 273 ms and again from 347 ms to 372 ms, and the left temporo-occipital area from 223 ms to 298 ms. An area of positivity was noted in the left posterior temporal region from 347 ms to 372 ms.
Figure 6. Part A depicts Block 5 grand-averaged brain maps of ERPs for the filtered /u/ F2 stimulus pair. Orange/red indicates areas of greatest positivity ($\geq 2.0 \mu V$) and blue indicates areas of greatest negativity ($\leq -2.0 \mu V$). Part B depicts a reconstructed image of the brain taken from axial, coronal, and sagittal views. Locations of the dipole for Block 5 are shown in red. The axial view depicts a source in the precuneus and paracentral lobule. A source could not be localized from the coronal or sagittal views.
Figure 7. Part A depicts Block 6 grand-averaged brain maps of ERPs for the filtered random /u/ stimulus pair. Orange/red indicates areas of greatest positivity (≥ 2.0 μV) and blue indicates areas of greatest negativity (≤ -2.0 μV). Part B depicts reconstructed image of the brain taken from axial, coronal, and sagittal views. Locations of the dipole for Block 6 are shown in red. The axial view depicts a source in the middle occipital gyrus. The coronal view shows a source in the thalamus and ventral posterior lateral nucleus. The sagittal view yields a source in the lingual gyrus.
Examination of the scalp distribution of the grand average of all stimulus pairs showed greatest negativity in the left frontal area from 174 ms to 223 ms and again from 298 ms to 347 ms. Areas of greatest positivity were noted in the left frontal region at 124 ms and from 248 ms to 273 ms.

Average latencies and amplitudes of the MMN response were computed for each stimulus block (Table 5). The beginning amplitude of the MMN for the /u/ phoneme was greatest during the presentation of Block 6, in which a band of spectral energy which did not correspond to F1 or F2 was removed. Block 4, in which spectral energy corresponding with F1 was digitally deleted, yielded the smallest MMN beginning amplitude. Removal of spectral energy corresponding to F2 resulted in a beginning amplitude which was greater than that of Block 4 but less than that of Block 6. Results were similar for the /æ/ phoneme and for the /u/ phoneme. The greatest MMN beginning amplitude resulted from the presentation of Block 3. Block 1, in which spectral energy corresponding to F1 was removed, demonstrated a slightly smaller MMN beginning amplitude than that exhibited by Block 3. Block 2, in which spectral energy corresponding to F2 was removed, displayed the smallest MMN beginning amplitude.

Table 5

Average latency, in ms, and amplitude, in μV, of the MMN for each stimulus block

<table>
<thead>
<tr>
<th>Block</th>
<th>Beginning Latency</th>
<th>Maximum Latency</th>
<th>End Latency</th>
<th>Beginning Amplitude</th>
<th>End Amplitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>113.0</td>
<td>168.2</td>
<td>195.2</td>
<td>-0.624</td>
<td>-0.164</td>
</tr>
<tr>
<td>2</td>
<td>179.6</td>
<td>222.2</td>
<td>241.4</td>
<td>-0.380</td>
<td>0.237</td>
</tr>
<tr>
<td>3</td>
<td>222.2</td>
<td>265.0</td>
<td>276.4</td>
<td>-0.650</td>
<td>-0.346</td>
</tr>
<tr>
<td>4</td>
<td>181.8</td>
<td>203.2</td>
<td>275.2</td>
<td>-0.013</td>
<td>0.369</td>
</tr>
<tr>
<td>5</td>
<td>171.6</td>
<td>203.2</td>
<td>195.2</td>
<td>-0.674</td>
<td>-0.164</td>
</tr>
<tr>
<td>6</td>
<td>128.8</td>
<td>168.2</td>
<td>210.0</td>
<td>-1.031</td>
<td>-0.338</td>
</tr>
</tbody>
</table>


**Dipole localization and analysis.** Dipoles were reconstructed and analyzed from sagittal, axial, and coronal views on a three dimensional model for the purpose of determining source localization. Block 1 demonstrated a dipole in the areas of the precuneus, inferior parietal lobule, and postcentral gyrus (Figure 2). Localization analysis for Block 2 showed a dipole in the areas of the medial frontal gyrus and postcentral gyrus (Figure 3). A dipole was identified for Block 3 in the cortical areas in and surrounding the precentral gyrus, medial frontal gyrus, and paracentral lobule (Figure 4). Block 4 yielded a dipole in the areas of the cingulate gyrus, precuneus, and paracentral lobule (Figure 5). A dipole was identified within the cortex for Block 5 in the precuneus and paracentral lobule areas of the cortex (Figure 6). Analysis of Block 6 showed a dipole in the middle occipital gyrus, lingual gyrus, thalamus, and ventral posterior lateral nucleus areas of the cortex (Figure 7). The grand average file of all stimulus pairs yielded a dipole in the cingulate gyrus.

**Repeated Measures**

Behavioral data and event-related potentials were acquired a second time from three participants one week after the initial test to determine test-retest reliability. Individual and grand averaged ERP waveform files were created for each test-retest participant. A $t$ test was conducted on the grand averaged ERP waveform files for the Block 1 file. No significant differences, $t(3) \geq 2.00$, $p>.05$, were found in any of the three participants for repeated measures, indicating that an acceptable level of test-retest reliability was established.

**Discussion**

The aims of the current investigation were to determine if the MMN response is evoked by natural speech sounds and if the amplitude of the MMN changes significantly following the presentation of deviant stimuli from which spectral energy within formant bands critical to vowel
identification has been removed. Additionally, the study sought to determine whether localization of dipoles within the cortex yields information pertaining to the processing of formant-specific linguistic information. The results of the current investigation bring into question the methods of many previous studies and refute the assumption that auditory processing of linguistic information is not present preattentively. Conclusions drawn by the author increase the existing literature supporting a formant-based theory of linguistic processing and confirm that the MMN response occurs following the presentation of natural speech sounds. Furthermore, the results of the present investigation provide a foundation for future research regarding formant-based preattentive processing of natural speech sounds.

**Summary and Evaluation of Results**

Beginning amplitudes of the MMN response, or the peak amplitudes of negativity for each block, were greatest for Blocks 3 and 6. Large MMN amplitudes reflected cortical processing of auditory stimuli in which bands of spectral energy unimportant to vowel identification were deleted. Small MMN amplitudes reflected the presentation of auditory stimuli in which spectral energy corresponding to formants was removed. Consequently, it is possible that the brain separates vowel information into smaller auditory components, namely the first few formants, for preattentive processing; furthermore, the brain may identify gaps of spectral energy not corresponding with formant frequencies as more deviant than gaps directly corresponding to familiar formant frequencies. The data suggest that formant information and identification contributed to preattentive processing of auditory stimuli in the cortex.

Results of the current study support the findings of Edmonds et al. (2010), who reported that the MMN response reflects early speech processing and that formant information in steady-state vowels contributes to speech processing. Edmonds and colleagues suggested that
subcortical structures may be involved in the processing of formant information from vowels, thereby supporting a notion of formant-based preattentive speech processing. The current investigation added to these results by demonstrating that the MMN response is elicited for natural speech sounds and that the amplitude is dependent upon formant information for each steady-state vowel.

Localization of dipoles within the cortex yielded important information regarding the processing of formant-specific linguistic information. One cortical site in which a dipole was identified, the precuneus, is located in the superior parietal lobule between the two cerebral hemispheres. While it is typically associated with conscious experience and visuospatial processing, the posterior region of the precuneus is thought to play a significant role in episodic memory retrieval (Cavanna & Trimble, 2006). Jacobsen (2004) found that there may be a long-term memory-based categorical processing of formants. It is possible that the precuneus plays a role in categorical processing of formants if and when memory is involved.

The cingulate gyrus was identified as a dipole source for the single grand averaged file of all stimulus blocks. The cortex of the cingulate gyrus is located in the cingulate cortex, which is directly superior to the corpus callosum (Binder et al., 1997). Previous imaging studies have identified the cingulate gyrus as an essential part of language processing (Binder et al., 1997; Yetkin et al., 1995), although the cingulate gyrus as a whole is typically examined in the context of its integral role in the limbic system (Binder et al., 1997). Because of its role in memory and learning, there is a possibility that the cingulate gyrus, like the precuneus, may play an essential role in long-term memory-based categorical processing of formants (Jacobsen, 2004).

A dipole was identified in Blocks 3, 4, and 5 within the paracentral lobule, which is posterior to the precuneus and superior to the cingulate gyrus. The precentral gyrus, in which a
A dipole was found for Block 3, and the postcentral gyrus, in which a dipole was found for Blocks 1 and 2, are continuous with the paracentral lobule. Although the areas identified are usually affiliated with sensory and motor processes in the body, the paracentral lobule, precentral gyrus, and postcentral gyrus have been identified previously as cortical sites essential to language processing (Vitacco, Brandeis, Pascual-Marqui, & Martin, 2002).

Vitacco and colleagues (2002) also identified the middle occipital gyrus, lingual gyrus, thalamus, and ventral posterior lateral nucleus, sites at which the current author identified dipoles for Block 6, as cortical areas important to language processing. The lingual gyrus, which is located in the occipital lobe, has been recognized for its importance in dreaming and vision, particularly for its role in recognizing words. Occipital areas important to language processing have also been examined in the context of letters and sounds rather than just words (Jessen et al., 1999). The ventral posterior lateral nucleus is the nucleus of the thalamus. The thalamus, which is situated between the midbrain and cerebral cortex, may not serve a primary function in language processing, but is activated for semantic language tasks (Binder et al., 1997).

Activation of the inferior parietal lobule during the presentation of Block 1 may be attributed to the button press task, as the cortical region is often activated during tool use. However, the inferior parietal lobule and surrounding cortical areas have been shown to be affiliated with language tasks (Vitacco, 2002; Yetkin, 1995). The medial frontal gyrus, which was identified as a dipole for Blocks 2 and 3, plays an essential role in executive decisions. Its activation may also be attributed to the button press task, although a recent study noted that the medial frontal gyrus is involved in language-related tasks (Talati & Hirsch, 2005).

The findings of the study were considerable. First, the MMN response may be evoked by natural speech sounds. Second, amplitudes of the MMN were greatest for pairs of auditory
stimuli in which spectral energy not corresponding to formant frequencies was removed. Last, dipoles reconstructed from temporal ERP data were located in cortical areas known to support language and auditory processing.

**Recommendations for Future Research**

Findings from the current investigation may have considerable theoretical implications for future research. Further studies are recommended for determining temporal and cortical differences in pure tone and natural speech processing. It may benefit the ERP and acoustical community at large to replicate previous studies examining electrophysiological responses to speech information in the cortex (Aaltonen, 1994; Aaltonen & Suonpaa, 1983; Alho et al., 1990; Cheour et al., 1998; Diesch & Luce, 2000; Edmonds et al., 2010; Jacobsen, 2004; Miner & Danhauer, 1977; Näätänen et al., 1997; Neff, 2010; Winkler et al., 1990; Winkler et al., 1999) by replacing compounded pure tones with natural speech stimuli. Future researchers may also wish to expand upon the current investigation by examining processing of vowels in words rather than steady-state vowels in isolation. Comparisons between words with filtered vowels and words in which vowels have been replaced by synthesized vowels may yield additional information regarding differences in auditory processing of pure tones versus natural speech sounds.

Future investigators may also wish to examine the degree of linguistic processing which occurs preattentively by differentiating between acoustic and linguistic processes. It may be informative to compare preattentive linguistic processing to later linguistic processing, as reflected by the P300 and N400 responses. The nature of the MMN response may be examined in further detail as well. Dialectical studies using vowels may shed light on how the MMN response changes depending on familiarity of the vowel. Decoupling of the voice bar from the
first and second formants may also elicit an MMN response, which could lead to new
information regarding how formants play a role in preattentive linguistic processing.

