J Am Acad Audiol 15:469–485 (2004)

Subtitled Videos and Mismatch Negativity (MMN) Investigations of Spoken Word Processing

Catharine M. Pettigrew*
Bruce E. Murdoch*
Curtis W. Ponton†
Joseph Kei*
Helen J. Chenery*
Paavo Alku‡

Abstract

The purpose of this study was to determine whether the presence of subtitles on a distracting, silent video affects the automatic mismatch negativity (MMN) response to simple tones, consonant-vowel (CV) nonwords, or CV words. Two experiments were conducted in this study, each including ten healthy young adult subjects. Experiment 1 investigated the effects of subtitles on the MMN response to simple tones (differing in frequency, duration, and intensity) and speech stimuli (CV nonwords and CV words with a /d/-/g/ contrast). Experiment 2 investigated the effects of subtitles on the MMN response to a variety of CV nonword and word contrasts that incorporated both small (e.g., /d/ vs. /g/) and/or large (e.g., /e:/ vs. /eI/) acoustic deviances.

The results indicated that the presence or absence of subtitles on the distracting silent video had no effect on the amplitude of the MMN or P3a responses to simple tones, CV nonwords, or CV words. In addition, the results also indicated that movement artifacts may be statistically reduced by the presence of subtitles on a distracting silent video. The implications of these results are that more "engaging" (i.e., subtitled) silent videos can be used as a distraction task for investigations into MMN responses to speech and nonspeech stimuli in young adult subjects, without affecting the amplitude of the responses.

Key Words: Consonant-vowel tokens, distractor task, event-related potential, mismatch negativity, simple tones, subtitled videos, words

Abbreviations: CV = consonant-vowel; ERP = event-related potential; fMRI = functional magnetic resonance imaging; ISI = interstimulus interval; MMN = mismatch negativity; MMNm = magnetic mismatch negativity; SOA = stimulus onset asynchrony; SSG = semisynthesized speech generation

Sumario

El propósito de este estudio fue determinar si la presencia de subtítulos en un video silencioso de distracción, afecta la respuesta automática de negatividad desigual (mismatch negativity: MMN) a tonos simples, sonidos consonante-vocal (CV) sin sentido lingüístico, o palabras C-V. Se condujeron dos experimentos en este estudio, incluyendo a diez adultos jóvenes sanos en cada uno. El experimento 1 investigó los efectos de los subtítulos en la respuesta MMN a tonos simples (con diferencias en frecuencia, duración e intensidad) y a estímulos de lenguaje (sonidos consonante-vocal sin sentido lingüístico y palabras CV con el contraste /d/-/g/). El experimento 2 investigó los efectos de los subtítulos en la respuesta MMN con una variedad de sonidos consonante-vocal (CV) sin sentido lingüístico y contrastes de palabras que

*Division of Speech Pathology, University of Queensland, Queensland, Australia; †Neuroscan, El Paso, Texas; ‡Laboratory of Acoustics and Audio Signal Processing, Helsinki University of Technology, Finland

Reprint requests: Catharine M. Pettigrew, School of Clinical Therapies, University College Cork, Cork, Ireland

incorporaban desviaciones acústicas pequeñas (p.e., /d/ vs. /g/) y grandes (p.e., /e:/ vs. /eI/.

Los resultados indicaron que la presencia o ausencia de los subtítulos en el video silencioso de distracción, no tuvo efecto sobre la amplitud de las respuestas MMN o P3a ante tonos simples, sonidos consonante-vocal (CV) no lingüísticos o palabras CV. Además, los resultados indicaron también que los artefactos de movimiento pueden ser estadísticamente reducidos por la presencia de los subtítulos en un video silencioso de distracción. Las implicaciones de estos resultados sugieren que el uso de videos silenciosos más "involucrantes" (p.e., subtitulados) puede ser utilizado como una tarea de distracción en la investigación de respuestas MMN ante estímulos lingüísticos y no lingüísticos, en sujetos adultos jóvenes, sin afectar la amplitud de dichas respuestas.

Palabras Clave: Fichas consonante-vocal, tarea de distracción, potencial relacionado con el evento, negatividad desigual, tonos simples, videos subtitulados, palabras

Abreviaturas: CV = consonante-vocal; ERP = potencial relacionado con el evento; fMRI = imágenes por resonancia magnética funcional; ISE = intervalo inter-estímulo; MMN = negatividad desigual; MMNm = negatividad desigual magnética; SOA = asincronía en el inicio del estímulo; SSG = generación de lenguaje semi-sintético

The mismatch negativity (MMN) is a fronto-central negative component of the auditory event-related potential (ERP) that is elicited by any discriminable change ("deviant") in some repetitive aspect of auditory stimulation ("standard"), and usually peaks 100–200 msec from the onset of the stimulus change (Näätänen, 2000). For example, the MMN is elicited when a tone changes in frequency, intensity, or duration, or when a phoneme is replaced by another phoneme. The mismatch negativity has been described as an objective measure of central auditory processing in the human brain (Näätänen, 2000), and recent studies have shown that the MMN can index auditory discrimination of linguistic stimuli such as phonemes (Aaltonen et al, 1993; Titova and Näätänen, 2001), consonant-vowel (CV) syllables (Sams et al, 1990; Kraus et al, 1992; Aulanko et al, 1993; Kraus et al, 1993, 1996), and pseudowords (Ceponiene et al, 1999). Most importantly, the MMN may be elicited in the absence of attention, which makes it an ideal tool for investigating auditory processing in clinical populations (Näätänen and Escera, 2000; Näätänen, 2001).

In fact, it is well established that the optimal condition in which to elicit the MMN is the passive condition, in which the subject's attention is directed away from the auditory stimuli, in order to prevent the contamination of the MMN by other ERP components such as the N2b or P3a components (Näätänen, 2000). That is, if the subject attends to the stimuli, the superimposition of these other ERP components can make the MMN difficult to recognize (Picton, Bentin et al, 2000; Muller-Gass and Campbell, 2002). Traditionally, subjects are usually asked to watch an interesting video, play a visual computer game, or read a book of their choice, and are usually asked to ignore the auditory input, in order to distract them from attending to the stimuli (Näätänen, 2000; Sinkkonen and Tervaniemi, 2000; Muller-Gass and Campbell, 2001, 2002). Furthermore, it is important that the chosen distractor "task" is interesting to the subject, as acquiring sufficient data to extract the MMN across a number of stimulus contrasts may require 10–30 minutes per contrast (depending on the amount of sweeps per average), even for normal subjects. Moreover, the MMN may be attenuated by decreased vigilance to the distractor task (Lang et al, 1995) and increased fatigue (Sallinen and Lyytinen, 1997; Picton, Alain et al, 2000).

As with any study of the alert human, subjects may not be able to completely ignore stimuli, even if specifically instructed to do so (Muller-Gass and Campbell, 2002), which highlights the importance of using effective distractor tasks in MMN investigations. It is reasonable to assume that subjects would need to be more attentive to a difficult distractor task rather than a comparatively easy distractor task, which would leave fewer resources available for also attending the auditory stimuli, either voluntarily or involuntarily (Muller-Gass and Campbell, 2002). Interestingly, previous studies that have investigated the effects of attention-dependent visual tasks on the auditory MMN response have found that there is no effect of visual processing load on the amplitude of the MMN response to either small or large frequency deviances (Alho et al, 1992, 1994), or to stimuli deviating in frequency or duration (Kathmann et al, 1999). For example, Alho et al (1994) presented subjects with auditory stimuli containing a frequency deviance (delivered randomly to the attended and unattended ears), and visual stimuli consisting of vertical line gratings being randomly presented on a monitor. During auditory attention, subjects attended to the auditory stimuli in one ear and responded to the deviant tones. During visual attention, subjects responded to the occasional visual stimuli. MMNs were elicited by the deviant tones, whether attended or unattended, and it was found that the MMN amplitudes were not affected by attention. Furthermore, in a study by Otten et al (2000), it was found that visual processing load has no effect on the MMN responses to large frequency deviants, at either short or long interstimulus intervals (ISIs).

However, only two studies have directly investigated the effects of certain distractor tasks on MMN recordings (Muller-Gass and Campbell, 2001; McArthur et al, 2003). For example, Muller-Gass and Campbell (2001) investigated the effects of reading on the MMN response to simple tones, while subjects were asked to (1) read material (book/magazine) of their choice, (2) read material provided by the experimenter, (3) read material provided by the experimenter, with the knowledge that they would be asked questions later, and (4) sit passively (no reading material). The authors reported that the MMN tended to be larger in the imposed reading conditions and that there was no evidence of increased distractibility in either the self-selected reading or the passive conditions. These results suggest that perhaps the MMN response was facilitated by the additional "effort" required to sustain experimenter-imposed reading.

