Effects of semantic matching and of semantic category on reaction time and N400 that resist numerous repetitions

J. Bruno Debruille a,b,c,∗, Louis Renoult a,c

a Douglas Mental Health University Institute, Montréal, Québec, H4H 1K3, Canada
b Department of Psychiatry, McGill University, Montréal, Québec, H3A 1A1, Canada
c Department of Neurology and Neurosurgery, McGill University, Montréal, Québec, H3A 2B4, Canada

1. Introduction

One frequent way to study how the brain processes the meaning of words is to explore the semantic priming effect, that is the facilitation of the processing of a word (e.g., nurse) which is produced by the prior presentation of a semantically related word (e.g., doctor) (Meyer & Schvaneveldt, 1971). Behavioral priming effects consist in reduced reaction times (RTs) to primed words compared to words preceded by semantically unrelated words. Priming effects also appear in the modulation of an event-related brain potential (ERP): the N400, a negative electrical field that peaks around 400 ms after word onset and constitutes a response of the brain to any potentially meaningful stimulus (Kutas & Federmeier, 2000). Primed words elicit smaller N400 potentials than unprimed words. Such RT and N400 priming effects have been observed in hundreds of experiments (see Kutas & Federmeier, 2000 for a review) including in tasks where access to the meaning of words is not necessary, such as lexical decision tasks (e.g., Bentin, McCarthy, & Wood, 1985). However, these effects are stronger in tasks that require subjects to focus on the semantic properties of words (Chwilla, Brown, & Hagoort, 1995; Holcomb, 1988; Mitchell et al., 1991), such as when subjects are asked to decide if the meaning of a word is related to the meaning of the preceding word (e.g., Debruille et al., 2007; Kiang & Kutas, 2005; Kiang, Kutas, Light, & Braff, 2007). In these tasks, RT and N400 effects are named semantic matching effects.

RTs and N400s are used not only to measure the extent to which some words are related to others, but also to study the processing of words according to their semantic category. Shorter RTs to concrete words (e.g., ‘apple’) than to abstract words (e.g., ‘idea’) have been consistently found (reviewed in Paivio, 1991). This is referred to as the concreteness effect. Moreover, within concrete words, shorter RTs for natural categories (e.g., names of animals) than for names of artifacts (e.g., names of tools) have been frequently described (Kiefer, 2001; Paz-Caballero, Cuetos, & Dobarro, 2006; Price & Humphreys, 1989; Proverbio, Del Zotto, & Zani, 2007). Similarly, N400s to concrete words have larger amplitudes than those evoked by abstract words (Kounios & Holcomb, 1994; West & Holcomb, 2000). Within concrete words, different scalp distributions have been associated with words from living vs. artifactual categories (Antal, Keri, Kovacs, Janka, & Benedek, 2000; Kiefer, 2001, 2005; Paz-Caballero et al., 2006; Proverbio et al., 2007; Sim & Kiefer, 2005; Sitnikova, West, Kuperberg, & Holcomb, 2006). The topography of these category effects is consistent with the results of neuropsychological studies of semantic priming with concrete words.
seems to interact with semantic matching (Besson, Kutas, & Van Petten, 1992; Mitchell, Andrews, & Ward, 1993). As a matter of fact, a single repetition is already associated with an important decrease in RTs. This is known as the repetition priming effect (Monsell, 1985; Scarborough, Cortese, & Scarborough, 1977). After multiple repetitions, it seems that a ‘floor’ is reached, although there are some discrepancies between studies. Not considering the ‘special case’ of immediate repetition for which sensory memory effects may intervene (Henson, 2003), certain studies found no further RT facilitation after only two (Besson et al., 1992; Forbach, Stanners, & Hochhaus, 1974; Smith & Halgren, 1989) or three stimuli repetitions (Kaznerski & Friedman, 1997). In contrast, others found a linear decrease in RTs using five (Jiang, Hasby, Martin, Ungerleider, & Parasuraman, 2000) or nine repetitions (Van Strien, Hagenbeek, Stam, Rombouts, & Barkhof, 2005).

For N400 amplitudes, the major reduction with repetition is part of the so-called old/new ERP effect (for a review see Rugg & Curran, 2007). It can still be seen with lags up to 30 min between the two presentations (Duzel, Yonelinas, Mangun, Heinze, & Tulving, 1997). Some studies reported that this amplitude reduction was proportional to the number of repetitions (Finnigan, Humphreys, Dennis, & Geffen, 2002; Rugg, Brovedani, & Doyle, 1992; Van Petten, Kutas, Klunder, Mitchiner, & McIsaac, 1991). Others found floors in N400 amplitude as early as the second or third presentation (Besson et al., 1992; Kaznerski & Friedman, 1997; Kouinos & Holcomb, 1992; Van Strien et al., 2005; Young & Rugg, 1992). Therefore N400 contrasts with other ‘semantic ERPs’, such as the recognition potential (RP) and the mismatch negativity (MMN) elicited by words, which seem to be impervious to repetition (for the RP see Martin-Loeches et al., 2002; for the MMN see Pulvermüller & Shtyrov, 2006).

Nevertheless, only a few studies have assessed the impact of repetition on the effects of category and of semantic matching. For the effect of category, Kiefer (2005) and Sim and Kiefer (2005) reported no interaction between repetition and category on RTs. In other words, the semantic category of the words (natural or artificial) appeared to have the same impact on RTs at the first and second presentation. In contrast, when considering ERPs in the N400 time window, interactions were found. The category effects were smaller (Sim & Kiefer, 2005) or non-existent (Kiefer, 2005) for stimuli that had been already presented. However, the use of a shape judgment task in Sim and Kiefer (2005), and that of a lexical decision task in Kiefer (2005) could be responsible for this lack of resistance. Category effects are more robust in tasks where semantic processing is explicitly required (for a review see Devlin et al., 2002).