The current investigation leaves several questions to be answered: Assuming that aspects
of linguistic processing do occur preattentively, how early on after the onset of the auditory
stimulus does the brain begin to separate out pure tones and nonspeech stimuli from natural
speech sounds? Where in the brain does this occur? Future investigations are recommended for
resolving queries and exploring areas pertaining to preattentive processing of linguistic
information.

In essence, research intended to examine cortical processing of natural speech sounds
should use natural speech sounds as stimuli. Jacobsen (2004) found that synthetic tones
representing formant values do not trigger speech perception. There may be some long-term
memory-based categorical processing of formants, but synthetic representations of speech are not
adequate to elicit this information. Because pure tones and natural speech sounds are processed
differently, the use of synthesized speech, particularly two or three pure tones compiled to
represent formants in human speech, may confound the study and produce inaccurate results.
References


Annotated Bibliography


**Objective:** The authors investigated the perception of vowel pitch and formant frequency by acquiring the mismatch negativity (MMN), an auditory event-related potential. **Methods:** Recorded vowels and pure tones were presented through earphones after the participant was fitted with 21 Ag/AgCl electrodes. The electrodes were fitted onto an electrode cap to 16 adult participants (nine females and seven males between 21 and 42 years of age). Sixteen pairs of standard and deviant vowels, differing in slight alterations of the voice bar (F0) and second formant (F2) frequencies, were matched with pure tone frequencies as the control. The standard response was subtracted from the deviant response for each stimulus block. The MMN onset latency, peak latency, and peak amplitude were manually scored from the difference curves. **Results:** MMN amplitude increased between all paired standard and deviant stimuli, although more dramatic results were obtained from pure tone presentations. **Conclusions:** The authors concluded that the MMN was elicited by all manipulated parameters in the experiments. Additionally, vowels produced lower MMN amplitudes than the pure tones, but only within specific (4% and 20%) frequency differences. **Relevance to the current work:** This study was one of the first to examine vowel perception through human auditory psychophysiology rather than employing traditional behavioral measures. It demonstrated that even small and phonetically irrelevant changes in F2 frequency result in a heightened MMN response.


**Objective:** The present experiment was designed to investigate the perceptual magnet effect using behavioral and psychophysiological (mismatch negativity, or MMN) discrimination data. **Methods:** The authors produced synthetic vowels with varying second formant (F2) values from 1520 Hz to 2966 Hz in steps of 30 mels. Participants, 13 Finnish speaking young adults between the ages of 21 and 32 years (nine females and four males), were presented with a same/different discrimination task in which they were instructed to label the vowels with differing F2 values as either /i/ or /y/. Each of 19 vowel variants was presented in random order 15 times. Results were averaged and all data with significant artifacts were removed before averaging. **Results:** “Good” categorizers consistently identified the synthetic vowels with low F2 values as /i/, or the prototype, whereas “poor” categorizers performed in an opposite manner and labeled these vowels as belonging to the nonprototype grouping. MMN values for the poor categorizers did not differ significantly between the prototype and nonprototype groupings, but the good categorizers demonstrated lower MMN values for the prototypes than the nonprototypes. **Conclusions:** This reveals that preattentive processing (the MMN) is sensitive to differences within a vowel category. Psychophysiological data indicated that both groups of listeners behaved similarly in attentive discrimination at about the same F2 location, meaning that the P300 results were similar across all participants. **Relevance to the current work:** Results from
this study indicate that the MMN may be attributable to processes other than acoustical stimulus differences.


**Objective:** The purpose of the study was to determine whether the violation of any phonological constraints in an individual’s native language elicits a mismatch negativity (MMN) response. **Methods:** Stimuli (the prototypical, highly ambiguous [æ] and the nonprototypical, highly ambiguous [ε]), consisted of synthetic vowels in isolation and the same vowels in a non-word context to avoid lexical effects in both languages. The only difference was in the second formant. The stimuli were presented in two experimental conditions to participants, including 10 Finnish (22-35 years of age) and 10 Estonian (20-43 years of age) females with normal hearing. All participants were fitted with Sn electrodes with two reference electrodes and the resulting data were averaged. **Results:** Stimuli presented to the participants elicited different MMN responses based on native language. The MMN differed between Finnish and Estonian listeners when the stimuli were the same. There was an intergroup difference at the latency of approximately 400 ms due to a late negative waveform in the Finnish group. **Conclusions:** Based on the results, it was concluded that the violation of any phonological constraint in an individual’s native language does elicit an MMN response. Alterations in the MMN are language-based depending on familiarity with phonological elements of stimuli. **Relevance to the current work:** The notion that the MMN response has some language component and phonological elements indicates that the MMN cannot be solely attributed to acoustic processing and perception.


**Objective:** The aim of this study was to use event-related potentials (ERPs) to examine competing phoneme discrimination predictions, namely the acoustic and categorical models of perception. **Methods:** Nine right-handed, normal hearing, native Finnish-speaking individuals (eight female, one male, 20-34 years of age) participated in the study. Synthesized stimuli, endpoints and defined intermediate points along the Finnish /i/-/y/ continuum, were presented in 12 blocks of 300 total stimuli to participants in an oddball (standard-deviant) paradigm. Results were computed at most electrode sites, particularly at the Fz and Cz sites. All data were analyzed via averaging and comparison of stimuli. **Results:** The mismatch negativity (MMN) response increased in amplitude and decreased in latency when the pure vowels /i/ and /y/ were presented to participants. A smaller MMN with a longer latency occurred following the presentation of a pure vowel as the standard and a boundary, or intermediate, vowel as the deviant. **Conclusions:** The results of this study support a more hybrid model of discrimination. It suggested the notion that phoneme discrimination occurs at a more acoustic, basic physiological level. A cognitive categorical component did emerge, however, given the different latencies and P300 responses during the presentation of /i/ and /y/. **Relevance to the current work:** The current study relies on both the acoustic and categorical models of perception. Aaltonen et al. show insight into both predictions of phoneme discrimination.

**Objective:** The intent of this study was to create a Finnish vowel chart by plotting the frequency locations of the first and second formants (F1 and F2) of each of eight vowel phonemes on a two-dimensional plane. **Methods:** Included in the study were 32 participants (native Finnish speakers) with normal hearing. Listeners were instructed to listen to 511 synthetically produced steady state vowel-like stimuli and indicate in writing which vowels they heard. F1 and F2 varied from 250-800 Hz and 800-2400 Hz, respectively. Each stimulus lasted in duration for 300 ms with an onset time of 5 ms. All data were averaged and results were derived from the subtraction of standard stimuli responses from deviant stimuli responses. **Results:** Results indicate that listeners were capable of vowel identification based solely on the presentation of the first two formants. Additionally, the participants identified vowels at specific formant frequencies. **Conclusions:** There may be some neural correlate linking formant frequency and vowel identification. **Relevance to the current work:** The significance of this study extends beyond the creation of a Finnish vowel chart. The importance of the first two vowel formants cannot be underestimated, as vowel discrimination and identification are based upon these two formants.


**Objective:** The purpose of this study was to examine whether the discrimination of synthetic vowels is impaired in aphasic individuals with left posterior brain damage. **Methods:** The authors measured the mismatch negativity (MMN) response in four aphasic participants (right-handed Finnish-speaking males with no hearing deficits) because of the non-attentive nature of the MMN. All participants were fitted with an electrode cap with a three-channel derivation. They were asked to listen to speech stimuli consisting of three vowels (along the Finnish /i/-/y/ continuum) produced by a speech synthesizer. **Results:** The two participants with more anterior lesions demonstrated an MMN response, although it was not as robust as might be expected in a healthy non-aphasic individual. The two participants with more posterior lesions demonstrated no MMN response at all. **Conclusions:** This study demonstrated that the MMN may be used to investigate automatic sensory processes underlying auditory and speech perception. It was also concluded that deviant MMN responses arise from participants with brain lesions. **Relevance to the current work:** This study is significant to further academic endeavors in that it demonstrated that participants in event-related potential studies must be healthy with no background of head trauma or neurological impairment. Inclusion of individuals like the ones in the current study may yield a deviant MMN response or no MMN at all.

Objective: The goal of this investigation was to examine the perceptual effect elicited by varying the amplitude of formants in a synthesized, two-formant vowel. Methods: The stimuli, two-formant synthesized vowels varying in duration between 240 and 480 ms, were presented over a series of two sessions in isolated conditions and also within the /h-d/ context. Participants were asked to identify the sound stimuli as vowels. Analysis was performed by calculating the center and the standard deviation around the center of each vowel area according to responses. Results: Data analysis suggested that as the relative amplitude of the second formant was reduced, identity of the vowel was independent of the formant amplitude in dB up to a certain point. After passing a certain point, however, the identity of the vowel changed such that it aligned to a different vowel with the same first formant. Additionally, subtle changes in amplitude which did not affect identity of the vowel in an isolated context did alter the identity of vowels in the /h-d/ context. Conclusions: It was concluded that when formant amplitudes are attenuated more than 28 dB, various biases of listeners as well as “a possible subsidiary mechanism” controls the perceived response. Relevance to the current work: When formant amplitudes were attenuated less than 28 dB, a less stark contrast was demonstrated in the listeners’ perceptual differentiation between vowels. Additionally, some vowels were mistaken for others. It is important, then, to completely eliminate selected formant frequencies from a recording rather than attenuate them for a true test of formant perception rather than a test of vowel identification.


Objective: The aim of this study was to determine whether an attentional-trace theory of selection attention is both correct and active during event-related potential (ERP) selective listening tasks. Methods: Tone stimuli, “relevant” stimuli and more common “irrelevant” stimuli differing in pitch, were produced and presented to nine participants (19-32 years of age, six males, three females). Relevant stimuli were varied in position and occurrence in each block. The electroencephalogram (EEG) was recorded using AgCl electrodes at four scalp locations and at the left mastoid as a reference. ERPs were averaged separately for all participants. Results: Processing negativity was elicited for all presentations of relevant stimuli. The amplitudes were, however, larger when relevant stimuli were presented more frequently amidst the irrelevant stimuli. Conclusions: It was concluded that an attentional-trace theory of selective attention is active when relevant stimuli are presented to the listener. Relevance to the current work: The current study utilizes an active mismatch negativity (MMN) task in which the listener is asked to selectively listen for deviant stimuli. An attentional trace cannot be found in the MMN.


Objective: The purpose of the published set of guidelines was to establish norms and standards for screening hearing impairments and middle-ear disorders. By publishing the guidelines, ASHA ensures that screenings are conducted ethically and that results obtained during the screening reflect true values throughout the nation. Relevance to the current work: All participants were required to undergo screening procedures to ensure that they qualified for the
behavioral and ERP data collection. The screening procedures were in accordance with ASHA guidelines and specifications.


**Objective:** The author investigated the formant center of gravity (FCOG) hypothesis, which states that if two formant vowels are in close proximity, the listener collapses them into one formant during perceptual processes and vowel quality is dependent upon where the collapsed formant lies along the frequency spectrum. **Methods:** Five participants with normal hearing listened to phonetic stimuli with varying formant amplitudes. The vowels presented had close spacing of the first and second formants (F1 and F2, respectively) and were presented in a two-formant matching paradigm. The listeners were instructed to find the best match in vowel quality by changing the formant amplitudes of F1 when F2 was lower and F2 when F1 was lower. They did so by adjusting a lever. Eight participants (four men and four women) were asked to listen to the same stimuli and identify the vowel, which changed depending on frequencies attenuated. **Results:** Results were inconsistent with the FCOG hypothesis in that formant amplitude significantly affected the vowel quality. **Conclusions:** The FCOG hypothesis was disproved in that two formant vowels in close proximity were not collapsed into one formant during perceptual processes. **Relevance to the current work:** Any decrease of formant amplitude of elimination of the formant entirely may be a good determiner of perceived or non-perceived vowel recognition capabilities.