Interestingly, a recent study by McArthur et al (2003) found that allowing subjects to watch an audible video during EEG recordings can degrade the auditory ERPs to standard and deviant tones. More specifically, video sound impaired the size, latency, and split-half reliability of the MMN response, as well as decreasing the size of P2. This result is not surprising given that previous studies have indicated that certain auditory attentional tasks can affect the morphology of the MMN response (Dittman-Balcar et al, 1999; Müller et al, 2002). For example, in the study by Dittman-Balcar et al (1999), subjects were required to first perform visual discrimination tasks, and then auditory discrimination tasks, both in a series of increasing task difficulty, while the MMN response to task-irrelevant auditory stimuli were simultaneously recorded. The results indicated that visual task difficulty had no effect on the MMN response but that a significantly smaller MMN response was elicited during the most demanding auditory condition.

Several recent studies that have obtained robust MMN responses to meaningful and nonmeaningful speech stimuli have used a variety of distractor tasks, such as videos with a soundtrack (Wunderlich and Cone-Wesson, 2001), silent videos (Pulvermüller et al, 2001; Shtyrov and Pulvermüller, 2002), and silent videos with subtitles (Pettigrew et al, 2004). However, the MMN results of these studies have been conflicting; therefore, it is important to consider the possible effects of a language-based distractor task (e.g., reading, subtitled videos, videos with soundtracks) on the MMN response to linguistically meaningful and nonmeaningful speech stimuli. Interestingly, several studies have successfully obtained robust MMN responses to nonspeech stimuli (e.g., tones) while using subtitled, silent videos as a distraction (Paavilainen et al, 1999; Jaramillo et al, 2001; Pettigrew et al, 2004). However, although Jaramillo et al (2001) also obtained MMN responses to vowel stimuli using subtitled videos as a distraction, Pettigrew et al (2004) were unable to obtain such responses to nonmeaningful CV nonwords or meaningful CV words using a /d/-/g/ fine acoustic contrast, using the same distraction method.

The possible effects of subtitles on the MMN response to speech and nonspeech stimuli have not yet been investigated; therefore, the aim of the present study is to investigate the possible enhancing/attenuating effects of reading video subtitles on the MMN response to simple tones, nonword speech stimuli, and real English word stimuli, in two experiments on healthy young adult subjects. The first experiment was designed in accordance with the study by Pettigrew et al (2004), and investigated the possible effects of reading video subtitles on the amplitude and latency of the MMN response to simple tones deviating in frequency, duration, and intensity, as well as nonwords and real words. The second experiment investigated the possible effects of reading video subtitles on the amplitude and latency of the MMN response to a wider variety of word versus nonword speech stimulus contrasts contained within a multiple deviant paradigm.

METHODS

Experiment 1

Subjects

Ten healthy native-English-speaking young adults (six females, four males) participated in Experiment 1 (mean age of 24.4 years; SD = 2.01, range = 21–28 years). All subjects were right-handed, had normal hearing thresholds (≤20 dBHL at 500–4000 Hz), and gave their informed consent prior to testing.

Stimuli

All subjects were presented with simple tone stimuli and speech stimuli. The deviant tone stimuli differed from the standard tone stimuli in frequency, duration, or intensity. The speech stimuli included a nonword CV syllable contrast (/de:/ vs. /ge:/) and a real word CV syllable contrast (/deI/ vs. /geI/). The stimulus blocks described below were randomly generated twice, so that two versions of the paradigm were presented to the subject. Each block was presented to the subject once with subtitles provided on the silent video, and once without the subtitles provided on the video, with all tone and speech stimulus blocks (as well as the presence/absence of subtitles) being presented in a random order across subjects.

The features of the simple tones were adapted from Paavilainen et al (1991) and were generated using Neuroscan® STIM software. The standard tones were 600 Hz, 75 dB SPL, sinusoidal tone bursts with 50 msec duration, including a 5 msec rise/fall time. The frequency deviant stimuli were 650 Hz tones, the duration deviant stimuli were 20 msec tones, and the intensity deviant stimuli were presented at 60 dB SPL. For each condition (subtitles present vs. subtitles absent), two blocks of 816 simple tone stimuli were presented to each subject using a multiple deviant oddball paradigm. The use of a multiple deviant paradigm for the tone stimulus presentation enabled the testing time to be reduced without affecting the amplitude of the MMN response to each deviant stimulus type (Deacon et al, 1998). Each of the deviant stimuli (frequency, duration, and intensity deviants) had a 12.5% probability of occurrence, resulting in a total of 204 sweeps obtained in response to each deviant stimulus. The stimuli were presented with a constant stimulus onset asynchrony (SOA) of 610 msec and in a pseudorandom fashion, so that no less than two standards were presented between deviant stimuli. In addition, each deviant stimulus was presented alone in a single block of 250 stimuli to obtain responses for the "deviant alone" condition.

The speech stimuli consisted of four CV stimuli (the nonwords /de:/ "de" and /ge:/ "ge," and the words /deI/ "day" and /geI/ "gay") and were created using the semisynthesized speech generation (SSG) technique (Alku et al, 1999). SSG makes it possible to synthesize natural sounding, yet controllable, speech stimuli by exciting an artificial vocal tract model with a real excitation of the human voice production mechanism, the glottal waveform. The formant values required for synthesis of the speech stimuli were computed with SSG from natural utterances produced by a native male speaker of Australian English (see Table 1 for values of the lowest four formants).

All the stimuli were synthesized using the same glottal excitation waveform, which implies that the fundamental frequency was the same across all of the stimuli and the acoustic differences between the stimuli were caused solely by the formant settings. More

Table 1. Center Frequencies of the Lowest Four Formants of the Speech Stimuli

	Formant Frequencies (Hz)			
Phoneme	**F1**	**F2**	**F3**	**F4**
/e/	689	1615	2562	3854
/I/$_1$	517	2046	2562	4000
/I/$_2$	409	2239	2606	4350
/d/	474	1701	2778	3941
/g/	495	1938	2498	3704

Note: The phonemes /I/$_1$ and /I/$_2$ are representatives of the same phoneme /I/ in English.

specifically, the CV stimuli were created so that the formant settings shifted continuously in time from one phoneme to another. For example, the start of the CV word "day" was in the formant settings of the phoneme /d/. After 45 msec, the settings were of the phoneme /e/, after which there was a time span of 91 msec during which the settings shifted to the phoneme /I/. During the last part of the word (duration 168 msec), the formant settings were in the positions of the phoneme /I/. However, in order to make the vowel sound more natural, the formant settings of /I/ were shifted from one setting (phoneme /I/$_1$ in Table 1) to another (phoneme /I/$_2$ in Table 1). For the word "gay," the lengths of the segments were the same as those for the word "day." For the two nonword CV stimuli, the length of the first segment (i.e., the shift from /d/ to /e/ for the nonword "de" and the shift from /g/ to /e/ for the nonword "ge") was also 45 msec. However, during the remainder of the nonwords, the settings were in a constant position (i.e., the phoneme /e/). Since all four stimuli began with a voiced plosive, a low amplitude segment of 28 msec was added to the beginning of each word. This segment, called the voice bar, reflects the fundamental frequency of phonation with no formant structure (Kent and Read, 1992). Thus, the total duration was 332 msec for all the stimuli.

The fundamental frequency (F_0) of the stimuli changed over time according to the intonation used by the speaker, as the glottal excitation used in the synthesis of the stimuli was computed from a natural utterance. The value of F_0 was 130 Hz at the beginning of the stimuli and was 105 Hz at the end of the stimuli. Finally, normalizing the energies of the stimuli equalized sound intensity. All speech stimuli were presented at an intensity level of 75 dB SPL.

The CV syllables /de:/ and /ge:/ were reported to be "nonwords" and the CV syllables /deI/ and /geI/ reported to be "real words" by ten native speakers of Australian English, prior to the commencement of testing (these native speakers did not participate in the ERP recordings). As the procedures used to generate the CV syllables /de:/ and /ge:/ involved the truncation of the diphthong vowel within the CV words /deI/ and /geI/, and the subsequent lengthening of that vowel to match the stimulus duration of the CV words, the resultant vowel within the CV syllables comprised formant values that do not easily fall within the range of a traditional singular vowel of English (Kent and Read, 1992). More specifically, the first and second formant values for the pseudovowel /e:/ bordered the outer formant value possibilities for the English vowels "u" in "up," "e" in "herd," and "a" in "at." As a result, the CV syllables /de:/ and /ge:/ were not considered to follow the phonotactic rules of English and were therefore labelled as nonwords rather than pseudowords.