As for semantic matching, Hanze and Meyer (1995) found a reduction of the effect on RTs when words were repeated in a naming task. Behavioral studies using lexical decision tasks have sometimes reported additive effects of semantic priming and repetition (den Heyer, Goring, & Dannenbring, 1985; den Heyer, 1986), and sometimes, a reduction of priming with repetition (Carroll & Kirsner, 1982; den Heyer & Benson, 1988). One interpretation of these discrepancies is that the two effects interact only when the number of trials between repetitions is relatively limited (i.e., between 0 and 7, see den Heyer & Benson, 1988). As for the N400, two studies using words in sentences found that repetition reduced semantic matching effects (Besson et al., 1992; Mitchell et al., 1993).

In Besson et al. (1992), where two levels of repetitions were contrasted, the robust N400 difference found between semantically congruous and incongruous sentence endings was attenuated at the second presentation and eliminated by the third presentation of the sentences. Their results suggest that semantic matching effects disappear at about the same time as the amplitude of the N400 deflection reaches floor levels. However in this study, complete sentences were repeated, making it likely that their final words, to which N400s were recorded, could be anticipated (see also Olichney et al., 2000). These words may thus have been ‘completely primed’, hence the quasi absence of N400 matching effect. This possibility is reinforced by the instruction to remember sentences’ final words that was given to the subjects. In contrast, in another study using pairs of words, Radeau, Besson, Fonteneau, & Castro (1998) reported that the effect of semantic matching on N400s and RTs at the second presentation was the same as that observed at the first presentation. However, the lag between the two presentations of a word was long (20 min) and the effect of repetition on N400s was small and only marginally significant, which is quite unusual.

In fact, the large reduction of RTs and of N400 amplitudes found in many studies with just one repetition suggests that little or nothing remains to be modulated after several repetitions. Accordingly, no effect of semantic matching and of semantic category would be expected. On the other hand, the evidence showing an absence of semantic matching and category effects on RTs and N400s for stimuli that have already been presented is scarce. The persistence of these effects for highly repeated stimuli does not appear impossible. In fact, such effects would be consistent with the results of a recent fMRI study that suggest that repetition modulates the activity of brain areas different from the areas whose activity is modulated by semantic matching (Rapoport, Moss, Stamatakis, & Tyler, 2006). Repetition could thus suppress the activity of only a part of the N400 generators. Others could still be modulated by semantic matching and category, notably when semantic processing is explicitly required.

The objectives of the present study were thus to test whether effects of category and of semantic matching on RTs and N400s could still be found at high repetition rates. To achieve this goal, we chose an explicit semantic task in the hope of maintaining the semantic processing of repeated words. Subjects had to decide if the meaning of a target word matched that of a prime word. Two versions of this task were used. The first version was a ‘classic’ design in which two category words were used as primes: ‘vivant’ and ‘chose’, which mean ‘living’ and ‘thing’ in French. Numerous non-repeated exemplars of these two categories were used as targets. Match and category effects were thus expected on RTs and N400s. The second version of our task was the critical one. Primes and targets were inverted: the primes were the numerous non-repeated category exemplars and the target words, to which N400s were recorded, were simply the two category words, each presented 60 times. These target words occurred at random with equal probability (0.5).

Note that the category words ‘vivant’ and ‘chose’, that is, ‘living’ and ‘thing’, were chosen not only to obtain category effects on RTs and N400s but also to obtain differential semantic matching effects. To achieve this goal, in both versions of the task, the category ‘vivant’ was only associated with animal exemplars. Meanwhile, the category ‘chose’ was associated with varied category exemplars: kitchen/desk objects, clothes, tools, furniture, buildings and transportation means. According to Collins and Loftus (1975), the greater the number of concepts that are primed in such a broad context, the less each of them can be primed. We thus expected the category ‘chose’ to be a less efficient prime than ‘vivant’ in the first version. Similarly, in the second version, ‘chose’ was less likely to be primed by the artifact exemplars than the category ‘vivant’ was to be primed by the animal exemplars. The use of these words could...
thus allow the assessment of semantic matching for two different strengths of priming. Meanwhile, it also permitted the assessment of the category effect for unprimed targets, for which it is usually observed (Antal et al., 2000; Kiefer, 2005; Paz-Caballero et al., 2006; Provero et al., 2007; Smitaka et al., 2006).

To characterize category effects, we compared their scalp distribution between versions. To analyze the semantic matching effects of the two versions, we compared their amplitudes, their time of onset, the latencies of their peaks and their distribution on the scalp. Finally, to see to what extent these match effects could be associated with similar brain generators, we submitted the subtraction ERPs to an independent component analysis (ICA) and a dipole source modeling of the independent components obtained.

If the semantic processes indexed by N400s and RTs are still present for the highly repeated target words of version 2, these indexes should be modulated by semantic category and semantic matching. Alternatively, if these semantic processes are abolished in the second version, the processing of the two repeated target words might be limited to the most salient perceptual differences existing between them. Indeed, in the second version, subjects may anticipate the correct answer after the presentation of the prime (for instance ‘vivant’ if the prime is an animal name) and then rapidly consider whether the string of letters that occurs is ‘vivant’ or whether it is ‘chose’. In this case of physical matching, target processing would be impervious to the semantic category of the target word and no differential matching effect should be seen despite the unequal semantic priming strengths of the two categories used. Results were published in an abstract form in Renoult & Debruille (2008).

2. Methods

2.1. Participants

32 right-handed native French speakers (14 women and 18 men) took part in the two versions of the experiment. They had normal or corrected-to-normal vision and no history of neurological or psychiatric disorders. They were recruited by newspaper advertisements among people aged between 18 and 50 years (mean: ± 10), and had at least a college level of education. They signed an informed consent form accepted by the Douglas Institute Research and Ethics Board. Data of five subjects were excluded from analyses due to excessive ocular artifacts in the EEG recordings.

