

Do video sounds interfere with auditory event-related potentials?

G. M. McARTHUR, D. V. M. BISHOP, and M. PROUDFOOT
University of Oxford, Oxford, England

To make the electroencephalogram (EEG) recording procedure more tolerable, listeners have been allowed in some experiments to watch an audible video while their auditory P1, N1, P2, and mismatch negativity (MMN) event-related potentials (ERPs) to experimental sounds have been measured. However, video sounds may degrade auditory ERPs to experimental sounds. This concern was tested with 19 adults who were instructed to ignore standard and deviant tones presented through headphones while they watched a video with the soundtrack audible in one condition and silent in the other. Video sound impaired the size, latency, and split-half reliability of the MMN, and it decreased the size of the P2. However, it had little effect on the P1 or N1 or on the split-half reliability of the P1–N1–P2 waveform, which was significantly more reliable than the MMN waveform regardless of whether the video sound was on or off. The impressive reliability of the P1 and N1 components allows for the use of video sound during EEG recording, and they may prove useful for assessing auditory processing in listeners who cannot tolerate long testing sessions.

Sounds are processed in the brain over a series of stages from the brainstem to the cortex. The operation of these stages can be measured noninvasively using late auditory event-related potentials (ERPs), which represent the average pattern of electrical activity, measured at the scalp with electrodes, produced by groups of neurons in response to a sound. Because auditory ERPs can be measured without a listener's attention, they are becoming a popular alternative to psychoacoustic tasks for measuring auditory discrimination, particularly for populations with poor attention, such as children with specific language impairment and adults with schizophrenia.

To make the electroencephalogram (EEG) recording procedure more tolerable for these listeners, and to divert their conscious attention away from the auditory stimuli to minimize confounding ERPs (e.g., N2b or P3a; Lang et al., 1995; Sinkkonen & Tervaniemi, 2000), many researchers ask their listeners to concentrate on a video while the experimental auditory stimuli are being presented. Most experiments use a silent video (e.g., Korpi-lahti, Krause, Holopainen, & Lang, 2001; Novitski et al., 2001; Pang & Taylor, 2000; Tervaniemi et al., 1999; Uwer & Suchodoletz, 2000). However, some experiments have allowed the video sound to be left on at a low level (40–50 dB SPL; e.g., Bellis, Nicol, & Kraus, 2000; Kraus et al., 1993; McArthur & Bishop, 2002; McGee, Kraus, & Nicol, 1997; Sharma, Kraus, McGee, Carrell, & Nicol, 1993; Todd, Michie, & Jablensky, 2001).

To our knowledge, no one has established that playing a low-level video soundtrack does not degrade ERP responses to the experimental auditory stimuli. Kraus, McGee, Carrell, and Sharma (1995) claim that leaving the video soundtrack on at 40 dB SPL has no effect on the auditory mismatch negativity component (MMN; the difference between a listener's ERP to a common standard stimulus and their ERP to a rarer deviant stimulus at around 200 msec; see waveforms in Figure 1 that were measured in this experiment). However, they provide no data to support this claim.

The results of experiments that have tested the effect of other types of background noise on ERPs are indeterminate. Novitski et al. (2001) found that the presence of 54-dB SPL bursts of fMRI scanner noise presented every 2,500 msec did not affect the MMN, but did increase the latency and decrease the amplitude of the P1, N1, and P2 auditory ERP components (the first positive, first negative, and second positive deflections in the auditory ERP, respectively; see Figure 1A). This is similar to the masking effect that broadband white noise has on ERPs (Novitski et al., 2001). Interestingly, in an unpublished experiment by researchers in the same laboratory found that continuous white background noise did affect the MMN (Novitski et al., 2001).

The aim of this experiment was to assess whether playing the soundtrack of a video degrades the morphology or reliability of auditory P1, N1, P2, and MMN ERP components to tones. We elicited these components by using tones that varied in frequency rather than duration or intensity because of our interest in the association between auditory ERPs and psychoacoustic tests of frequency discrimination (McArthur & Bishop, 2002), and the link between frequency discrimination and language impairments (McArthur & Bishop, 2001).

