

Cross-modal symbolic processing can elicit either an N2 or a protracted N2/N400 response

OREN GRIFFITHS, MIKE E. LE PELLEY, BRADLEY N. JACK, DAVID LUQUE, AND THOMAS J. WHITFORD
School of Psychology, UNSW Australia, Sydney, Australia

Abstract

A cross-modal symbolic paradigm was used to elicit EEG activity related to semantic incongruence. Twenty-five undergraduate students viewed pairings of visual lexical cues (e.g., DOG) with congruent (50% of trials) or incongruent (50%) auditory nonlexical stimuli (animal vocalizations; e.g., sound of a dog woofing or a cat meowing). In one condition, many different pairs of congruent/incongruent stimuli were shown, whereas in a second condition only two pairs of stimuli were repeatedly shown. A typical N400-like pattern of incongruence-related activity (including activity in the N2 time window) was evident in the condition using many stimuli, whereas the incongruence-related activity in the two-stimuli condition was confined to differential N2-like activity. A supplementary analysis excluded stimulus characteristics as the source of this differential activity between conditions. We found that a single individual performing a fixed task can demonstrate either a protracted N400-like pattern of activity or a more temporally focused N2-like pattern of activity in response to the same stimulus, which suggests that the N2 may be a precursor to the protracted N400 response.

Descriptors: EEG, ERPs, N400, N2, Semantic memory

The N400 is a well-known and well-examined component of ERPs used to study language processing. This negative-going component is typically elicited by presenting an incongruent word at the end of a coherent sentence (Kutas & Hillyard, 1980), e.g., “I take coffee with cream and *dog*.” The electrophysiological response to this incongruent word is then compared to that for a final word that is well predicted by the preceding sentence context: in this case, *sugar* (among other control conditions; see Kutas & Federmeier, 2011, for review). When the electrocortical activity evoked by *sugar* is compared to that for *dog*, the typical finding is that there is a differential negative-going component between 200–500 ms after the incongruent stimulus (Kutas & Federmeier, 2000), primarily at centroparietal and right hemisphere electrode sites.

It has been shown that the N400 can be measured even in the absence of meaningful sentences. Sanquist, Rohrbaugh, Syndulko, and Lindsley (1980) showed people pairs of either semantically related or unrelated words (e.g., BED–REST vs. BED–FILE). They found that ERPs evoked by the second word in unrelated pairs included an N400-like component relative to the ERPs evoked by the second word in a related pair (see also Boddy &

Weinberg, 1981; Harbin, Marsh & Harvey, 1984). This finding shows that the N400 can be observed without using whole sentences as primes, and that individual words can produce a semantic context of sufficient strength to evoke differential N400 activity.

Interestingly, only a small range of processes related to stimulus meaning appear to elicit an N400. Violations of a perceptual pattern do not elicit N400s, but instead produce different components (typically, a parietal positivity). Examples include the abrupt termination of a lowercase sentence with an uppercase word or a picture (Kutas & Hillyard, 1980) or playing a wrong note in a well-known melody (Besson & Macar, 1987). Van Petten and Riefelder (1995) pushed this idea further by considering semantic priming of nonword sounds by word cues (see also Orgs, Lange, Dombrowski, & Heil, 2006). The target sounds were auditory samples of easily nameable sounds, such as the sound of footsteps or of glass breaking. In the critical experiment, a spoken word describing each sound (e.g., FOOTSTEPS or SHATTER) was presented shortly before the auditory sample was played. On other trials (the majority), the auditory samples were preceded by unrelated spoken words or nonwords. A comparison of ERPs to target sounds on semantically congruent trials (e.g., the word FOOTSTEPS followed by the sound of footsteps) versus incongruent trials (e.g., the word SHATTER followed by the sound of footsteps) revealed a sustained negative-going component around 400 ms after sound onset on incongruent trials, which had a right posterior scalp distribution. In other words, this study showed that nonword sounds can elicit an N400. Orgs et al. (2006) extended this task by demonstrating the same effects when the priming stimulus was a word presented on screen, rather than a word spoken aloud. Cumulatively,

This work was supported by the Australian Research Council, with a Discovery Project grant (DP140104394) awarded to TJW, MELP, and DL, a Discovery Early Career Research Award awarded to OG (DE150100667), and a Future Fellowship awarded to MELP (FT100100260). TJW is also supported by Career Development Fellowship from the National Health and Medical Research Council of Australia (APP1090507).

Address correspondence to: Dr. Oren Griffiths, School of Psychology, UNSW Australia, Sydney, NSW 2052, Australia. E-mail: oren.griffiths@unsw.edu.au

these results suggest that the N400 component can index a type of symbolic semantic priming that extends beyond lexical stimuli: a word (e.g., FOOTSTEPS) can prime attributes of its referent (the sound produced by footsteps) in addition to its semantic associates.

It is notable that Van Petten and Rheinfelder's (1995) procedure for eliciting the N400 closely resembles tasks that have previously been used to examine the N2 component. This is especially notable given that it has been argued that the N400 may be closely related, or even identical, to the N2 (Deacon, Breton, Ritter, & Vaughan, 1991; Herning, Speer, & Jones, 1987; Polich, 1985). N2 components are widely studied, negative-going peaks that typically occur 200–350 ms after stimulus onset. They typically have a frontocentral scalp distribution, but this varies with the eliciting task (e.g., oddball N2 tasks typically elicit a posterior component; Conci, Gramann, Müller, & Elliott, 2006; Hickey, McDonald, & Theeuwes, 2006). While a precise taxonomy of the cognitive functions underlying N2 is still a matter for debate, a common procedure to elicit N2 activity is the sequential matching task. In a sequential matching task, people are shown two visual stimuli, one after the other, and are asked to note whether the second stimulus is the same as the first (usually with respect to a prespecified dimension, such as color). An N2 of greater amplitude is seen when the target mismatches the preceding stimulus, compared to when they match (Wang, Cui, Wang, Tian, & Zhang, 2004; Wang et al., 2003). It appears that the first stimulus prepares the cognitive apparatus for the second stimulus with respect to a particular feature dimension (e.g., color), such that a deviation from this primed value results in activity within the N2 time window.

Perhaps this process is shared with N400 tasks, such as that performed by Van Petten and Rheinfelder (1995). In their original task, the first word may have primed specific feature values of a subset of possible sounds (the semantically related sounds), such that a deviation from the expected value prompts additional processing in the N2 time window. However, the members of the class of sounds that are semantically related to a given word are perceptually distinct from each other, so priming any one feature value would be ineffective. That is to say, the word FOOTSTEPS is related to any sound that is recognizable as footsteps, not just one particular sound. Indeed, in Van Petten and Rheinfelder's study, a large number (99) of semantically associated word–sound stimulus pairs were shown infrequently (once), so participants could not anticipate the particular, exact sound that would follow any given word. This means that perceptual processing is insufficient to detect that the target sound is incongruent. Participants needed to additionally access the semantic features associated with the target stimulus to determine incongruence, and it may be this additional activity that produces the sustained activity across the 200–500 ms time period indicative of an N400 response.