**Objective:** The investigators sought to examine the topographic distribution of the event-related brain potential mismatch negativity (MMN) in aphasic patients versus the MMN distribution in healthy non-aphasic controls. **Methods:** Harmonically rich tones differing in duration and consonant-vowel (CV) syllables were presented as unattended oddball paradigms to 33 (18 aphasic and 11 control) right-handed Norwegian-speaking participants. All participants were fitted with a 19-electrode cap and two mastoid-placement electrodes to record event-related potentials continuously throughout testing. Data were filtered for artifacts, which were discarded, and all results were averaged. **Results:** Findings showed that the distribution of the MMN was less lateralized and centralized throughout the cortexes of the aphasic participants, specifically in CV syllables. MMN-amplitudes were attenuated in the brain-damaged hemisphere in the aphasic group. **Conclusions:** The topographic distribution of the MMN in aphasic patients was significantly different than that of the healthy non-aphasic controls. The non-aphasic controls demonstrated a more lateralized and centralized distribution for the MMN across the cortexes, particularly in CV syllables. **Relevance to the current work:** These findings are significant in that they indirectly oppose the findings of the study by Tampas, Harkrider, and Hedrick (2005) which concluded that an MMN response could not be elicited from the presentation CV syllables.

**Objective**: The purpose of the study by Beddor and Hawkins was to examine the factors that govern the center of gravity (COG) effect. The authors hypothesized that the influence of the formants on perceived vowel quality increases as low-frequency formants become less well-defined. **Methods**: A series of three experiments were conducted. In the first, five sets of multiformant nasal and nonnasal vowels were examined. Participants included 20 students between the ages of 18 and 25 years, all of whom were native speakers of American English. **Results**: The authors found that when two or more formants resided within 3.5 Bark of one another, the boundaries within which the perceptual COG lies were roughly determined. Additionally, when formant bandwidths are narrow, formant frequencies are apparent. Overall spectral shape is more influential in the spectrogram when spectral prominences are wide. **Conclusions**: Vowel quality is determined by two factors: the frequency of the most prominent harmonics in the low-frequency region, and the slopes of the skirts in the vicinity of these harmonics. **Relevance to the current work**: All of the spectral energy residing within a specific band of frequencies was removed from recordings in the current study.


**Objective**: This study set out to confirm the phenomenon of duplex perception by demonstrating that double dissociation, a concept which describes the perceived differences in sound signals when formant transitions and alterations in the harmonic structure of a transition are presented to separate ears, exists. **Methods**: Participants were presented with speech and nonspeech (chirp) stimuli over a series of five experiments. The speech stimuli were derived from a single formant transition in isolation. All five experiments examined different combinations of the speech and chirp stimuli with varying masking components. The same 10 participants were used for experiments I and II; 12 undergraduates were enlisted for experiment III and IV in which synthetic constructions of the consonant-vowel (CV) syllables /da/ and /ga/ were presented with varying noise masking conditions. The final experiment utilized nine undergraduates who correctly identified the /da/ and /ga/ syllables and chirps. **Results**: Findings indicated that nonspeech chirps were identified better when formant transitions were asynchronous. **Conclusions**: The study confirmed that participants achieve a higher level of performance when formant transitions are presented in isolation. **Relevance to the current work**: Formant transitions are identified more accurately when presented in isolation, so it is informative to employ isolated sound segments utilizing these formant differences.


**Objective**: The purpose of the study was to use functional magnetic resonance imaging (FMRI) to examine and identify language processing areas in the intact human brain. **Methods**: Thirty
right-handed subjects (15 women and 15 men, all of whom were healthy with no history of neurological, psychiatric, or auditory symptoms, ranging in age from 18 to 29 years) listened to words (“semantic decision”) requiring phonetic and semantic analysis. Stimuli in the semantic decision task consisted of English nouns for different animals with an average use frequency of 9.3 per million. Participants were instructed to press a button in response to stimuli based on whether they believed the animal to be native to the United States and used by humans. The participants also listened to sets of nonlinguistic sounds (“tone decision”) which served as a control task involving perceptual analysis. The tone decision task consisted of 500 Hz and 750 Hz pure tones played in a sequence of three to seven tones. Functional maps of the entire brain were acquired for every participant and were averaged using stereotaxic space. A split-half procedure was used to determine how well each task would generalize to other subject samples.

Results: Left hemisphere temporoparietal language areas were identified in the middle temporal, inferior temporal, fusiform, and angular gyri. Several prefrontal language areas were also found, all of which participated in tasks emphasizing receptive language functions. Conclusion: The findings refuted earlier works supporting a more classical model of language localization, although still supported current lesion data. Relevance to the current work: Several cortical areas examined in the study by Binder and colleagues, including the cingulate gyrus and other temporoparietal gyri, were identified in the current study as dipoles.


Objective: Boersma and Weenink created Praat, a free scientific software program, for speech and phonetic analysis. Relevance to the current work: Praat 5.2.26 was used in the current study to create spectrograms of the /æ/ and /u/ vowels.


Objective: The purpose of this review was to summarize current knowledge about the macroscopic and microscopic anatomy of the precuneus, its connections with other portions of the cortex, and its roles in a spectrum of highly integrated tasks. Conclusion: The precuneus is involved in a series of different tasks including visuo-spatial imagery, episodic memory retrieval, and consciousness. The consciousness aspect can be more clearly defined as first-person perspective taking and the agency experience. Additionally, the precuneus is highly connected with other cortical regions and is indirectly involved in several higher-order cognitive functions. Relevance to the current work: The current examination identified a dipole in the cortical region surrounding the precuneus.


Objective: The authors used electroencephalograms (EEGs) to investigate the development of memory traces specific to language in the brain. Methods: Estonian- and Finnish-speaking families provided consent for their six to 12 month olds to participants in the study. Nine
Estonian (six male and three female) and nine Finnish (seven male and two female) monolingual infants without any observed hearing loss were included. The infants were fitted with frontal (Fz), central (Cz), left-mastoid, and eye-placement electrodes. They were presented with a same-different task in which a standard /e/ phoneme was at random replaced with the deviant /ö/ or /õ/ stimulus. 

**Results:** Speech-recorded deviant stimuli caused significant mismatch negativity (MMN) responses with increased amplitudes. These findings were variable depending on the age and primary language of the infant. 

**Conclusions:** Findings indicated that language-dependent memory traces are created in the brain before 12 months of age. 

**Relevance to the current work:** Only native speakers of English may participate in the data collection portion of the current examination because of the language-dependent memory traces found by the authors.


**Objective:** The purpose of this study was to determine whether the amplitude and latency of the mismatch negativity (MMN) response was the same for healthy individuals across all ages. 

**Methods:** Synthesized stimuli, the Finnish vowels /y/ and /i/ in which the only difference was the second formant, were presented to participants in an oddball paradigm. Participants included 12 newborns (one to five days of age), three-month-old infants (six infants, 89-104 days of age), and 11 pre-term infants (gestational age 25-34 weeks, conceptional age 30-35 weeks). ERPs were averaged separately for standard and deviant stimuli, and only data including at least 100 acceptable EEG epochs were analyzed. 

**Results:** The infant MMN amplitude appeared to be the same as the MMN amplitude elicited in adults. The only difference seen was in latency, which increased in participants older than three months of age. 

**Conclusions:** The amplitude and latency of the MMN response is the same in healthy individuals across all ages. The only difference which can be seen is the latency response, which lengthens slightly once the infant reaches three months of age. 

**Relevance to the current work:** This study indicates that the MMN can be traced in all individuals, regardless of age.


**Objective:** Prior to this study, little research had investigated the brain mechanisms of vowel perception in infants. This study examined whether the mismatch negativity (MMN) response existed and could be identified in the event-related potentials (ERPs) of infants. 

**Methods:** Participants, including 12 healthy newborns (four female, eight male, one to five days old), were fitted with silver-chloride electrodes at the F3, F4, C3, C4, P3, and P4 scalp sites to acquire ERPs. Stimuli included synthesized Finnish vowels along the /i/-/y/ continuum as well as a boundary /y/i/ stimulus identified as either /i/ or /y/. All formants were fixed. The ERPs were separately averaged for standard and deviant responses during quiet sleep lasting for 20-30 minutes. 

**Results:** Difference waves were obtained by subtracting ERPs from standard stimulus /y/ from those of the deviant stimulus /i/. Consequently, MMN responses were identified in the ERPs of infants, and the MMNs found resembled those of adults. 

**Conclusions:** The MMN is
present in infants and may be used to examine brain mechanisms of vowel perception in infants.

*Relevance to the current work:* This study indicates that the MMN can be traced in all individuals, regardless of age.


*Objective:* CURRY 6 software was designed to overlay temporal electrophysiological data onto brain images and reconstruct sources of electrical activity. *Relevance to the current work:* CURRY 6 computer software was used in the current study for the purpose of dipole localization.


*Objective:* NeuroScan 4.5 software records EEG data as they are collected and has built-in analysis capabilities to filter and edit data after collection. *Relevance to the current work:* NeuroScan 4.5 computer software was used in the current study during EEG data collection. EEG data were streamed onto a computer using the software.


*Objective:* Stim 2 software was designed to stream auditory stimuli from the computer to the participant. *Relevance to the current work:* Stim 2 computer software was used in the current study during stimulus presentation.


*Objective:* The aim of this study was to determine whether the mismatch negativity (MMN) is elicited by acoustically different speech contrasts which are not perceived as being different by participants in a behavioral task. *Methods:* Synthesized speech syllable contrasts along the /da-ga/ continuum were presented to 12 female participants (23-31 years of age, normal hearing bilaterally). The stimuli presented included the continuum endpoints /da/ and /ga/, those synthetic recordings along the listener’s categorical boundary, and a non-perceivably different within-category contrast. They included five formants and were presented in a two-alternative, forced-choice, same/different discrimination task. All responses were averaged, and average responses to standards were subtracted from average responses to deviants. *Results:* There was no obvious relationship between behavioral performance and the presence of the MMN response in all participants. Mean onset latency was shortest for the 7-9 contrast condition. *Conclusions:* It was concluded that the MMN response may be used to examine the neurophysiology underlying the ability to discriminate acoustic speech parameters; it is not a correlate or determiner of ability to behaviorally discriminate speech. *Relevance to the current work:* The current study records the MMN response to determine whether speech formants contribute to vowel recognition.
Dalebout reports that the MMN response will reflect any acoustic changes in the stimuli, including the absence of specific vowel formants.


**Objective:** The purpose of the study was to investigate whether a mapping exists for the vowels [a], [æ], [u], and [i], between a hypothetical vowel space and cortical response space. The authors interpreted the mapping data to determine the spatial distribution of vowel evoked equivalent current dipoles. **Methods:** Synthetic vowel stimuli were presented to 11 native German-speaking participants (24-41 years of age, four females and seven males, right-handed, normal hearing) who wore a sensor array between the T3 and the C3 scalp positions. The auditory evoked neuromagnetic field was measured sequentially over both the left and right hemispheres, and the N100m and SF responses were recorded. Data analysis included equivalent current source analysis in which the model of a single dipole was utilized. **Results:** Some minor differences appeared in hemispheric vowel processing. Ordering of distances between the most disparate vowels (at each corner of the vowel quadrilateral) appeared to correspond highly with the ordering of distances between N100m and SF dipole locations. **Conclusions:** It was concluded that all tested spatial distributions of equivalent current dipole sources reflect an intermediary stage in processing between phonetic and auditory representation. **Relevance to the current work:** This study pertains to the current work due to its findings pertaining to the relationship between the vowel quadrilateral and auditory and phonetic processing.


