

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

## Journal of Neurolinguistics

journal homepage: [www.elsevier.com/locate/jneuroling](http://www.elsevier.com/locate/jneuroling)

# Metaphorical meaning learning in contexts: An event-related potential study

Wenjuan Liu<sup>a,b,c</sup>, Jinfeng Ding<sup>a,b</sup>, Lin Li<sup>a,b,d</sup>, Yufang Yang<sup>a,b,\*</sup><sup>a</sup> CAS Key Laboratory of Behavioral Science, Institute of Psychology, Beijing, China<sup>b</sup> Department of Psychology, University of Chinese Academy of Sciences, Beijing, China<sup>c</sup> School of Education and Psychology, University of Jinan, Jinan, China<sup>d</sup> School of Foreign Studies, Minzu University of China, Beijing, China

## ARTICLE INFO

## Keywords:

Metaphorical meaning  
 Word learning  
 Event-related potentials  
 N400  
 Contextual learning

## ABSTRACT

The current study explored whether the metaphorical meaning of novel words can be learned and how the metaphorical meaning was acquired during contextual reading. The novel words were embedded in two types of learning contexts: one supported the metaphorical meaning (metaphorical meaning learning condition: ML condition), and the other supported the literal meaning (literal meaning learning condition: LL condition). The learning effect was assessed via a semantic-relatedness judgment task with event-related potentials (ERPs) being recorded. In this task, the novel words served as primes, the literal corresponding concepts of the novel words (CC), the metaphorically related words (MR) and unrelated words (NR) served as targets. The results showed that the N400 amplitudes in response to the targets varied as a function of targets type in a graded manner increasing from the CC targets, to the MR targets, and to the NR targets in both learning conditions over the lateral electrodes. Furthermore, over the midline electrodes, the N400 difference between the CC and MR targets was gone in the ML condition; the N400 effect between the MR and NR targets was absent in the LL condition. These results suggested that learners could get the metaphorical meaning of newly learned words, and this process was influenced by the learning contexts. The study broadens our understanding on the learning of the lexical meaning of words.

## 1. Introduction

Metaphor is a crucial aspect of human language, which expands the meaning making resources available to the speakers (Halliday & Matthiessen, 1999). Conceptual metaphors involve establishing relationships between semantic domains. Conventional conceptual metaphors are encoded in the lexicon, in which a word has a literal meaning and a metaphorical meaning from different semantic domains (Ahrens et al., 2007; Lu & Zhang, 2012). In the domain of novel word learning, most previous studies mainly focused on the acquisition of the literal meaning, with the acquisition of the metaphorical meaning remaining to be explored. The current study aimed to investigate the learning of the metaphorical meaning of new words embedded in context during reading.

\* Corresponding author. Institute of Psychology, Chinese Academy of Sciences, Chaoyang District, Lincui Road 16, Beijing, 100101, China  
 E-mail address: [yangyf@psych.ac.cn](mailto:yangyf@psych.ac.cn) (Y. Yang).

<https://doi.org/10.1016/j.jneuroling.2018.08.004>

Received 30 January 2018; Received in revised form 17 August 2018; Accepted 23 August 2018

Available online 10 September 2018

0911-6044/ © 2018 Published by Elsevier Ltd.

### 1.1. Metaphorical meaning

A metaphor is made up of a topic and a vehicle. The topic is the object of a metaphor intends to describe, and the vehicle is used to understand the topic more easily (Camp, 2006). For example, in the metaphor “*the lawyer is a shark*”, the *lawyer* is the topic word, and the *shark* is the vehicle word. By establishing mapping relations between them based on their similarities, the words which served as the vehicles carry the metaphorical meaning (Lakoff & Johnson, 1980). The metaphorical meaning can be expressed in sentences or phrases (Gernsbacher, Keysar, Robertson, & Werner, 2001; Grauwe, Swain, Holcomb, Ditman, & Kuperberg, 2010; Jones & Estes, 2005, 2006). Previous studies on metaphor processing mainly focused on two questions (Gibbs, Bogdanovich, Sykes, & Barr, 1997). The first one is whether the processing of the literal meaning and the metaphorical meaning is different or not. The second question is how the factors, such as conventionality, aptness, and context, influence metaphor processing.

Based on lots of experimental evidence, researchers developed different models regarding the metaphor processing. According to traditional theories of pragmatics such as Grice's indirect access model, the literal meaning and the metaphorical meaning are processed in qualitatively different manners, with the precedence for the automatic processing of the literal one (Grice, 1975). The metaphorical meaning is obtained only after the literal meaning are computed and refuted by the context. However, more and more evidence found that there was no qualitative difference between the processing of the literal and metaphorical meanings (Gibbs et al., 1997). Researchers further proposed the direct access model (Gibbs, 1994; Glucksberg, 2001, 2003). This model suggests that the metaphorical meaning is as readily available as the literal meaning. The two meanings are processed concurrently and involve the same mechanisms.

In addition, a growing body of ERP research revised the indirect and direct access models and further developed new models, such as conceptual blending model (Fauconnier & Turner, 1998), the graded salience model (Giora, 1997), and so on. These heterogeneous models share the assumption that the processing of literal and metaphorical meanings relies on the same mechanisms (Coulson & Van Petten, 2002; Giora, 1997; Kintsch, 2000; Wolff & Gentner, 2000) which is in line with the direct access model. At the same time, these models also admit that there is difficulty in metaphor processing which is consistent with the indirect access model. However, different from the indirect access model which attributes the processing difficulty to the precedence of the literal meaning, these models postulate that difficulties in comprehension of the metaphorical sentences may be better explained in terms of mapping difficulty (Fauconnier & Turner, 1998) or salience of meaning (Giora, 1997). For instance, the conceptual blending theory proposed that people construct mental spaces continuously during language processing. A general cognitive operation, namely mapping processing among different spaces, builds networks of connected spaces. In metaphor processing, the connected mental spaces are built between the source and target domains which are distantly related. For the literal meaning processing, the mapping occurs between relatively close domains even in the same domain. For instance, in the literal use of “shark” in “the fish is a shark”, mapping is established between “fish” and “shark” on the basis of category membership. Therefore, the processing difficulty of metaphors could be attributed to the larger mapping difficulty between concepts in the metaphor comprehension than that in the processing of the literal meaning (Arzouan, Goldstein, & Faust, 2007; Coulson & Van Petten, 2002). Evidence came from ERP studies focusing on the N400 which is a well-established component reflecting the integration difficulty (Kutas & Federmeier, 2011). For instance, Coulson and van Petten found that words elicited larger N400s when they conveyed the metaphorical meaning relative to the literal meaning, indicating that the processing of the metaphorical meaning is more demanding than that of the literal meaning.

Besides the difference of processing between the literal meaning and the metaphorical meaning, researchers are also interested in the factors influencing the processing of metaphors. Previous studies have found that the conventionality (Giora, 1997; Laurent, Denhières, Passerieux, Iakimova, & Hardybaylé, 2006), aptness (Damerall & Kellogg, 2016), and context (Pynte, Besson, Robichon, & Poli, 1996) influence metaphor processing. Conventionality refers to the strength of association between a metaphor vehicle and its figurative meaning, namely a metaphor topic (Bowdle & Gentner, 2005; Giora, 1997; Wolff & Gentner, 2000). Based to this factor, metaphors can be divided into conventional metaphors and novel metaphors. For example, “*the lawyer is a shark*”, widely used in communication, is a conventional metaphor. While “*his sister is a megaphone*”, understood as his sister has a very loud voice or maybe she cannot keep secrets, is a novel metaphor made for occasional use (Dulcinati, Mazzarella, Pouscoulous, & Rodd, 2014). Lai, Curran, and Menn (2009) found that both conventional and novel metaphorical sentence-final words elicited larger N400 amplitudes compared to literal words, with a longer-lasting effect to the novel metaphorical words. These results indicated that the processing of novel metaphors was more taxing than that of conventional metaphors. Aptness refers to the degree to which the vehicle expresses an important feature of the topic (Blasko & Connine, 1993; Chiappe & Kennedy, 1999). For instance, “*that fashion model is a rail*” is a highly apt metaphor because the property (i.e., extremely thin) expressed by the vehicle is an important feature of the topic. In contrast, “*my filing clerk is a butcher*” is a less apt metaphor, because the property implied by the vehicle is irrelevant to the topic (Jones & Estes, 2006). Jones and Estes (2006) found that aptness could predict the speed and ease of metaphor comprehension. Context refers to the discourse in which the metaphor embedded. Previous studies have shown that the supportive context which explicitly described the relation between the topic and vehicle in a metaphor promotes metaphor comprehension (Bambini, Bertini, Schaeken, Stella, & Russo, 2016). In this study, they asked participants to process metaphors in two experiments. In Experiment 1, participants were presented with metaphors (e.g., “*Do you know what that lawyer is? A shark*”) and equivalent literal expressions (e.g., “*Do you know what that fish is? A shark*”) in a minimal context. In Experiment 2, the context was a supportive one, which contained cues expressing the relation between the metaphor's topic and vehicle (e.g., metaphorical condition: “*That lawyer is really aggressive. He is a shark*”; literal condition: “*That fish is really aggressive. It is a shark*”). Results showed that the critical nouns (i.e., *shark*) elicited different N400s between the two conditions in the first experiment but not in the second experiment. These results indicated that the supportive context can promote metaphor comprehension.

Understanding a metaphor and learning a metaphor may involve different cognitive processes, but little is known about how

people learn metaphors. Up to now, many researchers summarized the metaphorical meaning learning in terms of teaching experience. There were no empirical studies focusing on the learning of the metaphorical meaning of novel words. Exploring the metaphorical meaning learning can help us understand the word learning and conceptual representation extensively. Given the role of contexts in metaphor processing, the aim of this study is to explore whether the lexical metaphorical meaning of words can be learned through contextual reading.