2.2. Task and procedure

Subjects were seated comfortably in a dimly lit room in front of a computer screen placed 1 m from their eyes. Black words which length varied from 2.3 to 3 cm (mean: 2.6 ± 0.1) and height from 0.5 to 0.8 cm (mean: 0.68 ± 0.1) were presented on a white background at the center of this screen. The mean number of letters of the words was 7 ± 2. Each trial was made of three serially presented French words, each occurring for a 1 s duration and followed by a 1 s blank interval. The word ‘attention’, which has the same spelling and meaning in French and English, initiated each trial, followed by the prime word and the target word. In the first version of the experiment, the prime was the category word ‘vivant’ (= ‘living’) or ‘chose’ (= ‘thing’), and the target was one of 60 possible exemplars of animals or one of 60 possible exemplars of artifacts. In the second version, primes and targets were inverted. The prime was one of 60 possible exemplars of animals or one of 60 possible exemplars of artifacts, and the target was one of the category words, each presented 60 times. In both versions, in half of the trials the target word semantically matched the prime word and in the other half, it did not. When the target word occurred, subjects had to respond yes for matches or no for mismatches as rapidly and as accurately as possible by pressing one of two keyboard keys with their right index finger. Then, they were requested to blink in response to the ‘clignez’ (= ’blink’) instruction, which appeared 1–1.5 s after the offset of each target stimulus. The next trial began after a time interval that randomly varied between 1.5 and 2 s.

Both French category words ‘vivant’ and ‘chose’ are very frequent words, occurring (mean: 7 ± 2.6) for names of animals. There was no significant difference in frequency between the two categories (t = 0.34, p = .73).

The 120 animals’ names were used to make four lists of 30 words matched on average frequency and average word length (A1, A2, A3, A4). Similarly, four lists were also created out of the 120 exemplars of artifacts (T1, T2, T3, T4). Each exemplar word was used equally often as a target (in the first version) or as a prime (in the second version) and in a matching or in a mismatching trial. Half of the subjects had the first and then the second version, while the order was reversed for the other half of the subjects. Subjects were divided into eight subgroups according to the list of words they were presented with in each experimental condition and according to whether they started by version 1 or 2. Each exemplar word was presented only once during the experiment but each category word appeared 60 times in each version of the task, as a prime or as a target. Note that the term ‘animal targets’ in the text that follows refers not only to the trials in which 60 different animal exemplars were used as targets in the first version but also to the trials of the second version in which the corresponding category word (i.e., the word ‘vivant’) was presented 60 times as a target (30 times in the match and 30 times in the mismatch condition). Similarly, the term ‘artifact targets’ refers not only to the trials of the first version in which 60 different artifacts exemplars were used as targets but also to the trials of the second version in which the corresponding category word (i.e., the word ‘chose’) was presented 60 times.

2.3. Data acquisition

Accuracy and reaction time were recorded for each trial. The EEG was recorded with tin electrodes mounted in an elastic cap (Electrocap International) from 30 active points all referenced to the right ear lobe. 28 of these points were placed according to the extended International 10-20 System (Electrode Nomenclature Committee, 1991). These electrode sites could be grouped in a sagittal montage, which comprised Fz, FCz, Cz, and Pz; a para-sagittal montage, comprising FP1(2), F3(4), FC3(4), C3(4), CP3(4), P3(4), and O1(2); and a lateral montage, comprising F7(8), FT7(8), T3(4), T7(8), P7(5), and T5/6. The remaining two active electrodes were placed below each eye in order to allow the monitoring of vertical eye movements by comparing their EEG signals to those derived from Fp1 and Fp2. The monitoring of horizontal eye movements was done by comparing F8 to F7 signals. The EEG was amplified 20,000 times by Contact Precision amplifiers. High- and low-pass filter half-amplitude cut-offs were set at 0.01 and 100 Hz, with an additional 60 Hz electronic notch filter. Signals were then digitized on-line at a sampling rate of 256 Hz and stored along with stimulus and response codes for subsequent averaging using the Instep (version 4.3) software package.

EEG epochs contaminated by eye movements, excessive myogram, amplifier saturations or analog to digital clipping were removed offline by setting automatic rejection criteria. Trials for which analog to digital clipping exceeded a 100 ms duration, and electrodes for which amplitude exceeded ± 100 μV were excluded from averaging.

2.4. Data processing and measures

Mean RTs for each condition were computed using only the correct responses and excluding the trials where subjects took more than 2000 ms to respond. ERPs to target words were computed by averaging the 1000 ms EEG epochs of these trials of each experimental condition, using a − 200 to 0 ms baseline before target onset. For measuring the N400s, the mean voltage amplitude for each subject in each condition was computed in time windows that were centered on the latencies of the negative peaks that appeared between 200 and 500 ms on the grand averages of mismatch trials. These negative peaks culminated at 400 ms (N400) in version 1, as is classical, and at about 290 ms (N290) in version 2. Mean voltage amplitudes in each subject were thus computed in a 350–450 ms time window in version 1 and in a 240–340 ms time window in version 2. The short duration of the time windows (i.e., 100 ms) was chosen to focus on the negatives and to prevent the inclusion of the late positive complex (LPC), the deflection that immediately follows the negativities of interest.

The latency of the peak of the N400 in version 1 and of the N290 in version 2 were measured in each subject at Cz by assessing the timing of the maximum of the negative peak observed between 200 and 500 ms. These measures were done in the mismatch conditions, where the amplitudes of these negative waves were maximal. The waveforms were smoothed until only one negative peak appeared in this epoch. On the other hand, the latency of the onset of the N400 and of the N290 deflections and the onset latency of the match effects were determined using the method proposed by Hagoort, Hald, Bastiaansen, & Petersson (2004) (see Section 2.5).

2.5. Statistical analyses

For analyzing RTs, we ran a repeated-measures ANOVA with a multivariate approach which had, as within-subject factors, version (1st vs. 2nd), target category (animal vs. target) and match (match vs. mismatch).