**Objective:** The authors hypothesized that a phonetic mode of processing makes use of nontonotopic aspects of auditory cortex organization. The purposes of the study were to determine whether the neuromagnetic N1m and SF responses elicited by single vowel formants are related to formant and tone frequency, and how single tones and formants correlate to composite tones and vowels. **Methods:** A two-formant vowel heard as either /æ/ or /ɛ/ was synthesized with frequencies at 600 Hz and 2100 Hz and presented to 14 normal-hearing participants (22-42 years old, six females). The neuromagnetic field was measured with a 37-channel magnetometer device. Dipoles were determined with summation rule analysis and N1m peak latency was analyzed with an ANOVA design. **Results:** Data analysis revealed that linear summation sources were roughly twice as strong in dipole moment. The single vowel formants elicited strong responses pertaining to both formant and tone frequency. **Conclusions:** The vowels were determined by their respective spectral frequencies of their most prominent harmonics, at least at the level of the auditory cortex. Additionally, virtual frequency was determined by the spacing of the harmonics. **Relevance to the current work:** Diesch’s work suggests a more formant-based theory of auditory speech processing in which spectral frequencies and harmonics take the forefront.
Objective: It was not known previously whether spatial location or strength of the source of a mismatch response could be used as a predictor of loss and preservation of the mismatch response in aphasic patients. The goal of this study was to examine spatial location and strength elicited by vowels and plosive stops in neurologically healthy participants to determine which may be responsible for the mismatch response. Methods: The experiment was divided into two separate investigations. The first required 13 native German-speaking participants (22-44 years of age, right-handed with no history of hearing loss or neurological impairment, two women) to determine vowel-consonant difference on location, dipole moment, and peak latency of the auditory-evoked mismatch field (MMF). Stimuli included specific phonemes as well as the syllables /da/, /ba/, and /ga/, all of which were synthesized. Orthogonal spatial coefficients were used in data analysis to find field averages. Results: Deviant waveforms elicited distinct MMFs at a latency of roughly 100 ms. In Experiment 2, the /da-ba/ vowel contrasts did not differ in MMN dipole orientation, whereas the /da-ga/ vowel contrast did. The /da-ba/ vowel contrast elicited a more anterior response, whereas the /da-ga/ vowel contrast elicited a more posterior response. Conclusions: Magnetic MMFs were elicited through the presentation of vowel contrasts and plosive-stop consonant place-of-articulation contrasts. Both spatial location and strength could be used as a predictor of loss and preservation of the mismatch response in aphasic patients. Relevance to the current work: The authors identified cortical locations of the mismatch response using dipole source analysis, a method which may prove useful in the current examination. Furthermore, this study of MMFs catalogued auditory-evoked mismatch responses elicited by vowel contrasts and provided an excellent review of previous research regarding electroencephalographic data acquired from the presentation of vowel contrasts.

Objective: The current study investigated neural substrates of spectral envelope information and organization by examining details of spectral composition. More specifically, it determined whether there are stimulus-specific contributions to the vowel-evoked response from auditory fields. Methods: Stimuli, single vowel formants and two-formant vowels at 200, 400, 800, and 2600 Hz, were created by additive superposition of formants and were presented to participants. Fifteen normal hearing individuals (age range 22-43 years, seven women) were asked to identify a target vowel amidst seven nontarget vowels. The P50m and N100m responses were recorded and inversely correlated with the formant frequency of single formants. Results: The authors found a substantial effect of formant frequency of single formants. Higher-frequency formants yielded more anterior N100m sources, which decreased as a function of latency. Additionally, a positive correlation was drawn between amount of deceleration and formant frequency. Conclusions: It was concluded that multiple formant frequencies presented together such as in a true vowel interact in at least one stage of the afferent auditory pathway. Activity in these auditory fields varies with formant frequency or may reflect sharpness of tuning. Relevance to the current work: The investigators studied the formation of invariant phonetic percepts.
frequencies which represent peaks in the spectral envelope may be neurally encoded even though there is no apparent tonotopic mapping of the first formant frequency.


**Objective:** The purpose of this study was to examine the identification of digitally-altered front vowels in normal-hearing listeners. **Methods:** Twelve American vowels were generated in a /bVt/ context. The vowels were digitally altered such that high frequencies typically lost in sensorineural hearing were not present, thereby mimicking sensorineural hearing loss in normal-hearing participants. Each of the test items was 315 ms in duration. The words were organized into blocks consisting of 2400 stimulus trials and separate experiments. The stimuli were presented to 10 undergraduate and graduate students at Arizona State University. None of the students were familiar with the stimuli. In the second experiment, vowel identification was determined for stimuli with broadened higher formant peaks and increasingly wide first formants. The same 12 stimuli were used. **Results:** The participants identified vowels in which only one formant was present as back vowels, since back vowels have closely spaced first and second formants. Patterns of responses were related to the frequency of the second formant in that vowels with lower second formants were chosen most often. Additionally, the vowel type appeared to influence the response pattern. **Conclusions:** The authors found that front vowels may be identified correctly when a broad region of energy in the higher frequencies is present. Additionally, when the first formant is widened, vowel identified remains the same until the first formant is widened to six times its normal value. **Relevance to the current work:** The current study removed a band of spectral energy corresponding directly to the frequencies typically associated with either the first or second formant of each vowel. No higher frequencies were eliminated so that vowel identification could be maximized and examined.


**Objective:** The current study used electroencephalography (EEG) to examine the processing of vowel information pertaining to vocal-tract length (formant scale) and vowel type (formant ratios) in synthetically produced unvoiced vowels. **Methods:** Neural elements with sensitivity to vowels, specifically those in the anterior and posterior auditory areas, were recorded and isolated. Five canonical vowels (/a/, /i/, /u/, /o/, /e/) were presented in four different conditions to 10 healthy, right-handed participants (six male, four female, 23-38 years of age) with normal hearing and no known neurological disease. The participants, all wearing an electrode cap fitted with 64 Ag/AgCl ring electrodes, were instructed to report on the naturalness of the vowel. All data lacking significant artifacts were averaged to determine results. **Results:** The authors found that formant information from vowels may be processed completely by the level of the non-primary auditory cortex, suggesting that subcortical structures may be involved. **Conclusions:** It was concluded that the processing of vowel information does utilize information pertaining to vocal-tract length and vowel type in synthetically produced unvoiced vowels. **Relevance to the**
current work: Some components of formant processing may still emerge beyond the unimodal auditory cortex.


Objective: An Electro-Cap is a 32-channel cap which fits closely against the scalp. The electrodes measure electrical activity (in μV) at specific locations around the scalp. Relevance to the current work: An Electro-Cap was used to collect electrophysiological data from all participants.


Objective: The purpose of this study was to examine the relationship between acoustic stimulus and phonetic percept. Methods: The study was split into two experiments. In Experiment 1, two acoustic cues (the perceived contrast slit versus split) produced by the articulation of stop consonants, one spectral and one temporal, were varied. Participants (12 native American English speakers with normal hearing bilaterally) were asked to identify the syllables and responses were analyzed for phonetic content. In Experiment 2, the two acoustic cues produced by the articulation of stop consonants were varied again, although participants were asked to discriminate when the stimuli were mapped on top of each other. Results: For both experiments, less silence was necessarily to hear [split] when correct transitions in the sound stimuli were present. In the first experiment, every participant demonstrated a shift in phoneme boundary. Interestingly, the shifts occurred in the same direction. Conclusions: It was concluded that the equivalence of the acoustic cues was attributed to phonetic perception. Relevance to the current work: The current work utilizes acoustic stimuli to examine specific aspects of the phonetic percept. Fitch and colleagues demonstrated that acoustic cues may be attributed to phonetic perception.


Objective: Flanagan conducted a psycho-acoustic experiment to determine the difference limens, or the just-noticeable differences, for the amplitude of the second formant (F2) in a vowel. Methods: Vowel sounds were synthesized with varying F2 amplitudes to represent the American English vowel [æ] as spoken by an American English male speaker. These stimuli were presented to participants, who were asked to identify the sound heard. Results were obtained by making plots and marking coordinates of perceived sound. Results: It was found that the percentages of different judgments between the sound stimuli were not influenced significantly by whether the amplitude of the F2 was increased or decreased. Rather, the magnitude of the change in amplitude determined the difference limens. Conclusions: The author concluded that the results were attributed to responses to change in the distribution of spectral energy rather than to an overall change in amplitude of the stimulus. Relevance to the current work: Although the current work does not aim to determine difference limens, the study does investigate the F2 of
the steady-state [æ] vowel. Change in F2 amplitude will not influence the mismatch negativity response as greatly as general alterations in the distribution of spectral energy.


**Objective:** The purpose of this study was to use event-related brain potentials (ERPs) to examine the processes involved in the structural analysis of sentences. The properties of the processes were also examined. More specifically, the early left anterior negativity (ELAN) and the late parietally distributed positivity (P600) were observed in regard to their functions in the cortex during an analysis of sentences with phrase structure violations. **Methods:** The authors created sentences with differing degrees of phrase structure violations. In the first condition, only 20% of the sentences created contained phrase structure violations. In the second condition, 80% of the sentences contained phrase structure violation. The number of participants was not reported. **Results:** The ELAN was elicited for both the low and high violation proportion sentences. The P600, however, was elicited only for incorrect sentences with a low proportion of phrase structure violations. **Conclusions:** The results suggest that first-pass parsing processes, as reflected by the ELAN response, are automatic. Second-pass parsing processes, reflected by the P600 response, are not automatic. **Relevance to the current work:** ERPs reflect perceptual processes in the cortex. The ELAN and P600 responses are examples of such processes.


**Objective:** The intent of this study was to acquire electrophysiological correlates of speech voice onset time (VOT) and formant onset time (FOT) to determine whether voicing perception may be attributed to innate temporal sensitivity to voicing boundaries or linguistic experience. **Methods:** Participants included 14 normal-hearing native Hebrew speakers. They were presented with two sets of stimuli, one of which was naturally produced and digitally altered. The natural speech, /ba/ and /pa/ syllables, had their VOTs altered. The synthesized nonspeech sounds were composed of 2 formants varying in their FOTs. Electrophysiological data were acquired through scalp and face placement electrodes. All acquired data were averaged and subtracted from deviant responses to determine results. **Results:** Results showed that the P300 (P3) response was elicited in all speech stimuli but missing from all standard nonspeech stimuli. It appeared in deviant nonspeech stimuli. While the VOT affected N1 latency and N1 and P2 amplitudes, the FOT value significantly affected the P2 amplitude. **Conclusions:** It was concluded that voicing perception is influenced by linguistic experience rather than innate temporal sensitivity. Additionally, it was found that speech and nonspeech signals are processed differently even at the earliest stages of auditory processing. **Relevance to the current work:** Once again, a theory of linguistic perception contributing to cortical activity is supported over a theory of auditory stimuli detection.

**Objective:** The authors discussed neural plasticity in adulthood (specifically that in the central auditory system) as well as the extent to which single neurons and neural groupings respond to changes in frequency selectivity and organization. These alterations are attributed to either partial hearing loss or procedures which induce hearing loss at specific frequencies within an organism. Finally, they looked at cochlear implant users and aimed to quantify improvements during the post-implantation period which may be attributable to neural plasticity. **Conclusions:** Irvine et al. argue that increased sensitivity to specific frequencies or hearing loss at these frequencies will affect neural organization within the cortex. **Relevance to the current work:** This implies that the brain responds differently to specific frequencies and, consequently, is formant sensitive when processing speech in the central auditory system.