### 1.2. Contextual learning

There are many different manners to get the meaning of novel words, such as direct definition learning (Perfetti, Wlotko, & Hart, 2005), direct mapping learning (Carey & Bartlett, 1978; Markson & Bloom, 1997), cross-situation learning (Smith, Smith, & Blythe, 2011; Trueswell, Medina, Hafri, & Gleitman, 2013; Yu & Smith, 2007), associative learning (Dobel et al., 2010; Vadillo, Bárcena, & Matute, 2006) and so on. For adults, the main approach to acquire a new word is to learn it during reading (Jenkins, Stein, & Wysocki, 1984; Nagy, Anderson, & Herman, 1987; Nagy, Herman, & Anderson, 1985), namely contextual learning. In the contextual learning, the meaning of novel words can be inferred from the meanings of other words, or the integrative meaning of the whole context.

Previous contextual learning studies usually contain a learning phase and a testing phase. In the learning phase, participants read the context (sentences or discourses) to infer the meaning of novel words. In the testing phase, a semantic relatedness judgment task or a lexical decision task in a semantic priming paradigm is used to test the priming effect between the newly learned words and the target words (Borovsky, Elman, & Kutas, 2012; Borovsky, Kutas, & Elman, 2013; Mestres-missé, Rodrigue-Fornells, & Münte, 2010). For example, Borovsky et al. (2012) asked participants to infer the meaning of novel words which were presented at the end of the second sentence (e.g., *Peter sat gaping at the centerfold. He asked his friend if he could borrow the YERGE*), and then finish a lexical decision task in a semantic priming paradigm, with the novel words (YERGE) serving as the prime stimuli. The target stimuli included: the corresponding concepts of the novel words (e.g. *magazine*), the semantically related words (e.g. *novel*) and unrelated words (e.g. *accident*). The results showed that the corresponding concepts and the semantically related words elicited smaller N400s than the unrelated words. This result indicated that the meaning of novel words can be acquired and associated with known words in semantic memory immediately after the contextual learning.

Using the semantic priming paradigm, researchers have found that the learned word can establish different semantic relationships with known words, such as thematic relations and taxonomic relations (Ding, Chen, et al., 2017; Zhang, Chen, Wang, Yang, & Yang, 2017). Thematic relation refers to the relation between concepts which participate in the same scenario or event (Lin & Murphy, 2001). Thematically related concepts usually appear simultaneously and perform complementary roles in the same episodes or events. For example, “dogs” and “bones” are thematically related because they usually appear in an “eating” event. Taxonomic relation refers to the relation between concepts of the same category. Taxonomically related concepts usually share common features (Rogers & McClelland, 2004). For instance, “dogs” and “cats” are taxonomically related since they belong to the same category of mammal. It should be noted that both the thematic and taxonomic relations are constructed based on the literal meaning of words. The topic and the vehicle are associated in metaphors through mapping from the source domain to the target domain, and establish metaphorical relationship (Mcglone & Manfredi, 2002). For example, in the metaphor “*the lawyer is a shark*”, the vehicle “*shark*” has a literal meaning and a metaphorical meaning. The construction of the metaphorical relationship between these two meanings remains to be explored. The current study aimed to investigate whether the metaphorical meaning of novel words could be learned through the contextual learning approach.

### 1.3. The current study

In line with previous contextual learning studies (Mestres-Missé, Càmarà, Rodríguez-Fornells, Rotte, & Münte, 2008), the current study contained a learning phase and a testing phase. In the learning phase, participants read contexts to infer the meaning of novel words. The contexts included two types: one supported the learning of the metaphorical meaning of the novel word (ML condition), and the other supported the learning of the literal meaning (LL condition). Table 1 presents an example of the experimental stimuli. The two-character pseudo-words in Chinese which can be pronounced but without meaning served as novel words. Their literal meaning and metaphorical meaning are already present in the Chinese conceptual system. For instance, “*children are flowers*” is a common metaphorical mapping in Chinese, but we replaced the vehicle “*flowers*” with a novel word “*fangfen*”, which was learned during contextual reading.

In the testing phase, the learned novel words were paired with three types of words respectively: the literal meaning of the novel words, namely the corresponding concepts (CC), the metaphorically related words (MR), and the unrelated words (NR) (The novel words were two-character pseudo-words (eg. *fangfen*). Its corresponding concept was the vehicle words, such as “*flowers*” in corresponding metaphors. MR targets were the topic “*children*” in the metaphor “*children are flowers*”). In the semantic priming paradigm, lexical decision task and semantic-relatedness judgment task were used to test the priming effect. Some researchers hold that the lexical decision task reflects automatic spreading activation processing, while the semantic-relatedness judgment task reflects controlled processing. In previous studies, researchers have found that the metaphorical meaning processing was more demanding than the literal meaning processing (Coulson & Van Petten, 2002). Participants had to make an effort to process the metaphorical meaning. The semantic-relatedness judgment task which encourages participants to focus on the semantic aspect of the learned novel words may be more suitable to test the learning effect than the lexical decision task. For this reason, we adopted the semantic-relatedness judgment task to explore the metaphorical meaning learning.

**Table 1**

Example of the learning text and testing targets.

Learning text in the learning phase		
LL condition	一株完整的植物包括根、茎、叶、果实、种子和房份。一些植物的房份颜色非常漂亮,靠这样来吸引昆虫完成授粉。(A complete plant includes roots, stalks, leaves, fruits, seeds and fangfen. The fangfen of some plants have bright color, which can attract insects to finish transferring pollen.)	
ML condition	他们正处于天真烂漫无忧无虑的年代,被称为祖国的房份。他们在校园里高兴地跑来跑去,非常活跃,这些祖国的房份正健康地成长着。(They are in the best time during all life, and has been regarded as the fangfen of motherland. They are always running happily in campus, the fangfen of motherland was growing up healthily.)	
Primes and targets in testing phase		
Target conditions	Primes	Targets
Corresponding concept(CC):	房份(fangfen)	花朵(flower)
metaphorically related word(MR):	房份(fangfen)	孩子(child)
Unrelated word1(NR1):	房份(fangfen)	习惯(habit)
Unrelated word1(NR2):	房份(fangfen)	桂圆(longan)

Note: The examples were originally in Chinese, followed by the English paraphrased translations in parenthesis. The novel words in the contexts are in boldface. LL: learning of the literal meaning of the novel words; ML: learning of the metaphorical meaning of the novel words.

EEG was recorded while participants read the contexts and performed the semantic relatedness judgment task. Previous studies reported the N400 and late positive component (LPC) effects during the contextual learning (Chen, Wang, & Yang, 2014; Ding, Chen, et al., 2017; Ding, Liu, & Yang, 2017; Zhang et al., 2017). The N400 has been regarded as an index of the difficulty of retrieving concepts from semantic memory or integrating conceptual information into contexts (Kutas & Federmeier, 2011; Lau, Phillips, & Poeppel, 2008). The easier it is to retrieve or to integrate information in contexts, the smaller the N400 is. The LPC is proposed to reflect the semantic retrieval from long-term memory (Batterink & Neville, 2011; Van Petten, Kutas, Kluender, Mitchiner, & Mcisaac, 1991) and the conscious awareness of semantic relationship between words (Bouaffre & Faïta-Ainseba, 2007; Chen et al., 2014; Juottonen, Revonsuo, & Lang, 1996). Therefore, the N400 and LPC effects in response to the novel words during the learning phase, and to the target words in the testing phase could reflect the process of learning and the establishment of semantic representations for the words in metaphors, respectively.

In the learning phase, previous studies have found that people could rapidly infer the literal meaning of novel words as indicated by the smaller N400s and LPCs elicited by the novel words when they appeared at the second time than the first time (Zhang et al., 2017). Hence, we expected the N400 and LPC effects (i.e., the difference waveforms between the first and second presentations) in response to the novel words in the LL condition. According to the conceptual blending model (Coulson & Van Petten, 2002), the acquisition of the metaphorical meaning of the novel words would be more difficult than that of the literal meaning, the N400 and LPC effects would be larger in the ML condition than in the LL condition.

In the testing phase, the CC targets would elicit smaller N400s/LPCs relative to the NR targets in the LL condition, because previous studies have found that the novel words could prime their literal corresponding concepts after a short contextual learning phase (Borovsky et al., 2012). Since the discourses in the LL condition do not involve metaphor processing, we expected that the N400s/LPCs elicited by the MR targets and the NR targets would not be significantly different. According to the conceptual blending model, both the literal and metaphorical meaning of the novel words would be available during metaphorical discourses reading, the CC and MR targets would elicit smaller N400s/LPCs than the NR targets. Meanwhile the processing of the metaphorical meaning is more demanding than that of the literal meaning, the MR targets would elicit larger N400s than the CC targets.

## 2. Method

### 2.1. Participants

Twenty-nine university students (mean age 23 years, range 19–28 years; 15 males) participated in the ERP study. They were right-handed native speakers of Chinese with normal or corrected to normal vision. None of them had dyslexia or neurological impairments. Before the experiments, they signed a consent form. After the experiment, they were paid according to the response accuracy in the experiment. The study was approved by the Institutional Review Board of the Institute of Psychology, Chinese Academy of Science. Five participants' data were excluded because of poor signal-to-noise ratio (three males and two female). The final data from 24 participants (mean age 23 years, range 19–28 years; 12 males) were included in the statistical analysis.

### 2.2. Materials and design

We chose 60 metaphors, such as “children are flowers”, from either the dictionary or previous studies. In order to guarantee the credibility and validity of the experimental stimuli, twelve paid volunteers read all the metaphors and judge the plausibility on a 7-point Likert scale (1 indicates “unable to understand”, 7 indicates “the easiest to understand”). Then another group of 12 participants were asked to read these metaphors and rate their familiarities on a 7-point Likert scale (1 indicates unfamiliar, 7 indicates the most

familiar). The results showed that the metaphors used in the learning phase were highly plausible (mean (SD) = 6.03 (0.46)) and highly familiar (mean (SD) = 5.95 (0.48)).