In the CV nonword and CV word stimulus conditions (for both subtitles present vs. subtitles absent conditions), subjects were presented with two blocks of 800 stimuli per stimulus condition, using a single contrast oddball paradigm. In the CV nonword condition, the stimulus /de:/ acted as the standard (87.5% probability), and the stimulus /ge:/ acted as the deviant (12.5% probability). In the CV word condition, the stimulus /deI/ ("day") acted as the standard (87.5% probability) and the stimulus /geI/ ("gay") acted as the deviant (12.5% probability). A total of 200 sweeps were obtained to each deviant stimulus type. In all speech stimulus blocks, the stimuli were presented at an SOA of 900 msec and were presented in a pseudorandom order, so that no less than two standards were presented between deviant stimuli. In addition, each deviant stimulus was presented alone in a single block of 250 stimuli (the "deviant alone" condition). Following the testing sessions, each subject was asked to report on his/her levels of distractibility when subtitles were present or absent on the silent video.

Procedure

Subjects were tested in an acoustically attenuated and electrically shielded room. During the ERP recordings, subjects were instructed to ignore the auditory stimuli and watch a silent, subtitled video of their choice from a limited selection of foreign language videos. None of the subjects were familiar with the plots of the chosen videos prior to testing. Throughout the experiment, stimuli were presented using the Neuroscan STIM system and EAR insert earphones. For stimulus blocks that allowed subtitles, the subtitles for the video being watched were visible to the subject. For stimulus blocks that did not allow subtitles, the subtitles of the ongoing video were unobtrusively covered (this did not affect the viewing of the video on the screen). The ERP measurements lasted approximately two and a half to three hours per subject, including breaks.

Electroencephalographic (EEG) Recording

A 32-channel Neuroscan Quikcap was used to record the nose-referenced EEG (500 Hz sampling rate). All EEG measurements were made using a Neuroscan SynAmps amplifier and Neuroscan Scan 4.2 acquisition software. Recordings were obtained from the following locations using sintered electrodes, FP1, FP2, FPZ, FZ, FCZ, CZ, CPZ, PZ, F7, F3, F4, F8, FT7, FC3, FC4, FT8, T3, C3, C4, T4, TP7, CP3, CP4, TP8, T5, P3, P4, and T6, which were positioned according to the 10-20 International Electrode System (Jasper, 1958). The quality of the EEG recordings was monitored during data acquisition, and the continuous EEG data were stored on the computer for off-line averaging. Electrodes attached to the right and left outer canthi monitored horizontal eye movements, and vertical eye movements were monitored by electrodes attached above and below the orbit of the left eye. Reference electrodes were also placed on the left and right mastoids (M1 and M2, respectively) for later rereferencing of averaged data.

Experiment 2

Subjects

Ten healthy native-English-speaking young adults (five females, five males), none of whom participated in Experiment 1, participated in Experiment 2 (mean age of 25.70 years; SD = 4.16, range = 19–33 years). All subjects were right-handed, had normal hearing thresholds (≤20 dB at 500–4000 Hz), and gave their informed consent prior to testing.

Stimuli

All subjects were presented with the four CV stimuli (the nonwords /de:/ "de" and /ge:/ "ge," and the words /deI/ "day" and /geI/ "gay") in a multiple deviant paradigm. Similar to Experiment 1, the stimulus blocks described below were randomly generated twice, so that two versions of the paradigm were presented to the subject. Each block was presented to the subject once with subtitles provided on the silent video, and once without the subtitles provided on the video, with all stimulus blocks (including the presence/absence of subtitles) being presented in a random order across subjects.

For each condition (subtitles present vs.

Table 2. Stimulus Block Types

Type 1		Type 2		Type 3		Type 4	
Standard	Deviants	Standard	Deviants	Standard	Deviants	Standard	Deviants
de	ge	day	ge	ge	de	gay	ge
	day		gay		gay		de
	gay		de		day		day

subtitles absent), 16 blocks of 510 stimuli were presented to each subject using a multiple deviant oddball paradigm. Blocks were presented in a random order across subjects. Each of the four speech stimuli acted as a standard (p = 0.7) in four of the 16 blocks, with the remaining stimuli acting as deviants (p = 0.1 each) for those block types (see Table 2).

The stimuli were presented with an SOA of 900 msec and in a pseudorandom fashion, so that no less than two standards were presented between deviant stimuli. A total of 204 sweeps were obtained in response to each deviant stimulus against each standard stimulus. In addition, each deviant stimulus was presented alone in a single block of 200 stimuli (the "deviant alone" condition).

Procedure and EEG Recording

The procedure and EEG recording specifications were the same for Experiment 2 as for Experiment 1, except that the recordings were obtained over two testing sessions for Experiment 2 due to the number of trials presented. For each subject, the block types presented were split over the two testing sessions. The order in which the two testing sessions were presented (Blocktypes 1 and 2 vs. Blocktypes 3 and 4) was random across subjects, such that four subjects received Blocktypes 1 and 2 in the first session, and six subjects received Blocktypes 3 and 4 in their first session. Splitting the Blocktypes in this way over the two testing sessions enabled the subtitle vs. nonsubtitle condition comparisons to be made within stimulus contrasts, without the presence of confounding variables deriving from comparisons across testing sessions. The ERP measurements lasted approximately two and a half to three hours per subject, including breaks, and the two sessions were conducted within an average of 5.5 (SD = 5.54) days of each other.

Data Analysis

For both Experiment 1 and Experiment 2, the continuous EEG recordings were analyzed off-line. Eye-blink artifacts were removed from the data, and the analysis period was 700 msec, including a 100 msec baseline. Individual epochs were subjected to a linear detrend over the entire epoch. This was followed by baseline correction based on the pre-stimulus interval. Artifact rejection (±100 µV), averaging, and digital bandpass filtering (1–30 Hz; 12 dB/octave) were then carried out on the standard and deviant stimulus epochs in each block. The first 10 sweeps of each epoched file were excluded from the averaging process to reduce N1 amplitude variation associated with the start of the stimulation sequence (Pekkonen et al, 1995; Lavikainen et al, 1997; Sinkkonen and Tervaniemi, 2000). The same analyses were carried out on the "deviant alone" epochs. All of the processed average files were then individually rereferenced to the mastoids and baseline corrected. As the MMN inverts in polarity at electrodes below the level of the Sylvian fissure, arithmetic rereferencing of the contributing ERP waveforms to the average of the mastoids maximizes the MMN amplitudes at the frontal electrodes (Näätänen, 1995), thereby increasing the signal-to-noise ratio (Schröger, 1998).

Difference waveforms for all stimulus contrasts were calculated by subtracting the averaged response to the deviant when it was presented alone, from the averaged response to the deviant when it was presented in the oddball paradigm. For comparison purposes, difference waveforms were also calculated by subtracting the averaged response to the standard from the averaged response to the deviant. Grand average waveforms for all the stimulus contrasts were calculated across all subjects in each experiment. Grand average waveforms were rereferenced to the mastoids before visual inspection and statistical analyses.

In order to measure the MMN amplitudes, the grand average peak latency for each contrast was first determined from the grand average difference waves as the largest negative peak between a specified time period, at the electrode site at which the grand average MMN responses tended to be largest. For Experiment 1, the time period for grand average MMN peak detection was 100–300 msec (similar to Pettigrew et al, 2004). The electrode site at which the responses tended to be largest was FZ for the tone stimuli but was more difficult to determine for the speech stimuli due to the poor responses. Therefore, FZ was chosen as the electrode site for peak detection for both the tone and speech stimuli (similar to

Wunderlich and Cone-Wesson, 2001). For Experiment 2, the grand average peak latency for each contrast was determined from the grand average difference waves as the largest negative peak between 150 msec and 350 msec, at FCZ (the site at which the MMN response tended to be largest). For each individual average difference waveform, for all subjects in both Experiments, the MMN mean amplitude was measured over a 40 msec time window centered at the grand average peak latency, for all scalp locations (Winkler et al, 1998). The latency of the MMN for each individual was determined using this same time window.

The P3a response tended to achieve maximum amplitude at FCZ for all conditions in both experiments. Therefore, the peak latency of the P3a response was measured from the grand average difference response at FCZ, at the latency of the most positive peak within 170 msec of (or immediately following) the MMN response. The individual P3a mean amplitudes and latencies were measured over a 40 msec interval centered at the P3a peak latency (Winkler et al, 1998).