For the mean voltages of ERPs within the 350–450 ms time window for the 1st version and within the 240–340 ms time window for the 2nd version, three repeated-measures ANOVAs with multivariate approaches were performed with the same
within-subjects factors and the electrodes factor. For the para-sagittal and the lateral montage, another within-subject factor, hemisphere (right vs. left), was added. The Geisser and Greenhouse (1959) procedure was used to compensate for the heterogeneity of variance of the electrodes factor which had more than two levels. In this case, the original degrees of freedom are reported together with the corrected probability level and the epsilon (E) correction factor.

The latency of the onset of the match effect in both versions was determined by using the statistical method proposed by Hagoort et al. (2004). The difference waveform (mismatch – match) was tested against zero in a series of running one-tailed t-tests. These tests were performed on the mean amplitudes in 25 ms bins, which shifted in time by 10 ms (e.g., 150–175; 160–185; 170–195 ms, etc.), in the latency range of 150–500 ms after the onset of the critical word. This series of running t-tests was performed at the midline central electrode (Cz).

Finally, to measure the onset of the N400 and N200 negative deflections themselves, similar analyses were done in the mismatch conditions, where the amplitudes of these negative waves were maximal. These analyses were done in each version on an average of animal and artifact targets. This time, the waveforms were tested against the mean voltage of the ERPs in a 200–225 ms time window, that is, a time window closely centered on the peak of the positive wave, the P2, that preceded the negative deflection under focus.

2.6 Independent component analysis and dipole source modeling

In an attempt to compare the ERP generators of the N400 match effect of version 1 to those of version 2, subtraction ERPs1 were submitted to an ICA and a dipole source modeling. These analyses were restricted to animal targets, that is, to the targets for which we found significant N400 effects (see Section 3).

ICA decomposes the unaveraged EEG (Jung et al., 2001) or the event-related potentials (Dien, Khoe, & Mangun, 2007; Makeig et al., 1999) into a sum of sparsely activated independent components with fixed scalp maps and maximally independent time course (Makeig, Jung, Bell, Ghahremani, & Sejnowski, 1997). For biologically plausible components, these component maps have been shown to nearly fit the projection of a single equivalent current dipole, allowing hypotheses as to the localization of this dipole (Onton & Makeig, 2006).

The ICA was conducted with EEGLAB 6.01b (Delorme & Makeig, 2004), a freely available open source toolbox (http://www.sccn.ucsd.edu/eeglab) running under Matlab 6.5 (The Mathworks). Individual subjects subtraction ERPs (mismatch-match) of 1 s (−200 to 800 ms) from our 30 active electrodes were concatenated and submitted to ICA. Decompositions of version 1 and version 2 ERPs were performed separately. We applied infomax ICA (Bell & Sejnowski, 1995) with the EEGlab function runica (Delorme & Makeig, 2004), which produced 30 independent components.

Source localizations were computed for all the 30 components with DIPFIT2, an EEGlab plug-in that performs component localization by fitting an equivalent current dipole model using a non-linear optimization technique (Scherg, 1990). The head model used for the analyses was the spherical model (BESA), that is, re-registered with the average Montreal Neurological Institute (MNI) brain. The brain structures where the dipoles were localized were identified using the Talairach atlas coordinates (Talairach & Tournoux, 1988). The location solution was restricted to the gray matter, within a search range of 3 mm. Components for which the scalp projections of the dipoles left more than 15% of the variance of the component scalp maps unexplained, were excluded, as well as components associated to dipoles that were outside of the head.

We then used the envtopo/ function of EEGlab to find the components that made the largest potential contribution to the ERP matching effect (Delorme & Makeig, 2004; Onton, Westerfield, Townsend, & Makeig, 2006). Two 150 ms time windows centered on the negative deflections observed in the subtractions were used: a 400–550 ms window for version 1 and a 235–385 ms window for version 2. Independent components, which, taken together, accounted for at least 80% of variance (PVAF) in these windows were selected in each version. These components were then clustered according to similarity in scalp maps, dipole locations and contribution to the ERPs using the K-means algorithm implemented in EEGlab (Onton et al., 2006).

3. Results

3.1 Behavioral data

The mean number of errors was six, both in version 1 and 2 (=5% of trials). Given the risk of floor effects, no further accuracy analyses were made.

The ANOVA conducted for RTs revealed a main effect of version \( (F(1,26) = 21.9, p < .001) \) showing that RTs were shorter in version 2 (1029 ms) than in version 1 (1165 ms). A main effect of semantic matching was found \( (F(1,26) = 54.41, p < .001) \) indicating, as expected, that mismatch trials were associated with longer reaction times (1133 ms) than match trials (1063 ms). Importantly, there was no significant interaction between match and version. Version 2 was thus associated with a match effect that was similar in amplitude to that of version 1. As expected, the effect of semantic matching was significantly greater for animal targets than for artifact targets \( (F(1,26) = 27.67, p < .001) \), with no version x target x match interaction. Post hoc analyses revealed that the match effect was generally significant for animal targets \( (F(1,26) = 164.06, p < .001) \) but not for artifacts targets.

A main effect of target category was also obtained \( (F(1,26) = 15.81, p < .001) \), which indicated that animal targets were categorized faster (1074 ms) than artifact targets (1122 ms). This effect was just significantly \( (F(1,26) = 4.69, p = 0.04) \) greater in the 2nd (995 ms vs. 1063 ms) than in the 1st version (1152 ms vs. 1182 ms) but was significant in both versions \( (F(1,26) = 7.5, p = .01) \) for version 1 and \( (F(1,26) = 13.73, p = .001) \) for version 2. Given that, as mentioned, the target category interacted with match, post hoc analyses were run. They revealed that, in the match conditions, RTs were longer for artifact targets (1126 ms) than for animal targets (1002 ms) \( (F(1,26) = 28.19, p < .001) \). In the mismatch conditions, there was a version x target interaction \( (F(1,26) = 6.46, p = .017) \). Further post hoc analyses showed that, in mismatch trials, RTs for artifact targets were longer than for animal targets in version 1 \( (F(1,26) = 10.34, p = .003) \) but not in version 2.