According to this interpretation, the cognitive processes that elicit the N400 and the N2 do not differ in kind, but rather in complexity and duration. The N2 response to physical incongruence may act as a precursor to the N400 response to semantic incongruence. This implies a straightforward hypothesis: If it were possible to “sharpen” one's expectancies in a semantic priming task, such that a specific stimulus was anticipated rather than a semantic class of stimuli, this would reduce the need for participants to process the target stimuli to the level of meaning, and perhaps the protracted N400 response to incongruence would reduce to a more focused N2 response.

The present experiment tested this hypothesis using a variant of Van Petten and Rheinfelder's (1995) N400 paradigm. Consider what might happen in this task if the particular sensory characteris-

tics of the target stimulus were readily predictable on each trial. Notice that the prototypical serial-matching N2 task (e.g., Wang et al., 2003) uses only a few stimuli and repeats them frequently. Perhaps if Van Petten and Rheinfelder's N400 paradigm were performed with only a small number of stimuli that were highly frequent, this might reduce the need to process those highly frequent stimuli to the level of meaning in order to determine whether they match or mismatch. Here, with sufficient repetition the prime word FOOTSTEPS might come to anticipate a particular, highly familiar sound stimulus (that happens to be of footsteps), such that the sound itself need not be processed semantically. Indeed, the “semantic satiation” effect (Black, 2004; Smith, 1984) demonstrates that repetition reduces people's capacity to access the meaning of the target stimulus. Specifically, it has been shown that frequent repetition of a target word reduces the speed with which that word can subsequently be semantically categorized and, critically, it also reduces the degree to which incongruent pairings involving that word elicit differential N400 activity (Balota & Black, 1997; Kounios, Kotz, & Holcomb, 2000).

Thus, according to the view that N2 and N400 components both index similar processes of differing complexity, a restricted version of Van Petten and Rheinfelder's N400 task in which only two stimulus pairs are repeated frequently might elicit an N2 in response to incongruent word–sound pairs, rather than a protracted N400. The present experiment investigated this hypothesis by manipulating the number of stimuli shown during a version of Van Petten and Rheinfelder's N400 task. In one condition, participants performed the task with 10 word–sound pairs (which we expected would elicit an N400 for incongruent word–sound pairs), and in a second condition only two word–sound pairs were used, but were each experienced more frequently (which might instead produce an N2). Because the evoking stimuli necessarily differ between conditions under this manipulation, an additional analysis controlling for these differences was also included. In this additional analysis, we limited our analysis to only those trials that included the two evoking stimuli that were common to both conditions.

Method

Participants and Apparatus

Twenty-five undergraduate students from UNSW Australia participated for course credit. All participants gave written informed consent prior to the experiment and reported having normal hearing in both ears. Data of two participants were excluded from further analyses, because of technical errors in the EEG recording. There were 14 women and 11 men, and the mean age was 19.28 (range: 17–30, $SD = 2.53$). Participants were tested individually in a sound-attenuated cubicle, using a standard PC with a 23-inch monitor (1,920 × 1,280 resolution, 120 Hz refresh rate) positioned ~60 cm from the participant. Stimulus presentation was controlled by MATLAB using Psychophysics Toolbox extensions (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007; Pelli, 1997). Auditory stimuli were presented via headphones (Sennheiser HD201), and responses were made using a Cedrus RB-530 response key. Apparatus for EEG recording is described below.

The experiment was approved by the Human Research Ethics Advisory Panel (Psychology) at UNSW Australia, and was conducted in accordance with the ethical standards laid down in the Declaration of Helsinki (World Medical Association, 2001).

Auditory Materials

Auditory stimuli were 10 sound samples that were obtained from the Internet (from www.freesound.org). These sounds were all of animal vocalizations: specifically, a bird singing, a cow mooing, a duck quacking, a frog croaking, a cat meowing, a dog woofing, an owl hooting, a lion roaring, a sheep baaing, and a snake hissing. The stimuli were adjusted (using Audacity audio creation software) to have similar peak volumes and ramp-up times, and were altered to be as brief as possible while retaining their ease of identification. Two stimuli (the cat meow and the dog woof) were particularly closely matched on volume (75 dB SPL), duration (386 ms), and rise time (45 ms to 75% peak volume). These stimuli were selected for use in the two-stimuli condition (described below).

EEG Recording

A BioSemi ActiveTwo EEG system was used to continuously record (at 2048 Hz) from 64 scalp electrode sites positioned according to the extended 10-20 system. The signal was filtered online to limit the signal to 0.16–100 Hz. Each electrode was individually amplified. All EEG channels were referenced online to internal sensors (CMS/DRL) located in the parietal region of the cap, and rereferenced offline to the average of the two mastoid electrodes. Vertical and horizontal electrooculograms (EOGs) were recorded from electrodes 2 cm above and below the left eye, and 1 cm beyond the outer canthi of both eyes.

Procedure

The general procedure is summarized in Figure 1. All participants completed two conditions: the ten-stimuli and two-stimuli conditions. The order of these conditions was counterbalanced between participants. There were 120 trials per condition, thus all participants completed 240 trials in total. The ten-stimuli condition is described first. Screen background was black throughout the experiment. Each trial began with a white fixation cross shown for 1 s in the center of the screen (50 pixels; $1.98^\circ \times 1.40^\circ$ visual angle). Immediately after the fixation cross disappeared, a cue word was presented in the centre of the screen, written in white (size 36 font, 1.21° vertical). This word was always the name of an animal (BIRD, COW, DUCK, FROG, CAT, DOG, OWL, LION, SHEEP, and SNAKE). The word disappeared after 1 s, and 3 s later a sound stimulus was presented. The sound was always an animal vocalization. On 50% of trials, the vocalization was congruent with the preceding cue word (e.g., the sound of a snake's hiss followed the word SNAKE), and on the remainder an alternative sound was randomly selected from the other nine sounds (e.g., the sound of a lion roaring followed the word SNAKE). After a further 2 s, a response window was shown, in which MATCH was written on the left and MISMATCH was written on the right. Once participants made their response, visual corrective feedback was provided ("correct" or "incorrect" was shown). The next trial began after 1.5 s of a blank screen. The congruent and incongruent trials occurred in a random order, with one exception. The first 20 trials of both conditions consisted of only congruent trials so as to reduce any possible ambiguity about which sounds were congruent with which words. The remaining trials in each condition were equally divided between congruent and incongruent trials. Each of the 10 word stimuli was presented equally often (12 times) in the ten-stimuli condition. Participants were offered a brief break after every 40 trials.