Statistical Analyses

First, the percent of total sweeps rejected for each individual, due to artifact, was calculated for each subtitle condition (i.e., subtitles present vs. absent) for each stimulus type (i.e., tones and speech for Experiment 1, and speech for Experiment 2), and the average percentages of artifacts rejected across all subjects, per stimulus type, were then compared using paired t-tests. Second, for both experiments, two-tailed t-tests were used to determine if the mean amplitudes of the MMN grand average responses and P3a grand average responses (i.e., the average of the mean amplitudes and latencies across all subjects) significantly differed from zero for each stimulus contrast.

The amplitudes and peak latencies of the grand average MMN and P3a responses were then compared across subtitle conditions using Two-Way Analysis of Variance (ANOVA) for dependent measures (within-subject factors: subtitles, contrast). Greenhouse-Geisser corrections were applied when appropriate.

RESULTS

Total Sweeps Rejected

The mean percentage of total sweeps rejected due to artefact within each Experiment are presented in Table 3. The results indicate that in both the tone and speech conditions of Experiment 1, there was no significant difference in percent of sweeps rejected between subtitle conditions (present vs. absent). However, for Experiment 2, the results show that the subtitles "present" condition resulted in a significantly smaller percentage of sweeps rejected than the subtitles "absent" condition [$t(9) = 2.289$, $p < 0.05$).

Experiment 1

The grand average ERP waveforms at FZ for the tone stimulus contrasts are presented in Figure 1, and the grand average ERP waveforms at FZ for the speech stimulus contrasts are presented in Figure 2. The grand average ERP waveforms across all the fronto-central electrodes were of a similar

Table 3. Mean Percentage (%) of Sweeps Rejected Due to Artifact, per Stimulus Type, with Standard Deviation

Experiment 1:	Subtitles Present	Subtitles Absent
Tone Stimuli	1.17 ± 0.79	1.72 ±1.00
Speech Stimuli	1.72 ±1.12	2.27 ±1.98
Experiment 2:	**Subtitles Present**	**Subtitles Absent**
Speech Stimuli	2.15[a] ±1.10	2.84[a] ±1.81

[a] Mean percentages are significantly different ($p < 0.05$).

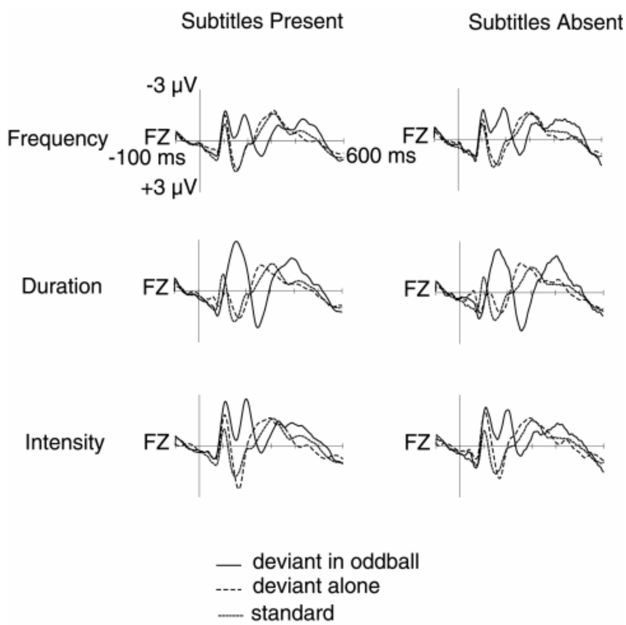

Figure 1. Experiment 1: Rereferenced grand average ERP waveforms at FZ for each tone contrast, including frequency (top), duration (middle), and intensity (bottom), for the subtitles present condition (left column) and subtitles absent condition (right column).

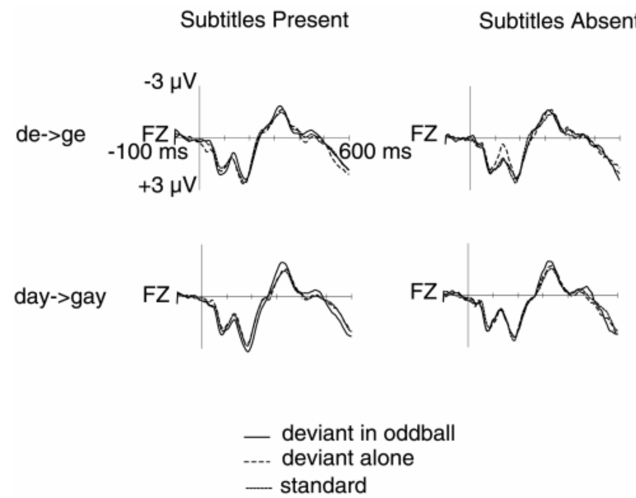

Figure 2. Experiment 1: Rereferenced grand average ERP waveforms at FZ for each speech contrast, including CV nonwords "de" ➡ "ge" (top) and CV words "day" ➡ "gay" (bottom), for both the subtitles present condition (left column) and subtitles absent condition (right column).

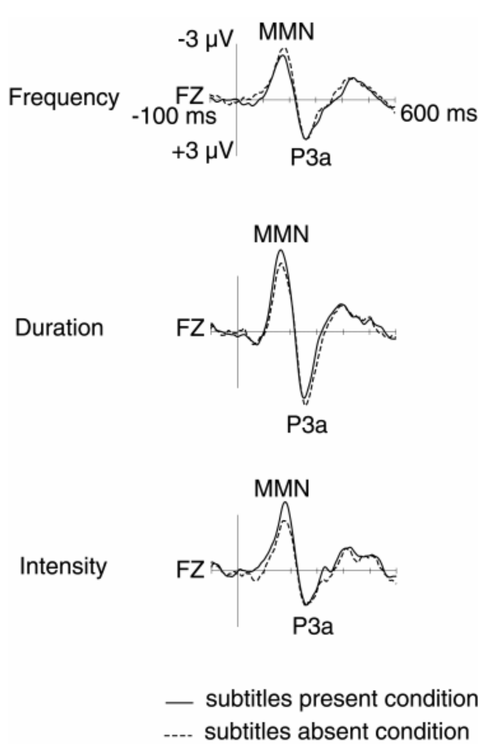

Figure 3. Experiment 1: Rereferenced grand average difference waveforms (deviant minus deviant alone) at FZ for each tone contrast, including frequency (top), duration (middle), and intensity (bottom), for both the subtitles present and absent conditions.

pattern to those shown at FZ.

MMN and P3a Responses to Simple Tones

The grand average difference waveforms (deviant minus deviant alone) at FZ in response to the tone stimuli are shown in Figure 3. All the deviant stimulus types elicited visually well-defined MMN responses, which inverted their polarity at the mastoid electrodes, below the Sylvian fissure. The grand average mean amplitudes and peak latencies of the MMN and P3a components at FZ for each tone contrast are shown in Table 4.

Grand Average MMN Response. The mean amplitudes of the MMN response in each tone contrast, with and without subtitles, were significantly different from zero ($p < 0.01$, two-tailed t-test). As expected, the Two-Way ANOVA revealed a significant main effect of contrast on both the amplitude [$F(2, 18) = 3.902$, $p < 0.05$] and latency [$F(2, 18) =$

Table 4. Experiment 1: Grand Average Mean Amplitudes and Peak Latencies of MMN and P3a Responses (with Standard Deviations) at FZ, for the Tone Conditions

	Subtitles Present			Subtitles Absent		
MMN	Frequency	Duration	Intensity	Frequency	Duration	Intensity
Mean Amplitude (µV)	-2.10** ± 2.02	-3.92*** ±1.16	-3.21*** ±1.42	-2.46** ±1.50	-3.22*** ±1.69	-2.36*** ±1.22
Peak Latency (msec)	175.80 ±14.16	165.80 ±13.65	181.40 ±13.17	181.00 ±16.42	170.20 ±11.87	175.80 ±14.89
P3a						
Mean Amplitude (µV)	1.91** ±1.41	3.38** ± 2.19	1.61* ±1.88	1.87** ±1.55	4.03*** ± 2.04	1.77** ±1.37
Peak Latency (msec)	263.30 ± 9.97	257.80 ±13.55	254.00 ±12.54	258.00 ±13.50	256.40 ± 9.51	252.40 ±11.35

* $p < 0.05$
** $p < 0.01$
*** $p < 0.001$

6.509, $p < 0.01$] of the MMN response. Importantly, however, there was no main effect of subtitles on the amplitude or latency of the MMN responses, for any contrast, nor was there a significant interaction effect between subtitles and contrast.