Correspondence should be addressed to G. M. McArthur, Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford, OX1 3UD, England (e-mail: genevieve.mcarthur@psy.ox.ac.uk).

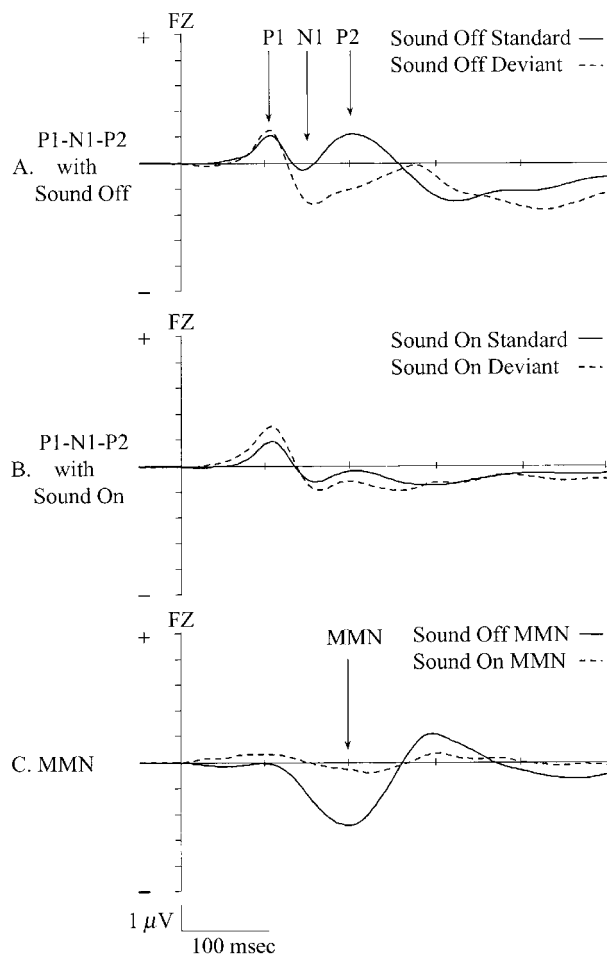


Figure 1. P1, N1, and P2 components to standard and deviant stimuli in the sound-off condition (A) and the sound-on condition (B). For each condition, the standard waveform is subtracted from the deviant waveform to produce the MMN (C).

METHOD

Methods were approved by the Department of Experimental Psychology's Ethics Committee at the University of Oxford. Informed consent was obtained from each listener.

Participants

The participants were 19 adults (13 females) from 18 to 50 years of age, from the university community. All had normal hearing thresholds (i.e., below 15 dB HL) for a 1-sec, 700-Hz pure tone.

Procedure

The participants were seated in an armchair in an electrically shielded testing booth, 2 m away from a small (21 × 28 cm) video television. Experimental auditory stimuli were presented diotically through headphones in two conditions: one with the video soundtrack set at approximately 50 dB SPL (sound-on condition; sound level was measured at the listener's head) and one with the video soundtrack turned off (sound-off condition). In both conditions, the listeners were instructed to concentrate on the video and ignore the stimuli presented through the headphones. Each listener selected his/her own video, which was a mainstream production such as

Mission Impossible 2 (Cruise & Woo, 2000) or Chicken Run (Park & Lord, 2000). Thus, the video soundtrack consisted of a continuous stream of irregular noise that comprised speech, music, and environmental sounds.

Each condition comprised 10 blocks of 200 stimuli (i.e., 2,000 trials total). In each block, either the standard stimulus was a 25-msec, 80-dB SPL 600-Hz tone and the deviant stimulus was a 25-msec, 80-dB SPL 700-Hz tone, or the standard and the deviant were reversed (i.e., a 700-Hz standard and a 600-Hz deviant). The standard and deviant stimuli were presented on 85% and 15% of the trials, respectively (i.e., 1,700 standard stimuli and 300 deviant stimuli total). The deviant was presented every 5th to 10th stimulus so that no 2 deviants were presented consecutively. The gap between trials was randomly jittered from 320 to 420 msec to avoid anticipatory ERP artifacts. The blocks were presented in random order.