Figure 1. Experimental procedure. A fixation cross was first shown for 1 s, before a visual word was presented for 1 s. The screen was then blank for 3 s prior to the presentation of a sound (an animal vocalization). The screen remained blank throughout the sound, and stayed blank for a further 2 s. Then, a response prompt was given and remained on screen until the participant selected either the left or right button. Corrective feedback was provided for 0.5 s, and then a 1.5-s blank screen was shown prior to the commencement of the next trial.

The only difference between the two-stimuli and ten-stimuli conditions was that, in the two-stimuli condition, only two words (CAT and DOG) were shown and only two sounds (woof and meow) were played. This meant that the two stimuli used in this task were presented 60 times each in the two-stimuli condition (half within congruent trials, and half within incongruent trials). The procedure took 45 min to complete, including time taken in break periods between blocks.

Data Quantification and Analysis

Data were collated and analyzed using BrainVision Analyzer (version 2.1). All EEG channels were rereferenced to the average of the two mastoids. A 1–30 Hz phase-shift free Butterworth filter was applied to the data. The data were epoched relative to the onset of the auditory stimulus (−100 to 800 ms) and baseline corrected (−100 to 0 ms). Ocular correction was applied to all channels using the procedure of Gratton, Coles, and Donchin (1983). To exclude artifacts, segments were subjected to four exclusion criteria: a gradient greater than 50 microvolts/second, a peak-to-peak difference of 200 microvolts, an amplitude of greater than ± 50 microvolts, or sustained activity of less than 0.5 microvolts for more than 100 ms. Overall, 2.95% of trials were excluded on the basis of these criteria.

For both conditions, all segments containing a congruent word–sound pair were averaged (congruent trials) separately from those that contained an incongruent word–sound pair (incongruent trials). The grand averages of the congruent and incongruent trials were calculated separately, for each individual.

Two time windows were used. The N2 was indexed by differential activity in the 250–300 ms period. The N400 is often indexed by differential activity in the 200–500 ms window (e.g., Orgs et al., 2006), but in order to keep measurement of the N400 as separate as possible from that of N2-related activity, we used a later window

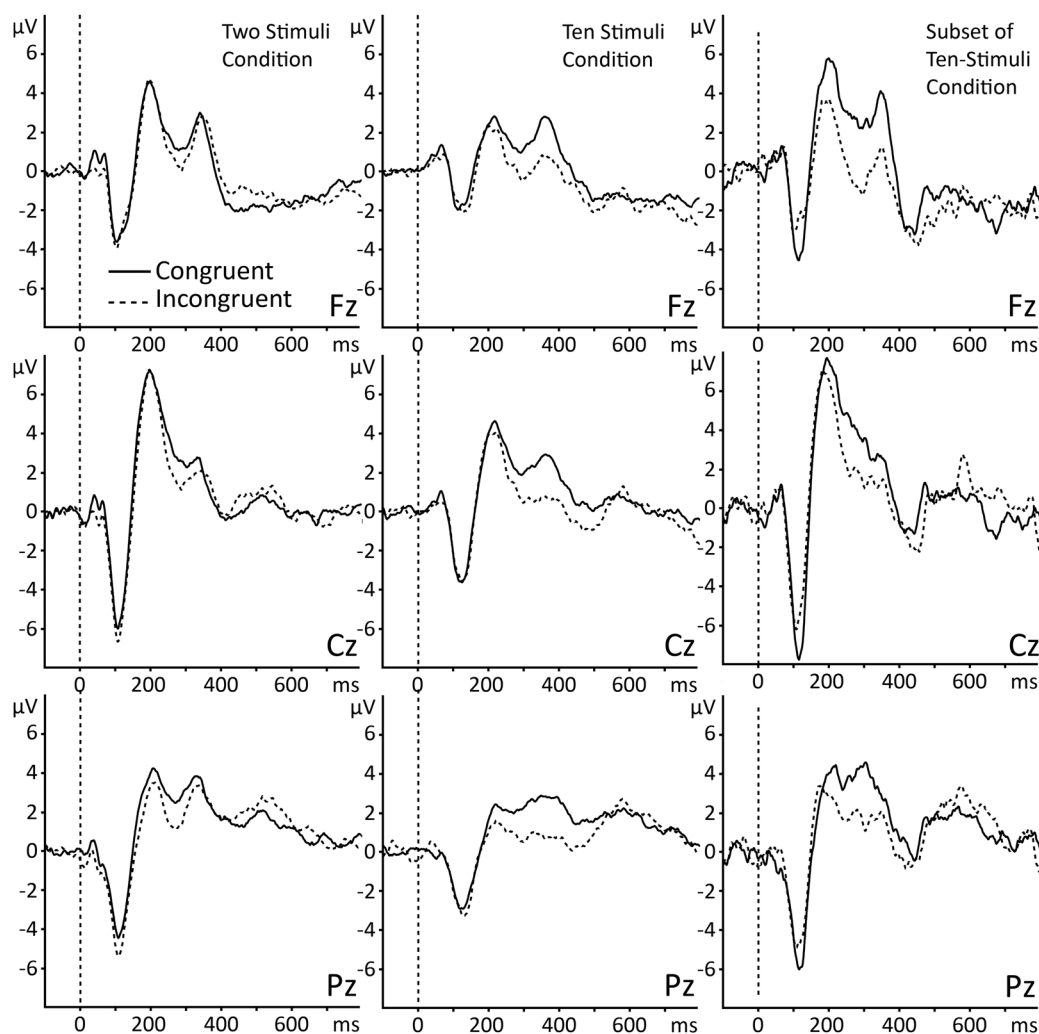


Figure 2. Auditory ERPs in the ten-stimuli (left), two-stimuli (middle), and the subset analysis of the ten-stimuli condition (right). In all the panels, the top ERP was measured at Fz, the middle ERP at Cz, and the lower ERP was measured at Pz. All ERPs were 50 Hz notch filtered for illustrative purposes only. Both vertical and horizontal scales were identical for all ERPs. The solid line indicates activity following a congruent trial (e.g., DOG–woof) and the broken line indicates activity following an incongruent trial (e.g., DOG–meow). The horizontal axis is measured in milliseconds. The vertical scale is measured in microvolts. As shown on the y axis, negative is down.

(300–500 ms) that was distinct from the N2 window. Mean amplitude values within both windows were analyzed using two identical analyses of variance (ANOVAs). Specifically, the N2 and N400 time windows were each individually analyzed using a $2 \times 2 \times 3$ ANOVA procedure, with within-subject factors of congruence (congruent trial or incongruent trial), condition (two-stimuli or ten-stimuli condition), and region of interest (ROI). There were three midline ROIs in total: anterior (average of AFz, FPz, Fz), central (FCz, Cz, CPz), and posterior (Pz, Poz, Oz). The same set of orthogonal planned contrasts were used in both ANOVAs. Effect size (partial eta-squared) statistics and 95% confidence intervals were provided for each contrast.