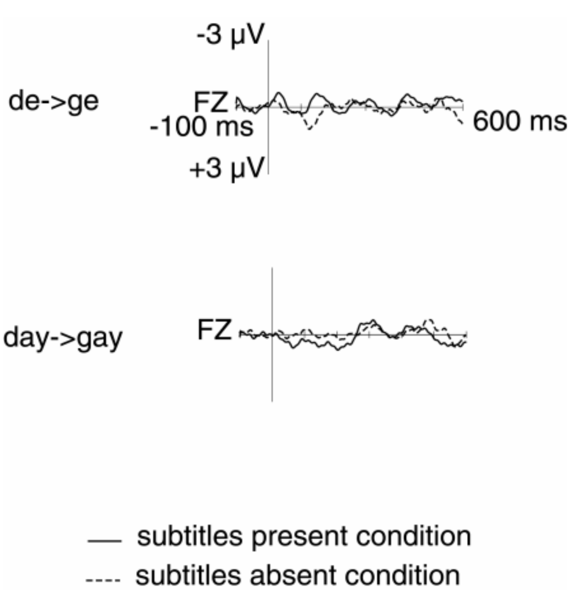

Figure 4. Experiment 1: Rereferenced grand average difference waveforms at FZ for each speech contrast, including CV nonwords "de" ➡ "ge" (top) and CV words "day" ➡ "gay" (bottom), for both the subtitles present and absent conditions.

Grand Average P3a Response. Visual inspection of the grand average difference waves revealed the presence of a positive component immediately following the MMN, which was identified as a P3a response. The mean amplitudes of the P3a response in each tone contrast, with and without subtitles, were significantly different from zero ($p < 0.05$, two-tailed t-test). The Two-Way ANOVA revealed that although there was a significant main effect of contrast on the amplitude of the P3a responses [$F(2, 18) = 11.596$, $p < 0.01$], there was no main effect of subtitles on the amplitude or latency of the P3a response for each contrast, nor was there any interaction effect.

MMN Responses to Speech Stimuli

There was no clear maximum electrode for the grand average speech data for either the CV nonword or CV word contrast. The analysis of the speech data was, therefore, based on measurements at FZ, similarly to the tone data, and as used by Wunderlich and Cone-Wesson (2001). The grand average difference waveforms (deviant minus deviant alone) in response to the speech stimuli at FZ are shown in Figure 4, and the grand average mean amplitudes at FZ for each speech contrast are shown in Table 5.

Grand Average MMN Response. The mean amplitudes of the MMN responses to

Table 5. Experiment 1: Grand Average Mean Amplitudes and Peak Latencies of MMN Responses (with Standard Deviations) at FZ, for the Speech Conditions

	Subtitles Present		Subtitles Absent	
MMN	de ➡ge	day ➡gay	de ➡ge	day ➡gay
Mean Amplitude (µV)	-0.46 ±0.89	-0.50 ±0.78	-0.22 ±1.06	-0.29 ±1.21
Peak Latency (msec)	-----	-----	-----	-----

the stimuli were not significantly different from zero for either contrast ($p > 0.05$, two-tailed t-test). Similar measurements of mean amplitude at FCZ also did not yield significant results. As the MMN responses to these contrasts were not significant, peak latency measurements are not reported, nor were analyses carried out on the peak latencies of these responses. Two-way ANOVAs of the results at FZ indicated that there were no significant main effects of subtitles or contrast on the MMN amplitudes ($p > 0.05$). Furthermore, as there were no apparent P3a responses upon visual inspection of the waveform data, P3a measurements were not carried out on these data.

Experiment 2

For the multiple deviant conditions, data are grouped in Figure 5 as blocks based on the four CV stimuli used as standards (the nonwords /de:/ "de" and /ge:/ "ge," and the words /deI/ "day" and /geI/ "gay"). The grand

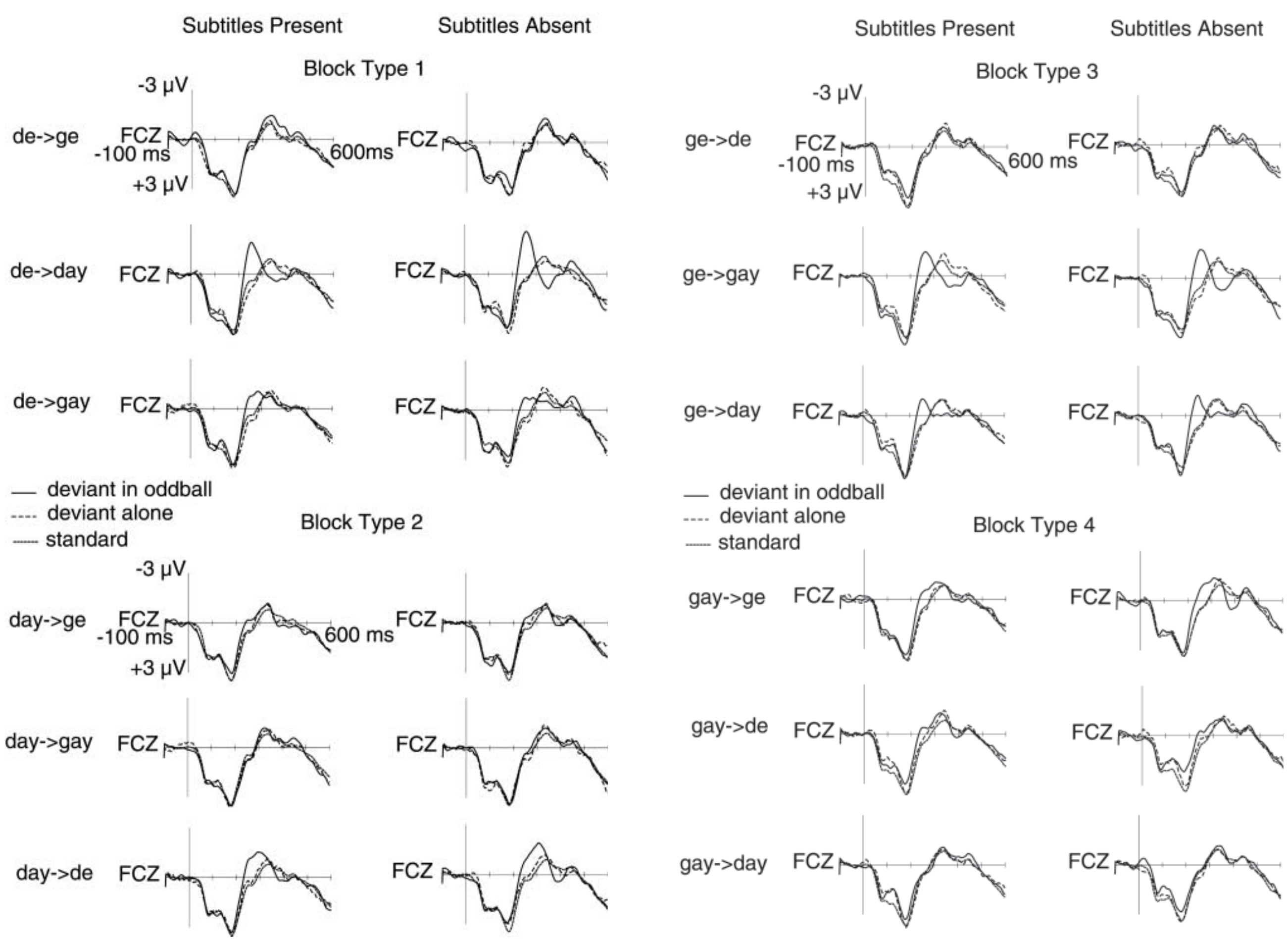

Figure 5. Experiment 2: Rereferenced grand average ERP waveforms at FCZ for each stimulus contrast within Block Types: (a) Block Type 1 (top) and Block Type 2 (bottom), (b) Block Type 3 (top) and Block Type 4 (bottom), for both the subtitles present condition (left column) and subtitles absent condition (right column).