Nonpolarized sintered electrodes, positioned according to the 10–20 International system, were used to record the EEG from 10 frontal sites, 6 temporal sites, 6 central sites, 5 parietal sites, and 1 occipital site. Vertical eye movements (VEOG) were measured with electrodes placed above and below the left eye; horizontal eye movements (HEOG) were measured with electrodes on the outer canthi of each eye. The ground electrode was positioned between FPz and Fz. Linked mastoids were used as the on-line reference. The signal was amplified 20,000 times and sampled at 250 Hz.

VEOG activity was removed from the EEG sites using a standard ocular reduction algorithm (Neurosoft, 1999). EEG activity was band-pass filtered (10-Hz low-pass and 0.1-Hz high-pass; 12 dB per octave roll-off) and divided into 550-msec epochs with a 50-msec prestimulus interval. Epochs were baseline corrected from –50 to 0 msec. Epochs with changes in HEOG or EEG activity greater than 150 µV were rejected.

For each listener, P1, N1, and P2 responses were measured from the ERP to the standard stimuli, which was calculated by averaging epochs of all standard stimuli (600 and 700 Hz) except those that fell immediately after a deviant stimulus. The ERP to deviant stimuli was calculated by averaging epochs to all (i.e., 700 and 600 Hz) deviant stimuli. Each listener's standard ERP was subtracted from his/her deviant ERP to compute the MMN response.

Standard peak detection procedures that return the maximum or minimum value in an interval were not used, because some listeners did not have an N1 or P2 peak. Instead, true P1, N1, P2, and MMN peaks were identified as the largest point within an interval (50–115, 115–200, 170–250, and 150–350 msec, respectively) flanked by two increasingly negative (for P1 and P2) or two increasingly positive (for N1 and MMN) values on each side. Absence of a true peak returned a missing value.

We assessed the split-half reliability of the P1–N1–P2 and MMN waveforms by using intraclass correlation coefficients, which measure the degree of overlap between the shape and absolute voltage of two waveforms. Coefficients range from 0 (dissimilar waveforms) to 1.0 (identical waveforms). A coefficient of .5 indicates that one waveform accounts for 50% of the variance in the other waveform (Neurosoft, 1999). We calculated intraclass correlation coefficients between each listener's P1–N1–P2 waveform (–50 to 500 msec) to 600-Hz standards and his/her P1–N1–P2 waveform to 700-Hz standards. Similarly, we calculated intraclass correlation coefficients between each listener's MMN waveform (–50 to 500 msec) in blocks that used 600-Hz standards and 700-Hz deviants and their MMN waveform in blocks that used 700-Hz standards and 600-Hz deviants.

RESULTS

Activity at Fz was used to represent the P1, N1, P2, and MMN components, because this was the site with

the largest response and it is commonly used to represent auditory ERPs. Figure 1 illustrates the morphology of the P1, N1, and P2 components to standard and deviant stimuli in the sound-off condition (A) and the sound-on condition (B). For each condition, the standard waveform is subtracted from the deviant waveform to produce the MMN (C).

There was little difference between the number of epochs accepted for each listener's P1–N1–P2 and MMN waveforms in the two conditions (see Table 1). Table 1 also presents the associated means, standard deviations, and paired samples *t*-test statistics for the peak amplitude and latency of the P1, N1, P2, and MMN components, and for the split-half reliabilities of the P1–N1–P2 and MMN waveforms. A difference was considered statistically significant if $p < .0004$ (a Bonferroni correction of $p < .05$ was used to account for the higher probability of finding a significant result across 12 *t* tests).

There was little difference between P1 peak amplitude or peak latency in the sound-on and sound-off conditions. Similarly, there was no reliable difference between the mean peak amplitude of N1 in the sound-on and sound-off conditions (3 listeners were missing an N1 peak in the sound-on condition, and 1 in the sound-off condition). The latency of the N1 peak tended to be later in the sound-on than in the sound-off condition. However, the difference of 4.2 msec was trivial in terms of individual differences in auditory ERPs; it was not statistically significant.