**Objective:** The purpose of this study was to investigate how language-specific perceptual processing, particularly that influenced by early language experience, can interfere with second language acquisition. **Methods:** The authors presented the English /r/ and /l/ phonemes to native monolingual speakers of German (12 participants), Japanese (24 participants), and American English (19 participants). The stimuli were 18 /ra/ and /la/ syllables in which the F2 and F3 frequencies varied. The perceptual results were mapped two dimensionally and compared to native-language categorization judgments. **Results:** Results indicated that perceptual processing of within- and between-category acoustic stimuli change between speakers of different languages, as seen in the Japanese neural encoding for F2 cues and German attentiveness to more specific auditory information. **Conclusions:** It was concluded that early language experience does alter low-level perceptual processing and interferes with the formation of higher-level linguistic information. **Relevance to the current work:** Iverson et al. demonstrated that all participants in studies utilizing linguistic and auditory information must be native speakers due to language-specific perceptual processing differences which develop later in life.


**Objective:** This investigation of categorical speech perception observed the mismatch negativity (MMN) event-related potential (ERP) to determine whether vowel-defining first and second formant (F1 and F2, respectively) information is perceived and processed using speech-like categorical processing. It was assumed that if the tones were recognized by the speech perception system, a smaller MMN decrease in amplitude would result from complex tones outside of the vowel space. Conversely, tones representing formants within the vowel space would cause a significant increase in MMN amplitude due to long-term memory-based categorical processing of the tones. **Methods:** The F1 and F2 frequencies were compiled together but did not present as
speech because of missing articulatory complexity. An /a/-like tone and two tones differing by 50 Hz higher or lower than the original frequency were presented in four separate oddball paradigms to eight native German-speaking participants (one male, seven females, ranging in age from 20 to 36 years) to elicit the MMN response. The non-speech tones consisted of sinusoid partials directly corresponding to F0, F1, and F2 frequencies with a fundamental frequency value of 100 Hz, for a total of nine stimuli. The participants were fitted with an electrode cap prior to data collection, and five electrodes (Fz, FC3, FC4, Cz, Pz) were used for analysis after data acquisition. 

Results: Findings showed that the MMN was elicited by deviant stimuli, but there was no decrease in MMN amplitude for F1-F2 formant structure presentations outside of the vowel space. 

Conclusions: Synthetic tones representing formant values (but not sounding like speech) do not trigger speech perception. 

Relevance to the current work: The author’s findings mean that either 1) there is no long-term memory-based categorical processing of formants, or 2) there may be some long-term memory-based categorical processing of formants, but synthetic representations of speech are not adequate to elicit this information.


Objective: The purpose of this study was to determine the functional relationship between various parts of the temporal region with subcortical structures. 

Methods: Eighty individuals, 23 women and 57 men, participated in the study. These individuals suffered from temporal lobe seizures and ranged in age from 10 to 50 years. Different areas of the temporal lobe were stimulated invasively using the Grass depth electrode. All participants were subjected to electrical stimulations spaced 1 cm apart at a depth of 4 to 5 cm. They reported feelings associated with stimulation. 

Results: Participant responses to electrode stimulation were varied depending on placement, but responses from the same subcortical areas were generally more uniform. 

Conclusions: It was concluded that local activation of cortical regions at either superficial or deep locations gave rise to motor, sensory, psychical, or emotional responses. Additionally, there was limited or no spread of activity reaching beyond the immediate location stimulation. 

Relevance to the current work: All electroencephalographic research utilizes the 10-20 system of electrode placement, first developed by Jasper.


Objective: The purpose of this study was to unify audiometric standards. A physical discrepancy existed between thresholds for a 1000-cps (cycles per second) and the authors set out to eliminate this discrepancy through standardization. Method: The authors defined the difference in intensity between thresholds for pure tones and for speech. They measured the relative acuity for speech and for 1000-cps (Hz) and measured thresholds for each in normal-hearing subjects. Participants, three groups ranging in age from 20 to 31 years, listened to spondee words. The first group consisted of 10 sophisticated listeners, the second group consisted of 10 undergraduate students with no previous experience as listeners within the context of an auditory
test, and the third group (made up of 96 subjects) also had no prior experience with auditory tests. Their thresholds were measured using the ‘up-and-down’ method as well as the ‘clinical audiometric technique.’

**Results:** Results were measured for threshold, order of test administration, sex, ear, and prior knowledge of test vocabulary. Few differences were found between results obtained by a clinical versus a laboratory procedure. There was very little difference found between sexes. A slight advantage was found in the right ear. It was also found that the effect of test order is not critical and that prior knowledge of the spondee vocabulary had a significant (2.7 dB difference) effect on spondee threshold.

**Conclusions:** The authors concluded that the multiplicity of conditions must be specified before a norm can be established for speech audiometry. The intensity difference between normal hearing for spondee words and a 1000-cps pure tone is greater than 6 dB. Standardization for the two is necessary in order to acquire true hearing thresholds.

**Relevance to the current work:** The authors set a precedent for standardizing audiometric equipment and techniques for testing hearing. The audiometric methods used in the current study were based upon the research conducted by Jerger and his colleagues.


**Objective:** The purpose of the study was to determine the extent of processing of meaningless random strings of letters as compared to real words. The authors hypothesized that language-related brain regions are involved significantly in the processing of meaning letter strings, assuming that the letter strings are evaluated in the context of potential lexical meaning.

**Method:** Participants, 12 healthy right-handed volunteers (five females, seven males, 26.5 years of age on average with a range of 22 to 35 years) were asked to respond to random stimulus presentation while in an fMRI machine. Stimuli included 20 unpronounceable letter strings and 20 highly imaginable German nouns taken directly from a standard word list. All items were presented for 10 seconds.

**Results:** Activation was demonstrated in the left inferior frontal gyrus, left superior temporal gyrus, and left parietal and occipital regions.

**Conclusions:** The left angular gyrus, bilateral precuneus and left posterior cingulate gyrus were activated when participants were presented with real words contrasted with random letter strings. This may indicate that the individual accesses higher semantic associations when identifying strings of letters. A connectionist model of word processing was supported.

**Relevance to the current work:** The current study used a program designed to construct three dimensional images and locate dipoles from ERP data. Sites of activity were noticed in all regions mentioned in the study by Jessen and colleagues.


**Objective:** The purpose of the book was to discuss the acoustic analysis of speech, including acoustic properties of speech and equipment used in the analysis of speech.

**Relevance to the current work:** Kent and Read discussed and expounded upon results of the earlier study by
Peterson and Barney (1952) which identified the formants most important to vowel identification.


Objective: This study examined language lateralization in healthy right-handed subjects and sought to determine whether the notion of the rarity of right hemisphere language dominance in healthy right-handed individuals is correct. Methods: 188 participants (111 females and 77 males ranging in age from 21-50 years) agreed to take part in the study. The degree of their right-handedness was determined using a word generation task, fMRI, and intracarotid amobarbital injection. Changes in cerebral blood flow velocity (CBFV) in the basal arteries were also measured during language tasks. Results: Language is lateralized along a bimodal continuum, and at 7.5%, right hemisphere dominance is more common than previously thought. Conclusions: Results of the study indicate that lateralization is equivalent in men and women and right hemisphere dominance is not very common. Relevance to the current work: Only right-handed participants are used in the current examination.


Objective: The authors designed this study to effectively characterize the mismatch negativity (MMN) event-related potential and to determine whether the MMN can be used to diagnose central auditory system pathologies in school-age children and adults. Methods: Participants included 10 adults ranging from 19-29 years of age as well as 10 children ranging from seven to 11 years of age. All participants were healthy and had normal hearing. Synthesized speech stimuli (/da/ and /ga/) were presented in an oddball paradigm along a nine-item continuum in which the second and third formant frequencies were altered to transition from /da/ to /ga/. Electrodes were placed on each participant’s head with recordings for analysis coming from the Fz/A2 locations using a forehead electrode as the ground. Results: The MMN was detectable in all participants when presented with a deviant speech stimulus, meaning that it can be elicited from synthesized speech sounds from every point along a nine-item continuum. Conclusions: The researchers concluded that the speech-elicited MMN can be used as both a diagnostic measure in adults and school-age children for central auditory system pathologies as well as in research. Relevance to the current work: The nature and description of the MMN is invaluable to the current examination.


Objective: The authors moved beyond temporal studies of the mismatch negativity (MMN) response in this study and traced the auditory event-related potential’s cortical localization in the medial geniculate nucleus of a guinea pig. Methods: Fifteen properly anesthetized guinea pigs
were fitted with epidural silver bead electrodes to record surface auditory event-related potentials. The speech contrasts /ba/-/wa/ and /ga/-/da/ (with F2 and F3 frequency values chosen to accommodate optimal guinea pig hearing thresholds) were presented to the guinea pigs through insert phones in an oddball paradigm at 75 dB SPL to evoke an MMN response. Results: Results demonstrated larger changes in the MMN amplitude in the /ba/-/wa/ paradigm.

Conclusions: An MMN response could be stimulated from the midline surface of the medial geniculate, but other regions including the ventral portion or the surface over the temporal lobes did not elicit any response. Relevance to the current work: These findings are significant for two reasons: 1) the spectrotemporal changes of formant transitions may be processed systematically in a hierarchical fashion, and 2) nonprimary portions of the auditory pathway do contribute to the MMN response.


Objective: This paper provides a thorough overview of issues pertaining to making recordings of the mismatch negativity (MMN) and interpreting results. It first reviews technical insight into acquiring clean, unflawed recordings of the MMN. This includes using the nose as a reference rather than ear or mastoid, as well as the method of taking averages and superposing the MMN on exogenous and endogenous waveforms. Specific aspects of MMN parameters and responses are reported next, followed by a discussion of physiological factors and amplitudes associated with the MMN. The smallest measureable MMN response was reported in the literature as the difference limen (DL) of a pure tone frequency, which changes at each frequency. In conclusion, some clinical findings and electrodiagnostic applications are discussed, such as the pathological implications of lacking an MMN response after the age of 3 during the presentation of pure tone stimuli. Relevance to the current work: The correct process in acquiring an MMN recording is documented in full.


Objective: The human central auditory system was examined in this study through functional magnetic resonance imaging (fMRI) techniques. Methods: Eight individuals with normal hearing and five individuals with unilateral deafness were recruited to participate in the study. Their hearing thresholds between 250-8000 Hz were evaluated prior to fMRI data collection. The participants received the MRI test in a supine position and all results were analyzed using MATLAB before any judgments regarding activation and lateralization were made. Only one participant was left-handed, but his cortical information did not deviate from right-handed results. Results: It was found that contralateral activation was detected in the temporal lobe, thalamus and midbrain in normal hearing subjects. Conclusions: Auditory information crosses over at the base of the brainstem and travels ipsilaterally to the auditory cortex. Relevance to the current work: This study is very informative regarding cortical localization of auditory stimuli and the pathway which auditory stimuli take when processed in the brain.

**Objective:** The purpose of this study was to examine the additivity of the mismatch negativity (MMN) while looking at how vowels and pitch interact during a singing activity. **Methods:** Twelve native French speakers (five males, seven females, ranging in age from 18-38 years) took part in the study and were fitted with an electrode cap covered with 64 Ag-AgCl electrodes with a reference electrode on the tip of the nose. Participants were instructed to listen to sung vowels, namely /ɛ/ and /ɔ/, and rare deviant stimuli differing in pitch only, vowel only, or in both pitch and vowel. Data were filtered and averaged. Responses in which significant artifacts were identified were discarded. **Results:** An MMN and a P3a, a subcomponent of the P300 response, were elicited for all deviant stimuli. While the P3a component was larger in double deviant stimuli, the MMN amplitudes remained statistically similar enough across all deviant stimuli and could not to be considered additive. **Conclusions:** It was concluded that vowels and pitch are not additive components of the MMN response. **Relevance to the current work:** The significance of this study includes the notion that pitch and vowel are processed as integrated units. Pitch does not have to be considered an entirely separate entity in studies specifically examining vowels and vowel processing.


