Brain & Language 125 (2013) 47–53

Contents lists available at SciVerse ScienceDirect

Brain & Language

journal homepage: www.elsevier.com/locate/b&l

# Concreteness in word processing: ERP and behavioral effects in a lexical decision task

Horacio A. Barber [a,*], Leun J. Otten [b], Stavroula-Thaleia Kousta [c], Gabriella Vigliocco [c]

[a] Department of Cognitive Psychology, University of La Laguna, Spain
[b] Institute of Cognitive Neuroscience, University College London, UK
[c] Department of Cognitive, Perceptual, and Brain Sciences, University College London, UK

ARTICLE INFO

*Article history:*
Accepted 14 January 2013

*Keywords:*
Lexical representation
Word recognition
Concreteness
Imageability
N400–N700

ABSTRACT

Relative to abstract words, concrete words typically elicit faster response times and larger N400 and N700 event-related potential (ERP) brain responses. These effects have been interpreted as reflecting the denser links to associated semantic information of concrete words and their recruitment of visual imagery processes. Here, we examined whether there are ERP differences between concrete and abstract stimuli controlled for a large number of factors including context availability (i.e., richness of semantic associations) and imageability. We found that abstract words elicited faster behavioral responses but that concrete words still elicited larger N400 and N700 responses. We propose that once all other factors, including imageability and context availability are controlled, abstract words may trigger a larger number of superficial linguistic associations that can be quickly used for response decisions. The ERP differences, however, would index the greater semantic processing (integration of multimodal information) for concrete than abstract words during meaning activation.



## 1. Introduction

Concrete words are usually responded to faster than abstract words (see Schwanenflugel, 1991 for an early review) and this behavioral advantage is also mapped into differences in electrophysiological responses. In particular, it has been found that event-related brain potentials (ERPs) for concrete words show a long-lasting negativity compared to ERPs for abstract words (Holcomb, Kounios, Anderson, & West, 1999; Huang, Lee, & Federmeier, 2010; Kanske & Kotz, 2007; Kellenbach, Wijers, Hovius, Mulder, & Mulder, 2002; Kounios & Holcomb, 1994; West & Holcomb, 2000). The initial part of this negativity, between 300 and 500 ms after word onset, has been related to the N400 component. The N400 is a negative deflection peaking at around 400 ms, associated with the processing of potentially meaningful stimuli. In sentence reading, N400 amplitude correlates with how well a word fits with a prior context; the better the word fits, the smaller the N400 amplitude (see review in Kutas & Federmeier, 2011). Moreover, studies using single words suggest that the N400 is associated with the difficulty in identifying a string of letters: the more difficult the recognition process, the larger the N400 amplitude (Barber & Kutas, 2007). However, in contrast to the centroparietal distribution of the classical N400 effect for written words, the N400 concreteness effect usually extends to frontal electrodes, although its specific scalp topography varies across studies (see review in Adorni & Proverbio, 2012). Therefore it is a matter of debate whether the N400 concreteness effect shares exactly the same neural generators as the typical N400 component (Huang et al., 2010). Additionally, the larger negativity to concrete than abstract words usually persists beyond the standard N400 time window. The later segment, between 500 and 1000 ms, has been labeled N700, and is largest over frontal scalp sites. West and Holcomb (2000) obtained ERPs to concrete and abstract words in two experimental tasks: a semantic judgment task and a task designed to evoke mental images. While concreteness effects were larger in the N400 time window in the semantic task, in the imagery task these effects were larger in the N700 time window at especially frontal electrodes. In light of these results, the authors proposed that the N700 reflects the activation of an image-based processing mechanism (see also Nittono, Suehiro, & Hori, 2002). A recent study tested this hypothesis by using an explicit mental-image generation task versus a word association task (Welcome, Paivio, McRae, & Joanisse, 2011). Although the N400 concreteness effect was only found in the more linguistic word association task, the comparison between concrete and abstract words in the mental imagery task resulted in a difference at late latencies (756–1500 ms) and with the opposite polarity than the previously reported N700 effect. These results therefore do not provide strong support for the

* Corresponding author. Address: Departamento de Psicología Cognitiva, Universidad de La Laguna, Campus de Guajara s/n, La Laguna, Tenerife 38205, Spain.
*E-mail address:* hbarber@ull.es (H.A. Barber).

0093-934X/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.bandl.2013.01.005

proposal that the N700 reflects imagery processes. More recently, Adorni and Proverbio (2012) used source localization analyses to associate the late concreteness effect with the ventrolateral prefrontal cortex, an area functionally related to top-down control of semantic memory.