3.2 Electrophysiological data

3.2.1 General description

Fig. 1 represents the grand averages of the ERPs elicited by the target words in the two versions of the task, separating animal targets (at the top) from artifacts targets (at the bottom). Negative is plotted upwards in all figures. After a positive wave peaking around 200 ms post-onset, these ERPs include a negative deflection, whose onset in version 2 (continuous lines) appears to coincide with its onset in version 1 (dashed lines). For animal targets ERPs, this negative deflection also appears larger for mismatches (in blue) than for matches (in red) not only in the 1st but also in the 2nd version. For these mismatches, the negative deflection reaches its maximum earlier in the 2nd (around 290 ms at Cz) than in the 1st version (about 400 ms at Cz). The positive wave that follows the negative deflection appears to have, consequently, an earlier occurrence in the 2nd version of the task. In the mismatch condition at Cz, this late positive wave peaks around 675 ms in version 1 and around 400 ms in version 2. In Fig. 2, which includes the mismatch–match ERP subtractions, the match effect appears to be substantially earlier in the 2nd than in the 1st version. The effect peaks around 480 ms for animal targets (top of Fig. 2) and 400 ms for artifact targets (bottom of Fig. 2) in version 1 at Cz electrode. In version 2, it peaks around 300 ms for animal targets and at about 335 ms for artifact targets.

3.2.2 Peak latency analyses

Statistical analyses of the peak latencies of the post-P200 negativities confirmed that these deflections occurred significantly earlier in version 1 than in version 2 \( (F(1,26) = 78.49, p < .001) \).

3.2.3 Mean voltage amplitudes

Statistical analyses of the mean voltage amplitudes revealed a main effect of version for all montages \( (F(1,26) = 75.71, p < .001) \) for the sagittal; \( F(1,26) = 45.01, p < .001 \) for the para-
Fig. 1. Grand average ERPs (n = 27) to target words in the two versions of the semantic categorization task. Negativity is up. Animal targets (top) refer to the trials in which different animal exemplars were used as targets in the 1st version (dashed lines), and to the trials in which the corresponding category word, that is, the word ‘vivant’ (‘living’ in English), was used repeatedly as a target in the 2nd version (continuous lines). Artifact targets (bottom) refer to the trials in which different artifacts exemplars were used as targets in the 1st version, and to the trials in which the corresponding category word, ‘chose’ (‘thing’ in English), was repeatedly used as a target in the 2nd version.

A main effect of match was found on all montages (F(1,26) = 9.55, p = .005 for the sagittal; F(1,26) = 6.30, p = .019 for the para-sagittal, and F(1,26) = 4.26, p = .049 for the lateral montage). Most importantly, there was no interaction between match and version, nor any interaction between these 2 factors and electrodes or hemisphere, showing that the match effect was similar in both versions of the task and did not differ in its scalp distribution. As expected, the effect of match was significantly more important for animal than for artifact targets (as illustrated by Fig. 1) at the sagittal (F(1,26) = 6.34, p = .018) and the para-sagittal montage (F(1,26) = 4.46, p = .044). This difference was similar in both versions as no multiple interaction implying this factor was found.

Analyses restricted to animal targets (Fig. 1, top) showed that the match effect was significant at all montages (F(1,26) = 16.99, p < .001 for the sagittal; F(1,26) = 9.95, p = .004 for the para-sagittal and F(1,26) = 4.49, p = .044 for the lateral montage). There was no interaction between match and version or between these factors and electrodes or laterality (see Fig. 3A for the voltage maps of version 1 and 2).
Analyses restricted to artifact targets (Fig. 1, bottom) showed an absence of main effect of match, except for a trend at the sagittal montage ($F(1,26) = 3.65, p = .067$). For this montage, there was a significant interaction between version, match and electrodes ($F(3,78) = 4.27, E = 0.66, p = .02$) illustrating that the match effect was somewhat more important in the 2nd than in the 1st version and more so at centro-parietal electrodes than at frontal sites. Post hoc analyses for the sagittal montage confirmed that the match effect was never significant for artifact targets in the 1st version. It only marginally reached significance at the midline central electrode (Cz) in the 2nd version ($F(1,26) = 4.55, p = .043$) while there was a trend at Pz ($F(1,26) = 3.21, p = .085$). For the lateral montage, a significant interaction between match and hemisphere was found ($F(1,26) = 9.48, p = .005$) illustrating that the match effect was more pronounced on the right than on the left side (see Fig. 1). Post hoc analyses for the lateral montage showed that the match effect was significant over the right hemisphere ($F(1,26) = 8.38, p = .008$), and tended to be greater in the 2nd version than in the 1st ($F(1,26) = 3.02, p = .094$).

A main effect of target category was also obtained on all montages ($F(1,26) = 22.29, p < .001$ for the sagittal; $F(1,26) = 8.12, p = .008$ for the para-sagittal, and $F(1,26) = 9.66, p = .005$ for the lateral montage). Most importantly, this effect was similar in both versions as there was no interaction between target category and version or between these factors and the electrodes or hemisphere factors. There was however, as previously mentioned, a target category × match interaction on the sagittal ($F(1,26) = 6.34, p = .018$), and the para-sagittal montage ($F(1,26) = 4.46, p = .044$). Post hoc
analyses for the match conditions showed that ERPs to artifact targets were more negative than those to animal targets on the sagittal ($F(1,26) = 20.21$, $p < .001$) and on the para-sagittal montage ($F(1,26) = 9.06$, $p = .006$). In the mismatch conditions, target ERPs differed only in interaction with the electrodes factor ($F(3,78) = 4.77$, $E = 0.89$, $p = .006$ for the sagittal and $F(6,156) = 4.93$, $E = 0.32$, $p = .012$ for the para-sagittal montage). Further analyses for the sagittal montage showed that ERPs to artifact targets were significantly more negative than those to animal targets at Cz ($F(1,26) = 8.22$, $p = .008$) and tended to be so at Pz ($F(1,26) = 1.36$, $p = .068$). For the para-sagittal montage, ERPs to artifact targets were more negative than those to animal targets at C4 ($F(1,26) = 8.3$, $p = .008$), and P4 ($F(1,26) = 6.29$, $p = .019$). In contrast, at prefrontal sites, artifact tended to be more positive than animal targets ($F(1,26) = 3$, $p = .09$). This effect just missed significance at Fp2 ($F(1,26) = 3.52$, $p = .072$).