After we finished the above evaluations, we extended the metaphors into two-sentence discourses in which the vehicles were replaced by the novel words (the metaphorical meaning learning condition, ML condition). The novel words were two-character pseudo-words in Chinese which can be pronounced but without meaning. For each novel word, we constructed discourses where they only convey the literal meaning (the literal meaning learning condition, LL condition). In both types of discourses, each novel word was presented twice. The first presentation was located at the end of the first sentence, and the second presentation was located in the middle of the second sentence. We divided the 120 discourses (60 discourse in each condition) into two versions, each of which contained 30 ML discourses and 30 LL discourses. The two discourses for the same novel word were not presented in the same version. We asked 16 participants (nine for each version) to guess the literal meaning of the novel word in each discourse. The results showed that the guessing rates of the novel words were equally high in the ML and LL conditions (ML condition: mean (SD) = 85.83% (17.60%), LL condition: mean (SD) = 85.21% (19.05%); item analysis:  $t_{(59)} = 0.19, p = .850$ ; subject analysis:  $t_{(15)} = 1.46, p = .164$ ).

Furthermore, we tested the cloze probability of the literal meaning of novel words in the learning discourses. Another 12 subjects were recruited to complete the cloze probability test, in which the discourses were truncated before the critical words and the subjects were asked to finish the sentences with words that came to mind. The results showed that the cloze probability of the corresponding concepts of the novel words in the first sentences (mean (SD) = 69.31% (2.65%)) was lower than that of the second sentences (mean (SD) = 80.56% (2.39%), item analysis:  $F_{(1,59)} = 28.34, p < .001, \eta_p^2 = 0.324$ ; subject analysis:  $F_{(1,11)} = 12.96, p = .004, \eta_p^2 = 0.541$ ). There was no difference between the ML condition (mean (SD) = 78.11% (3.18%)) and the LL condition (mean (SD) = 77.65% (3.49%)); item analysis:  $F_{(1,59)} = 1.80, p = .184, \eta_p^2 = 0.030$ ; subject analysis:  $F_{(1,11)} = 0.51, p = .489, \eta_p^2 = 0.044$ ). We also asked 16 participants to evaluate the coherence of the learning materials on a 7-point Likert scale (1 indicates incoherent, 7 indicates the most coherent). The results showed that the ML discourses and the LL discourses were equally highly coherent (ML condition: mean (SD) = 5.53 (0.56); LL condition: mean (SD) = 5.45 (0.60); item analysis:  $t_{(59)} = 0.76, p = .449$ ; subject analysis:  $t_{(15)} = 1.23, p = .292$ ).

In the testing phase, the novel words served as primes in the semantic relatedness judgment task, paired with three types of real-word targets (see Table 1 for one group of examples). In order to counterbalance participants' responses, we selected two NR targets for each novel word (NR1 targets and NR2 targets).

Because metaphors usually use concrete concepts to represent abstract senses (Forgács, Bardolph, Amsel, Delong, & Kutas, 2015; Hopkins & Fishwick, 2001; Lakoff & Johnson, 1980), we asked 11 participants to judge the concreteness of the target words in a 7-point Likert scale (1 indicates the most abstract, 7 indicates the most concrete). A repeated measures ANOVA, with target condition (CC, MR, NR1 and NR2) was performed. There was a significant main effect of target condition (item analysis:  $F_{(3,177)} = 13.19, p < .001, \eta_p^2 = 0.183$ ; subject analysis:  $F_{(3,30)} = 25.84, p < .001, \eta_p^2 = 0.721$ ). The results showed that the concreteness of the CC targets was significantly higher than that of the MR (CC vs. MR:  $t_{(59)} = 8.46, p < .001$ ) and NR (CC vs. NR1:  $t_{(59)} = 4.58, p < .001$ ; CC vs. NR2:  $t_{(59)} = 3.57, p = .006$ ) targets. There were no significant difference between the concreteness of the MR and NR1 or NR2 (MR vs. NR1:  $t_{(59)} = -1.95, p = .333$ ; MR vs. NR2:  $t_{(59)} = -2.36, p = .130$ ; NR1 vs. NR2:  $t_{(59)} = -0.43, p \approx 1$ ) targets. Subject analysis showed the similar data mode (CC vs. MR:  $t_{(10)} = 7.03, p < .001$ ; CC vs. NR1:  $t_{(10)} = 5.27, p = .002$ ; CC vs. NR2:  $t_{(10)} = 5.10, p = .003$ ; MR vs. NR1:  $t_{(10)} = -1.99, p = .447$ ; MR vs. NR2:  $t_{(10)} = -3.16, p = .061$ ; NR1 vs. NR2:  $t_{(10)} = 2.32, p = .250$ ).

Moreover, we asked 12 participants to rate the relatedness between the CC targets and the MR, NR1, as well as NR2 targets on a 7-point Likert scale (1 indicates unrelated, 7 indicates the most related). The participants rated the MR targets to be more related to the CC targets than the NR targets (item analysis: MR vs. NR1:  $t_{(59)} = 33.74, p < .001$ ; MR vs. NR2:  $t_{(59)} = 37.20, p < .001$ ; NR1 vs. NR2:  $t_{(59)} = 1.37, p = .520$ ; subject analysis: MR vs. NR1:  $t_{(11)} = 14.42, p < .001$ ; MR vs. NR2:  $t_{(11)} = 14.58, p < .001$ ; NR1 vs. NR2:  $t_{(11)} = 0.40, p \approx 1$ ). These results were in line with the results of the latent semantic analysis (Landauer & Dumais, 1997), which also represented the semantic association between the CC targets and the MR targets, as well as the NR targets (website: <http://www.lsa.url.tw/modules/lsa>, statistic cosine Similarity of LSA). The statistical analysis showed that the score of the MR targets was significantly larger than that of the NR targets (MR vs. NR1:  $t_{(59)} = 3.64, p = .003$ ; MR vs. NR2:  $t_{(59)} = 3.30, p = .007$ ). There was no significant difference between the NR targets (NR1 vs. NR2:  $t_{(59)} = 0.30, p \approx 1$ ). Last but not least, the word frequency ( $F_{(3,177)} = 1.75, p = .175, \eta_p^2 = 0.029$ ) (Cai & Brysbaert, 2010) and the number of strokes ( $F_{(3,177)} = 0.73, p = .501, \eta_p^2 = 0.013$ ) of all the target words were matched. Table 2 presents the data of aforementioned evaluations.

**Table 2**  
Means (SDs) of the stimuli properties.

Condition	Semantic relatedness	LSA score	Concreteness	Frequency	Number of stroke
CC	–	–	5.62(1.11)	31.02 (102.92)	17.43(3.93)
MR	5.71(0.80)	0.11(0.11)	3.91(1.40)	73.16(132.20)	16.13(4.64)
NR1	1.65(0.49)	0.03(0.05)	4.47(1.67)	44.20(187.65)	16.72(4.85)
NR2	1.53(0.41)	0.04(0.07)	4.61(1.76)	22.44(46.79)	16.55(4.22)

### 2.3. Procedure

The participants were seated in a comfortable chair in front of the screen of a computer at a distance of about 80 cm. The words were presented in white color on a black background. The font size was 32 in both the learning and the testing phases. Similar to the learning phase procedure of previous studies (Mestres-missé, Rodrigue-Fornells, & Münte, 2007; Zhang et al., 2017), a fixation was presented on the screen lasting for randomly from 800 to 1200 ms. Then the learning discourse was presented one phrase for 500 ms at a time except for the novel word, which was always presented in isolation for 1000 ms. The interval between two phrases was 300 ms. After the last phrase, the whole passage was presented on the screen. The participants were asked to read the passage again, and press the space button until they had learned the meaning of the novel words. A resting screen of 2000 ms was presented before the next trial began.

After the learning phase and a short break, participants performed the semantic relatedness judgment task. A fixation cross was presented in the center of the screen lasting for randomly from 800 to 1200 ms. The priming word was presented for 1000 ms, followed by a 500-ms blank. Then the target word appeared and lasted for 1000 ms. Until the target word disappeared, participants judged the semantic relatedness between the two words (7 indicates the most related, and 1 indicates unrelated). After that a resting screen appeared for 2000 ms and then the next trial started.

Twenty-four participants were divided evenly among the two versions of learning materials (12 participants per version). Each participant read one type of discourses for the same novel word. Therefore, it was not possible that the novel word was learned as a lexically ambiguous word that contains two meanings (e.g., flower and child). We divided the discourses into ten blocks, each of which contained six learning contexts (three LL contexts and three ML contexts). The learning phase and the testing phase were interleaved. Participants read six discourses in the learning phase in a pseudo-random order with no more than three contexts of the same condition being presented in succession. In the testing phase, the word pairs were also presented in a pseudo-random order. Word pairs containing the same novel words were separated by at least five other word pairs, and the same response was not repeated more than three times consecutively. All these manipulations guaranteed that the priming effects were not due to the repetition of the novel words. Each block lasted for 8–10 min, followed by a short break. The whole experiment, including the preparation, instructions, a short practice consisting of 4 novel words, and the 10 experimental blocks, took about two and a half hours.

### 2.4. Electroencephalogram (EEG) recording and preprocessing

EEG data were collected throughout the learning phase and the test phase. The EEG data were recorded through an elastic cap, which was equipped with 64 Ag/AgCl electrodes according to the International 10–20 system. The online reference electrode was placed on the right mastoid electrode. And an electrode placed between FPZ and FZ served as the ground. The vertical eye movements and blinks were monitored through a supratro-suborbital bipolar montage. A right-to-left canthal bipolar montage was used to monitor the horizontal eye movements. Most electrode impedances were kept below 5 K $\Omega$ . EEG data were recorded at a sampling rate of 500 Hz, with a band pass filter from 0.05 Hz to 100 Hz, and amplified with AC amplifiers.