It is important to note that the evoking stimuli differed between the two- and ten-stimuli conditions; that is, the evoking stimuli in the two-stimuli condition were always woof or meow sounds, whereas there were 10 different evoking stimuli in the ten-stimuli condition. To control for this difference, an additional analysis was conducted using a subset of the items from the ten-stimuli data set. Specifically, in this additional analysis—which we dub the *subset* analysis—only the trials featuring the two auditory stimuli used in

the two-stimuli condition (i.e., the sounds meow and woof) were considered. Consequently, the ten-stimuli condition's statistics in this subset analysis are based on fewer trials per individual (22.08 on average, after exclusions due to artifacts) than the two-stimuli data set (109.60 trials, after exclusions). In all other respects (e.g., structure and type of inferential statistics used), the subset analysis was identical to that used to compare performance between the two- and ten-stimuli conditions in the main analysis. Finally, an additional ANOVA with one extra factor (position: first or last trials) was performed on the data from the two-stimuli condition to examine whether neural responding changed as a function of stimulus repetition.

Results

Behavioral Data

Participants' match and mismatch judgments were generally very accurate, indicating that participants were easily able to identify congruent and incongruent trials: mean accuracy was 95.45% with

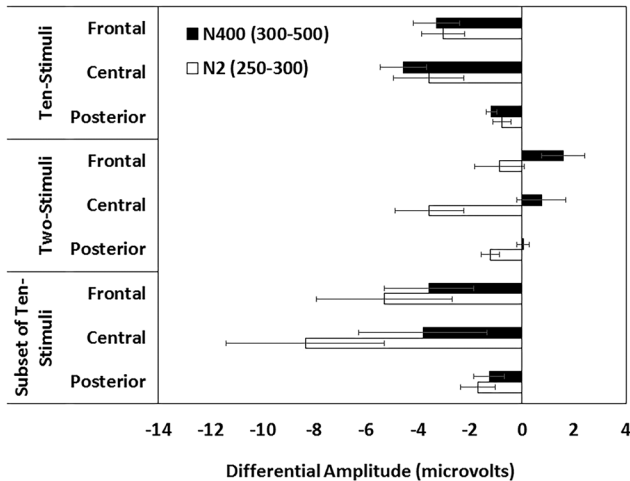


Figure 3. Mean differential amplitude (incongruent minus congruent trials) averaged across the midline electrode ROIs: posterior (AFz, FPz, Fz), central (FCz, Cz, CPz), and posterior (Pz, POz, Oz). These amplitudes were calculated separately for the N2 time window (250–300 ms, white column) and the N400 time window (300–500 ms, black column). The data are further divided between conditions (two-stimuli, ten-stimuli, and subset of ten-stimuli condition; see Subset Analysis). Error bars indicate standard error of the mean.

a standard deviation of 5.4%. All trials in which an incorrect response was made were omitted from further analyses.

ERP Analysis

ERPs for the congruent and incongruent trials in the two-stimulus condition (left panel) and ten-stimulus condition (middle panel) are shown in Figure 2 (positive polarity is plotted upward). The mean differential amplitudes (i.e., the amplitude for incongruent minus congruent trials) for each ROI separated by time window and condition are shown in Figure 3. Head maps depicting the voltage maps of mean amplitude (incongruent minus congruent trials) for the two- and ten-stimuli conditions are shown in the upper panels of Figure 4.

N2. Within the N2 time window (250–300 ms), a main effect of congruence was observed, $F(1,24) = 16.44$, $p < .001$, $\eta_p^2 = .41$, $CI[6.66, 20.48]$, whereby more negative mean amplitudes were seen for the incongruent trials compared to congruent trials. That is, an overall N2 effect was observed, collapsing across condition. The other contrasts of primary interest were those that interacted with this incongruent/congruent contrast. There was no significant interaction between congruency and condition within the N2 time window, $F(1,24) = 1.53$, $p = .23$, $\eta_p^2 = .06$, $CI[-8.33, 2.08]$, indicating no significant differences in overall N2 magnitude between the two- and ten-stimuli conditions. Two omnibus tests showed that congruence interacted with ROI, $F(2,48) = 3.25$, $p = .047$, $\eta_p^2 = .12$, but that the three-way interaction between ROI, condition, and congruence was not significant, $F(2,48) = 1.28$, $p = .29$, $\eta_p^2 = .05$. This suggests that there was some variation in the magnitude of N2 activity across the midline regions, but that this did not differ between conditions. The head maps show a diffuse pattern of negativity, and this is supported by the observation that all of the simple effects of incongruent versus congruent at each ROI were individually significant, minimum $F(1,24) = 7.31$, $p = .01$, $\eta_p^2 = .23$, $CI[0.93, 6.91]$.

N400. Within the 300–500 ms time window, a main effect of congruence was observed, $F(1,24) = 7.79$, $p = .01$, $\eta_p^2 = .25$, $CI[3.47, 23.13]$. That is, an overall N400 effect was observed, collapsing across condition. As previously, the other contrasts of primary interest were those that interacted with this incongruent/congruent

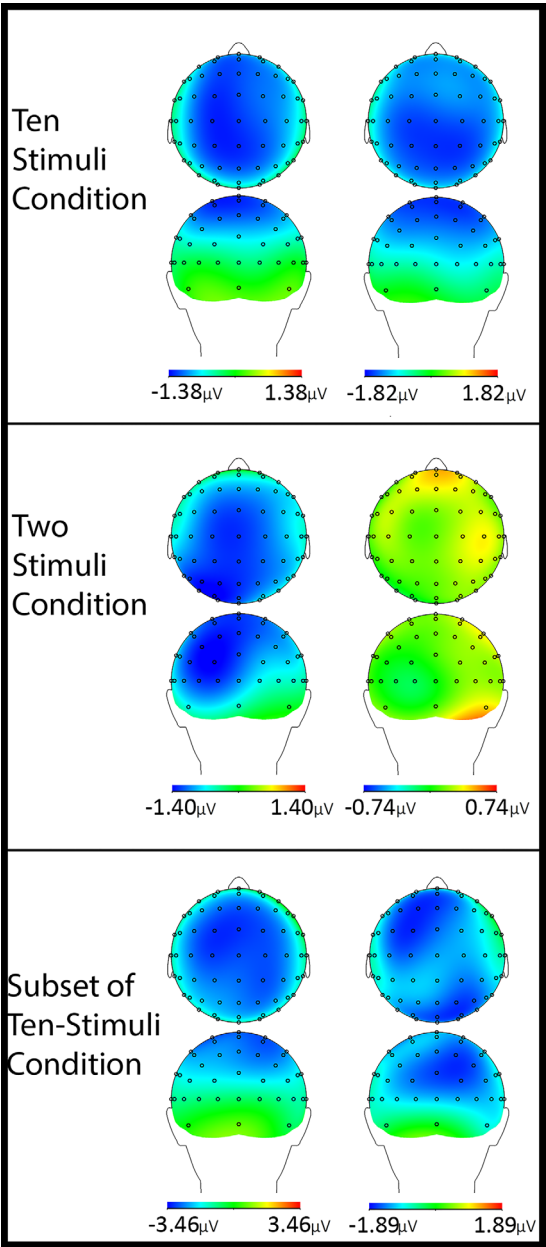


Figure 4. Activity maps plotting the location of differential activity between congruent and incongruent auditory stimuli. The upper panels depict activity in the ten-stimuli condition, the middle panel depicts activity in the two-stimuli condition, and the lower panel depicts activity following the stimuli shared between conditions but presented in the ten-stimuli condition (i.e., for the subset analysis of the ten-stimuli condition). Within each panel, the upper row of images shows top views whereas the lower row shows back views. Within each panel, the left images depict activity during the N2 time window (250–300 ms) and the right panel depicts activity during the N400 time window (300–500 ms). The color scale differs between conditions and time windows, so as to maximize sensitivity to any localizable patterns of activity within each condition and time window.