Fig. 3B shows the scalp voltage maps that were obtained by subtracting ERPs to animal targets from those to artifact targets in version 1 and version 2 in mismatch conditions (since the effect of category is usually studied in unprimed targets). Compared to ERPs to animal targets, ERPs to artifact targets were generally more negative over centro-parietal sites and more positive over frontal sites. This effect was slightly and non-significantly more focused on central sites in version 2 than in version 1, likely due to the use of only one repeated category in the former case and multiple target words in the latter.

Finally, to investigate if the match effect varied with the number of repetitions, we compared this effect across the number of presentations in version 2. These analyses were restrained to the animal targets, that is to the targets for which we found significant N400 matching effects. Mean voltage amplitudes for the first 30 presentations (first 15 match trials and first 15 mismatch trials) were compared to those of the last 30 presentations (last 15 match trials and last 15 mismatch trials). The analysis run for each montage revealed no interaction between repetition and match. There was a triple interaction of match $\times$ number of repetitions $\times$ hemisphere at the para-sagittal ($F(1,26) = 11.57$, $p = .003$) and at the lateral montage ($F(1,26) = 6.92$, $p = .015$). Post hoc analyses made for each hemisphere revealed that the repetition $\times$ match interaction was never significant.

3.2.4. Analyses on onset latencies

Statistical analyses were made on the latency of the onset of the match effect (mismatch–match conditions, Fig. 2). These analyses made at Cz showed that, for animal targets, this effect started between 250 and 275 ms after target onset in version 2 ($t = 2.26$; $p = .02$), whereas in version 1, it started between 410 and 435 ms ($t = 1.82$; $p = .04$). For artifact targets, in version 2, the match effect started between 290 and 315 ms ($t = 2.18$; $p = .04$), whereas, in version 1, no significant onset was found.
The superposition of the ERPs from version 1 to those from the version 2 on Fig. 1 suggests that the negative ERP deflections themselves onset approximately at the same time. Analyses made with an average of the ERPs to animal and artifact targets in mismatch conditions, where these deflections were the largest, confirmed this observation. 

The effect of the semantic category of target words was similar in both versions. Compared to ERPs evoked by animal targets, ERPs evoked by artifact targets were more negative over centro-parietal sites and tended to be more positive over frontal sites. These scalp distributions are consistent with the results of a number of studies (Antal et al., 2000; Kiefer, 2001; Paz-Caballero et al., 2006; Proverbio et al., 2007). Meanwhile, in both versions, RTs to artifact targets were longer than those to animal targets, with the exception of mismatching targets in version 2 where no difference was observed. Nevertheless, even in this condition, the scalp distribution of ERPs depended on category. Longer RTs to artifact targets compared to animal targets have been frequently described (Kiefer, 2001; Paz-Caballero et al., 2006; Price & Humphreys, 1989; Proverbio et al., 2007). They are sometimes interpreted as reflecting an easier categorization of names of natural categories because of a greater overlap in semantic features (e.g., Saffran & Scholl, 1999). Here, the use of a cohesive category (‘living’) associated with names of animals and a non-cohesive category (‘thing’) associated with various sets of artifacts (i.e., kitchen/desk objects, clothes, tools, furniture, buildings and transportation means) was also accompanied by smaller semantic matching effects for the artifact category in both versions. This manipulation allowed us to compare two different levels of priming across repetition conditions. In both versions, the match effect on N400s and RTs was significant for animal targets but not for artifact targets, with the exception of Cz electrode where the match effect was marginally significant for artifact targets in the second version. As this difference between versions was restricted to one electrode site and not supported by reaction time data, no further attempt to interpret it was made.

Our finding of a maintenance of the effect of category over numerous repetitions is in accordance with the observation that the distinction between living and man-made items is the most common and robust category dichotomy reported in the literature (Devlin et al., 2002). However, it contrasts with the marked reduction of this effect with one repetition reported by Sim and Kiefer (2005) and Kiefer (2005). Category-related modulations of brain activity have more often been observed in tasks relying on explicit semantic categorization (Devlin et al., 2002). It is thus possible that the use of implicit tasks in these studies was responsible for the suppression of the effect with repetition. Accordingly, category effects

3.2.5. ICA and dipole source modeling

The ICA and the dipole source modeling were limited to the animal targets for which match effects were significant. In version 1, 12 independent components (ICs) accounted for 83% of the variance of the match effect in the N400 time window, while in version 2, 8 ICs accounted for 81% of the effect. These ICs were then clustered according to similarity in scalp maps, dipoles, and contribution to the ERPs, using the K-means algorithm implemented in EEGLAB. We chose six as the initial number of clusters for version 1 and four for version 2, that is half the number of ICs in each version. After inspection of the initial clustering results, the number of clusters was adjusted progressively until the resulting clusters were stable.

Fig. 4 shows dipole locations for all clusters and their mean scalp maps. In both versions, the match effect was associated with a first cluster that projected most strongly to the posterior scalp at occipito-temporal locations, with a clear left-sided asymmetry in scalp maps. In both versions, the match effect was associated with cluster 1 distribution on the other side of the scalp.