We used NeuroScan software 4.3 to process the raw data of testing phase. After automatic ocular artifacts correction (Semlitsch, Anderer, Schuster, & Presslich, 1986), the data were filtered off-line with a 0.1–30 Hz band-pass filter, and then were segmented into 1200 ms epochs starting 200 ms before the onset of the critical words. A baseline correction was applied from 200 to 0 ms preceding the critical words onset. The artifact rejection was performed with  $\pm 75 \mu\text{V}$  artifact rejection criterion. On average, 11% of all trials were rejected. The rejections were equally distributed across the eight conditions. After that, the ERP data were re-referenced off-line to the average of both mastoids. At last, trials were averaged in each condition at each electrode for each participant.

In the learning phase, there may be ERP waveform differences prior to the onset of the critical word between the first and second presentations due to the distinct lexical information. In such cases, the use of a pre-stimulus baseline may result in baseline artifacts (Osterhout, McLaughlin, Kim, Greenwald, & Inoue, 2004; Steinhauer & Drury, 2012). According to previous studies (e.g., Wolff, Schlesewsky, Hirotsani, & Bornkessel-Schlesewsky, 2008), the 0.1–30 Hz band-pass filter procedure can remove the slow voltage shifts and avoid causing the baseline artifact at the same time. Hence, we did not apply a baseline correction for the critical words in the learning phase.

## 3. Results

### 3.1. Behavior data

We conducted a 2 (Learning condition: LL and ML) by 4 (Target condition: CC, MR and NR1 and NR2) repeated measures ANOVA, and we did not find any difference between the NR1 and NR2 targets. So in order to ease data processing, we picked up the NR1 targets as the NR targets for further analysis. We conducted 2 (Learning condition: LL and ML) by 3 (Target condition: CC, MR and NR targets) repeated measures ANOVAs for behavioral data and ERP data in the following analysis. The evaluation scores and the RT results of the semantic relatedness judgment task were shown in Table 3.

#### 3.1.1. The evaluation scores

Prior to the analysis, we removed items to which participants did not make responses. Furthermore, we deleted the data with erroneous responses and extreme data which were 2.5 standard deviations (SDs) away from the mean in each condition of each participant in the semantic relatedness judgment task (0.90% among all data). The repeated measures ANOVA revealed a significant

**Table 3**

Mean evaluation scores and RT in milliseconds with standard deviations (SDs) for CC, MR and NR targets in the semantic evaluation task.

	Evaluation scores		RT	
	ML	LL	ML	LL
CC	6.64(0.05)	6.70(0.06)	655(164)	605(167)
MR	4.47(0.16)	3.28(0.21)	1360(330)	1241(365)
NR	1.50(0.10)	1.45(0.10)	833(275)	732(287)

Note: CC: literal corresponding concepts of the novel words; MR: metaphorically related words; NR: unrelated words; ML: learning the metaphorical meaning; LL: learning the literal meaning.

main effect of Learning condition ( $F_{(1,24)} = 93.12, p < .001, \eta_p^2 = 0.802$ ), with the targets in the ML condition (mean (SD) = 3.62(0.09)) being evaluated more semantically related to the novel words than the LL condition (mean (SD) = 3.27(0.10)). There was a significant main effect of Target condition ( $F_{(2,46)} = 59.16, p < .001, \eta_p^2 = 0.957$ ). Pair-wise comparisons revealed that the evaluation score for the CC targets was higher than those for the MR ( $t_{(23)} = 15.63, p < .001$ ) and NR ( $t_{(23)} = 44.80, p < .001$ ) targets. In addition, the evaluation score for the MR targets was higher than that for the NR targets ( $t_{(23)} = 15.79, p < .001$ ). We also found a significant interaction between the two factors ( $F_{(2,46)} = 66.51, p < .001, \eta_p^2 = 0.656$ ). Further simple effect analysis revealed that the evaluation score for the CC targets ( $t_{(23)} = -1.21, p = .24$ ) or the NR targets ( $t_{(23)} = 1.27, p = .21$ ) in two learning conditions were not significantly different. The score for the MR targets in the ML condition was significantly higher than that in the LL condition ( $t_{(23)} = 9.54, p < .001$ ).

### 3.1.2. The RT data

Before calculating the mean RT data for each condition (see Table 3), we deleted the data with erroneous responses and the extreme data which were 2.5 standard deviations (SDs) away from the mean in each condition of each participant (3.92% among all data). The repeated measures ANOVA showed significant effects of both Learning condition ( $F_{(1,23)} = 15.27, p < .001, \eta_p^2 = 0.399$ ) and Target condition ( $F_{(1,23)} = 83.75, p < .001, \eta_p^2 = 0.785$ ). Participants needed more time to respond in the ML condition than in the LL condition. Pair-wise comparisons revealed that the participants responded faster for the CC targets than for the MR ( $t_{(23)} = -11.26, p < .001$ ) and NR ( $t_{(23)} = -3.05, p = .017$ ) targets. In addition, the RT for the NR targets was shorter than that for the MR targets ( $t_{(23)} = -9.77, p < .001$ ).

## 3.2. ERP data

We presented the ERP responses to the novel words in the learning phase as well as to the target words in the testing phase. In order to explore the scalp distributions of the detected effects and according to previous studies investigating the novel word learning during contextual reading (Chen et al., 2014; Ding, Chen, et al., 2017; Ding, Liu, et al., 2017; Zhang et al., 2017), we analyzed the ERP waveforms for the midline and lateral electrodes separately, combining with another position factor, namely Anteriority (anterior, central, and posterior). Fig. 1 shows the selected electrodes and the division for analysis. For the midline electrodes, repeated measures ANOVAs were performed with Learning condition, Novel word position (first and second presentation), and Anteriority (anterior, central, and posterior) serving as three within-subject factors. For the lateral electrodes, Hemisphere (left and right) was included as another within-subject factor in the repeated measures ANOVAs. Only the (marginally) significant effects containing the critical manipulations (Learning condition and Novel word position for the learning phase; Learning condition and Target condition for the testing phase) were reported.

### 3.2.1. The learning phase

Fig. 2 presented the grand average waveforms elicited by the novel words at the two positions in the learning phase at nine representative electrodes (surrounded by circles in Fig. 1). Based on previous studies (Zhang et al., 2017) and in view of the effects, two time windows were selected for statistical analysis: (1) The standard N400 time window: 300–500 ms; (2) A late positivity time window: 600–800 ms.

In the 300–500 ms time window, for the midline electrodes and lateral electrodes, we did not find any significant effects ( $ps > .05$ ). In the 600–800 ms time window, for the midline electrodes, the main effect of Novel word position was significant ( $F_{(1,23)} = 8.05, p = .009, \eta_p^2 = 0.259$ ). The novel word presented for the first time elicited more positive-going waveforms than the second time. The main effect of Learning condition was significant ( $F_{(1,23)} = 4.59, p = .043, \eta_p^2 = 0.161$ ), with the novel word presented in the ML condition eliciting more positive-going waveforms than those presented in the LL condition. Other effects did not reach statistical significance ( $ps > .05$ ). For the lateral electrodes, there was a significant main effect of Novel word position ( $F_{(1,23)} = 6.06, p = .022, \eta_p^2 = 0.209$ ), revealing that the novel words presented for the first time elicited more positive-going waveforms than those presented for the second time. The main effect of Learning condition was significant ( $F_{(1,23)} = 4.60, p = .043, \eta_p^2 = 0.167$ ), with the novel word presented in the ML condition eliciting more positive-going waveforms than those presented in the LL condition. Other effects did not reach statistical significance ( $ps > .05$ ).

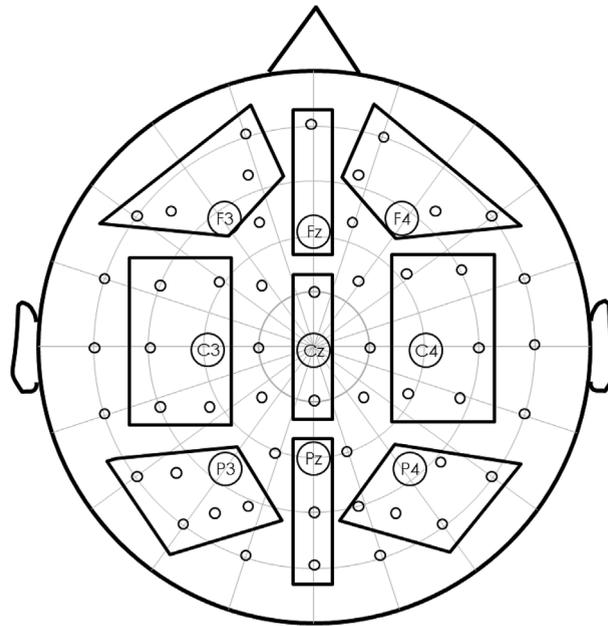


Fig. 1. Electrode layout and the selected electrodes for statistical analysis. The average waveforms over the midline and lateral electrodes were analyzed separately. The nine electrodes circled by the line were used to present the ERP waveforms.

### 3.2.2. The testing phase

The grand average waveforms elicited by the target words in the ML and LL conditions were presented in Fig. 3 and Fig. 4, respectively, at nine representative electrodes. In view of the effect and based on previous semantic priming studies which reported long-lasting N400 effects (300–600 ms post-onset) for the target words (Coulson & Severens, 2007; Wiese & Schweinberger, 2008), we chose 300–600 ms to analyze the N400 effect.