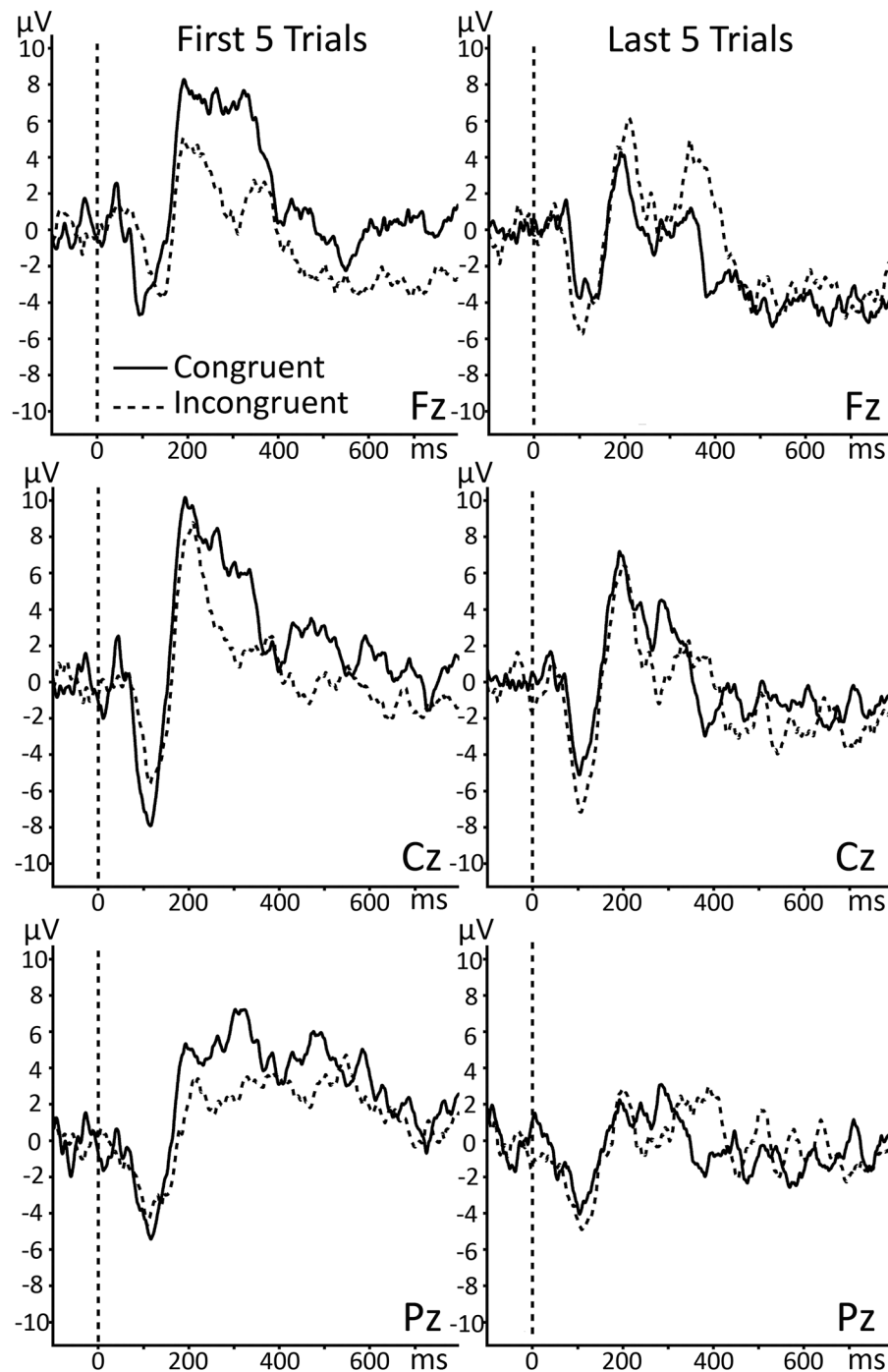


Figure 5. Auditory ERPs in the two-stimuli condition of those participants who completed this condition first ($n = 15$, see Repetition Analysis). In all panels, the solid line indicates activity following a congruent trial (e.g., DOG–woof) and the broken line indicates activity following an incongruent trial (e.g., DOG–meow). The left panels show ERPs for the first five trials in the block, while the right panels show ERPs for the final five trials. The top panels show activity at site Fz, the middle panels show Cz activity, and the lower panels show activity at Pz. All ERPs were 50 Hz notch filtered for illustrative purposes only. Both vertical and horizontal scales were identical for all ERPs. The horizontal axis is measured in milliseconds. The vertical scale is measured in microvolts. As shown on the y axis, negative is down.

contrast. Congruence interacted with condition, $F(1,24) = 29.05$, $p < .001$, $\eta_p^2 = .55$, $CI[14.09, 31.6]$, indicating that there was more differential activity in the ten-stimuli than the two-stimuli condition. A significant omnibus interaction between congruence, ROI, and condition suggested that this difference between the conditions was not equally distributed across all ROIs, $F(2,48) = 9.33$, $p < .001$, $\eta_p^2 = .28$. More specifically, planned contrasts further

revealed that the N400 effect was larger at each of the midline sites in the ten-stimuli condition than the two-stimuli condition, minimum $F(1,24) = 19.20$, $p < .001$, $\eta_p^2 = .44$, $CI[5.63, 15.65]$.

Simple effect contrasts revealed that there was significant N400 differential activity (more negative voltage on incongruent than congruent trials) at each of the midline sites in the ten-stimuli condition when considered in isolation, minimum $F(1,24) = 13.97$,

$p = .001$, $\eta_p^2 = .37$, $CI[2.94, 10.19]$, but no significant N400 differential activity at any ROI in the two-stimuli condition, maximum $F(1,24) = 3.96$, $p = .06$, $\eta_p^2 = .14$, $CI[2.94, 10.19]$ (note that this was in the opposite direction, indicating a trend toward more positive values on incongruent trials).