The pattern of ERPs associated with the faster responding to concrete than abstract words has been accounted for by a ''context-extended dual coding hypothesis'' (Holcomb et al., 1999). This account integrates two of the dominant cognitive theories concerning differences between concrete and abstract words: the dual-coding theory (Paivio, 1986) and the context-availability hypothesis (Schwanenflugel, 1991). Dual coding argues that concrete words are represented in a verbal as well as non-verbal code. Abstract words, in contrast, are only represented in a verbal code. According to the context-availability hypothesis, differences between concrete and abstract words would instead arise in a single verbal-semantic store because of stronger and denser associative links for concrete than abstract words. Whereas either the dual-coding or context-availability theory can account for the reaction time (RT) advantage for concrete over abstract words, they cannot also account for the ERP results. Thus, their integration is proposed by the context-extended dual coding hypothesis (Holcomb et al., 1999). According to this account, the larger N400 for concrete words represents post-lexical processing efforts associated with integrating denser associative networks, in line with the context-availability hypothesis. Concrete words would also benefit from activation of visual imagery at a later stage, as indexed by the N700 differences, in line with the dual coding view. The fact that the stimuli that elicit the shorter RTs (concrete words) also elicit larger N400s (rather than smaller as one would expect from the literature) can be accounted for by proposing that the behavioral advantage for concrete words arises because greater semantic integration demands are used strategically as an index of lexicality in lexical decisions. However, Kounios and colleagues (Kounios et al., 2009) found larger N400 amplitudes for concrete words with a lower number of associates (i.e. lower semantic richness) relative to those with a higher number of associates and abstract words. Similarly, a recent study that independently manipulated concreteness and semantic richness (number of listed semantic features) reported larger N400 amplitudes for concrete than abstracts words, but smaller N400 amplitudes for words associated with richer than poorer semantic concepts (Amsel & Cree, in press; but see Rabovsky, Sommer, & Abdel-Rahman, 2012 for the opposite result). These studies contrast with predictions from the context-extended dual coding hypothesis, thus leaving open the possibility that the ERP correlates of the concreteness effect may be linked to different cognitive processes.

It is critical to note that comparisons between concrete and abstract words involve a contrast between items that can differ along other lexical and sublexical dimensions in addition to concreteness. First, the ''context-extended dual coding'' hypothesis assumes that differences in concreteness come about because of differences in other dimensions, namely differences in context availability (which can be operationalised in terms of ratings of how many contexts one can think of for a given word, Schwanenflugel, 1991) and differences in imageability between concrete and abstract words. Hence, if these dimensions are controlled, no RT and ERP differences between the words would be expected. Second, it is also the case that most often studies did not control for additional lexical variables. For instance, concrete and abstract words matched for written frequency can still greatly differ in familiarity to the advantage of concrete words (e.g., *artichoke* vs *heresy*), and familiarity has been shown to modulate the behavioral concreteness effect (Levy-Drori & Henik, 2006). In the present study we matched items for a large number of lexical and sub-lexical variables including familiarity and, critically, both imageability and context availability. In the literature, it is invariably assumed that the psycholinguistic constructs of concreteness and imageability tap into the same underlying theoretical construct. Thus, concreteness and imageability ratings have been used interchangeably in most of the recent literature in the field (e.g. Binder, Westbury, McKiernan, Possing, & Medler, 2005; Fliessbach, Weis, Klaver, Elger, & Weber, 2006; Giesbrecht, Camblin, & Swaab, 2004; Richardson, 2003). However, concreteness and imageability tap into, at least partially, different aspects of semantic representations. Most of the concrete words are considered imageable, whereas abstract words show higher variability in imageability ratings. Moreover, Kousta, Vigliocco, Vinson, Andrews, and Del Campo (2011) analyzed ratings for more than 4000 words in the MRC Psycholinguistic Database and showed that the frequency distribution of concreteness ratings is bimodal (with two distinct modes for abstract and concrete words), capturing the categorical ontological distinction between concrete (spatiotemporally-bound concepts) and abstract concepts (non-spatiotemporally bound concepts). On the contrary, the distribution of imageability ratings is unimodal, indexing the graded amounts of sensory (primarily visual) information associated to the words.

Kousta et al. (2011) have further shown that when all known lexical and sublexical dimensions (including familiarity, imageability and context-availability) are controlled, or partialled out in regression analyses, abstract words elicit faster RTs than concrete words in lexical decision tasks, resulting therefore in an *abstractness effect* (Kousta et al., 2011). In the experiments and regression analyses by Kousta et al. (2011) it was found that emotional valence differed between concrete and abstract words, with abstract words being more valenced than concrete ones, a difference that could account for the RT advantage for abstract words reported. However, it is important to note that ERPs and behavioral concreteness effects are not necessarily related to the same cognitive processes. In fact, Holcomb et al. (1999) proposed the opposite on the basis of the paradoxical larger N400 amplitudes associated with concrete words. Therefore, it is an open question if a standard N400–N700 concreteness effect can be found after controlling for imageability and context-availability. If so, concreteness ERP effects would be difficult to account for in terms of the context-extended dual coding hypothesis (Holcomb et al., 1999). Alternatively, concreteness ERP effects could be related to the greater amount of multimodal information activated for concrete than abstract words. While the activation of concrete words implies the activation of sensory and motor representations associated with their meanings, abstract words (especially when presented out of context) would only activate superficial asociations with related concepts. The differences in type of activated information could also modulate the control and integration mechanisms associated with lexico-semantic retrieval (as proposed by Adorni & Proverbio, 2012). Differences between concrete and abstract words would be due more to the type of information that is activated in a single lexico-semantic system than to the number of contextual asociations or mental images triggered by them.