For the midline electrodes, there was a significant main effect of Target condition ( $F_{(2,46)} = 29.72, p < .001, \eta_p^2 = 0.564$ ). Pair-wise comparisons showed that the N400 amplitudes to the targets varied in a graded manner increasing from the CC targets to the MR targets and the NR targets (CC vs. MR:  $t_{(23)} = 4.20, p = .001$ ; CC vs. NR:  $t_{(23)} = 6.69, p < .001$ ; MR vs. NR:  $t_{(23)} = 6.82, p = .001$ ). In addition, the interaction between Learning condition and Target condition was marginally significant ( $F_{(2,46)} = 2.68, p = .079, \eta_p^2 = 0.104$ ). Though a generalizable interaction might be difficult to infer ( $p = .079$ ), we hypothesized that the learning condition can affect the metaphorical meaning learning according to previous studies revealing contextual promotion on the metaphorical meaning processing (Bambini et al., 2016; Pynte et al., 1996). The processing of MR targets could be facilitated in the ML condition than in the LL condition. Based on this hypothesis, we directly reported the pair-wise comparison results of three type of target words in two learning conditions separately. The results showed that in the ML condition, the CC targets and the MR targets elicited smaller N400s than the NR targets (CC vs. MR:  $t_{(23)} = 2.17, p = .122$ ; CC vs. NR:  $t_{(23)} = 6.95, p < .001$ ; MR vs. NR:  $t_{(23)} = 4.54, p < .001$ ). In the LL condition, the CC targets elicited a smaller N400 than the MR targets and the NR targets (CC vs. MR:  $t_{(23)} = 4.10, p = .001$ ; CC vs. NR:  $t_{(23)} = 5.05, p < .001$ ; MR vs. NR:  $t_{(23)} = 2.06, p = .151$ ).

For the lateral electrodes, there was also a significant main effect of Target condition ( $F_{(2,46)} = 43.61, p < .001, \eta_p^2 = 0.655$ ). Pair-wise comparison showed that the N400 amplitudes to targets varied in a graded manner increasing from the CC targets to the MR targets and the NR targets (CC vs. MR:  $t_{(23)} = 4.65, p < .001$ ; CC vs. NR:  $t_{(23)} = 8.22, p < .001$ ; MR vs. NR:  $t_{(23)} = 5.39, p = .001$ ). In addition, the interaction between Learning condition, Target condition, Hemisphere and Anteriority was marginally significant ( $F_{(2,46)} = 2.28, p = .087, \eta_p^2 = 0.090$ ). The following three-way ANOVAs revealed significant interaction among Learning condition, Target condition, and Hemisphere in the central region (anterior:  $F_{(2,46)} = 0.85, p = .430, \eta_p^2 = 0.036$ ; central:  $F_{(2,46)} = 3.53, p = .037, \eta_p^2 = 0.133$ ; posterior:  $F_{(2,46)} = 1.63, p = .210, \eta_p^2 = 0.036$ ). However, over the central electrodes, the two-way interaction in each hemisphere did not reach significance (left:  $F_{(2,46)} = 0.52, p = .597, \eta_p^2 = 0.022$ ; right:  $F_{(2,46)} = 2.15, p = .128, \eta_p^2 = 0.085$ ).

## 4. Discussion

The current study examined whether the metaphorical meaning of novel words can be learned and how the metaphorical meaning was acquired during contextual reading. In the learning phase, the novel words elicited larger LPCs when they were presented for the first time than the second time. During the testing phase, the behavioral data showed that the responses to the CC targets were faster than those to the MR and NR targets. In addition, the RTs for the MR targets were longer than those for the NR targets. The ERP data showed that the N400 amplitudes elicited by the CC, MR and NR targets increased gradually in both learning conditions over the lateral electrodes tested. For the midline electrodes, the N400 difference between the CC and MR targets was gone in the ML

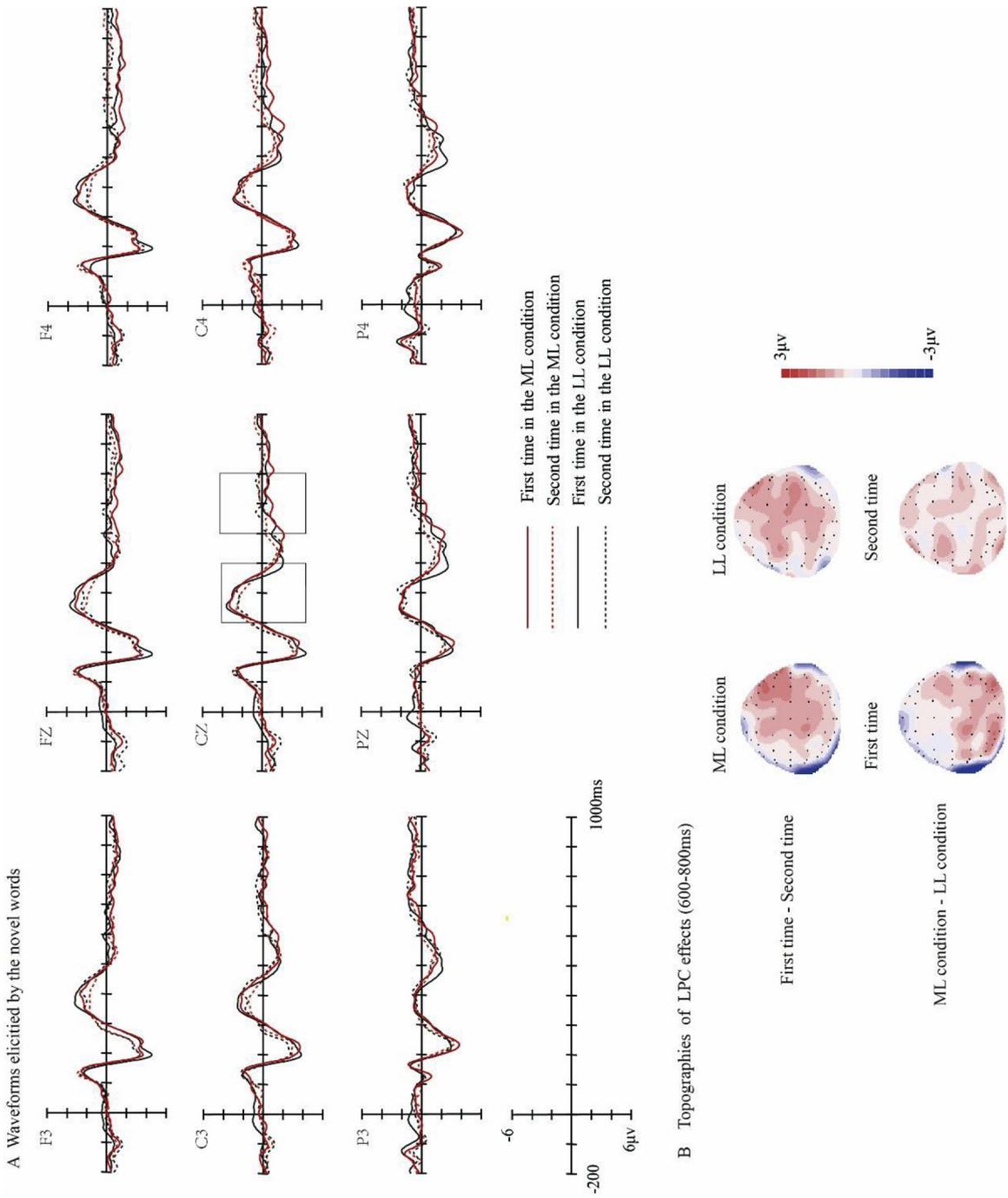
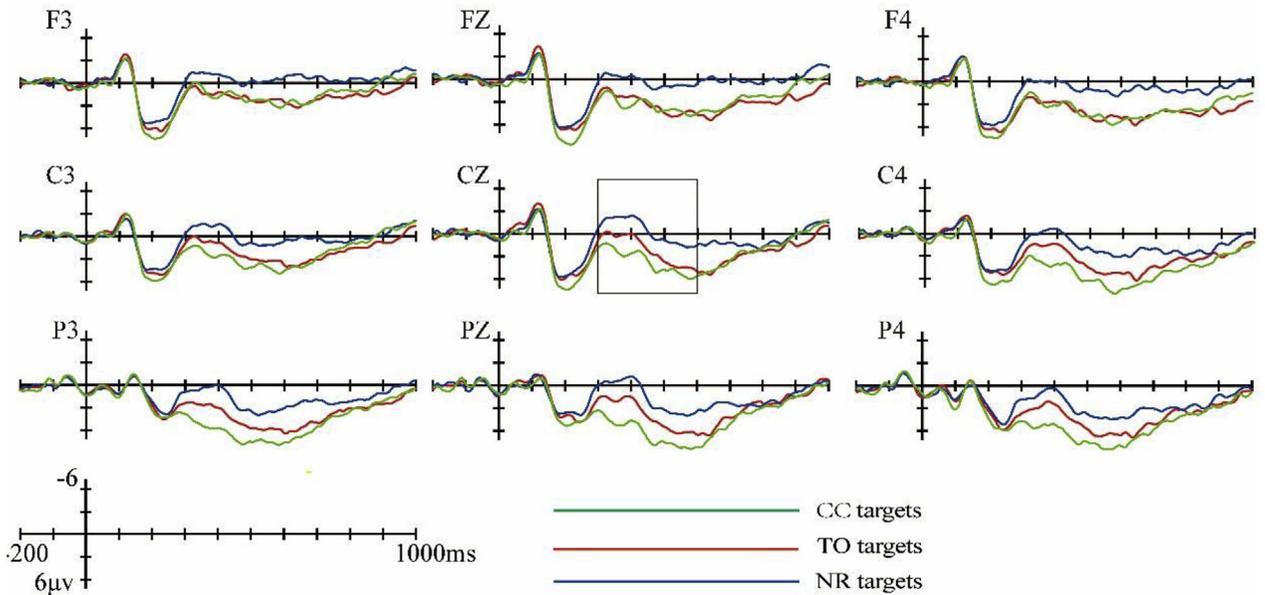


Fig. 2. ERP waveforms (A) and topographies (B) for the novel words in the learning phase. ML: learning the metaphorical meaning; LL: learning the literal meaning.