In contrast, there was a nontrivial and reliable reduction in the peak amplitude of P2 in the sound-on as opposed to the sound-off condition. This effect was observed in 15 of the 19 listeners (the remaining 4 listeners did not have a P2 peak in the sound-on condition). However, there was no difference between the latency of the P2 peak in the sound-on and sound-off conditions.

Similarly, the mean peak amplitude of MMN was significantly smaller in the sound-on condition than in the sound-off condition. This effect was observed in 15 of the 19 listeners. The mean latency of the MMN tended to be delayed in the sound-on condition in comparison with the sound-off condition. However, the difference was not statistically significant.

Figure 2 illustrates the reliability of the group's P1–N1–P2 (A) and MMN (B) waveforms, comparing the waveforms measured in the stimulus blocks that had 600-Hz standard tones (with 700-Hz deviant tones for the MMN wave) with the waveforms measured in the blocks that had 700-Hz standard tones (with 600-Hz deviant tones for the MMN wave).

The difference between the intraclass correlation coefficients of the MMN waveforms in the sound-on as opposed to the sound-off condition was large and statistically significant (see Table 1). There was only a small, nonsignificant difference between the intraclass correlation coefficients of the P1–N1–P2 waveforms in the sound-on and sound-off conditions. Further, the mean intraclass correlation coefficient of the P1–N1–P2 waveform was (1) high in both conditions, (2) higher in the sound-on and sound-off conditions than that of the MMN response in both the sound-on and sound-off conditions [$t(18) = 10.50, p < .001$, and $t(18) = 5.41, p < .001$, respectively], and (3) significantly higher in the less optimal sound-on condition than the coefficient of the MMN waveform in the optimal sound-off condition [$t(18) = 3.10, p = .006$].

DISCUSSION

The aim of this simple, yet apparently unique, experiment was to assess whether the presence of low-level video sound would interfere with the morphology or re-

Table 1
Statistics for the Auditory P1, N1, P2, and MMN Components in the Sound-On and Sound-Off Conditions, and for the Split-Half Correlation Coefficients and Number of Epochs That Composed Waveforms in the Sound-On and Sound-Off Conditions

	Sound On		Sound Off		Comparison
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	Paired Samples <i>t</i> Test
Amplitude					
P1	1.37	0.81	1.68	1.10	$t(18) = 1.55, p = .14$
N1	-1.07	0.89	-0.91	1.44	$t(15) = 0.25, p = .81$
P2	0.24	1.21	1.90	1.58	$t(14) = 6.29, p < .001^*$
MMN	-0.70	0.84	-2.68	1.37	$t(18) = 5.31, p < .001^*$
Latency					
P1	102.21	12.68	102.63	22.52	$t(18) = 0.12, p = .92$
N1	155.75	14.00	151.55	23.40	$t(15) = 2.99, p = .009$
P2	210.00	21.48	213.33	22.97	$t(14) = 0.00, p = 1.00$
MMN	230.84	53.31	195.05	16.70	$t(18) = 2.69, p = .02$
Split-Half Coefficients					
P1–N1–P2	.81	0.15	.92	0.08	$t(18) = 2.82, p = .01$
MMN	.36	0.17	.70	0.17	$t(18) = 6.39, p < .001^*$
Number of Epochs					
P1–N1–P2	1,378.05	23.63	1,362.63	92.47	$t(18) = 0.74, p = .47$
MMN	295.00	5.11	292.11	16.79	$t(18) = 0.78, p = .45$

*Statistically significant difference ($p < .004$; Bonferroni correction).