**Objective:** The purpose of this study was to determine whether there are physiological correlates of categorical perception of speech. **Methods:** A nine-stimulus continuum between the computer-modified speech syllables /ba/ and /da/ was presented to participants in 4 different experiments. The first experiment required attentiveness during the same-different paradigm presentation, while the second experiment required that participants read a book and ignore the sound signals. The fourth study examined responses to stimuli with regard to phonemic category or intensity. Subjects in the experiments included 10 English-speaking, right-handed males between 21 and 45 years of age. They were fitted with an electrode cap and data were collected from 10 different locations on the scalp with a reference electrode on the right mastoid. **Results:** An MMN response was elicited in all deviant stimuli across all experiments except during the presentation of the /da/ stimulus, which had a wider frequency-content compared to the other stimuli. **Conclusions:** Results indicated that the MMN response evoked by speech sounds may be more reflective of acoustic changes rather than phonetic changes in stimuli. **Relevance to the current work:** The current examination also evaluates the MMN response evoked by speech sounds and whether it is reflective of phonetic changes rather than acoustic changes in stimuli.


**Objective:** This paper provides a thorough overview of speech-evoked auditory event-related potentials (ERPs), specifically the P1-N1-P2 complex, acoustic change complex (ACC),
mismatch negativity (MMN), and P3 responses. It also discusses several applications for ERPs including effects of hearing loss and amplification on neural encoding of speech, behavioral detection and discrimination, and the impact of auditory training on the neural processing of speech. **Conclusions:** An important conclusion drawn from the discussion was the notion that ERPs making it possible to differentiate whether perceptual confusions are a result of inability to detect specific aspects of an acoustic signal. Additionally, ERPs indicate which aspects of speech signals are not neurally coded.


**Objective:** The present experiment was designed to investigate whether there is a relationship between the traditional hypothesis of vowel perception based on formant frequencies and optimal octaves (frequency bandwidths with enough acoustic information to be perceived correctly as a specific vowel). **Methods:** All 48 participants (33 for the first half of the study and 15 for the second half) were presented with three individual vowels, /i/, /a/, and /u/, with eight filter and two nonfilter control conditions. They were instructed to transcribe, using the International Phonetic Alphabet, the stimuli presented, and to indicate whether the sound was distorted. **Results:** Results indicate that there are specific bandwidths which permit the correct identification of each vowel, and these bandwidths are compatible with both optimal octaves and formant frequencies. **Conclusions:** The perceptual existence of optimal octaves was confirmed through filtered vowel similarity rating and identification. **Relevance to the current work:** Application for future investigations includes the notion that specific frequency bandwidths may be isolated or filtered in a sound segment without affecting the intelligibility of the vowel.


**Objective:** This paper served as a review of literature regarding the mismatch negativity (MMN) event-related potential component and its pertinence to both healthy and pathological central auditory function. The author cites previous studies and characterizes the MMN as elicited by discriminable changes in a listening task utilizing a same/different paradigm, an objective measure of individual discrimination ability for different sound features, free from attentional variation, a measure of auditory short-term memory, and an indicator of neurophysiological processes underlying hearing. The author also discussed some dipoles and general locations within the auditory cortex which may be involved in the MMN response, supposing that the MMN has auditory-cortex sources and subcortical sources. **Relevance to the current work:** The significance of this paper largely resides in its thorough discussion of the MMN as well as its descriptions of previous notable research.

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

**Objective:** The objective of this study was to discuss current understandings of the perception of speech sounds in the human brain based on information acquired from the mismatch negativity
(MMN) and its magnetic equivalent (MMNm). MMNs can be acquired for both speech and nonspeech stimuli and with nearly any auditory variable or feature (pitch, duration, frequency, phoneme) because of the nature of same/different tasks. Recent MMN studies have confirmed that phonemes as well as larger linguistic units are perceived and processed based on language-specific phonetic traces, which are developed before an infant transitions into childhood. These traces are short-term, meaning that the MMN response does not decrease with increased familiarity to a listening task. Relevance to the current work: Dipoles can be traced to the posterior portion of the auditory cortex in the left hemisphere. Furthermore, acoustic invariances specific to individual speech sounds may explain how speech is perceived as the same across different speakers.


**Objective:** This paper provides a thorough overview of the mismatch negativity (MMN) and its role in discrimination thresholds and accuracy. A large focus of discussion was brain plasticity and how it changes over time. The MMN response is dependent upon short-term memory traces. The amplitude of the MMN response decreases as an individual grows older or develops a neurodegenerative disease because of shortening of memory duration. Because the amplitude of the MMN response directly correlates to the intensity of the deviant stimuli, the MMN provides an accurate index of an individual’s ability to discriminate between stimuli. The response also provides information on his or her degree of accuracy. Relevance to the current work: The significance of this paper can be summarized in its description of how the MMN can be used to show a participant’s ability to discriminate complex sound stimuli such as different phonemes. Healthy, young individuals with normal hearing and no history of neurological impairment should yield typical MMN responses.


**Objective:** The purpose of this study was to determine whether language-dependent memory traces, activated in the processing of speech, exist. Furthermore, they sought to locate the sources of these language-dependent memory traces within the cortex. **Methods:** Thirteen normal-hearing Finnish participants (18-29 years old, right-handed, nine males and four females) and 11 normal-hearing Estonian participants (19-31 years of age, right-handed, five males and six females) agreed to be involved in the study. The participants were fitted with an electrode cap in which the mismatch negativity (MMN) response was acquired. They were instructed to listen to a series of auditory stimuli, namely the prototype /e/ phoneme as the nondeviant stimulus and the /ö/ phoneme as the deviant stimulus. F1, F3, and F4 remained the same while F2 was altered in the deviant stimulus. The deviant stimulus was categorized as a prototype when presented as the Finnish /ö/ and a nonprototype when presented as the Estonian /ö/. **Results:** Results showed that amplitudes of the MMN increased during the presentation of the prototype deviant stimulus as opposed to the presentation of the nonprototype deviant stimulus. These memory traces were
traced to the auditory cortex of the left hemisphere. **Conclusions:** This indicates that language-dependent memory traces exist and can be traced to the left hemisphere. **Relevance to the current study:** Although the MMN response does not decrease over time due to memory traces, it is indicative of language-dependent memory traces. Localization studies will yield results centralized to the left hemisphere.


**Objective:** This study is a replication of a previous examination (Tampas et al., 2005), which examined auditory event-related potentials (mismatch negativity and P300) during the presentation of synthetically generated consonant-vowel (CV) speech and nonspeech contrasts. The current author sought to determine whether the mismatch negativity (MMN) response changes in response to the presentation of phonetic or acoustic stimuli. **Methods:** Twenty-one adults, 11 males and 10 females between the ages of 18 and 30 years, participated in the study. Participants were presented with a same/different discrimination task of contrasts. Event-related potentials were acquired from 270 trials of each common and deviant stimulus contrast pair, and the FCZ electrode was used to investigate all MMN responses. **Results:** Localized distributions of data acquired from electrode placement on the scalp yielded information on processing responses to within- and across-category speech and nonspeech stimuli. **Conclusions:** Results confirm that speech sounds and nonspeech sounds are processed differently in the cortex. However, contrary to previous findings, the MMN response is influenced by linguistic information. MMN responses elicited by nonspeech contrasts display longer latencies and more negative peak amplitudes than those seen in speech contrasts, but marked MMN responses do appear given the presentation of speech stimuli. Relevance to the current work: Neff’s findings opened the door for further investigations of the MMN in relation to the processing of speech stimuli.


**Objective:** The aim of this study was to investigate whether the mismatch negativity (commonly referred to as the MMN, or the PMN, corresponding to the term “phonological mismatch negativity” in this particular study) is a reflection of prelexical or lexical stages of spoken word recognition. **Methods:** Twenty right-handed English-speaking individuals (five males, 15 females, average age of 22.5 years) participated in the study. They were fitted with an electrode cap measuring electrical activity at 15 cortical sites using Ag/AgCl electrodes. They were screened for further participation with the Woodcock Diagnostic Reading Batter and the Rosner Auditory Discrimination test. Participants were instructed to segment 160 English monosyllabic words (all beginning with a consonant cluster) by omitting the initial phonemes. They were then presented with a correct option as well as incorrect options, which acted as deviant stimuli. **Results:** The PMN was elicited equally for all incorrect items, including word and nonword options. **Conclusions:** These results indicate that the PMN is a function of prelexical phonemic
expectations. 

Relevance to the current work: Because of the results of this study, the current work does not examine lexical phonemic units.


**Objective:** The purpose of the study was to determine how well vowel segment durations in spoken sentences are represented in auditory sensory storage and to evaluate the extent to which syntactic and/or auditory-phonetic context affects phoneme boundaries in the identification of phonemic vowel length. **Methods:** Participants, 12 native Dutch-speaking listeners, were asked to identify five different test utterances in a binary forced choice task. The just-noticeable difference (JND) was calculated. **Results:** The authors found that the perceptual accuracy of vowel duration perception was approximately 90 ms with a JND of 5 ms. They also found that phoneme boundary values are affected by syntactic context as well as the auditory-phonetic environment. Phoneme boundaries were also influenced by monosyllabic phoneme utterances following the presentation of the Dutch vowel. **Conclusions:** It was concluded that the auditory-phonetic environment makes a difference regarding vowel length identification. The amount of time necessary for vowels to be identified, as well as the noise in which it is presented and the steady state of the vowel, impact overall identification. 

Relevance to the current work: The study by Nooteboom and Doodeman pertains to the current work in that the vowel segments played to participants were of an appropriate length to maximize identification of the vowel. Additionally, only parts of the initially recorded segment were played to eliminate any prosodic alterations which may affect perception or brain response.


**Objective:** The purpose of the study was to determine whether the mismatch negativity (MMN) response reflects modality-specific processes or non-specific attentive phenomena. Previous research had used the MMN solely for auditory testing. **Methods:** Nine listeners (19-42 years of age, two females, seven males) participated in the experiment. Five different stimulus conditions utilizing both auditory and visual stimuli were presented to participants. All blocks involved 400 stimuli, and the event-related potentials (ERPs) were collected from three scalp locations with a reference electrode at the right mastoid. Grand averages of the responses to the standards were compared to those of the deviants. **Results:** Only the auditory stimuli evoked an MMN response. Deviant visual stimuli presented amidst more common standard stimuli did not elicit an MMN response. **Conclusions:** The MMN is only elicited by the auditory modality. It is not a manifestation of an attentional mechanism elicited through many different modalities. 

Relevance to the current work: Nyman et al. confirm that the MMN cannot be elicited by any attentional mechanism save the auditory modality. Attentional tasks employed during data acquisition will not influence the MMN amplitude or frequency of occurrence if auditory stimuli remain unaltered.

**Objective:** This paper discusses the Praat acoustics program (Boersma, 2001). More specifically, it discusses GSU Praat Tools, which simplifies the Praat program for users and offers new functions such as editing, filtering, and otherwise changing sounds. The author provides a succinct overview of additional features such as labels, windows, commands, analysis tools, and directories. **Relevance to the current work:** This free script package has implications for further research in that it may be used in conjunction with programs such as Adobe Audition to digitally alter speech sounds.


**Objective:** The present study was designed to determine whether phonological status plays a role in maintaining discrimination of phonetic information. The authors hypothesized that if phonological status plays a role in maintaining discrimination of phonetic information, then younger babies in the six to eight month range should be the only group to correctly discriminate. **Methods:** Participants, healthy native English speakers with normal hearing, ranged in age from six months to adults and were split into three groups (six to eight months, 10 to 12 months, and adults). They were presented in three experiments with the phonetically, rather than phonemically, different consonant vowel segments [də] and [tʰə]. In experiment 1, 30 male adults ranging from 18-30 years of age were presented with stimuli in a sound booth at 65 dB over a loudspeaker and asked to discriminate between [də] and [tʰə]. They did not significantly discriminate between the two. Ten stimuli were used in experiment 2 (five [də]s and five [tʰə]s), in which 18 participants ranging from 18 to 25 years of age were asked to discriminate the syllables. In the last experiment, the aforementioned syllables were played to the infant groups. **Results:** Results indicated that adults and older infants (10 to 12 months) could not discriminate as well as the youngest participant group. **Conclusions:** Phonological status does play a role in maintaining discrimination of phonetic information. **Relevance to the current work:** Digitally filtered vowel recordings in the current examination may influence phonological status, but the study by Pegg and colleagues confirmed that this possible change in phonological status will have to bearing on discrimination of phonetic information.