Dipole source modeling revealed two sources that were common to version 1 and 2. In version 1, the first cluster grouped five ICs and was located in the anterior part of the left parahippocampal gyrus (Brodmann Area: BA 28) (x: −21, y: −21, z: −6). The residual variance of the dipole model (RVDM) was 11%. The second cluster, grouping three ICs, was located in the right medial frontal gyrus, in the vicinity of the anterior frontal–polar cortex (BA 10) (x: 17, y: 48, z: 10). The RVDM for this solution was 10%. The third cluster grouped four ICs and had a generator in the left medial frontal gyrus (BA 10 and 9) (x: −17, y: 42, z: 14). The RVDM was 9%.

In version 2, cluster 1 grouped four ICs and had a source in the anterior part of the left parahippocampal gyrus (BA 28) (x: −15, y: −10, z: −10). The RVDM for this solution was 8%. The second cluster grouped four ICs and was located in the right medial frontal gyrus, near the anterior frontal–polar cortex (BA 10) and the right anterior cingulate cortex (BA 32) (x: 9, y: 45, z: 6). The RVDM for this solution was 10%.

4. Discussion

The present study aimed at investigating the impact of massive repetition of target words on the effects of semantic matching and category, as assessed by N400s and RTs. To force semantic processing, subjects were given an explicit semantic categorization task: they had to decide whether target words belonged to the same category as prime words. The first version of the task was a ‘classic’ design in which two category words were used as primes and numerous non-repeated exemplars of these two categories as targets. In the second version of the task, primes and targets were inverted: the primes were the numerous non-repeated category exemplars and the target words, to which N400s were recorded, were simply the two category words, each presented 60 times. Importantly, as categories and exemplars can be associated with different N400 raw amplitudes (Kounios & Holcomb, 1992), only the persistence of N400 effects with repetition was analyzed. Note that, in any case, Kounios and Holcomb (1992) showed that the effect of category level does not interact with semantic congruity or repetition.

Both versions were characterized by effects of semantic matching and of semantic category on N400s and RTs. Such effects have, to our knowledge, never been found on an N400-like potential elicited by content words in conditions of massive repetition. The fact that the topography of their distributions on the scalp in version 2 was similar to that of version 1 strongly supports the idea that they are genuine N400 effects obtained with concrete words. This conclusion is reinforced by the finding that the match effect seemed to involve common brain regions in both versions: the anterior part of the parahippocampal gyrus and the medial frontal gyrus.

Behavioral and electrophysiological results are therefore consistent with the idea that, at least in certain circumstances, RT and N400 effects can be found in massive repetition conditions and thus that the semantic processes these measures index can persist in these conditions.

The effect of the semantic category of target words was similar in both versions. Compared to ERPs evoked by animal targets, ERPs evoked by artifact targets were more negative over centro-parietal sites and tended to be more positive over frontal sites. These scalp distributions are consistent with the results of a number of studies (Antal et al., 2000; Kiefer, 2001; Paz-Caballero et al., 2006; Proverbio et al., 2007). Meanwhile, in both versions, RTs to artifact targets were longer than those to animal targets, with the exception of mismatching targets in version 2 where no difference was observed. Nevertheless, even in this condition, the scalp distribution of ERPs depended on category. Longer RTs to artifact targets compared to animal targets have been frequently described (Kiefer, 2001; Paz-Caballero et al., 2006; Price & Humphreys, 1989; Proverbio et al., 2007). They are sometimes interpreted as reflecting an easier categorization of names of natural categories because of a greater overlap in semantic features (e.g., Saffran & Scholl, 1999). Here, the use of a cohesive category (‘living’) associated with names of animals and a non-cohesive category (‘thing’) associated with various sets of artifacts (i.e., kitchen/desk objects, clothes, tools, furniture, buildings and transportation means) was also accompanied by smaller semantic matching effects for the artifact category in both versions. This manipulation allowed us to compare two different levels of priming across repetition conditions. In both versions, the match effect on N400s and RTs was significant for animal targets but not for artifact targets, with the exception of Cz electrode where the match effect was marginally significant for artifact targets in the second version. As this difference between versions was restricted to one electrode site and not supported by reaction time data, no further attempt to interpret it was made.

Our finding of a maintenance of the effect of category over numerous repetitions is in accordance with the observation that the distinction between living and man-made items is the most common and robust category dichotomy reported in the literature (Devlin et al., 2002). However, it contrasts with the marked reduction of this effect with one repetition reported by Sim and Kiefer (2005) and Kiefer (2005). Category-related modulations of brain activity have more often been observed in tasks relying on explicit semantic categorization (Devlin et al., 2002). It is thus possible that the use of implicit tasks in these studies was responsible for the suppression of the effect with repetition. Accordingly, category effects

2 The scalp distribution of these N400 effects are known to be more anterior than the scalp distribution of the N400 effects obtained with abstract words (see for instance Kiefer, 2001, 2005; West & Holcomb, 2000).
may resist repetition only when category information is required to perform the task. Such a phenomenon was observed by Kounios and Holcomb (1994) for the effect of concreteness: the N400 difference between concrete and abstract words was eliminated by repetition in a lexical decision task but simply reduced in a concrete–abstract classification task.

While the conditions of massive repetition in version 2 resulted in shorter RTs than in version 1, these RTs were still much longer than those obtained in simple choice reaction time tasks using words as stimuli (e.g., Kutas, McCarthy, & Donchin, 1977). Together with the differential effect of semantic matching and the effect of category, this strongly suggests that the processing of the two repeated target words of the second version was not limited to a simple discrimination based on the physical features that differentiated them. This processing included a semantic stage.

In both versions, the localization technique used led us to associate the match effect with activity in the medial prefrontal cortex (Brodmann Area: BA 10) and in the anterior part of the parahippocampal gyrus (BA 28). Intracranial recordings (Grunwald, Lehnertz, Heinze, Helmstaedter, & Elger, 1998; Nobre & McCarthy, 1995; Smith, Stapleton, & Halgren, 1986) and high-density ERP recordings (Johnson & Hamm, 2000) suggest that one of the most prominent generators of the scalp N400 is in the anterior medial temporal lobe (AMTL), close to the collateraleral sulcus which separates the parahippocampal gyrus from the fusiform gyrus. This region, which includes BA 28, would have a critical role in accessing semantic information (Van Petten & Luka, 2006).