A Waveforms elicited by the target words in the ML condition



B Topographies

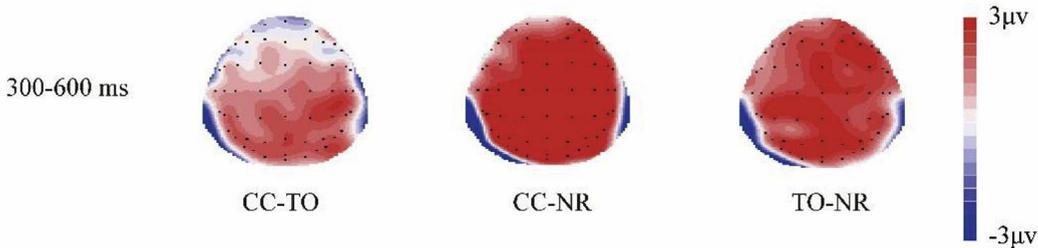


Fig. 3. ERP waveforms (A) and topographies (B) for the target words in the metaphorical meaning learning condition in the testing phase. CC: Corresponding concepts; MR: Metaphorically related words; NR: unrelated words.

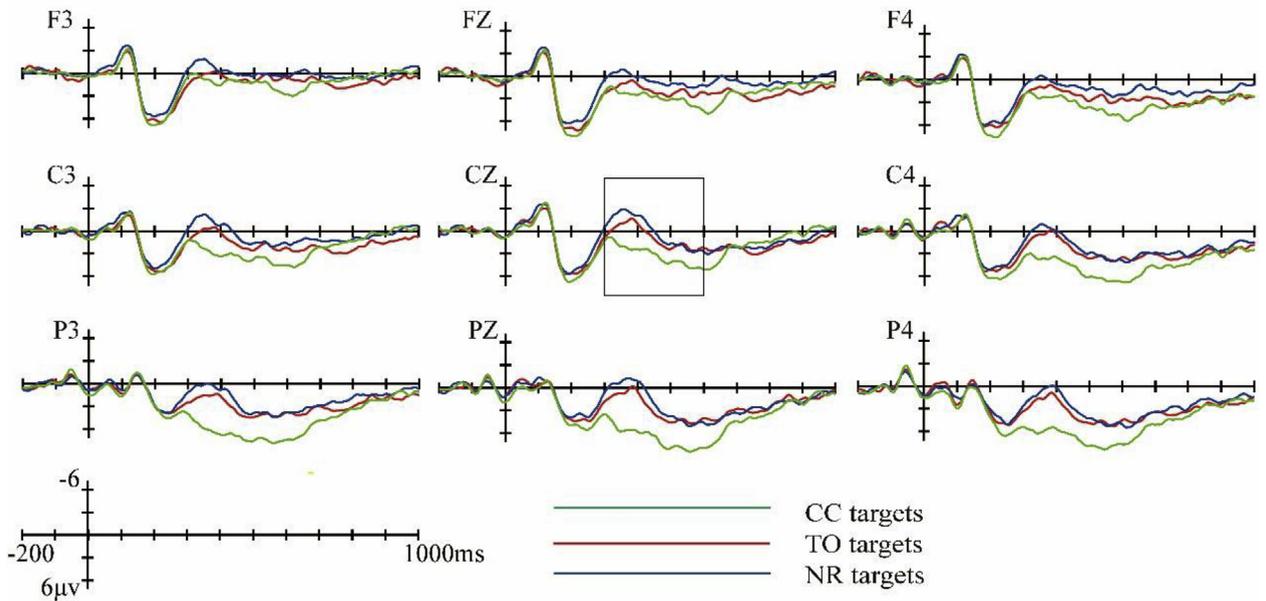
condition. In the LL condition, the N400 difference between the MR and NR targets was absent.

In the learning phase, the novel word in the first presentation elicited a significantly larger LPC compared with the second presentation. This result was consistent with previous studies (Zhang et al., 2017). The LPC effect is proposed to reflect the retrieval of semantic information from long-term memory and the integration of the retrieved information into working memory (Batterink & Neville, 2011; Van Petten et al., 1991). Participants were asked to infer the novel word's meaning on the basis of the learning discourses. When they first encountered the novel word, they had to retrieve semantic information from long-term memory according to the current context and integrate the information in the working memory to infer the meaning of the novel words. When the novel words were presented for the second time, the semantic retrieval was not as difficult as the first time. Therefore, the novel words in the first presentation elicited a significantly larger LPC comparing to the second presentation.

The novel words in the ML condition elicited a larger LPC compared with the LL condition. This result was consistent with previous studies revealing larger LPCs for the metaphor processing than for the literal expression processing (Bambini et al., 2016; Wilson & Carston, 2007). For instance, Bambini et al. (2016) recorded EEG activities while participants read metaphors (e.g., “That lawyer is really aggressive. He is a shark.”) and literal expressions (e.g., “That fish is really aggressive. It is a shark.”) Results showed that the critical nouns (i.e., shark) elicited larger LPCs in the metaphor condition than in the literal expression condition. They proposed that the LPC effect reflected a pragmatic interpretative process, namely a inferential processing from the literal meaning to metaphorical meaning. In the current study, the novel words in the ML condition elicited larger LPCs than those in the LL condition, suggesting that participants needed to process the pragmatic information in metaphorical meaning learning.

Interestingly, differing from the previous studies (Mestres-missé et al., 2007; Zhang et al., 2017), the N400s elicited by the novel

A Waveforms elicited by the target words in the LL condition



B Topographies

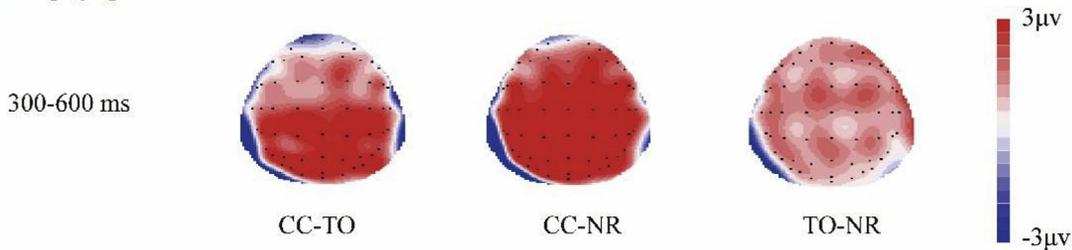


Fig. 4. ERP waveforms (A) and topographies (B) for the target words in the literal meaning learning condition in the testing phase. CC: Corresponding concepts; MR: Metaphorically related words; NR: unrelated words.

words were not significantly different between the first and second presentations. The N400 component has been usually regarded as an index for the difficulty of semantic integration (Kutas & Federmeier, 2000). The absence of the N400 effect in the current study might be attributed to the relatively low guessing rates of the novel words (the current study: 86% on average vs. Zhang et al.: 95% on average). Although the information for the inference has been mostly retrieved from long-term memory (as reflected by the LPC effect), the low guessing rates made participants also try to infer the meaning of the novel words at their second presentation and find coherence rather than detecting mistakes or violations. Therefore, the semantic integration of the novel words was not easier for the second time than the first time, resulting in the similar N400s in response to the novel words.

The ERP effects being not modulated by the learning contexts seems to contradict with the conceptual blending model during the learning of the metaphorical meaning of novel words. The conceptual blending model proposes that the processing of the metaphorical meaning of words is more demanding/difficult than that of the literal meaning, the LPC effect should be larger in the ML condition than in the LL condition. However, the absence of the modulation might be accounted for by the inferring difficulty. Participants inferred the meaning of novel words through the whole discourses as reflected by the similar N400 elicited by the novel words at the first and second positions. Therefore, the processing difficulty of the metaphorical meaning might be superimposed on the general inferring difficulty of the novel words.

During the semantic relatedness judgment task, the scores for the CC targets and MR targets were higher than that for the NR targets in both learning conditions, indicating that participants could be aware of the semantic relationships between the novel words and the CC as well as MR targets. This result suggested that participants acquired both the literal meaning and the metaphorical meaning of the novel words in both learning conditions. In addition, the score for the MR targets in the ML condition was higher than that in the LL condition. This indicated that the supportive contexts could promote the metaphor processing (Bambini et al., 2016;

Pynthe et al., 1996), which would further benefit the learning of the metaphorical meaning of novel words.

The RTs for the CC targets were shorter than those for the NR targets, which also supported the acquisition of the literal meaning of the novel words. However, although the MR targets were more semantically related to the novel words than the NR targets, the RTs for the MR targets were longer than those for the NR targets. This might be interpreted by the difficulty of the current task. Participants were asked to judge the degree of the semantic relatedness between the novel words and the target words. The CC targets were the corresponding concepts of the novel words, and the NR targets were completely unrelated to the novel words. It is easier for participants to decide the semantic relatedness between the novel words and the CC targets, as well as the NR targets. However, the MR targets were moderately related to the novel words, participants needed more time to make responses for the MR targets. That is, participants should pay extra attention to process the relationships between the novel words and the MR targets, which produced longer reaction time for the MR targets.

For the ERP data in the testing phase, we found that the N400s elicited by the targets increased in a graded manner from the CC targets, to the MR targets and to the NR targets in both learning conditions. The N400 amplitude is an indicator of the spread of semantic activation associated with the processing of prime words (Rolke, Heil, Streb, & Hennighausen, 2010), with target words preceded by semantically related words eliciting smaller N400s relative to those preceded by unrelated words in semantic relatedness judgment task (Batterink & Neville, 2011; Chen et al., 2014; Mestres-missé et al., 2007). The observed N400 effect for the CC targets reflected the spread of semantic activation from the prime stimuli to the target stimuli, suggesting that participants have learned the literal meaning of the novel words.