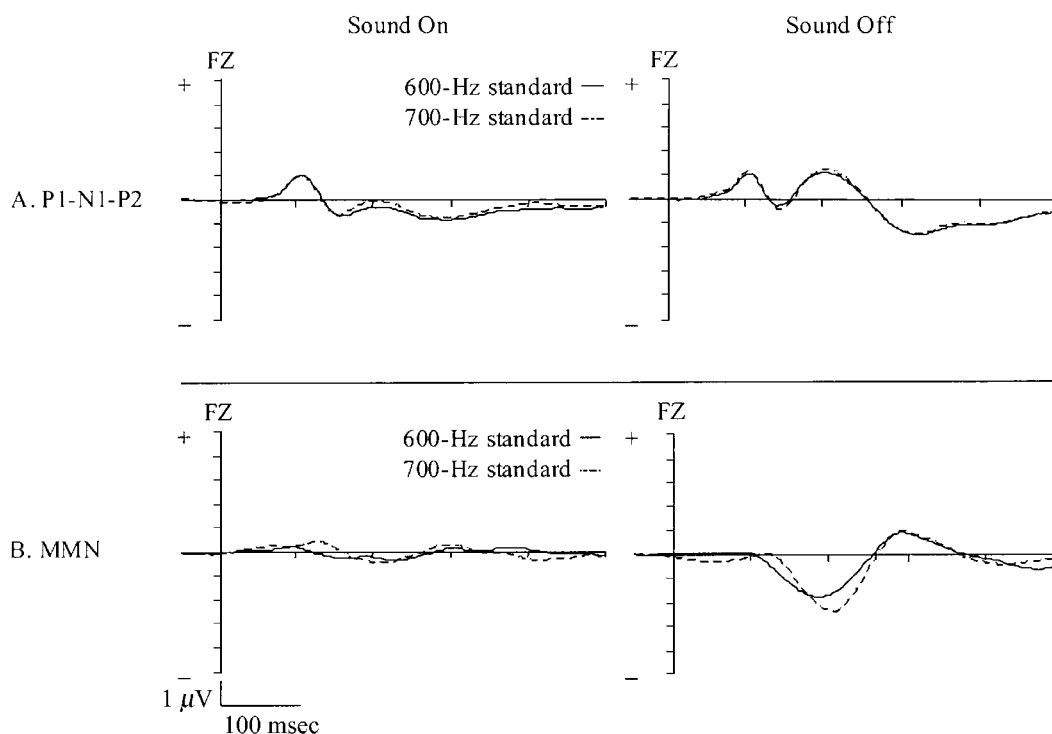


Figure 2. The (A) P1–N1–P2 (ERP to standard stimuli) and (B) MMN (difference between ERP to standard stimuli subtracted from ERP to deviant stimuli) waveforms measured in the two stimulus blocks with 600-Hz standard tones (with 700-Hz deviant tones for the MMN wave) compared with the waveforms measured in the two blocks with 700-Hz standard tones (with 600-Hz deviant tones for the MMN wave).

liability of the auditory ERPs. Video sound impaired the size, latency, and split-half reliability of the MMN. It also decreased the size of the P2. As mentioned above, broadband white noise has a similar effect on auditory ERPs (Novitski et al., 2001), suggesting that the video soundtrack masked the MMN and P2 responses. It is also noteworthy that the attenuation of the MMN and P2 responses by video noise are related. The smaller P2 responses to standards in the sound-on condition reduced the difference between P2 responses to standards and deviants, resulting in a diminished MMN. With these factors in play, researchers should be cautious about using an audible video soundtrack when they are interested in making accurate and reliable measurements of the MMN.

In contrast, the presence of low-level video sound had no practical effect on the size or the latency of the P1 or N1 responses. It did reduce the reliability of the P1–N1–P2 waveform. However, the reliability of this waveform was so robust in both conditions that the practical significance of this effect seems negligible. This contrasts with the reliability of the MMN, which was low in comparison with the P1 and N1 responses regardless of whether video sound was present or not.

These results contrast with Novitski et al.'s (2001) finding that scanner noise affects the P1, N1, and P2

ERP components. This inconsistency could be explained by the different stimuli-to-background noise ratios in the two experiments. Novitski et al. used a low ratio, setting their experimental stimuli and scanner noise at a similar intensity (around 57 dB SPL). In the present experiment, we used a higher ratio, with stimuli that were more intense than the video noise (80 and 50 dB SPL, respectively). The relatively soft video noise in this experiment may have exerted less masking than the relatively loud scanner noise in Novitski et al.'s experiment, causing less interference with the P1, N1, and P2 responses.