Here, we contrast concrete and abstract items that do not differ along a large number of lexical and sublexical factors, including imageability, context availability and also emotional valence to establish whether we can find a behavioral *abstractness* effect. Additionally, and considering that electrophysiological and behavioral measures can reflect different cognitive processes, we assess ERP differences in order to test the context-extended dual hypothesis. This hypothesis would predict that we should not observe ERP differences between concrete and abstract words given that items are matched for both context-availability and imageability. Thus, if we were to find ERP differences between concrete and abstract words we would have to conclude that the context-extended dual coding hypothesis does not provide a general account for EEG correlates of concreteness effects. Concrete and abstract words, however, can differ with regards to the amount of multimodal information that they activate when presented out of context (greater for concrete than abstract words). ERP effects could therefore index genuine differences in the level of meaning activation. Thus, we might expect to observe a greater N400 for concrete than abstract words *regardless* of whether the abstract words lead to slower or faster RTs than concrete words. Moreover, N700 effects after controlling for imagebility cannot easily be explained only by mental imagery processes. This effect would instead be more consistent with executive control mechanisms that operate during meaning activation and working memory operations.

## 2. Method

### *2.1. Participants*

Eighteen right-handed native English speakers (9 women; mean of age 24 years, range 18–33) participated voluntarily in exchange for payment (£7.50/h). Handedness was assessed using an abridged version of the Edinburgh Inventory (Oldfield, 1971). All participants gave informed consent and reported no neurologic or psychiatric history. Three participants were excluded from the analyses due to a high number of artefacts in their data (more than 50% of the relevant observations). The experimental procedures were approved by the joint University College London and University College London Hospitals ethics committee.

### *2.2. Materials*

Sixty abstract (e.g. *number*, *quest*, *joke*, *crime*, *fashion*) and 60 concrete (e.g. *weapon*, *lobby*, *belt*, *guest*, *material*) words were selected from a number of databases (see Table 1). Concreteness values were taken from the MRC Psycholinguistic Database (Coltheart, 1981), which uses norms from Gilhooly and Logie (1980). Concreteness ratings were obtained on a 7-point scale; words referring to objects, materials, or persons were to receive a high concreteness rating, and words referring to abstract concepts that could not be experienced by the senses were to receive a low concreteness rating. Words lists differed significantly in concreteness but were matched pairwise for context availability, imageability and a number of other (sub)lexical variables. For grammatical class, we made sure that for each pair, the words shared the same part of speech as defined in the MRC database (Coltheart, 1981). For context availability, we obtained ratings by asking 47 native English speakers to rate words on a 7-point Likert scale according to how easy it is to come up with a particular context or circumstance in which a word might appear. The instructions to participants were identical to those used by Schwanenflugel and Stowe (1989). We obtained norms for 650 words, with each word rated by 22 or 25 speakers[1]. For imageability, we used norms taken from Wilson (1988) and Stadthagen-Gonzalez and Davis (2006). These norms indicate the ease with which words induce mental images. High imagery ratings are given to words that quickly and easily provoke mental images, and low ratings to words that provoke mental images with difficulty or not at all. Finally, concrete and abstract words did not differ in terms of emotional valence ($p = .11$). Valence values were taken from Bradley and Lang (1999) and Kousta et al. (2011).

We used the same databases to create 120 pseudowords. Sixty concrete and 60 abstract words were selected and matched pairwise with the experimental items for concreteness, length, and mean positional bigram frequency. A single letter was then altered from each matched word. All but 11 pseudowords had one orthographic neighbor (i.e., the intended concrete or abstract word). No words or pseudowords were repeated in the experiment.

**Table 1**
Mean values for lexical and sublexical characteristics of stimuli.

| | Abstract | Concrete |
|---|---|---|
| Concreteness[a] | 341 | 547 |
| Imageability[a] | 494 | 495 |
| Context availability[b] | 556 | 558 |
| Emotional valence[e] | 5.58 | 5.11 |
| Age of acquisition[f] | 391 | 419 |
| Number of synsets[c] | 4.70 | 5.46 |
| Familiarity[a] | 494 | 492 |
| Log frequency[d] | 8.59 | 8.81 |
| Orthographic neighborhood density[d] | 3.33 | 3.54 |
| Number of letters | 5.95 | 5.98 |
| Number of syllables[d] | 1.83 | 1.83 |
| Number of morphemes[d] | 1.16 | 1.08 |
| Mean positional bigram frequency[d] | 1657 | 1635 |

[a] MRC psycholinguistic database (Coltheart, 1981).
[b] Kousta et al., 2011.
[c] WordNet (Fellbaum, 1998).
[d] The english lexicon project (Balota et al., 2007).
[e] Bradley and Lang (1999), and Kousta et al. (2011), scale from 0 to 10 (5 = neutral).
[f] The bristol norms (Stadthagen-Gonzalez & Davis, 2006).