In summary, no differential N400 activity was seen in the two-stimuli condition at any midline site (or overall), whereas robust differential N400 activity was seen at each midline site (and overall) in the ten-stimuli condition.

Subset Analyses

As noted above, a supplementary analysis was performed that compared the magnitude of the N2 and N400 components in the two- and ten-stimuli conditions on only those trials that had a shared evoking stimulus—namely, those trials that resulted in a woof or meow sound. Data from the ten-stimuli condition, when limited to the woof and meow trials, is hereafter referred to as the subset condition. ERPs for the subset condition are shown in the right panel of Figure 2, voltage maps are shown in the lower panel of Figure 4, and the mean differential amplitude values for each ROI are shown in the lower columns of Figure 3. The subset condition was compared against the two-stimuli condition using the same inferential analysis as used above.

N2. A main effect of congruency was observed in the N2 time window, $F(1,24) = 14.11$, $p < .001$, $\eta_p^2 = .37$, $CI[9.48, 32.59]$. That is, an overall N2 effect was observed, collapsing across condition. Critically, two omnibus tests show that there were no significant interactions between congruence and condition, $F(1,24) = 2.59$, $p = .11$, $\eta_p^2 = .10$, or between congruence, condition, and ROI, $F(2,48) = 1.68$, $p = .20$, $\eta_p^2 = .07$. No evidence was found for a difference in N2 magnitude between the two-stimuli and subset conditions.

A series of planned contrasts revealed that a significant N2 component was evident in the subset condition when considered alone. These contrasts showed that, within the subset condition, there was significantly more activity on incongruent than congruent trials at all ROIs, minimum $F(1,24) = 4.27$, $p = .049$, $\eta_p^2 = .15$, $CI[0.01, 10.61]$.

N400. No overall N400 effect was observed. That is, averaged across condition and ROI, the main effect of congruence was not significant, $F(1,24) = 3.11$, $p = .09$, $\eta_p^2 = .11$, $CI[-2.13, 27.21]$. However, this appears to be due to the N400 being present in the subset but not the two-stimuli condition. This conclusion was supported by a significant interaction between congruence and condition, $F(1,24) = 5.08$, $p = .03$, $\eta_p^2 = .17$, $CI[1.86, 42.28]$ and a significant simple effect of congruence in the subset condition, $F(1,24) = 4.73$, $p = .04$, $\eta_p^2 = .16$, $CI[0.88, 33.72]$. From the earlier analyses, we already know that no differential activity was present in the N400 time window in the two-stimuli condition. An omnibus test showed that this difference in N400 activity between conditions did not significantly differ across the three ROIs, $F(2,48) = 1.88$, $p = .16$, $\eta_p^2 = .07$.

Repetition Analysis

If stimulus repetition underlies the observed differences between the two- and ten-stimuli conditions, then the ERPs should change across the course of the two-stimuli condition, in which the same stimuli were frequently repeated. Specifically, one might expect

the initial trials to resemble responding in the ten-stimuli condition (i.e., a protracted negativity across N2 and N400 time windows), while the final trials ought to more closely resemble the pattern anticipated for frequently repeated stimuli (i.e., with incongruity effects specific to the N2 time window). A secondary analysis assessed this prediction using the same ANOVA structure as used previously but with one additional factor, position (start of block or end of block). The data entered into this analysis were the first five and last five trials of the two-stimuli condition of those participants who completed the two-stimuli condition first (and hence for whom the woof and meow sounds were relatively novel at the start of the condition, $n = 15$). One might expect a priori that any effect of stimulus repetition ought to be most evident when the stimuli have not been encountered previously.

The ERPs obtained at midline sites for participants in the first and last five trials of the two-stimuli condition are shown in Figure 5. As can be seen in the figure, there appears to be more prolonged incongruence-related negativity in the initial trials (left panels) than on the final trials (right panels). Our primary question concerned whether the congruent/incongruent difference decreased across trials in the N2 time window, or in the N400 time window, or in both. To that end, planned contrasts examined the effect of trial position (first vs. last trials), on the difference between the congruent and incongruent trials within (a) the N2 time window (250–300 ms) and (b) the N400 time window (300–500 ms). There was no significant difference in the magnitude of the N2 effect between the first and last trials, $F(1,14) = 1.94$, $p = .18$, $\eta_p^2 = .12$, $CI[-3.25, 16.70]$, but there was a nonsignificant trend toward the N400 incongruence activity being larger on the initial trials than on the final trials, $F(1,14) = 3.46$, $p = .08$, $\eta_p^2 = .20$, $CI[-1.66, 31.40]$. Follow-up simple effect contrasts found that significant differential (incongruent minus congruent) activity was evident in the N400 period in the initial five trials, $F(1,14) = 8.19$, $p = .01$, $\eta_p^2 = .37$, $CI[3.23, 20.04]$, but this activity was not evident during this time window on the final five trials, $F(1,14) = 0.24$, $p = .63$, $\eta_p^2 = .02$, $CI[-16.81, 10.34]$. Although the data from five trials was noisy, it appears that the magnitude of the N400 incongruence response reduced across the course of the two-stimuli condition.

Discussion

The present experiment investigated whether the semantic incongruence between a visual lexical cue (e.g., the written word CAT) and a nonlexical auditory stimulus (the sound of a dog woofing) would elicit a characteristic and well-studied ERP component, the N400. An N400-like pattern of activity was indeed observed, consistent with Van Petten and Riefelder's (1995) earlier related study, which used only auditory stimuli. Incongruent auditory stimuli in the ten-stimuli condition, like those in Van Petten and Riefelder's experiment, elicited a pattern of sustained negativity beginning around 200 ms and, in our data, terminating around 500 ms after stimulus onset. However, in an otherwise similar condition in which only two visual and two auditory stimuli were shown (rather than 10), a very different pattern of incongruence-related activity was seen. In the two-stimuli condition, a more focused pattern of activity was evident in responses to semantic incongruence. Specifically, an incongruence-related negativity was identified that centered around the N2 component, and was evident for only a brief window (approximately 250–300 ms). A secondary analysis showed that this more-focused pattern of activity was not evident at the onset of this condition, but developed only with exposure to the frequently repeated stimuli. Finally, a subset analysis

demonstrated that the difference in ERPs between conditions was not merely a consequence of differences in the evoking stimuli in the two conditions, since similar differences in ERPs were observed when analysis of the ten-stimuli condition was restricted to the sounds that occurred in the two-stimuli condition. That is, ERP differences between the two- and ten-stimuli conditions were observed even when the evoking stimuli were identical.