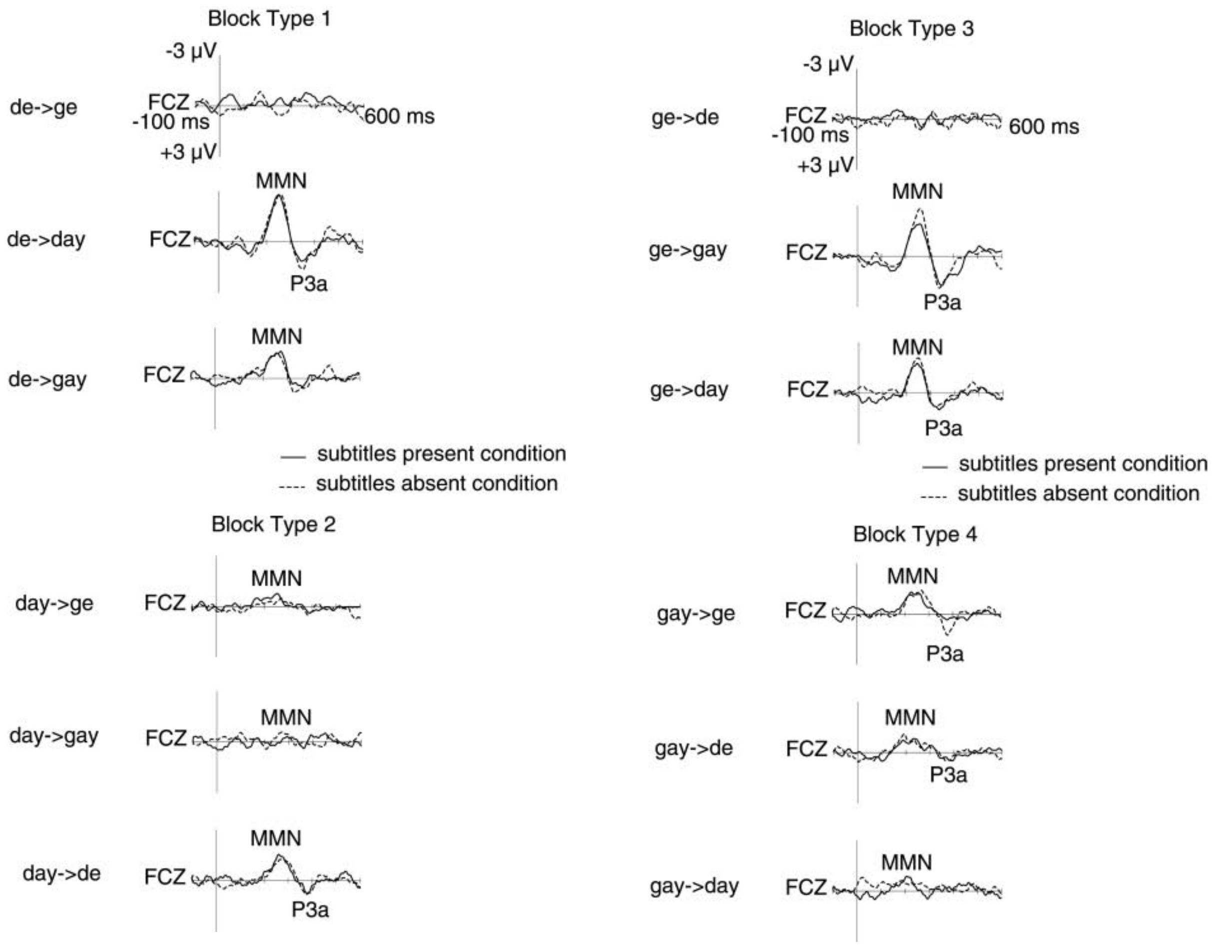

Figure 6. Experiment 2: Rereferenced grand average difference waveforms (deviant minus deviant alone) at FCZ for each stimulus contrast within Block Types: (a) Block Type 1 (top) and Block Type 2 (bottom), (b) Block Type 3 (top) and Block Type 4 (bottom), for both the subtitles present and absent conditions.

average ERP waveforms across all the fronto-central electrodes were of a similar pattern to those shown at FCZ.

Grand Average MMN Responses. The grand average difference waveforms (deviant minus deviant alone) at FCZ, for each experimental block of contrasts, are presented in Figure 6. The waveforms were of a similar pattern across all fronto-central electrode sites. The mean amplitude values for the grand average MMN responses at FCZ, and the peak latencies for those responses with amplitudes significantly different from zero, are presented in Table 6.

Two-Way ANOVAs indicated that there was, as expected, a significant main effect of contrast on the mean amplitude of the MMN response, [$F(11, 99) = 7.160$, $p < 0.001$]; however, there was no significant main effect of subtitles on the mean amplitude of the MMN responses, nor was there any interaction effect between subtitles and contrast. Only 6 of the 12 contrasts elicited significant MMN responses in both the subtitled and nonsubtitled conditions ("de" ➡ "day," "de" ➡ "gay," "day" ➡ "de," "ge" ➡ "gay," "ge" ➡ "day," "gay" ➡ "ge"). Therefore, paired t-tests were carried out for each of these contrasts on the peak latency of the MMN responses in both conditions. The results indicated that for the contrasts "ge ➡gay" and "gay ➡ge," the subtitled condition elicited MMN responses at earlier latencies than the nonsubtitled condition ([$t(9) = 2.329$, $p < 0.05$] and [$t(9) = 5.786$, $p < 0.001$], respectively). The remaining four

Table 6. Experiment 2: Mean Amplitudes and Peak Latencies of the Grand Average MMN Responses at FCZ for each Contrast (with Standard Deviations)

	Subtitles Present			Subtitles Absent		
Block Type 1	**de → ge**	**de → day**	**de → gay**	**de → ge**	**de → day**	**de → gay**
Mean Amplitude (µV)	-0.55 ± 0.80	-2.45** ±1.51	-1.37** ±1.10	-0.64 ±1.42	-2.42** ±1.48	-1.29*** ± 0.71
Peak Latency (msec)	-----	248.40 ±10.45	265.20 ±11.44	-----	258.20 ±10.69	257.80 ±15.56
Block Type 2	**day → ge**	**day → gay**	**day → de**	**day → ge**	**day → gay**	**day → de**
Mean Amplitude (µV)	-0.81* ± 0.98	-0.23 ± 0.87	-1.30** ± 0.83	-0.34 ±1.35	-0.49* ± 0.65	-1.13* ±1.15
Peak Latency (msec)	251.00 ±10.55	-----	266.60 ±9.38	-----	274.60 ±16.71	274.80 ±13.96
Block Type 3	**ge → de**	**ge → gay**	**ge → day**	**ge → de**	**ge → gay**	**ge → day**
Mean Amplitude (µV)	-0.47 ± 0.76	-1.65*** ± 0.85	-1.54** ± 1.33	-0.23 ± 0.89	-2.41*** ± 1.33	-1.82** ± 1.46
Peak Latency (msec)	-----	253.40 ±10.59	241.80 ± 14.53	-----	261.60 ± 8.21	250.20 ± 12.10
Block Type 4	**gay → ge**	**gay → de**	**gay → day**	**gay → ge**	**gay → de**	**gay → day**
Mean Amplitude (µV)	-1.12** ± 0.72	-0.69 ±1.01	-0.70* ± 0.79	-1.22** ± 1.15	-0.85** ± 0.70	-0.53 ± 0.96
Peak Latency (msec)	233.60 ±11.62	-----	203.60 ±11.77	271.00 ±14.40	194.00 ± 11.47	-----

*p < 0.05
** p < 0.01

contrasts did not elicit MMN responses of differing latencies across subtitle conditions (p > 0.05).

Grand Average P3a Responses. The mean amplitude values for the grand average P3a responses at FCZ, and the peak latencies for those responses with amplitudes significantly different from zero, are presented in Table 7.

Although there was an expected significant main effect of contrast on the mean amplitude of the P3a response [$F(11, 99) = 3.779$, $p < 0.001$], there was no significant main effect of subtitles, nor was there any significant interaction effect. Only 2 of the 12 contrasts elicited significant P3a responses in both the subtitled and nonsubtitled conditions ("de" → "day," "ge" → "gay"). Paired t-tests were carried out on the peak latencies of the MMN responses in both conditions for each of these two contrasts, and revealed that for the contrast "ge" → "gay," the subtitled condition elicited a P3a response at a longer latency than the nonsubtitled condition [$t(9) = -2.332$, $p < 0.05$]. The remaining contrast "de" →

Table 7. Experiment 2: Mean Amplitudes and Peak Latencies of the Grand Average P3a Responses at FCZ for each Contrast (with Standard Deviations)

	Subtitles Present			Subtitles Absent		
Block Type 1	**de → ge**	**de → day**	**de → gay**	**de → ge**	**de → day**	**de → gay**
Mean Amplitude (µV)	-0.31 ± 0.43	0.97** ± 0.74	0.36 ± 1.10	0.47 ±1.49	1.34* ±1.55	0.64 ±1.06
Peak Latency (msec)	-----	347.60 ±14.93	-----	-----	345.60 ±10.23	-----
Block Type 2	**day → ge**	**day → gay**	**day → de**	**day → ge**	**day → gay**	**day → de**
Mean Amplitude (µV)	0.27 ± 0.61	0.18 ± 0.64	0.53 ±1.16	0.33 ± 0.50	0.06 ± 0.97	0.68* ± 0.84
Peak Latency (msec)	-----	-----	-----	-----	-----	373.60 ±13.96
Block Type 3	**ge → de**	**ge → gay**	**ge → day**	**ge → de**	**ge → gay**	**ge → day**
Mean Amplitude (µV)	0.30 ±1.00	1.27* ±1.38	1.01* ±1.28	0.40 ± 0.68	1.39** ± 1.25	0.66 ± 1.19
Peak Latency (msec)	-----	350.60 ±15.44	336.20 ±12.45	-----	339.60 ±12.71	-----
Block Type 4	**gay → ge**	**gay → de**	**gay → day**	**gay → ge**	**gay → de**	**gay → day**
Mean Amplitude (µV)	0.24 ± 0.78	0.36* ± 0.47	0.28 ± 0.95	0.96** ± 0.90	0.37 ± 0.71	-0.21 ± 0.57
Peak Latency (msec)	-----	387.80 ±16.62	-----	381.60 ±11.50	-----	-----

*p < 0.05
**p < 0.01

"day" did not elicit MMN responses of differing latency across subtitle conditions ($p > 0.05$).