### *2.3. Procedure*

Stimuli were presented at eye level on a monitor at a distance of approximately 70 cm from the participant. Strings appeared in a black Arial font (size 24) in a gray rectangle ($183 \times 97$ pixels) against a black background. The sequence of events on each trial was as follows: a fixation crosshair was displayed for 700 ms, followed by a blank screen for 300 ms, a string for 300 ms, a question mark that remained on the screen for 2000 ms or until a response was given, and then finally another fixation crosshair for an interval that varied randomly between 1250 and 1750 ms. The phrase "TIME IS UP" followed the question mark for 1000 ms if no response was given. The experiment started with 10 practice trials, followed by four lexical decision blocks lasting between 3 and 5 min each. Stimuli were presented in a random order, and the sequence was randomized anew for every participant. To indicate whether or not a string was an English word, participants were asked to position their left and right index fingers on two response buttons marked yes/no Response hands were counterbalanced across participants. Both speed and accuracy were stressed.

### *2.4. EEG recordings and ERP analyses*

EEG was recorded from 29 scalp sites with Ag/AgCl electrodes, embedded in an elasticated cap according to an equidistant electrode montage (montage 10; www.easycap.de/easycap/e/electrodes/13_M10.htm). Two further electrodes were placed on the left and right mastoids. A midfrontal electrode (equivalent to Fz in the 10–20 system) was used as the online reference; the data were algebraically re-referenced to linked mastoids offline, reinstating the online reference site. Electrodes placed above and below the right eye and on the outer canthi recorded vertical and horizontal eye movements. All signals were amplified and band-pass filtered between 0.01–70 Hz and digitized at a rate of 200 Hz. The data were low-pass filtered offline with a 25 Hz cut-off (96 dB roll-off, zero phase shift filter).

[1] A related factor that has been shown to affect lexical decision times is a word's contextual distinctiveness (McDonald & Shillcock, 2001). Although CD values are not available for all our stimuli, the analysis of a representative subset of our word lists (45 (/60) abstract and 43 (/60) concrete words) did not show statistically significant differences ($P = .31$).

Epochs of 1280 ms length, starting at 100 ms before item onset, were extracted from the continuous record. Separate ERPs were calculated for each experimental condition, subject, and electrode site. Only trials with correct lexical decisions were included. A correction procedure was applied to minimize blink artefacts (Rugg, Mark, Gilchrist, & Roberts, 1997). Trials containing artefacts other than blinks were discarded (8.6% of the trials, evenly distributed across conditions). ERPs were quantified by computing mean amplitudes across the 250–450 and 700–1000 ms latency regions, relative to the 100 ms pre-stimulus baseline. These intervals were chosen in accordance with the previous ERP literature on concreteness effects (e.g. Kanske & Kotz, 2007; Kellenbach et al., 2002; West & Holcomb, 2000), and adjusted after visual inspection of the grand averages. The analyses incorporated amplitudes at 20 scalp sites (excluding electrodes from the midline and the central line) to partition the electrode grid according to the left/right and anterior/posterior position of electrodes. The data were analyzed with repeated measures analyses of variance (ANOVAs), incorporating the Greenhouse-Geisser correction for violations of sphericity (Keselman & Rogan, 1980). The ANOVAs included factors of CONCRETENESS (concrete/abstract word), ROSTRALITY (anterior/posterior), LATERALITY (left/right), and ELECTRODE (4 sites). For the 250–450 ms interval, an additional ANOVA was directed at effects of lexicality. This analysis compared the data elicited by pseudowords and words, collapsing across concrete and abstract words. This ANOVA incorporated the factor of LEXICALITY (words/pseudowords) rather than CONCRETENESS.

## 3. Results

### 3.1. Behavioral data

An ANOVA comparing lexical decision latencies across concrete words, abstract words, and pseudowords indicated a significant difference ($F(2,28) = 18.037$, $p < .01$). Pairwise comparisons revealed that lexical decisions were faster for abstract words (559 ms; sd = 73) than concrete words (577 ms; sd = 72) ($F(1,14) = 7.872$, $p = .014$) and faster for both abstract and concrete words relative to pseudowords (664 ms; sd = 123) ($F(1,14) = 19.616$ and 17.007, $p < .01$). There was a trend in the same direction in accuracy rates ($F(2,28) = 3.102$, $p = .061$: concrete words = 95.8 (sd = 2.5); abstract words = 95.0 (sd = 2.7); pseudowords = 93.8 (sd = 3.7)).

### 3.2. Electrophysiological data

Fig. 1 shows the grand average ERP waveforms for words (collapsed across word type) and pseudowords at a frontal electrode site. Visual inspection suggests differences in the time window of the N400 component, starting at around 250 ms and lasting for 200 ms. As expected in a task that engages semantic information, pseudowords elicited more negative amplitudes than words in this interval. The difference is broadly distributed across the scalp with a maximum over frontocentral sites (see spline map in Fig. 1). An ANOVA on the mean amplitudes in the 250–450 ms interval revealed a significant effect of LEXICALITY [$F(1,14) = 39.23$, $p < .001$, partial $\eta^2 = .737$]. LEXICALITY did not interact with ROSTRALITY or LATERALITY ($F(1,14) = 2.18$, $p > .20$; $F < 1$, respectively).