Comparisons with Prior Research

The present study is not the first to demonstrate an N400 effect to nonspeech sounds (see Orgs et al., 2006; Van Petten & Rheinfelder, 1995), but to our knowledge it is the first to demonstrate both an N2 and an N400 evoked by the same stimuli in the same task. So how could the same task using the same stimuli elicit two quite different components, which are generally thought to be associated with quite different cognitive processes? The N400 is generally thought to index semantic and linguistic processing, perhaps via access to semantic memory (Kutas & Federmeir, 2000), whereas the N2 is typically interpreted as indexing cognitive control in the visual domain (see Folstein & Van Petten, 2008, for review) or classifying and categorizing deviant stimuli (Mueller, Brehmer, von Oertzen, Li, & Lindenberger, 2008; Näätänen & Gaillard, 1983). However, if one considers the two components more broadly, then the processing attributed to each can be seen as similar. As discussed earlier, the N2 can be more broadly characterized as “resulting from a deviation in form or context of a prevailing stimulus” (Patel & Azzam, 2005, p. 147). The same can be said of the N400, as long as the word *context* here refers only to a semantic context.

Deacon et al. (1991) noted the porous boundaries between typical N400 and N2 tasks in their study. They demonstrated that the same stimuli (words) elicited either an N400 or an N2 when judged in different ways. Specifically, when participants were asked to compare two sequentially presented words on the basis of their category memberships (e.g., are they both kinds of animals?), incongruence from Stimulus 1 to Stimulus 2 elicited N400-like activity (relative to congruent pairs). However, when the target stimulus was compared to the preceding stimulus on the basis of visual characteristics (e.g., whether the two words were the same size), then incongruity between the two stimuli elicited N2-like activity. This finding suggests that the way participants treat the target stimulus leads to an N2 or N400 pattern of activity being evoked by stimulus incongruity. Specifically, if the target stimulus is treated as a rich and meaningful object whose meaning needs to be extracted and compared against stored semantic knowledge, then an N400 is observed. By contrast, if that same stimulus is instead processed superficially and perceptually, then incongruity is detected and signaled rapidly in N2-like differential activity.

The present data go beyond these prior studies in that they demonstrate that a protracted N400 or just an N2 can be differentially evoked by the same stimuli, in the same task. Thus, the difference in cognitive activity indexed by these components does not appear to be inherent to the task itself, but rather to participants' approach to the task. This suggests that the N2 and N400 have a common functional underpinning, but are differentially affected by aspects of the experimental context. Specifically, the number or frequency of stimuli presented during the task appears to strongly affect the N400 pattern of activity.

Sounds as Tokens

This pattern of data can be explained if one first assumes that both the N400 and the N2 index deviancy detection—or, more generally, cognitive processes involved in relating the target stimulus to that prior—but that these processes differ in complexity and duration (see Deacon et al., 1991; Herning et al., 1987; Polich, 1985). Notice that the N2 is typically elicited in oddball tasks (in which one stimulus occurs frequently, with an infrequent oddball secondary stimulus; Sams, Paavilainen, Alho, & Näätänen, 1985), sequential matching tasks (Wang et al., 2003), or go/no-go tasks (in which one stimulus requires a particular motor response, while a second requires suppression of that response; Pfefferbaum, Ford, Weller, & Kopell, 1985). Thus, N2 protocols typically use stimuli that are arbitrarily associated with each other and with the response options. For example, a square picture might indicate that a left response is needed. As such, these tasks primarily use symbols that are meaningfully related to each other only in the context of the experiment. One can therefore consider these stimuli as tokens rather than inherently meaningful stimuli. On this analysis, N2 activity indexes the detection of a physical discrepancy between an anticipated sensory event and an observed sensory event (the target stimulus). If the target stimulus is merely a token, and thus requires (or evokes) no further semantic processing, such as in go/no-go or sequential matching tasks, then only an N2 will be observed.

By contrast, N400 tasks typically use stimuli that are objectively linguistically/semantically related. That is, they use stimuli whose relationships with each other exist outside the experiment and are context general. For example, to an English speaker the word DOG and the sound woof are related to each other in any context. Indeed, the robust observation of semantic priming shows that presenting a meaningful stimulus primes other stimuli of the same meaning not only in the context of meaningful sentences, but also when presented in relatively context-free circumstances (e.g., in word lists). Moreover, a meaningful stimulus appears to prime not only the meaning of subsequent stimuli, but also their sensory properties, at least in some cases. For example, Van Petten, Coulson, Rubin, Plante, and Parks (1999) noted that when a highly anticipated word (e.g., captain) was replaced with a semantic deviant with a common initial phoneme (e.g., captive), the onset of N400 activity was delayed (to around 350 ms). This suggests that people were anticipating a particular sound, and it was only when a physical deviance was detected (as may typically be indexed by N2 activity) that further semantic processing is undertaken. In a semantically congruent trial, where the target stimulus is semantically related to the previously shown word (or a semantically constrained context), minimal additional semantic processing is required, as the semantic properties of the target stimulus (which are shared with the priming stimuli) are already activated. However, if the target stimulus is unrelated to that shown previously (i.e., an incongruent trial), significant additional processing is required to activate the semantic correlates of the target stimulus, and it is this additional activity that is indexed by the protracted N400 response. The N400 may always be a kind of N2/N400 response, whereby physical incongruence is detected first and then prompts a search for semantic congruence/incongruence.

Under this account, the N400 indexes a more complicated, diffuse, and context-general semantic matching process, whereas the N2 indexes a narrower, simpler, and context-dependent perceptual matching process. Perhaps when a small, highly repeated stimulus set is used, the former, more complicated semantic process reduces to a simpler perceptual process, thereby yielding only an N2 in the

two-stimuli condition. More specifically, if the semantic comparison task is repeated several times with the same stimuli, it might cause (a) the mental template against which the target stimuli are compared to become highly defined and perceptually rich (e.g., after seeing the word DOG a single specific woof sound is anticipated, rather than all things related to dogs), and (b) this would allow the target stimuli to be processed shallowly, as an auditory token that is meaningful only in the context of the experiment. This latter phenomenon is readily demonstrated out of the lab. If one repeatedly utters the same word 100 times, then the sound of the word (its perceptual features) becomes somewhat divorced from its meaning, and instead becomes perceived as a series of syllables: the semantic satiation effect (Smith, 1984). Perhaps that is what occurred in the two-stimuli condition. That is, after many repetitions, the woof and meow sounds ceased being processed for abstract, linguistic meaning (i.e., the vocalization that dogs and cats make, respectively) and instead were processed as arbitrary symbols that form part of the present task (i.e., the sound tokens that were paired with the DOG and CAT visual stimuli in this experiment). Thus, because these auditory tokens were no longer being processed for meaning, they became incapable of eliciting a

protracted N400 activity when they followed a semantically incongruent word.

Conclusion

The primary finding of the current study was that incongruence between visual lexical cues and nonlexical auditory sounds could elicit either a protracted N400-like pattern of activity or a more focal, N2-like pattern of activity during performance of the same task. The differential pattern of incongruence-related activity evoked when participants performed a semantic match/mismatch task depended on the frequency with which the target stimuli were repeated. Intriguingly, this implies that the components yielded by semantic incongruence primarily depend upon the participant's approach to the task, rather than the task, the stimuli, or the nature of the incongruence (semantic, perceptual) per se. These results pose interesting questions concerning the relatedness of the N2 and N400 components, and are broadly consistent with the suggestion that the N400 component reflects a more complicated and protracted N2 incongruence response (Deacon et al., 1991; Polich, 1985).