**Objective:** The current examination measured the mismatch negativity (MMN) response and used conscious behavioral discrimination techniques to study the development of a new vowel category. **Methods:** Participants were split into three groups. The first group, native speakers of Finnish with minimal knowledge of English, included 11 healthy adults, four being female, ranging in age from 18-21. The second group, 10 participants in total with six females, ranging in age from 19-25, were Finnish advanced students of English. The third group, native speakers
of English, included nine participants, five females, ranging in age from 20-21. They were presented with one pair of Finnish (/i/-/e/) and three pairs of English (/i/-/e/, /i/-/I/, /e/-/I/) synthetically-produced vowels over the course of two experiments. In Experiment 1, the MMN was recorded; in Experiment 2, the same stimuli were presented and the reaction time (RT) was recorded. The electroencephalogram (EEG) was recorded using Ag-AgCl electrodes at the Fz, Cz, and Pz locations. Results: Native-like memory traces were only present in the native speaking groups. Conclusions: Long-term classroom learning cannot replace or develop native-like memory traces. Relevance to the current work: Native-like memory traces cannot be recreated later in life. Consequently, all studies investigating cortical responses to language must employ native-speaking participants.


Objective: The purpose of the current study was to determine whether preattentive discrimination, as measured via the mismatch negativity, or MMN, could be enhanced through early exposure in language immersion. Methods: Two groups of Finnish children participated in the study. The first group was monolingual (nine Finnish-speaking right-handed children ranging in age from five to seven years, six females), while the second group (nine native Finnish-speaking right-handed children ranging in age from five to seven years, four females) learned French in school. All participants wore Sn electrodes with two face-placement reference electrodes placed below and above the eye. Stimuli, the Finnish phonemes /o/ and /u/, and a within-category pair in Finnish which would be considered phonologically different in French, were presented through an oddball paradigm. Results: Results indicated that the monolingual group did not register a difference in the within-category stimulus pair, whereas an amplified MMN response appeared in the immersion group. Conclusions: It was concluded that preattentive discrimination of vowels is enhanced through early exposure in language immersion programs. Relevance to the current work: All studies investigating cortical responses to language must employ native-speaking participants.


Objective: The purpose of this study was to examine the relationship between the vowel phoneme intended by a speaker and the vowel phoneme identified by a listener. Additionally, the authors investigated the formants of each phoneme and which were most important for vowel identification. Methods: Participants, a total of 76 speakers (33 men, 28 women, and 15 children) listened to word lists containing 10 monosyllabic words beginning with /h/ and ending with /d/. They were required to record themselves saying each word on the word list using a magnetic tape recorder. Each participant recorded two lists of 10 words. The test material was randomized and all acoustic measurements were made with a sound spectrograph. The words, now randomized, were presented to 70 listeners over a series of eight sessions. The participants were given a list of the 1520 words over 200 lines and were asked to cross out every word they heard. Results: Results of acoustic analysis of the first and second formants (F1 and F2, respectively) were charted for all 76 speakers (F1 values were placed on the x-axis, whereas F2 values were
plotted on the y-axis). **Conclusions:** The authors concluded that vowel identification is dependent upon the first two formants. Spectrograms differ for different consonants and vowels. **Relevance to the current work:** In determining which formants are most important for vowel identification, the authors contributed to a wealth of future research pertaining to the role of formants in vowel identification. The formants examined in the current study are dependent upon the results acquired by Peterson and Barney.


**Objective:** The purpose of this study was to measure the frequencies and level of the first three formants in a series of Dutch vowels. The second objective of the study was to present a specific set of formant data in accordance with the principal-components representation. **Methods:** The frequency spectra of 12 Dutch vowels were recorded in 50 male native-Dutch speakers. The sound-pressure levels were analyzed in a principal-components analysis. **Results:** Spectral differences were found among the 12 vowels. A speaker-dependent correction was applied, as well as a 1/3-octave filter for all vowel sounds. **Conclusions:** Dimensional analysis is preferred when performing automatic speech recognition. The authors determined that dimensional analysis is more attractive than the more traditional formant analysis methods because dimensional analysis is simpler and may be carried out in real time. **Relevance to the current work:** Pols, Tromp, and Plomp demonstrated through dimensional analysis that the middle of each vowel, regardless of the language being recorded, is steadier in pitch because prosodic alterations typically found at the beginning and end of a recorded vowel are eliminated.


**Objective:** This study was intended to examine the perception and production of vowels in individuals with cochlear implants who became deaf after developing speech. **Methods:** Five participants (age 29-55 years, three females, four with cochlear implants and one with normal hearing) used a speech synthesizer to alter the first and second formants (F1 and F2, respectively) and to create Finnish vowels. The authors determined discrimination of the standard /i/ stimulus from the deviant /i/ with an alteration in the F2 continuum by fitting each participant with an electrode cap (using the standard 10-20 system) to measure the mismatch negativity (MMN). **Results:** Only those participants with the best vowel perception abilities demonstrated an amplified MMN response. **Conclusions:** Memory traces developed early on in the process of language acquisition remain stable over time. **Relevance to the current work:** The implications of this study extend beyond the realm of cochlear implants into the wider academic discipline of memory traces for vowels. The MMN response is not indicative of memory traces and does not decrease in amplitude after multiple presentations.

**Objective:** The purpose of this study was to establish whether vowel familiarity and its phonemic status in a language affect conscious and automatic vowel discrimination. **Methods:** Two experiments were designed to collect event-related potentials (ERPs) and to elicit a mismatch negativity (MMN) response. The stimuli were six synthesized, steady-state Komi and Finnish vowels. The participants (10 students: six native Finnish-speaking females, 20-30 years of age) were instructed to indicate when they heard a deviant stimulus by pressing a button during Experiment 1. In Experiment 2, ERPs were recorded and the participants were asked not to attend to the stimuli. All results were averaged and standard stimuli responses were subtracted from deviant stimuli responses. **Results:** In both experiments, the standard stimuli presented did not elicit any differences in the MMN response over time. **Conclusions:** The researchers concluded that the phonemic status of the standard stimulus does not play a role at the pre-attentive level of processing, but it does at the attentive level. **Relevance to the current work:** The conclusion drawn by Savela et al. is in direct contrast with the results of previous studies which concluded that ERP results show effects of the phonetic status of the present stimuli.


**Objective:** The current study aimed to investigate whether a sound intensity is processed in the brain differently depending on whether it is vocal, conveying communicative intent, or nonvocal, conveying non-communicative intent. The investigators hypothesized that intensity changes in vocal stimulus sequences would receive primacy over such changes in nonvocal stimulus sequences. **Methods:** The study made use of event-related potentials and measured the amplitude of the mismatch negativity (MMN) and P300 for intensity change. Twenty-four nondeviant and deviant stimuli were presented to 20 male (average age of 20 years) and 20 female (average age of 19 years) participants in an oddball paradigm in which the deviant stimuli differed in intensity. The participants were fitted with an electrode cap covered by 64 electrodes, and four face-placement electrodes served to record the electro-occulogram (EOG). **Results:** The authors found that there was no significant change in MMN amplitude for vocal deviants, but there were significant changes in MMN for nonvocal stimuli. Additionally, the MMN and P300 results for vocal sounds were amplified in the female participants, indicating that women were more sensitive to the social relevance of speech. **Conclusions:** It was determined that a sound intensity is processed in the brain differently if it conveys no communicative intent, especially in women who are more sensitive to changes. **Relevance to the current work:** This is relevant to further research in that given the presentation of auditory material to acquire event-related potentials, results of male and female participants should be analyzed separately and then compared.

**Objective:** This investigation utilized two vowel and stop-consonant discrimination experiments to ascertain whether stop-consonant perception is categorical in nature. **Methods:** Stimuli (the /p/, /t/, and /k/ phonemes embedded in an /-ak/ context, as well as /i/, /e/, and /a/ phonemes embedded in a /p-t/ context) were presented to native Dutch speakers. Fourteen college-age participants were instructed to identify the phonemes and categorize them as one of three phoneme options. The first experiment involved five tests, whereas the second test involved six tests of phoneme detection and categorization. **Results:** Vowel discrimination was subject to range effects, whereas stop-consonant discrimination was not. This implies that different theories of speech sound discrimination may come into play depending on whether a consonant or a vowel is processed. **Conclusions:** It was concluded that stop-consonant perception is highly categorical, whereas vowel perception is far less categorical. **Relevance to the current work:** The current study utilizes vowels rather than stop-consonant syllables. This should aid in uniform categorization, detection, and identification.


**Objective:** The goal of this study was to determine how voice onset time (VOT) distinctions for the voiced and unvoiced stop-consonant syllables /da/ and /ta/ are encoded within the brain. **Methods:** The authors observed behavioral perception of the syllables as well as the mismatch negativity (MMN) response within the context of Cortical Auditory Evoked Potentials (CAEP). Sensory encoding and sensory discrimination were examined via analysis of the obligatory evoked response N1 CAEP and the MMN CAEP. Sixteen native English-speaking participants (11 females, five males, ranging in age from 20-30 years) with no reported speech or hearing problems took part in the study. Stimuli with VOT varying along a continuum in 10 ms steps were presented to the participants through headphones at 75 dB SPL. These stimuli had an overall duration of 200 ms with a fundamental frequency of 114 Hz. Participants were fitted with three silver-chloride electrodes along the midline of the scalp (Fz, Cz, and Pz placements) as well as a forehead placement reference electrode. Thus, the MMN was collected through the presentation of an oddball paradigm. **Results:** Findings indicated that a distinct category boundary exists between 50-60 ms for all listeners and that across-category discrimination is more accurate than within-category discrimination. **Conclusions:** The authors concluded that neurophysiologic correlates of categorical perception exist for VOT. **Relevance to the current work:** Neurophysiologic correlates of categorical perception and speech in general do exist, although many researchers would argue against any theories supporting a phonetic model of processing.

**Objective:** The purpose of this study was to establish how neurophysiologic levels of stimulus processing differ between speakers of different languages. **Method:** The perception of native and nonnative phonetic categories was examined via acquisition of electrophysiologic data. Participants (10 native Hindi speakers and 10 monolingual American-English speakers) were led through behavioral and electrophysiologic tests measuring the N1 and mismatch negativity (MMN) cortical evoked potentials. They were instructed to listen to nonsynthesized bilabial consonant-vowel (CV) stimuli in which the voice onset time (VOT) ranged from -90 to 0 ms. **Results:** Results indicated that the Hindi listeners discriminated between the CV stimuli, whereas the native English speakers who do not consider duration of pre-voicing to be phonemically relevant did not discriminate accurately. A significant MMN response was also noted in speakers of Hindi but not in speakers of English. **Conclusions:** Levels of stimulus processing are influenced by knowledge of specific languages. **Relevance to the current work:** All participants in the current work need to be native, monolingual speakers of American English. Otherwise, levels of stimulus processing may differ between subjects.