The waveforms separating concrete and abstract words are shown in Fig. 2. In the N400 time window, concrete words elicited more negative amplitudes than abstract words. This difference exhibited a widespread scalp distribution with a maximum over frontal scalp sites (see spline map in Fig. 2). An ANOVA on the mean amplitudes in the 250–450 ms interval showed a significant effect of CONCRETENESS ($F(1,14) = 4.94$, $p < .05$, partial $\eta^2 = .261$), with the interaction with ROSTRALITY just falling below significance ($F(1,14) = 3.98$, $p = .066$).

Differences as a function of concreteness are also evident at later latencies. Starting at around 700 ms and continuing until the end of the analysis epoch, concrete words again elicited a more negative deflection than abstract words. This deflection is largest over frontal scalp sites. In evidence, an ANOVA on the amplitudes in the 700–1000 ms interval demonstrated a significant interaction between CONCRETENESS and ROSTRALITY ($F(1,14) = 7.63$, $p < .05$, partial $\eta^2 = .353$). The interaction with the LATERALITY factor was not significant ($F < 1$).

## 4. Discussion

We examined the visual recognition processes associated with concrete and abstract words matched on a large number of lexical and sublexical variables, including imageability and context availability. Replicating recent findings with similarly matched stimuli, lexical decision responses were faster for abstract words (Kousta et al., 2011). This contrasts with most of the existing literature, which typically describes faster response times for concrete words (e.g. Kanske & Kotz, 2007; Kounios & Holcomb, 1994; Schwanenflugel, 1991). Crucially, despite the abstractness behavioral effect, we found the same pattern of ERP differences that has been

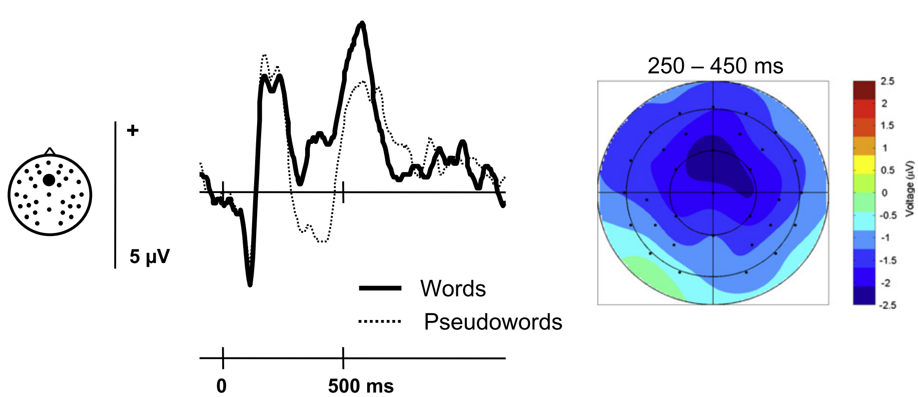


**Fig. 1.** Lexicality effect. Grand average ERP waveforms for words (collapsed across word type) and pseudowords at a midfrontal electrode site. In this and the next figure, vertical lines mark the onset of the letter strings ($t = 0$). Positive values are plotted upward. The two-dimensional voltage spline map on the right illustrates the distribution across the scalp of the lexicality effect in the 250–450 ms time window. The effect was computed by subtracting the values associated with the unweighted average of concrete and abstract words from those associated with pseudowords. Maps in this and the next figure are scaled symmetrically depending on the maximum negative value of each effect.

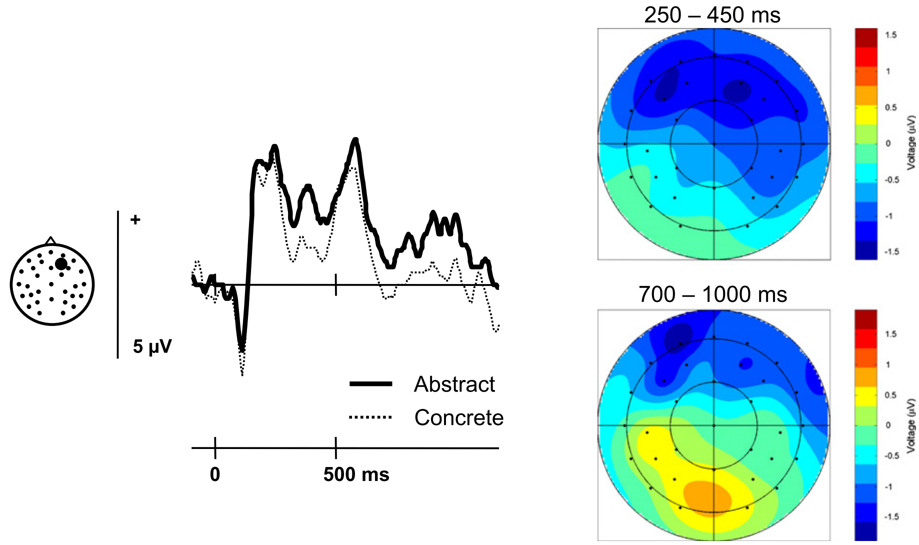


**Fig. 2.** Concreteness effect. Grand average ERP waveforms for concrete and abstract words at a right frontal electrode site. Shown on the right are the voltage spline maps illustrating the scalp distributions of the concreteness effects (amplitude differences between concrete and abstract words) in the 250–450 ms and 700–1000 ms intervals.

reported in previous studies. Namely, ERPs for concrete words were more negative-going than ERPs for abstract words between 250 and 450 ms after word onset. ERP concreteness effects with this polarity and latency have been interpreted as modulations of the N400 component (Kounios & Holcomb, 1994). Differences between concrete and abstract words were also observed in a later time window, between 700 and 1000 ms. This latter effect, which was largest over frontal sites, conforms to the previously reported N700 concreteness effect (e.g. West & Holcomb, 2000).