Different stimuli-to-background noise ratios cannot explain why video noise degraded the MMN in the present experiment, whereas scanner noise had no effect on the MMN in Novitski et al. (2001). We would predict that lower stimuli-to-background noise ratios would have a greater effect on the MMN. However, Novitski et al.'s low ratio had no effect on the MMN. Interestingly, these contradictory results could be explained by the unpublished findings of Novitski et al.'s colleagues that continuous white background noise does affect the MMN (Novitski et al., 2001). A video soundtrack, with its cacophony of continuous speech, music, and environmental sounds, may be more similar to continuous white background noise than Novitski et al.'s scanner noise, which was intermittent and contained only low-

frequency peaks from 50 to 1000 Hz. It may be that only continuous sounds composed of a wide range of frequencies affect the MMN but not the P1 and N1 responses.

These results contribute to a small but growing body of research on the reliability of auditory ERPs. Reliability is typically measured with correlation coefficients that represent the degree of similarity between ERP responses measured in two separate testing sessions (test-retest reliability) or in two halves of the same testing session (split-half reliability). Correlation coefficients range from 1.0 (the ERP measures are exactly the same) to 0 (the ERP measures are completely dissimilar) to -1.0 (the ERP measures are the inverse of each other). In past experiments, the P1, N1, and P2 have been very reliable, with test-retest or split-half correlation coefficients typically ranging from .7 to .9 (Escera & Grau, 1996; Escera, Yago, Polo, & Grau, 2000; Pekkonen, Rinne, & Näätänen, 1995; Uwer & Suchodoletz, 2000). However, the reliability of the MMN to a frequency deviant is considerably lower (.2 to .6; Escera & Grau, 1996; Kathmann, Frodl-Bauch, & Hegerl, 1999; Tervaniemi et al., 1999). In fact, the reliability of the MMN appears to approach that of the P1 and N1 only under specific conditions, such as when a duration deviant rather than an intensity or frequency deviant is used (Kathmann et al., 1999; Tervaniemi et al., 1999), when the duration deviant is 66% shorter than the standard rather than 33% (Tervaniemi et al., 1999), or when the MMN is measured using certain interstimulus intervals, amplitude measures (Escera et al., 2000), and electrodes (e.g., F4; Pekkonen et al., 1995). It has been suggested that the MMN is less reliable than the N1 component because it is based on fewer deviant stimuli (Escera & Grau, 1996; Pekkonen et al., 1995). However, Escera et al. (2000) recently found that the N1 is more reliable than the MMN even when it is based on the same number of stimuli as MMN deviants.

In conclusion, in this experiment we sought to determine whether video noise would interfere with the morphology or latency of auditory ERPs. The results suggest that a video soundtrack may be used with impunity when one is assessing auditory P1 and N1 ERP responses, but that the soundtrack may interfere with the accurate and reliable measurement of the MMN. Further, the impressive reliability of the P1 and N1 components suggests that they offer a paradigm for assessing low-level auditory processing in listeners who can tolerate only short testing sessions.