In the ML condition, the CC targets also elicited smaller N400s than the NR targets, suggesting that participants could learn the literal meaning of the novel words in the discourses supporting the metaphorical meaning. Importantly, we found that the MR targets elicited smaller N400 amplitudes than the NR targets. This result showed that the novel words primed the MR targets, indicating that participants have learned the metaphorical meaning of the novel words which is consistent with the behavioral results. In addition, the MR targets elicited a larger N400 than the CC targets, suggesting a larger facilitation of the novel words for the CC targets than for the MR targets. During reading the metaphorical contexts, participants infer the meaning of novel words which were the vehicles of the metaphors. According to the smaller N400s for the CC and MR targets than the NR targets, both the literal meaning and the metaphorical meaning of novel words are available during metaphor processing. Meanwhile, the larger N400s for the MR targets than the CC targets suggests that the processing of the metaphorical meaning is more demanding than that of the literal meaning. This result is in line with the conceptual blending model (Coulson & Van Petten, 2002).

Interestingly, the N400 effect in response to the MR targets was also observed in the LL condition, indicating that participants got the metaphorical meaning of novel words. Theoretically, participants should only learn the literal meaning of the novel words in the LL condition, because the discourses in this condition did not include metaphor processing. The N400 effect for the MR targets in the LL condition was probably due to the high familiarity of the metaphors in the present study. The literal meaning and the metaphorical meaning of the vehicles (the novel words) are highly bound together. Once participants learned the literal corresponding concepts of the novel words, the connections between the novel words and the metaphorically related words were built up, leading to the reduced N400s for the MR targets preceded by the novel words.

Over the midline electrodes tested, we found an influence of the learning contexts on the building of the semantic association between the novel words and the target words. In the ML condition, the N400 difference was gone between the CC and MR targets; in the LL condition, the N400 effect was absent between the MR and NR targets. These results indicated that the priming effect of the novel words for the MR targets was larger in the ML condition than in the LL condition, suggesting that the supportive contexts for the metaphors facilitated the acquisition of the metaphorical meaning of the novel words. Previous studies have shown that the supportive context which explicitly described the relation between the topic and vehicle in a metaphor promotes metaphor comprehension (Bambini et al., 2016). This promotion could be due to the strong association between the literal and metaphorical meanings of the vehicle because the contextual cues facilitated the mapping between the literal and metaphorical meanings. In the present study, the discourses in the ML condition expanded the metaphors and contained more information, resulted in strong connections between the novel words and the MR targets. Although participants were aware of the semantic relationships between the novel words and the MR targets in the LL condition, the semantic connections were not as strong as those in the ML condition.

In conclusion, the behavioral and ERP data provide convergent evidence that participants can get the metaphorical meaning of novel words during contextual learning. Consistent with metaphor processing, the learning of the metaphorical meaning is more difficult than the literal meaning. However, the supportive contexts for the metaphors facilitate the acquisition of the metaphorical meaning. The findings of the present study support the conceptual blending model in the domain of word learning.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant Numbers: 61433018), and Funded by CPSF-CAS Joint Foundation for Excellent Postdoctoral Fellows (Grant No.2016LH0015), as well as the Scientific Foundation of Institute of Psychology, Chinese Academy of Sciences (Grant No. Y6CX212007).

## Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.jneuroling.2018.08.004>.

## References

- Ahrens, K., Liu, H. L., Lee, C. Y., Gong, S. P., Fang, S. Y., & Hsu, Y. Y. (2007). Functional MRI of conventional and anomalous metaphors in Mandarin Chinese. *Brain and Language*, *100*(2), 163–171. <https://doi.org/10.1016/j.bandl.2005.10.004>.
- Arzouan, Y., Goldstein, A., & Faust, M. (2007). Brainwaves are stethoscopes: ERP correlates of novel metaphor comprehension. *Brain Research*, *1160*(1), 69–81. <https://doi.org/10.1016/j.brainres.2007.05.034>.
- Bambini, V., Bertini, C., Schaeken, W., Stella, A., & Russo, F. D. (2016). Disentangling metaphor from context: An ERP study. *Frontiers in Psychology*, *7*(559), <https://doi.org/10.3389/fpsyg.2016.00559>.
- Batterink, L., & Neville, H. (2011). Implicit and explicit mechanisms of word learning in a narrative context: An event-related potential study. *Journal of Cognitive Neuroscience*, *23*(11), 3181–3196. <https://doi.org/10.1162/jocn.a.00013>.
- Blasko, D. G., & Connine, C. M. (1993). Effects of familiarity and aptness on metaphor processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 295–308. <https://doi.org/10.1037/0278-7393.19.2.295>.
- Borovsky, A., Elman, J. L., & Kutas, M. (2012). Once is enough: N400 indexes semantic integration of novel word meanings from a single exposure in context. *Language Learning and Development*, *8*(3), 278–302. <https://doi.org/10.1080/15475441.2011.614893>.
- Borovsky, A., Kutas, M., & Elman, J. L. (2013). Getting it right: Word learning across the hemispheres. *Neuropsychologia*, *51*(5), 825–837. <https://doi.org/10.1016/j.neuropsychologia.2013.01.027>.
- Bouaffre, S., & Faïta-Aïnseba, F. (2007). Hemispheric differences in the time-course of semantic priming processes: Evidence from event-related potentials (ERPs). *Brain and Cognition*, *63*(2), 123–135. <https://doi.org/10.1016/j.bandc.2006.10.006>.
- Bowlde, B. F., & Gentner, D. (2005). The career of metaphor. *Psychological Review*, *112*, 193–216. <https://doi.org/10.1037/0033-295X.112.1.193>.
- Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLoS One*, *5*(6), e10729. <https://doi.org/10.1371/journal.pone.0010729>.
- Camp, E. (2006). Metaphor in the mind: The cognition of metaphor. *Philosophy Compass*, *1*, 154–170. <https://doi.org/10.1111/j.1747-9991.2006.00013.x>.
- Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Proceedings of the stanford child language conference: Vol. 15*, (pp. 17–29).
- Chen, S., Wang, L., & Yang, Y. (2014). Acquiring concepts and features of novel words by two types of learning: Direct mapping and inference. *Neuropsychologia*, *56*(1), 204–218. <https://doi.org/10.1016/j.neuropsychologia.2014.01.012>.
- Chiappe, D. L., & Kennedy, J. M. (1999). Aptness predicts preference for metaphors or similes, as well as recall bias. *Psychonomic Bulletin & Review*, *6*(4), 668–676. <https://doi.org/10.3758/BF03212977>.
- Coulson, S., & Severens, E. (2007). Hemispheric asymmetry and pun comprehension: When cowboys have sore calves. *Brain and Language*, *100*(2), 172–1787. <https://doi.org/10.1016/j.bandl.2005.08.009>.
- Coulson, S., & Van Petten, C. (2002). Conceptual integration and metaphor: An event-related potential study. *Memory & Cognition*, *30*(6), 958–968. <https://doi.org/10.3758/BF03195780>.
- Damerall, A. W., & Kellogg, R. T. (2016). Familiarity and aptness in metaphor comprehension. *American Journal of Psychology*, *129*(1), 49–64. <https://doi.org/10.5406/amerjpsyc.129.1.0049>.
- Ding, J. F., Chen, S., Wang, L., & Yang, Y. F. (2017). Thematic and taxonomic relations of novel words learned from action and perceptual features. *Journal of Neurolinguistics*, *41*, 70–84. <https://doi.org/10.1016/j.jneuroling.2016.10.002>.
- Ding, J. F., Liu, W. J., & Yang, Y. F. (2017). The influence of concreteness of concepts on the integration of novel words into the semantic network. *Frontiers in Psychology*, *8*. <https://doi.org/10.3389/fpsyg.2017.02111>.
- Dobel, C., Breitenstein, C., Klauke, B., Knecht, S., Pantev, C., & Zwitserlood, P. (2010). New names for known things: On the association of novel word forms with existing semantic information. *Journal of Cognitive Neuroscience*, *22*(6), 1251–1261. <https://doi.org/10.1162/jocn.2009.21297>.
- Dulcinati, G., Mazzarella, D., Pouscoulous, N., & Rodd, J. (2014). Processing metaphor: The role of conventionality, familiarity and dominance. *UCL Working Papers in Linguistics*, *72*–88.
- Fauconnier, G., & Turner, M. (1998). Conceptual integration networks. *Cognitive Science*, *22*(2), 133–187. [https://doi.org/10.1016/S0364-0213\(99\)80038-X](https://doi.org/10.1016/S0364-0213(99)80038-X).
- Forgács, B., Bardolph, M. D., Amsel, B. D., Delong, K. A., & Kutas, M. (2015). Metaphors are physical and abstract: ERPs to metaphorically modified nouns resemble ERPs to abstract language. *Frontiers in Human Neuroscience*, *9*, 28. <https://doi.org/10.3389/fnhum.2015.00028>.
- Gernsbacher, M. A., Keysar, B., Robertson, R. R., & Werner, N. K. (2001). The role of suppression and enhancement in understanding metaphors. *Journal of Memory and Language*, *45*(3), 433–450. <https://doi.org/10.1006/jmla.2000.2782>.
- Gibbs, R. W. (1994). *The poetics of mind: Figurative thought, language, and understanding*. Cambridge: Cambridge University Press.
- Gibbs, R. W., Bogdanovich, J. M., Sykes, J. R., & Barr, D. J. (1997). Metaphor in idiom comprehension. *Journal of Memory and Language*, *37*(2), 141–154. <https://doi.org/10.1006/jmla.1996.2506>.
- Giora, R. (1997). Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics*, *8*(3), 183–206. <https://doi.org/10.1515/cogl.1997.8.3.183>.
- Glucksberg, S. (2001). *Understanding figurative language*. New York: Oxford University Press.
- Glucksberg, S. (2003). The psycholinguistics of metaphor. *Trends in Cognitive Sciences*, *7*(2), 92–96. [https://doi.org/10.1016/S1364-6613\(02\)00040-2](https://doi.org/10.1016/S1364-6613(02)00040-2).
- Grauwe, S. D., Swain, A., Holcomb, P. J., Ditman, T., & Kuperberg, G. R. (2010). Electrophysiological insights into the processing of nominal metaphors. *Neuropsychologia*, *48*(7), 1965–1984. <https://doi.org/10.1016/j.neuropsychologia.2010.03.017>.
- Grice, H. P. (1975). *Logic and conversation, Vol. 3*. New York: Academic Press.
- Halliday, M. A. K., & Matthiessen, C. M. I. M. (1999). *Construing experience through meaning: A language-based approach to cognition*. London: Continuum.
- Hopkins, J. F., & Fishwick, P. A. (2001). A three-dimensional human agent metaphor for modeling and simulation. *Proceedings of the IEEE*, *89*(2), 131–147. <https://doi.org/10.1109/5.910850>.
- Jenkins, J. R., Stein, M. L., & Wysocki, K. (1984). Learning vocabulary through reading. *American Educational Research Journal*, *21*(4), 767–787. <https://doi.org/10.3102/00028312021004767>.
- Jones, L. L., & Estes, Z. (2005). Metaphor comprehension as attributive categorization. *Journal of Memory and Language*, *53*(1), 110–124. <https://doi.org/10.1016/j.jml.2005.01.016>.
- Jones, L. L., & Estes, Z. (2006). Roosters, robins, and alarm clocks: Aptness and conventionality in metaphor comprehension. *Journal of Memory and Language*, *55*(1), 18–32. <https://doi.org/10.1016/j.jml.2006.02.004>.
- Juottonen, K., Revonsuo, A., & Lang, H. (1996). Dissimilar age influences on two ERP waveforms (LPC and N400) reflecting semantic context effect. *Cognitive Brain Research*, *4*(2), 99–107. [https://doi.org/10.1016/0926-6410\(96\)00022-5](https://doi.org/10.1016/0926-6410(96)00022-5).
- Kintsch, W. (2000). Metaphor comprehension: A computational theory. *Psychonomic Bulletin & Review*, *7*(2), 257–266. <https://doi.org/10.3758/BF03212981>.
- Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*(12), 463–470. [https://doi.org/10.1016/S1364-6613\(00\)01560-6](https://doi.org/10.1016/S1364-6613(00)01560-6).
- Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, *62*(1), 1–47. <https://doi.org/10.1146/annurev.psych.093008.131123>.
- Lai, V. T., Curran, T., & Menn, L. (2009). Comprehending conventional and novel metaphors: An ERP study. *Brain Research*, *1284*(1), 145–155. <https://doi.org/10.1016/j.brainres.2009.05.088>.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*(2), 211–240. <https://doi.org/10.1037/0033-295X.104.2.211>.
- Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. *Nature Reviews Neuroscience*, *9*(12), 920. <https://doi.org/10.1038/nrn2532>.