The N400 concreteness effect took place in the same time window as the lexicality effect, an effect that has been shown to modulate also the N400 component, with larger amplitudes for pseudowords compared to words (Barber, Vergara, & Carreiras, 2004). The modulation of the N400 component by the lexical status of the stimuli can be taken as evidence of word meaning activation when performing the lexical decision task. The concreteness effect had a frontocentral distribution in line with previous studies that compared concrete and abstract words (e.g. Holcomb et al., 1999; Kanske & Kotz, 2007; West & Holcomb, 2000). Lexicality and concreteness effects showed similar but not identical scalp distributions. Concreteness effects showed a more frontal distribution than lexicality, though the distribution of the latter was still more frontal than the standard centroparietal N400 effect for words presented in context. These variations in the distribution of the effects are consistent with a general view of the N400 component as the sum of several neural generators, whose activation levels depend on both the available context-related information (and the resulting integration processes) and the lexico-semantic features of the stimuli that must be recovered from long-term memory.

Thus, whereas tightly matched stimuli resulted in faster responses for abstract words, ERPs to concrete words still showed the previously reported larger N400 and N700 deflections. These findings allow two main conclusions. First, the results do not support the predictions of the prevailing functional interpretation of ERP correlates of concrete and abstract word processing, namely the context-extended dual coding hypothesis. Second, concreteness/abstractness effects in lexical decision response times and electrical brain activity are dissociable. Distinct processes must therefore underlie each of them.

Previous research has argued that concrete words have richer semantic representations, with stronger and denser links to associated semantic information than abstract words (Holcomb et al., 1999; West & Holcomb, 2000). In such a view, the larger N400 for concrete words results from the greater effort that is required to integrate the more extensive semantic information into a higher-level representation. However, in our experiment, concrete and abstract words were matched on a number of semantic variables that measure semantic richness, critically including context availability (see footnote 1 in relation to contextual distinctiveness). This challenges an interpretation of the N400 concreteness effect in terms of the amount of semantic information associated with the words. Importantly, words in the experiment also did not differ with regards to emotional valence, considered to be a critical determinant of abstractness behavioral effects in Kousta et al. (2011).

The types of knowledge underlying abstract and concrete concepts can be significantly different. While concrete concepts rely on modality-specific features, abstract concepts rely on emotional associations as well as a variety of other situational and linguistic information (Andrews, Vigliocco, & Vinson, 2009; Kousta et al., 2011). Barsalou and Wiemer-Hastings (2005) have proposed that, although situations (i.e. contexts) are necessary for a complete representation of both abstract and concrete concepts, they are essential for abstract concepts. When words are presented out of context and under time pressure (as in the present experiment), abstract meanings will therefore receive minimal processing. Instead, abstract words will activate a number of superficial associations with other words, which cannot necessarily be integrated in a unified concept. Taken the N400 component as an index of meaning activation (Molinaro, Conrad, Barber, & Carreiras, 2010), its amplitude variations in our experiment may reflect differences in the level of meaning activation when words are presented in isolation; whereas concrete words would activate and integrate multimodal (sensory-motor) features from distributed cortical networks, processing of abstract words would result in a shallower activation process. Consistent with this possibility, neuroimaging studies suggest that concrete word processing engages a large number of networks linked with the specific sensory-motor properties of the item (see reviews in Binder, Desai, Graves, & Conant, 2009; Cappa, 2008; Martin, 2007). In contrast, studies converge in showing that abstract words engage perisylvian and prefrontal areas to a greater degree than concrete words (Binder et al., 2009; Goldberg, Perfetti, Fiez, & Schneider, 2007).

Concrete and abstract words were not only matched for context availability, but also for imageability. Nonetheless, we found a larger N700 for concrete words, just as in previous research (cf. Holcomb et al., 1999; Kanske & Kotz, 2007; West & Holcomb, 2000). West and Holcomb (2000) related the N700 concreteness effect to enhanced imagery-related processing for concrete words, consistent with the dual coding theory (Paivio, 1986). Although West and Holcomb (2000) did not specify the exact nature of such imagery processes, they suggested that it could be related to visual imagery (Kosslyn, Thompson, & Alpert, 1997). The fact that visual imageability was matched across word types in our experiment makes such an interpretation unlikely. Here, the greater N700 elicited by concrete words may reflect differences in the amount of sensory-motor features that need to be integrated and manipulated to build and maintain a mental representation. In a previous study, we reported a similar sustained frontal effect that differentiated nouns and verbs with different amounts of associated motor and sensory features (Barber, Kousta, Otten, & Vigliocco, 2010). Therefore although the N700 effect cannot just be reduced to visual imagery (which was matched in the present study), it could still be related to internal representations, and the effect may reflect the amount and characteristics of the multimodal information involved in a mental representation. This view is also compatible with the proposal that the N700 effect could be linked with activity in ventrolateral prefrontal cortex, which could control and modulate the retrieval of representations from the long-term memory system (Adorni & Proverbio, 2012).