References

- Balota, D. A., & Black, S. (1997). Semantic satiation in healthy young and older adults. *Memory & Cognition*, 25, 190–202. doi: 10.3758/BF03201112
- Besson, M., & Macar, F. (1987). An event-related potential analysis of incongruity in music and other non-linguistic contexts. *Psychophysiology*, 24, 14–25. doi: 10.1111/j.1469-8986.1987.tb01853.x
- Black, S. (2004). Review of semantic satiation. *Advances in Psychology Research*, 26, 63–74.
- Boddy, J., & Weinberg, H. (1981). Brain potentials, perceptual mechanisms and semantic categorization. *Biological Psychology*, 12, 43–61. doi: 10.1016/0301-0511(81)90019-3
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10, 433–436. doi: 10.1163/156856897X00357
- Conci, M., Gramann, K., Müller, H. J., & Elliott, M. A. (2006). Electrophysiological correlates of similarity-based interference during detection of visual forms. *Journal of Cognitive Neuroscience*, 18, 880–888. doi: 10.1162/jocn.2006.18.6.880
- Deacon, D., Breton, F., Ritter, W., & Vaughan, H. G. (1991). The relationship between N2 and N400: Scalp distribution, stimulus probability, and task relevance. *Psychophysiology*, 28, 185–200. doi: 10.1111/j.1469-8986.1991.tb00411.x
- Folstein, J. R., & Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology*, 45, 152–170. doi: 10.1111/j.1469-8986.2007.00602.x
- Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, 55, 468–484. doi: 10.1016/0013-4694(83)90135-9
- Harbin, T. J., Marsh, G. R., & Harvey, M. T. (1984). Differences in the late components of the event-related potential due to age and to semantic and non-semantic tasks. *Electroencephalography and Clinical Neurophysiology*, 59, 489–496. doi: 10.1016/0168-5597(84)90007-8
- Herning, R. I., Speer, M., & Jones, R. T. (1987). Event-related potentials to spoken equations: Is the N400 really a late N200? In R. Johnson, Jr., J. W. Rorhbaugh, & R. Parasuraman (Eds.) *Current trends in event-related potential research* (pp.394–398). Amsterdam, The Netherlands: Elsevier.
- Hickey, C., McDonald, J. J., & Theeuwes, J. (2006). Electrophysiological evidence of the capture of visual attention. *Journal of Cognitive Neuroscience*, 18, 604–613. doi: 10.1162/jocn.2006.18.4.604
- Kleiner, M., Brainard, D. H., & Pelli, D. G. (2007). What's new in Psychtoolbox-3? [ECVP Abstract Supplement]. *Perception*, 36.
- Kounios, J., Kotz, S., & Holcomb, P. (2000). On the locus of the semantic satiation effect: Evidence from event-related brain potentials. *Memory & Cognition*, 28, 1366–1377. doi: 10.3758/BF03211837
- Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, 4, 463–470. doi: 10.1016/S1364-6613(00)01560-6
- Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, 62, 621–647. doi: 10.1146/annurev.psych.093008.131123
- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207, 203–205. doi: 10.1126/science.7350657
- Mueller, V., Brehmer, Y., von Oertzen, T., Li, S. C., & Lindenberger, U. (2008). Electrophysiological correlates of selective attention: A lifespan comparison. *BMC Neuroscience*, 9, 18–39. doi: 10.1186/1471-2202-9-18
- Näätänen, R., & Gaillard, A. W. K. (1983). The orienting reflex and the N2 deflection of the ERP. In A. W. K. Gaillard & R. Naatanen (Eds.), *Tutorials in event-related potential research: Endogenous components* (pp. 119–141). Amsterdam, The Netherlands: Elsevier. doi: 10.1016/S0166-4115(08)62036-1
- Orgs, G., Lange, K., Dombrowski, J. H., & Heil, M. (2006). Conceptual priming for environmental sounds and words: An ERP study. *Brain and Cognition*, 62, 267–272. doi: 10.1016/j.bandc.2006.05.003
- Patel, S. H., & Azzam, P. N. (2005). Characterization of N200 and P300: Selected studies of the event-related potential. *International Journal of Medical Sciences*, 2, 147. doi: 10.7150/ijms.2.147
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437–442. doi: 10.1163/156856897X00366
- Pfefferbaum, A., Ford, J. M., Weller, B. J., & Kopell, B. S. (1985). ERPs to response production and inhibition. *Electroencephalography & Clinical Neurophysiology*, 60, 423–434. doi: 10.1016/0013-4694(85)91017-X
- Polich, J. (1985). Semantic categorization and event-related potentials. *Brain and Language*, 26, 304–321. doi: 10.1016/0093-934X(85)90045-8
- Sams, M., Paavilainen, P., Alho, K., & Näätänen, R. (1985). Auditory frequency discrimination and event-related potentials. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, 62, 437–448. doi: 10.1016/0168-5597(85)90054-1
- Sanquist, T. F., Rohrbaugh, J. W., Syndulko, K., & Lindsay, D. B. (1980). Electrocorical signs of levels of processing: Perceptual analysis and recognition memory. *Psychophysiology*, 3, 27–36. doi: 10.1111/j.1469-8986.1980.tb02299.x
- Smith, L. C. (1984). Semantic satiation affects category membership decision time but not lexical priming. *Memory & Cognition*, 12, 483–488. doi: 10.3758/BF03198310
- Van Petten, C., Coulson, S., Rubin, S., Plante, E., & Parks, M. (1999). Time course of word identification and semantic integration in spoken language. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 25, 394–417. doi: 10.1037/0278-7393.25.2.394
- Van Petten, C., & Rheinfielder, H. (1995). Conceptual relationships between spoken words and environmental sounds: Event-related brain

- potential measures. *Neuropsychologia*, 33, 485–508. doi: 10.1016/0028-3932(94)00133-A
- Wang, Y., Cui, L., Wang, H., Tian, S., & Zhang, X. (2004). The sequential processing of visual feature conjunction mismatches in the human brain. *Psychophysiology*, 41, 21–29. doi: 10.1111/j.1469-8986.2003.00134.x
- Wang, Y., Tian, S., Wang, H., Cui, L., Zhang, Y., & Zhang, X. (2003). Event-related potentials evoked by multi-feature conflict under different attentive conditions. *Experimental Brain Research*, 148, 451–457. doi: 10.1007/s00221-002-1319-y
- World Medical Association. (2001). World Medical Association Declaration of Helsinki. Ethical principles for medical research involving human subjects. *Bulletin of the World Health Organization*, 79(4), 373.

(RECEIVED November 11, 2015; ACCEPTED February 24, 2016)