**Objective:** The purpose of the study was to determine whether input pathways involved in executive and volitional functions converge or remain segregated at the executive levels of cortical information processing. **Methods:** Thirty healthy individuals were asked to participate in a perceptual go/no-go task during a brain imaging session. Stimuli included ten “go” responses and ten “no-go” responses. Instructions were based on spatial, temporal or object stimulus features. Participants performed this task multiple times over several modalities including executed and imagined paradigms. Ten participants performed motor and imagined tasks, ten performed visual, auditory, and tactile tasks, and ten performed the primary study of visual cues and motor responses. **Results:** Lateralization of hemispheres was noted on the tasks. Activity within the inferior and middle occipital gyri, middle temporal gyrus, and medial frontal gyrus was noted in the left hemisphere during the “what” and “when” tasks. **Conclusions:** The medial frontal gyrus and cingulate gyrus are distinct and separate in their organization and functions. The authors concluded that the executive mechanisms which operate within the medial frontal gyrus are involved in input processing streams. **Relevance to the current work:** Talati and Hirsch related function of the medial frontal gyrus to executive functioning and possibly to language-related tasks. The current study identified the medial frontal gyrus as a dipole.
Objective: The authors recorded and measured auditory event-related potentials (mismatch negativity and P300) to determine whether behavioral and P300 responses reflect a phonetic, categorical level of processing. Methods: Same/different and oddball paradigms were used in conjunction with the presentation of synthetically generated consonant-vowel (CV) speech and nonspeech contrasts. Ten participants (eight females and two males) from the ages of 19 to 35 were presented with the speech and nonspeech stimuli, and the ERPs were acquired using a two-channel electrode configuration. All data from standard and deviant stimuli were collected from the Cz-A1 and Cz-A2 electrodes. The recordings were averaged and data with significant artifact were discarded.

Results: A mismatch negativity (MMN) response was elicited by the frequency glide (nonspeech) stimuli, but absent for all speech stimulus contrasts. A P300 response was evoked in both the speech and nonspeech stimuli, although it was stronger in the nonspeech stimuli. The response was smaller in amplitude and longer in latency in the speech stimuli.

Conclusions: The authors concluded that the behavioral and P300 responses do not reflect a phonetic, categorical level of processing.

Relevance to the current work: The results of this study are in direct opposition with a more recent study by Neff (2010), which showed that an MMN response is elicited from both speech and nonspeech contrasts.


Objective: This study examined the role of behavioral “transfer of learning” and neural plasticity in underlying physiologic processes. Methods: Participants, 18 normal-hearing monolingual English speakers ranging in age from 18-28 years, were fitted with an electrode cap at nine active electrode locations to measure the mismatch negativity (MMN) response over a series of nine days. They were asked to discriminate and identify a prevoiced alveolar stop along the /da/-/ta/ continuum after being trained to identify a prevoiced labial stop sound (along the /ba/-/pa/ continuum) not phonemically relevant in the English language. The stimuli, 3500 standards and 1000 deviants, were presented in an oddball paradigm. All recordings with severe artifacts were discarded and the remaining recordings were averaged.

Results: Those participants who were trained to hear a prevoiced labial stop identified the prevoiced labial stop more adequately than the untrained participant group. Furthermore, the MMN response for the prevoiced labial stop was more amplified in the training group, indicating that training can generalize to listening situations not directly linked to the new stimulus.

Conclusions: Perceptual learning may be altered through auditory training.

Relevance to the current work: This study is significant to future studies involving the repeated presentation of auditory stimuli. Training effects must be taken into consideration, even though the MMN is not representative of memory traces.

Objective: The purpose of the study was to investigate statistical correspondence of ERP tomography and fMRI in a complex visual language task. Methods: Ten healthy subjects (21 to 41 years of age with a mean age of 26.8 years) with no history of neurological damage or reading disability participated in two separate sessions using the same experimental task. Six participants completed the MRI and four participants completed the EEG task before completing the other method. Each method was completed within three weeks of the other for each participant. Stimuli, well-known Swiss-German four-to-six letter words, were presented in blocks which were repeated five times. Results: Both ERP tomography and fMRI were shown to represent similar neural networks within the bilateral occipital gyrus, lingual gyrus, precuneus and middle frontal gyrus, left inferior and superior parietal lobe, middle and superior temporal gyrus, cingulate gyrus, superior frontal gyrus, and precentral gyrus. Statistical analysis of mean correspondence was significant. Conclusions: Both ERP and fMRI can be used to examine and trace cortical locations of activity in the brain during language-related tasks. However, because only half of all participants in the study demonstrated correlational statistical significance between ERP and fMRI, the two must be used in conjunction with caution and on a person-by-person basis. Relevance to the current work: The current work used ERP data to ascertain information regarding dipole localization in the cortex. Additionally, cortical activity was noted by the current author in all sites mentioned by Vitacco and colleagues.


Objective: This study aimed to investigate whether the pre-attentive change-detection neural process is influenced by auditory (sensory) and phonetic (categorical) representations of vowels. Methods: The authors measured event-related brain potentials (ERPs) and noted the mismatch negativity (MMN) response during the presentation of standard-deviant speech stimuli pairs to participants. Stimuli, 430 Hungarian and Finnish /e/ synthesized phonetic variants, were presented equally to all participants (Experiment 1 required 12 native Finnish speakers, 20-35 years of age, and 14 native Hungarian speakers, 18-24 years of age; experiment 2 required 10 native Finnish speakers, 20-35 years of age, and 10 native Hungarian speakers, 22-27 years of age). All participants were instructed not to pay attention to the same-different stimuli presented. Results: The authors found that the MMN was elicited by both language groups in across- and within-category contrasts. Conclusions: The authors concluded that the pre-attentive change-detection process is influenced by both auditory and phonetic representations of vowels. Relevance to the current work: This study supports both phonemic and auditory models of linguistic and phonetic perception. The current work investigates correlates of phonemic processing.
Objective: The aim of the study was to investigate whether the mismatch negativity (MMN) response is sensitive to a small variation in the intensity of the frequent stimulus. Methods: Synthesized tones were presented to 10 healthy participants (20-32 years of age, four females, six males) in a passive MMN examination. Each block presented varied in intensity or frequency within the block to differing degrees. The participants were fitted with an electroencephalography (EEG) cap and the EEG was recorded from four midline scalp locations with a reference electrode on the right mastoid. Tests were averaged and computed using the BMDP statistical software package. Results: The deviant stimuli presented in the intensity- and frequency-altered stimuli both elicited an MMN response. The amplitude of the MMN decreased as the variability of the frequent stimulus increased. Conclusions: It was shown that the MMN tolerates some variation in the intensity of the repetitive stimulus. Relevance to the current work: The speech recordings used in the current study must be set to the same intensity for listeners so that any differences in intensity do not affect results.


Objective: Winkler and colleagues investigated attention independent parts of event-related potentials (ERP) by examining the effect of the interstimulus interval (ISI) of paired stimuli on the mismatch negativity (MMN) response. Methods: Eight participants were fitted with eight Ag-AgCl scalp-placement electrodes and one reference electrode at the tip of the nose. They were instructed to ignore auditory stimuli, which were presented to them in a same-different paradigm through head phones to their right ears. All response data were averaged separately for each stimulus type and ERPs were filtered below 30 Hz. Results: The MMN response was not elicited by all deviant stimuli. Only the deviant stimuli with larger ISIs registered an increased MMN amplitude. Conclusions: It was concluded that the ISI of paired stimuli does have an effect on the MMN response. Relevance to the current work: The intertone intervals between stimuli and blocks of stimuli must be great enough such that MMN responses may be elicited.


Objective: The Declaration of Helsinki was developed as a statement of ethical principles for medical research in which human subjects are involved. This includes any human material or data collected at any time. Relevance to the current work: The Declaration of Helsinki pertains to the current investigation in that all testing of subjects was conducted in an ethical manner previously specified by the World Medical Association (WMA) in 2008. Additionally, the methods of screening and testing were approved through Brigham Young University’s Institutional Review Board (IRB), which was in harmony with the ethical principles published by the WMA.

**Objective:** The purpose of this study was two-fold. First, the investigators sought to observe the effect of stimulus frequency on mismatch negativity (MMN), the obligatory evoked response (N1), and P2. The investigators also examined the effect of stimulus complexity on the cortical auditory event-related potentials (ERPs) using words and consonant-vowel (CV) stimuli in which the formant frequencies were altered. **Methods:** Participants (seven females and five males ranging from 22-45 years, all with normal hearing between 500 and 4000 Hz) wore seven scalp-placement electrodes and three reference electrodes. They listened to three sets of tone bursts in the speech frequency range using the words /bæd/, /dæd/, and the CVs /bæ/ and /dæ/. Recordings with excessive or unusual artifacts were discarded and all others were averaged. **Results:** Results indicated that both the N1 and MMN responses reflect the tonotopy of the auditory cortex. A negative correlation was drawn between frequency and N1 amplitude and latency. Additionally, the MMN response demonstrated a significant increase in amplitude for some speech contrasts. **Conclusions:** The presence and area of the elicited MMN responses were found to be affected by stimulus parameters such as complexity and frequency. Discriminable contrasts of tones evoked the MMN more often than discriminable words, a tone complex, or consonant-vowel syllables. **Relevance to the current work:** The MMN response is more likely to respond to more simple auditory stimuli than complex stimuli such as words or syllables.


**Objective:** The purpose of the study was to compare word generation tasks performed silently and aloud using functional magnetic resonance imaging (fMRI). **Methods:** Nine healthy, English-speaking subjects (between the ages of 20 and 40 with no symptoms of neurologic disease) were asked to perform word generation paradigms aloud or silently while in an fMRI machine. The paradigms consisted of reciting a word list provided by the authors, first aloud, and then silently in 20-second intervals. Participants were given insert ear plugs and were asked to remain as still as possible while in the scanner. Echoplanar images were acquired superimposing cross-correlation techniques on anatomic reference images. **Results:** Activation in the inferior frontal lobes, sensorimotor cortex regions, supplementary motor areas, and anterior cingulate gyri was seen for both silent and aloud word generation tasks. **Conclusions:** Yetkin and colleagues found that for the purposes of mapping language functions with fMRI, word generation tasks may be performed aloud or silently. Silent word generation yields greater activation in the left frontal lobe. **Relevance to the current work:** Dipoles were located in several of the regions, namely the cingulate gyrus and other gyri, discussed in this study.

**Objective:** The authors of this study aimed to examine the authenticity of the phoneme boundary effect at the neural level. **Methods:** The authors measured the mismatch negativity (MMN) response for across- and within-category changes in the length of Finnish phonemes in three separate groups of participants (13 native Finnish speakers, 14 second-language users of Finnish, and 12 Russian non-Finnish-speakers). All participants listened to stimuli (the /u/ and /u:/ phonemes embedded into the words /tuku/ and /tu:ku/) presented in a two-alternative, forced-choice categorization task. All data were averaged and standard data were subtracted from deviant data. **Results:** The MMN amplitude was significantly greater for native Finnish speakers, most likely due to a language prototype developed in the early years of language acquisition, and lower in non-native speakers as well as non-Finnish-speaking participants. **Conclusions:** Because there were no across- and within-category changes, neural responses cannot be attributed to phoneme boundary crossing. **Relevance to the current work:** The results of this study are in direct opposition to those derived from later studies, such as an investigation conducted by Neff in 2010.
Appendix A

Appendix E (for general participants)
Informed Consent to Act as a Human Research Subject
Brain Mapping of the Mismatch Negativity Response in Vowel Formant 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 and pure tones 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 Amanda Fujiki and Elizabeth Perry, BA, students 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 liquid and are easily removed with water. Blunt needles will be used as a part of this study to help apply the electrode liquid. They will never be used to puncture the skin.

Acoustic and linguistic processing will be measured using an electrode cap, which simply measure 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 pure tones (like beeps) and speech vowel sounds, 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 liquid used in applying the electrodes. Allergic reactions to the liquid 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 including removing the electrodes and liquid 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 any time 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 94602; phone (801) 422-6458; e-mail: david_mcperson@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 relation 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:________________________________________________