REFERENCES

- BELLIS, T. J., NICOL, T., & KRAUS, N. (2000). Ageing affects hemispheric asymmetry in the neural representation of speech sounds. *Journal of Neuroscience*, *20*, 791-797.
- CRUISE, T. (PRODUCER), & WOO, L. (DIRECTOR) (2000). *Mission impossible 2* [Video]. United States: Paramount Pictures.
- ESCERA, C., & GRAU, C. (1996). Short-term replicability of the mismatch negativity. *Electroencephalography & Clinical Neurophysiology*, *100*, 549-554.
- ESCERA, C., YAGO, E., POLO, M. D., & GRAU, C. (2000). The individual replicability of the mismatch negativity at short and long inter-stimulus intervals. *Clinical Neurophysiology*, *111*, 546-551.
- KATHMANN, N., FRODL-BAUCH, T., & HEGERL, U. (1999). Stability of the mismatch negativity under different stimulus and attention conditions. *Clinical Neurophysiology*, *110*, 317-323.
- KORPILAHTI, P., KRAUSE, C. M., HOLOPAINEN, I., & LANG, A. H. (2001). Early and late mismatch negativity elicited by words and speech-like stimuli in children. *Brain & Language*, *76*, 332-339.
- KRAUS, N., MCGEE, T., CARRELL, T. D., & SHARMA, A. (1995). Neurophysiologic bases of speech discrimination. *Ear & Hearing*, *16*, 19-37.
- 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. *Electroencephalography & Clinical Neurophysiology*, *88*, 123-130.
- LANG, A. H., EOROLA, O., KORPILAHTI, P., HOLOPAINEN, I., SALO, S., & AALTONEN, O. (1995). Practical issues in the clinical application of mismatch negativity. *Ear & Hearing*, *16*, 118-130.
- MCCARTHER, G. M., & BISHOP, D. V. M. (2001). Auditory perceptual processing in people with reading and oral language impairments: Current issues and recommendations. *Dyslexia*, *7*, 150-170.
- MCCARTHER, G. M., & BISHOP, D. V. M. (2002). Event-related potentials reflect individual differences in age-invariant auditory skills. *NeuroReport*, *13*, 1079-1082.
- MCGEE, T., KRAUS, N., & NICOL, T. (1997). Is it really a mismatch negativity? An assessment of methods for determining response validity in individual subjects. *Electroencephalography & Clinical Neurophysiology*, *104*, 359-368.
- NEUROSOFT, INC. (1999). *SCAN: User guide*. Sterling, VA: Neurosoft, Inc.
- NOVITSKI, N., AHLÖ, K., KORSYUKOV, O., CARLSON, S., MARTINKAUPPI, S., ESCERA, C., RINNE, T., ARONEN, H. J., & NÄÄTÄNEN, R. (2001). Effects of acoustic gradient noise from functional magnetic resonance imaging on auditory processing as reflected by event-related potentials. *NeuroImage*, *14*, 244-251.
- PANG, E. W., & TAYLOR, M. J. (2000). Tracking the development of the N1 from age 3 to adulthood: An examination of speech and non-speech stimuli. *Clinical Neurophysiology*, *111*, 388-397.
- PARK, N. (PRODUCER), & LORD, P. (DIRECTOR) (2000). *Chicken run* [Video]. Bristol, U.K.: Aardman Animations, Ltd.
- PEKKONEN, E., RINNE, T., & NÄÄTÄNEN, R. (1995). Variability and replicability of the mismatch negativity. *Electroencephalography & Clinical Neurophysiology*, *96*, 546-554.
- SHARMA, A., KRAUS, N., MCGEE, T., CARRELL, T., & NICOL, T. (1993). Acoustic versus phonetic representation of speech as reflected by the mismatch negativity event-related potential. *Electroencephalography & Clinical Neurophysiology*, *88*, 64-71.
- SINKKONEN, J., & TERVANIEMI, M. (2000). Towards optimal recording and analysis of the mismatch negativity. *Audiology & Neuro-Otology*, *5*, 235-246.
- TERVANIEMI, M., LEHTOKOSKI, A., SINKKONEN, J., VIRTANEN, J., ILMONIEMI, R. J., & NÄÄTÄNEN, R. (1999). Test-retest reliability of mismatch negativity for duration, frequency and intensity changes. *Clinical Neurophysiology*, *110*, 1388-1393.
- TODD, J., MICHIE, P. T., & JABLENSKY, A. V. (2001). Do loudness cues contribute to duration mismatch negativity reduction in schizophrenia? *NeuroReport*, *12*, 4069-4073.
- UWER, R., & SUCHODOLETZ, W. VON (2000). Stability of mismatch negativities in children. *Clinical Neurophysiology*, *111*, 45-52.

(Manuscript received May 28, 2002;

revision accepted for publication September 15, 2002.)