Interestingly our results show a dissociation between electrophysiological and behavioral measures. In spite of the classical N400-N700 concreteness effect, response latencies were shorter for abstract than for concrete words. Although there is extensive evidence of semantic effects in lexical decisions (Balota et al., 2007), word versus non-word decisions can also take advantage of other linguistic cues (e.g. orthography, phonology or general levels of activation). In the present study, stimuli were matched for all those other variables that normally favor concrete words, such as familiarity, age of acquisition, imageability and context availability. Under these circumstances, the proposed larger number of superficial word associations triggered by abstract words (Barsalou & Wiemer-Hastings, 2005) can be exploited as a quick index for confirming lexicality. Therefore, we suggest that ERP concreteness effects can be taken as an index of meaning activation processes, modulated by the degree of multimodality of the semantic information being integrated (greater for concrete than abstract words). The behavioral abstractness effect may, instead, reflect decision and response-selection mechanisms that are extremely sensitive to control (Neely, 1991). Regardless of the validity of this interpretation, the current data clearly point to a dissociaton between behavioral and electrophysiological differences between concrete and abstract words.

In summary, we have shown that the adequate control of confounding lexical and sublexical variables is critical for studies on word concreteness effects using factorial designs. However, even after the control of several critical variables we show that the ERP concreteness effect can be found. These results are not easily explained in the framework of the context-extended dual coding hypothesis. Alternatively, we propose that differences between the processing of concrete and abstracts words emerge from the amount of sensory-motor information that is activated and integrated within a single lexico-semantic system involving extensive cortical networks in the brain.

## Acknowledgments

Support for this work was provided by European Union (FP6-2004-NESTPATH) Grant 028714 and by ESRC Grant RES-062-23-2012 to Gabriella Vigliocco, and the Grant PSI2010-19767 of the Spanish Ministry of Science to Horacio A. Barber.

## Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bandl.2013.01.005.

## References

Adorni, R., & Proverbio, A. M. (2012). The neural manifestation of the word concreteness effect: An electrical neuroimaging study. *Neuropsychologia, 50*, 880–891.

Amsel, B.D., & Cree, G.S. (in press). Semantic richness, concreteness, and object domain: An electrophysiological study. *Canadian Journal of Experimental Psychology*.

Andrews, M., Vigliocco, G., & Vinson, D. (2009). Integrating experiential and distributional data to learn semantic representations. *Psychological Review, 116*(3), 463–498.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. I., Kessler, B., Loftis, B., et al. (2007). The English lexicon project. *Behavior Research Methods, 39*, 445–459.

Barber, H. A., Kousta, S. T., Otten, L. J., & Vigliocco, G. (2010). Event-related potentials to event-related words: Grammatical class and semantic attributes in the representation of knowledge. *Brain Research, 21*(1332), 65–74.

Barber, H. A., & Kutas, M. (2007). Interplay between computational models and cognitive electrophysiology in visual word recognition. *Brain Research Reviews, 53*, 98–123.

Barber, H. A., Vergara, M., & Carreiras, M. (2004). Syllable-frequency effects in visual word recognition: Evidence from ERPs. *Neuroreport, 15*(3), 545–548.

Barsalou, L. W., & Wiemer-Hastings, K. (2005). Situating abstract concepts. In D. Pecher & R. Zwaan (Eds.), *Grounding cognition: The role of perception and action in memory, language, and thought* (pp. 129–163). New York: Cambridge University Press.

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex, 19*, 2767–2796.

Binder, J. R., Westbury, C. F., McKiernan, K. A., Possing, E. T., & Medler, D. A. (2005). Distinct brain systems for processing concrete and abstract words. *Journal of Cognitive Neuroscience, 17*, 905–917.

Bradley, M. M., & Lang, P. J. (1999). Affective norms for English words (ANEW): Stimuli, instruction manual and affective ratings. Technical report C-1, Gainesville, FL. The Center for Research in Psychophysiology, University of Florida.

Cappa, S. F. (2008). Imaging studies of semantic memory. *Current Opinion in Neurology, 21*(6), 669–675.

Coltheart, M. (1981). MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology, 33a*, 497–505.

Fellbaum, C. (1998). *WordNet: An electronic lexical database*. Cambridge, MA: MIT Press.

Fliessbach, K., Weis, S., Klaver, P., Elger, C. E., & Weber, B. (2006). The effect of word concreteness on recognition memory. *NeuroImage, 32*, 1413–1421.