The involvement of the anterior part of the parahippocampal gyrus in the match effect, independent of repetition, may be related to the response of neurons in homologous areas in awake monkeys. These neurons, located in the entorhinal and perirhinal cortices, have been reported to reach a plateau in activity after one repetition (Xiang & Brown, 1998). Our results could thus be used to suggest that the AMTL generator is still modulated by semantic matching even when its activity has been decreased by repetition.

The other generator obtained for the N400 match effect in both versions was located in the anterior medial frontal-polar cortex (AMFPC), in the vicinity of BA 10. Using magnetoencephalography and a distributed source modeling method, Halgren et al. (2002) found numerous N400 generators in the temporal and frontal lobes, including frontal-polar regions. Intracranial studies in humans have also described generators in several prefrontal areas, with sources located slightly posterior to the present findings in the anterior cingulate gyrus (BA 32) and near areas BA 45 and 46 (Guillem, Rougier, & Claverie, 1999; Halgren et al., 1994). Frontal-polar regions (BA 10) have been found to be part of a network of common activations in working memory, semantic memory and episodic-memory tasks (Nyberg et al., 2003; see also MacLeod, Buckner, Miezin, Petersen,
& Raichle, 1998). They have been notably implicated in tasks that require that information be maintained on line while exploring and processing other items (Koechlin, Basso, Pietrini, Panzer, & Grafman, 1999; Koechlin, Corrado, Pietrini, & Grafman, 2000). The medial part of BA 10, obtained in the present study, along with the adjacent anterior cingulate cortex (BA 32), have more specifically been found to be active in conditions where the task sequence triggers precise expectations (Koechlin et al., 2000). It is thus possible that this generator was related to the present task conditions which required, in both versions, that subjects held the identity of the prime and compared its meaning to that of the target. The reason why version 1 was associated with a bilateral activation while version 2 was associated with a unilateral activation is unclear, but could be due to the greater retrieval demands of version 1 where multiple non-repeated targets were presented.

The similarity in the effects of semantic matching on N400-like ERPs across the two versions of the task contrasts with the results of Besson et al. (1992), who showed a suppression of these effects after two repetitions. As mentioned, this discrepancy is likely due to the repetition of complete sentences used in that study, making it likely that final words, to which N400s were recorded, could be anticipated (see also Olichney et al., 2000).

While the peak latency of the N400 deflection of the first version is comparable to that of most N400 studies, the peak latency of the N290 deflection in the second version is reminiscent of that found by Simon, Bernard, Largy, Lalonde, & Rebai (2004). These authors used a group of three words with massive repetition (100) during a lexical decision task. They obtained a negativity peaking at 280 ms. The authors asked if their ‘N280s’ were N400 deflections that peaked earlier. Our results support this hypothesis not only because we found semantic effects on the N290 but also because our N290 and N400 deflections had the same onset latency. Nevertheless, further experiments will be necessary to test if it is actually the massive repetition of target words that directly influences the peak latency of the N400 and the latency of the onset of N400 effects. Indeed, massive repetition was accompanied by a reduced number of different targets in our protocol. This reduced number could thus be responsible for the earlier N400 deflections and the earlier N400 effects. Results of a recognition memory task by Van Strien et al. (2005) are compatible with this hypothesis, as the authors obtained a classic N400 peak latency using 30 target words presented 10 times each.

It appears unlikely that additional repetitions would produce further effects on the ERPs within the N400 time window. However, the possibility remains that while being reduced after a few repetitions, N400 effects “rebound” after massive repetition. In an attempt to look at this possibility in our data, we compared the amplitude of the N400s evoked by the first 30 presentations of our target words to those evoked by the last 30 presentations. This was made in version 2 for animal targets for which significant match effects were observed. However, we found no interaction between the number of repetitions and the match effect.

Together, the similarity of the match and semantic category effects between task versions suggests similar semantic processing and the involvement of the same ERP component. Despite the high level of repetition in version 2, it thus seems that semantic processing was still effective and, at least partly, similar to that occurring in non-repeated conditions. Future investigations should also determine if the robustness of the N400 semantic processing with repetition may be generalized to different experimental situations, notably when this processing is implicit, as in lexical decision tasks.

This resistance of semantic processes to numerous repetitions supports the idea that the processes indexed by RT and N400 could occur and thus be measured for one repeated word. Assessing semantic processing for single repeated words allows the study of semantic categories that subsume just a few exemplars, or just a few homogeneous exemplars. Moreover, with massive repetition, certain psycholinguistic properties of words, such as the frequency of their usage, are likely to have little or no impact. Indeed, while both repetition and higher frequency of usage reduce N400 amplitudes, previous studies using one to three repetitions (Smith & Halgren, 1987; Young & Rugg, 1992) showed that the repetition effect is larger for rare than for frequent words, resulting in no, or in much smaller, effects of frequency on N400s as soon as the third presentation. Note that, even if a small effect of frequency on raw N400 amplitudes still existed in massively repeated conditions, it should not have an impact on the N400 effects. The use of massive repetition should thus allow the comparison of the processing of words whose frequencies of usage differ. More generally, it provides us with the possibility of studying semantic associations between two meaningful stimuli, and therefore the link between two specific concepts. On the other hand, given the robustness of the semantic matching effect in version 2, it seems that these links could be studied in individual subjects. This could be of interest for case studies in psychiatry, for instance in the exploration of the neurocognitive mechanisms of delusion (Debruille et al., 2007; Neagoe, 2000; Silva & Leong, 1994).

Acknowledgements

This study was supported by the grant 194517-02 from the National Sciences and Engineering Research Council of Canada (NSERC) allocated to the first author who is supported by the scholarship 10084 from the Fonds de la Recherche en Santé du Québec (FRSQ). Louis Renoult is supported by the fellowship 13542 from the FRSQ.

References