Summary of Results. Overall, the results indicate that the use of silent, subtitled videos as a distractor task does not affect the amplitude of MMN responses to simple tones (frequency, duration, or intensity deviants). Furthermore, the results show that the use of subtitled videos as a distraction from the auditory speech input has no attenuating effect on responses to the fine acoustic contrast /d/-/g/ in CV nonwords and CV words (e.g., "de" ➡ "ge," "day" ➡ "gay"), nor any enhancing effect on responses to larger acoustic contrasts that also incorporate a lexical change (e.g., "de" ➡ "day"). The results also indicate that there may exist an effect of subtitles on the peak latency of the MMN response to CV speech stimuli with a large acoustic difference (i.e., "ge" ➡ "gay" and "gay" ➡ "ge"), such that the presence of subtitles may elicit an earlier MMN response. However, this effect was not consistent across similar stimulus contrasts (e.g., "de" ➡ "day," "day" ➡ "de"). Furthermore, the results indicate that there may also be a subtitled effect on the peak latency of the P3a response to CV speech stimuli with a large acoustic difference ("ge" ➡ "gay") such that the presence of subtitles elicits a slower P3a response. Once again, however, this effect was not consistent across similar stimulus contrasts (e.g., "de" ➡ "day"). Importantly, the results indicate that the presence of subtitles in a silent film may significantly contribute to the reduction of movement artifacts, most likely because the subject is more engaged in the distractor task, thus enabling "cleaner" EEG measurements to be recorded. In addition, all subjects reported that they were able to more effectively "ignore" the auditory input when subtitles were provided with the video.

DISCUSSION

The results of the present study indicate that the presence or absence of subtitles on the distracting video had no significant attenuating or enhancing effect on the amplitude of the MMN responses to simple tones, CV syllables, or CV words (with either small or large acoustic deviances). In addition, the amplitudes of the P3a responses (which reflect automatic attention switch to the auditory input) were not affected. Taken together, these results indicate that the allocation of attentional resources to a distractor task that involves both visual processing and language processing (silent video with subtitles), rather than just visual processing (silent video without subtitles), has no effect on the amplitudes of the automatic MMN and P3a responses to nonlinguistically meaningful or linguistically meaningful auditory input.

These results support the findings of previous studies that found no effect of visual processing load on the amplitude of MMN responses to tonal stimuli (Alho et al, 1992, 1994; Kathmann et al, 1999; Picton, Alain et al, 2000). More importantly, however, these results also provide new evidence that the high-level, attentional linguistic processing of video subtitles does not affect the simultaneous, early levels of automatic acoustic and/or linguistic processing of auditory speech input, as reflected by the amplitude of the MMN response. The apparent functional independence of these two levels of linguistic processing is not surprising, given that the MMN has been reported to occur relatively independently of cognitive processes (Kraus et al, 1995; Kraus and Cheour 2000), and the fact that two different processing modalities (visual vs. auditory) are involved in processing the auditory stimuli and the visual distraction. This proposal is further supported by previous functional magnetic resonance imaging (fMRI) findings that visual and auditory processing of sentences occur using different neural substrates, including differences in location of cortical activation, and differences in the total amount of activation in the two modalities in several regions (Michael et al, 2001). That is, language processing through the visual modality would not significantly affect the auditory processing of stimulus input or the subsequently generated MMN response, even when linguistic stimuli are involved in both instances.

Interestingly, the finding that the presence of subtitles in a video distraction task may sometimes elicit an earlier MMN response and slightly slower P3a response to CV speech stimuli with large acoustic contrasts, also incorporating a lexical change (e.g., "ge" ➡ "gay"), may provide further evidence to suggest that the subtitles provide a more effective distraction than the absence of subtitles. That is, it is reasonable to propose that the distracting presence of subtitles on the video may have facilitated a quicker MMN response from the automatic MMN generators, and a slower switch to attention (i.e., delayed P3a response), to the large acoustic deviance in the auditory input. However, this effect was not consistent across similar contrasts; therefore, the results must be interpreted with caution, and further studies are needed to determine the consistency of the effect of subtitles on the peak latency of MMN and P3a responses to a wider range of speech stimuli.

It is likely that the findings of the present study may be generalized to the use of reading material as a distractor task during MMN experiments using speech and nonspeech stimuli. That is, it may be proposed that the written language processing of words and sentences in a book would not significantly affect the amplitude of the simultaneously recorded MMN response to linguistically meaningful and/or nonmeaningful stimuli presented auditorily. However, the most pertinent difference between reading a book and reading subtitles on a video, as a distraction from the auditory input, is the visual processing demands imposed, or the easiness of the task. For example, when a subject is presented with subtitles on a screen, they have limited time in which to process the written information, and in addition, visual processing of the background scenes is presumably ongoing. In contrast, when a subject is provided with a book or magazine as reading material, they are usually allowed to read at their own pace. Given that the arduousness of the primary visual task may determine the attentional resources available for also attending to the auditory stimuli (either voluntarily or involuntarily), it is possible that the use of subtitled videos may be more effective than reading in distracting the subject from the auditory stimuli. Therefore, further studies are needed not only to confirm that the use of reading material has no significant effect on the amplitude of the MMN response to linguistically meaningful and nonmeaningful speech stimuli but also to further elucidate the advantages or disadvantages of using reading material versus silent videos with subtitles as a distractor task for MMN experiments.

The implications of the results are that silent, subtitled videos can serve as an appropriate distraction from auditory speech input, without affecting the amplitude of the MMN response. The results from the first experiment included in the current study indicate that this finding holds true for both simple tone stimuli, as well as non-linguistically meaningful and linguistically meaningful speech contrasts. Furthermore, the results of the second experiment of the current study indicate that subtitles do not affect the amplitude of MMN responses to speech contrasts that contain acoustic and/or lexical changes, further indicating the validity of using subtitled videos to make EEG recording sessions more tolerable for subjects. The artifact rejection results, and the subjects' reports that they paid less attention to the auditory input when subtitles were provided with the silent film, indicate that the subjects were more "engaged" by the silent video when subtitles were present, providing further evidence for the methodological usefulness of using subtitled videos in EEG recordings.

However, it must be noted that the appropriateness of this distraction task may not be applicable to the testing of clinical populations. That is, subjects with neurological impairments such as visual impairment, or language disorders from a stroke, traumatic brain injury or Alzheimer's disease, may experience more difficulty when attempting to process video subtitles than healthy controls. For example, it is well established that many aphasic patients have difficulties with reading (Caplan, 1992), as well as attention (Murray, 1999); therefore, it is highly likely that such subjects would become frustrated with the video or bored more quickly, which may affect the quality of the MMN recording. With any neurologically impaired subject, the way in which attentional resources may be allocated to various levels of linguistic processing during an MMN experiment will most likely differ to control subjects; therefore, further studies are needed to investigate the effects of a variety of distractor tasks on MMN responses to speech and nonspeech stimuli in neurologically impaired individuals, in order to determine the most effective distraction task for use in MMN experiments on these individuals.

CONCLUSION

The results of the present study indicate that the inclusion of subtitles on a silent video has no effect on the amplitude of the MMN or P3a responses to simple tones, CV nonwords, or CV words, incorporating acoustic and/or lexical changes. In addition, it was found that the presence of subtitles may facilitate a quicker MMN response and a slower switch to attention in response to CV speech stimuli with large acoustic deviances (i.e., "ge" ➡ "gay"); however, this result was not consistent across similar contrasts. Importantly, the results also indicate that movement artifacts may be considerably reduced by the presence of subtitles on a distracting silent video, which suggests that more "engaging" (i.e., subtitled) silent videos can be used as a distraction task for investigations into MMN responses to speech and nonspeech stimuli in young adult subjects, without affecting the amplitude of the responses. However, further studies are needed to investigate the effects of a wider variety of distractor tasks on speech and nonspeech MMN responses, as well as the most appropriate and effective distractor tasks to be applied in investigations of neurologically impaired individuals, in order to facilitate the most effective MMN data collection methods.