Giesbrecht, B., Camblin, C. C., & Swaab, T. Y. (2004). Separable effects of semantic priming and imageability on word processing in human cortex. *Cerebral Cortex, 14*, 521–529.

Gilhooly, K. J., & Logie, R. H. (1980). Age of acquisition, imagery, concreteness, familiarity and ambiguity measures for 1944 words. *Behaviour Research Methods and Instrumentation, 12*, 395–427.

Goldberg, R. F., Perfetti, C. A., Fiez, J. A., & Schneider, W. (2007). Selective retrieval of abstract semantic knowledge in Left Prefrontal Cortex. *The Journal of Neuroscience, 27*(14), 3790–3798.

Holcomb, P. J., Kounios, J., Anderson, J. E., & West, W. C. (1999). Dual-coding, context vailability, and concreteness effects in sentence comprehension: An electrophysiological investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 721–742.

Huang, H., Lee, C., & Federmeier, K. D. (2010). Imagine that! ERPs provide evidence for distinct hemispheric contributions to the processing of concrete and abstract concepts. *NeuroImage, 49*, 1116–1123.

Kanske, P., & Kotz, S. A. (2007). Concreteness in emotional words: ERP evidence from a hemifield study. *Brain Research, 1148*, 138–148.

Kellenbach, M. L., Wijers, A. A., Hovius, M., Mulder, J., & Mulder, G. (2002). Neural differentiation of lexico-syntactic categories or semantic features? Event-related potential evidence for both. *Journal of Cognitive Neuroscience, 14*, 561–577.

Keselman, H. J., & Rogan, J. C. (1980). Repeated measures F tests and psychophysiological research: Controlling the number of false positives. *Psychophysiology, 17*, 499–503.

Kosslyn, S. M., Thompson, W. L., & Alpert, N. M. (1997). Neural systems shared by visual imagery and visual perception: A positron emission tomography study. *Neuroimage, 6*, 320–334.

Kounios, J., Green, D. L., Payne, L., Fleck, J. I., Grondin, R., & McRae, K. (2009). Semantic richness and the activation of concepts in semantic memory: Evidence from event-related potentials. *Brain Research, 28*(1282), 95–102.

Kounios, J., & Holcomb, P. J. (1994). Concreteness effects in semantic processing: ERP evidence supporting dual-encoding theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 804–823.
Kousta, S. T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The representation of abstract words: Why emotion matters. *Journal of Experimental Psychology: General, 140*(1), 14–34.
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.
Levy-Drori, S., & Henik, A. (2006). Concreteness and context availability in lexical decision tasks. *American Journal of Psychology, 119*(1), 45–65.
Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology, 58*, 25–45.
McDonald, S. A., & Shillcock, R. C. (2001). Rethinking the word frequency effect: The neglected role of distributional information in lexical processing. *Language and Speech, 44*, 295–323.
Molinaro, N., Conrad, M., Barber, H. A., & Carreiras, M. (2010). On the functional nature of the N400: Contrasting effects related to visual word recognition and contextual semantic integration. *Cognitive Neuroscience, 1*, 1–7.
Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading* (pp. 264–336). Hillsdale, NJ: Lawrence Erlbaum Associates.
Nittono, H., Suehiro, M., & Hori, T. (2002). Word imageability and N400 in an incidental memory paradigm. *International Journal of Psychophysiology, 44*, 219–229.
Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*, 97–113.
Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford: Oxford University Press.
Rabovsky, M., Sommer, W., & Abdel-Rahman, R. (2012). The time course of semantic richness effects in visual word recognition. *Frontiers in Human Neuroscience, 6*, 11.
Richardson, J. (2003). Dual coding versus relational processing in memory for concrete and abstract words. *European Journal of Cognitive Psychology, 15*, 481–509.
Rugg, M. D., Mark, R. E., Gilchrist, J., & Roberts, R. C. (1997). ERP repetition effects in indirect and direct tasks: Effects of age and inter-item lag. *Psychophysiology, 34*, 572–586.
Schwanenflugel, P. J., & Stowe, R. W. (1989). Context availability and the processing of abstract and concrete words in sentences. *Reading Research Quarterly, 24*, 114–126.
Schwanenflugel, P. (1991). Why are abstract concepts hard to understand? In P. Schwanenflugel (Ed.), *The psychology of word meanings* (pp. 223–250). Hillsdale, NJ: Erlbaum.
Stadthagen-Gonzalez, H., & Davis, C. J. (2006). The Bristol norms for age of acquisition, imageability, and familiarity. *Behavior Research Methods, 38*, 598–605.
Welcome, S. E., Paivio, A., McRae, K., & Joanisse, M. F. (2011). An electrophysiological study of task demands on concreteness effects: Evidence for dual coding theory. *Experimental Brain Research, 212*, 347–358.
West, W. C., & Holcomb, P. J. (2000). Imaginal, semantic, and surface-level processing of concrete and abstract words: An electrophysiological investigation. *Journal of Cognitive Neuroscience, 12*, 1024–1037.
Wilson, M. (1988). MRC psycholinguistic database: Machine-usable dictionary, version 2.00. *Behavior Research Methods, Instruments, and Computers, 20*, 6–10.