- Laurent, J. P., Denhières, G., Passerieux, C., Iakimova, G., & Hardybaylé, M. C. (2006). On understanding idiomatic language: The salience hypothesis assessed by ERPs. *Brain Research, 1068*(1), 151–160. <https://doi.org/10.1016/j.brainres.2005.10.076>.
- Lin, E. L., & Murphy, G. L. (2001). Thematic relations in adults' concepts. *Journal of Experimental Psychology: General, 130*(1), 3e28. <https://doi.org/10.1037/0096-3445.130.1.3>.
- Lu, A., & Zhang, J. X. (2012). Event-related potential evidence for the early activation of literal meaning during comprehension of conventional lexical metaphors. *Neuropsychologia, 50*(8), 1730–1738. <https://doi.org/10.1016/j.neuropsychologia.2012.03.027>.
- Markson, L., & Bloom, P. (1997). Evidence against a dedicated system for word learning in children. *Nature, 385*, 813–815. <https://doi.org/10.1038/385813a0>.
- Mcglone, M. S., & Manfredi, D. A. (2002). Topic-vehicle interaction in comprehension. *Memory & Cognition, 29*(8), 1209–1219. <https://doi.org/10.3758/BF03206390>.
- Mestres-Missé, A., Càmarà, E., Rodríguez-Fornells, A., Rotte, M., & Münte, T. F. (2008). Functional neuroanatomy of meaning acquisition from context. *Journal of Cognitive Neuroscience, 20*(12), 2153–2166. [https://doi.org/10.1016/S0361-3682\(99\)00061-6](https://doi.org/10.1016/S0361-3682(99)00061-6).
- Mestres-missé, A., Rodrigue-Fornells, A., & Münte, T. F. (2007). Watching the brain during meaning acquisition. *Cerebral Cortex, 17*(8), 1858–1866. <https://doi.org/10.1093/cercor/bhl094>.
- Mestres-missé, A., Rodrigue-Fornells, A., & Münte, T. F. (2010). Neural differences in the mapping of verb and noun concepts onto novel words. *NeuroImage, 49*(3), 2826. <https://doi.org/10.1016/j.neuroimage.2009.10.018>.
- Nagy, W. E., Anderson, R. C., & Herman, P. A. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal, 24*(2), 237–270. <https://doi.org/10.3102/00028312024002237>.
- Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly, 20*(2), 233–253. <https://doi.org/10.2307/747758>.
- Osterhout, L., McLaughlin, J., Kim, A., Greenwald, R., & Inoue, K. (2004). Sentences in the brain: Event related potentials as real time reflections of sentence comprehension and language learning. In M. Carreiras, & J. Clifton (Eds.). *The on-line study of sentence comprehension: Eye tracking, ERP and beyond*.
- Perfetti, C. A., Wlotko, E. W., & Hart, L. A. (2005). Word learning and individual differences in word learning reflected in event-related potentials. *Journal of Experimental Psychology Learning Memory and Cognition, 31*(6), 1281–1292. <https://doi.org/10.1037/0278-7393.31.6.1281>.
- Pynte, J., Besson, M., Robichon, F., & Poli, J. (1996). The time-course of metaphor comprehension: An event-related potential study. *Brain and Language, 55*(3), 293–316. <https://doi.org/10.1006/brln.1996.0107>.
- Rogers, T. T., & McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach*. London: the MIT Press.
- Rolke, B., Heil, M., Streb, J., & Hennighausen, E. (2010). Missed prime words within the attentional blink evoke an n400 semantic priming effect. *Psychophysiology, 38*(2), 165–174. <https://doi.org/10.1017/S0048577201991504>.
- Semlitsch, H. V., Anderer, P., Schuster, P., & Presslich, O. (1986). A solution for reliable and valid reduction of ocular artifacts, applied to the P300 ERP. *Psychophysiology, 23*(6), 695–703. <https://doi.org/10.1111/j.1469-8986.1986.tb00696.x>.
- Smith, K., Smith, A. D. M., & Blythe, R. A. (2011). Cross-Situational learning: An experimental study of word-learning mechanisms. *Cognitive Science, 35*(3), 480–498. <https://doi.org/10.1111/j.1551-6709.2010.01158.x>.
- Steinhauer, K., & Drury, J. E. (2012). On the early left-anterior negativity (elan) in syntax studies. *Brain and Language, 120*(2), 135–162. <https://doi.org/10.1016/j.bandl.2011.07.001>.
- Trueswell, J. C., Medina, T. N., Hafri, A., & Gleitman, L. R. (2013). Propose but verify: Fast mapping meets cross-situational word learning. *Cognitive Psychology, 66*(1), 126–156. <https://doi.org/10.1016/j.cogpsych.2012.10.001>.
- Vadillo, M. A., Bárcena, R., & Matute, H. (2006). The internet as a research tool in the study of associative learning: An example from overshadowing. *Behavioural Processes, 73*(1), 36–40. <https://doi.org/10.1016/j.beproc.2006.01.014>.
- Van Petten, C., Kutas, M., Kluender, R., Mitchiner, M., & Mcisaac, H. (1991). Fractionating the word repetition effect with event-related potentials. *Journal of Cognitive Neuroscience, 3*(2), 131–150. <https://doi.org/10.1162/jocn.1991.3.2.131>.
- Wiese, H., & Schweinberger, S. R. (2008). Event-related potentials indicate different processes to mediate categorical and associative priming in person recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(5), 1246–1263. <https://doi.org/10.1037/a0012937>.
- Wilson, D., & Carston, R. (2007). Aunitary approach to lexical pragmatics: Relevance, inference and adhoc concepts. *Pragmatics, edN. Burton-Roberts* (pp. 230–259). Basingstoke: Palgrave Macmillan.
- Wolff, P., & Gentner, D. (2000). Evidence for role-neutral initial processing of metaphors. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(2), 529–541. <https://doi.org/10.1037/0278-7393.26.2.529>.
- Wolff, S., Schlesewsky, M., Hirotani, M., & Bornkesselschlesewsky, I. (2008). The neural mechanisms of word order processing revisited: Electrophysiological evidence from Japanese. *Brain and Language, 107*(2), 133–157. <https://doi.org/10.1016/j.bandl.2008.06.003>.
- Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science, 18*(5), 414–420. <https://doi.org/10.1111/j.1467-9280.2007.01915.x>.
- Zhang, M. C., Chen, S., Wang, L., Yang, X. H., & Yang, Y. F. (2017). Episodic specificity in acquiring thematic knowledge of novel words from descriptive episodes. *Frontiers in Psychology, 8*(488), <https://doi.org/10.3389/fpsyg.2017.00488>.