REFERENCES

Aaltonen O, Tuomainen J, Laine M, Niemi P. (1993) Cortical differences in tonal versus vowel processing as revealed by an ERP component called mismatch negativity (MMN). *Brain Lang* 44:139–152.

Alho K, Woods DL, Algazi A. (1994) Processing of auditory stimuli during auditory and visual attention as revealed by event-related potentials. *Psychophysiology* 31:469–479.

Alho K, Woods DL, Algazi A, Näätänen R. (1992) Intermodal selective attention. II. Effects of attentional load on processing of auditory and visual stimuli in central space. *Electroencephalogr Clin Neurophysiol* 82:356–368.

Alku P, Tiitinen H, Näätänen R. (1999) A method for generating natural-sounding speech stimuli for cognitive brain research. *Clin Neurophysiol* 110:1329–1333.

Aulanko R, Hari R, Lounasmaa OV, Näätänen R, Sams M. (1993) Phonetic invariance in the human auditory cortex. *Neuroreport* 4:1356–1358.

Caplan D. (1992) *Language: Structure, Processing, and Disorders*. London: MIT Press.

Ceponiene R, Service E, Kurjenluoma S, Cheour M, Näätänen R. (1999) Children's performance on pseudoword repetition depends on auditory trace quality: evidence from event-related potentials. *Dev Psychol* 35:709–720.

Deacon D, Nousak JM, Pilotti M, Ritter W, Yang C-M. (1998) Automatic change detection: does the auditory system use representations of individual stimulus features or gestalts? *Psychophysiology* 35:413–419.

Dittman-Balcar A, Thienel R, Schall U. (1999) Attention-dependent allocation of auditory processing resources as measured by mismatch negativity. *Neuroreport* 10:3749–3753.

Jaramillo M, Ilvonen T, Kujala T, Alku P, Tervaniemi M, Alho K. (2001) Are different kinds of acoustic features processed differently for speech and non-speech sounds? *Cogn Brain Res* 12:459–466.

Jasper HH. (1958) The ten-twenty electrode system of the International Federation. *Electroencephalogr Clin Neurophysiol* 10:371–375.

Kathmann N, Frodl-Bauch T, Hegerl U. (1999) Stability of the mismatch negativity under different stimulus and attention conditions. *Clin Neurophysiol* 110:317–323.

Kent RD, Read C. (1992) *The Acoustic Analysis of Speech*. San Diego: Singular Publishing Group, Inc.

Kraus N, Cheour M. (2000) Speech sound representation in the brain. *Audiol Neurootol* 5:140–150.

Kraus N, McGee T, Carrell TD, Sharma A. (1995) Neurophysiologic bases of speech discrimination. *Ear Hear* 16:19–37.

Kraus N, McGee TJ, Carrell TD, Zecker SG, Nicol TG, Koch DB. (1996) Auditory neurophysiologic responses and discrimination deficits in children with learning problems. *Science* 273:971–973.

Kraus N, McGee T, Micco A, Sharma A, Carrell T, Nicol T. (1993) Mismatch negativity in school-age children to speech stimuli that are just perceptibly different. *Electroencephalogr Clin Neurophysiol* 88:123–130.

Kraus N, McGee T, Sharma A, Carrell T, Nicol T. (1992) Mismatch negativity event-related potential elicited by speech stimuli. *Ear Hear* 13:158–164.

Lang AH, Eerola O, Korpilahti P, Holopainen I, Salo S, Aaltonen O. (1995) Practical issues in the clinical application of mismatch negativity. *Ear Hear* 16:118–130.

Lavikainen J, Tiitinen H, May P, Näätänen R. (1997) Binaural interaction in the human brain can be noninvasively accessed with long-latency event-related potentials. *Neurosci Lett* 222:37–40.

McArthur GM, Bishop DV, Proudfoot M. (2003) Do video sounds interfere with auditory event-related potentials? *Behav Res Methods Instrum Comput* 35 :329–333.

Michael EB, Keller TA, Carpenter PA, Just MA. (2001) fMRI investigation of sentence comprehension by eye and ear: modality fingerprints on cognitive processes. *Hum Brain Mapp* 13:239–252.

Müller BW, Achenbach C, Oades RD, Bender S, Schall U. (2002) Modulation of mismatch negativity by stimulus deviance and modality of attention. *Neuroreport* 13:1317–1320.

Muller-Gass A, Campbell K. (2001) (... and were instructed to read a self-selected book while ignoring the auditory stimuli): the effects of reading demands on the mismatch negativity. *Int J Psychophysiol* 41:234.

Muller-Gass A, Campbell K. (2002) Event-related potential measures of the inhibition of information processing: I. Selective attention in the waking state. *Int J Psychophysiol* 46:177–195.

Murray LL. (1999) Attention and aphasia: theory, research and clinical implications. *Aphasiology* 13:91–111.

Näätänen R. (1995) The mismatch negativity: a powerful tool for cognitive neuroscience. *Ear Hear* 16:6–18.

Näätänen R. (2000) Mismatch negativity (MMN): perspectives for application. *Int J Psychophysiol* 37:3–10.

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* 38:1–21.

Näätänen R, Escera C. (2000) Mismatch negativity: clinical and other applications. *Audiol Neurootol* 5:105–110.

Otten LJ, Alain C, Picton TW. (2000) Effects of visual attentional load on auditory processing. *Neuroreport* 11:875–880.

Paavilainen P, Alho K, Reinikainen K, Sams M, Näätänen R. (1991) Right hemisphere dominance of different mismatch negativities. *Electroencephalogr Clin Neurophysiol* 78:466–479.

Paavilainen P, Jaramillo M, Näätänen R, Winkler I. (1999) Neuronal populations in the human brain extracting invariant relationships from acoustic variance. *Neurosci Lett* 264:179–182.

Pekkonen E, Rinne T, Näätänen R. (1995) Variability and replicability of the mismatch negativity. *Electroencephalogr Clin Neurophysiol* 96:546–554.

Pettigrew CM, Murdoch BE, Kei J, Chenery H, Sockalingam R, Ponton CW, Finnigan S, Alku P. (2004) Processing of English words with fine acoustic contrasts and simple tones: a mismatch negativity study. *J Am Acad Audiol* 15:47–66.

Picton T, Alain C, Otten L, Ritter W, Achim A. (2000) Mismatch negativity: different water in the same river. *Audiol Neurootol* 5:111–139.

Picton T, Bentin S, Berg P, Donchin E, Hillyard SA, Johnson Jr. R, Miller GA, Ritter W, Ruchkin DS, Rugg MD, Taylor MJ. (2000) Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. *Psychophysiology* 37:127–152.

Pulvermüller F, Kujala T, Shtyrov Y, Simola J, Tiitinen H, Alku P, Alho K, Martinkauppi S, Ilmoniemi RJ, Näätänen R. (2001) Memory traces for words as revealed by the mismatch negativity. *Neuroimage* 14:607–616.

Sallinen M, Lyytinen H. (1997) Mismatch negativity during objective and subjective sleepiness. *Psychophysiology* 34:694–702.

Sams M, Aulanko R, Aaltonen O, Näätänen R. (1990) Event-related potentials to infrequent changes in synthesized phonetic stimuli. *J Cogn Neurosci* 2:344–357.

Schröger E. (1998) Measurement and interpretation of the mismatch negativity. *Behav Res Methods Instrum Comput* 30:131–145.

Shtyrov Y, Pulvermüller F. (2002) Neurophysiological evidence of memory traces for words in the human brain. *Neuroreport* 13:521–525.

Sinkkonen J, Tervaniemi M. (2000) Towards optimal recording and analysis of the mismatch negativity. *Audiol Neurootol* 5:235–246.

Sussman E, Winkler I, Wang J. (2002) MMN and attention: a new look at an old topic. *Int J Psychophysiol* 45:60–61.

Titova N, Näätänen R. (2001) Preattentive voice discrimination by the human brain as indexed by the mismatch negativity. *Neurosci Lett* 308:63–65.

Winkler I, Tervaniemi M, Schröger E, Wolff C, Näätänen R. (1998) Preattentive processing of auditory spatial information in humans. *Neurosci Lett* 242:49–52.

Wunderlich JL, Cone-Wesson BK. (2001) Effects of stimulus frequency and complexity on the mismatch negativity and other components of the cortical auditory-evoked potential. *J Acoust Soc Am* 109:1526–1537.