International Journal of Psychophysiology 109 (2016) 81–91

Contents lists available at ScienceDirect

International Journal of Psychophysiology

journal homepage: www.elsevier.com/locate/ijpsycho

Neural correlates of multimodal metaphor comprehension: Evidence from event-related potentials and time-frequency decompositions

Qingguo Ma [a,b,c,1], Linfeng Hu [b,c,1], Can Xiao [b,c], Jun Bian [b,c], Jia Jin [d], Qiuzhen Wang [b,c,*]

[a] Institute of Neural Management Sciences, Zhejiang University of Technology, Hangzhou, China
[b] School of Management, Zhejiang University, Hangzhou, China
[c] Neuromanagement Laboratory, Zhejiang University, Hangzhou, China
[d] Business School, Ningbo University, Ningbo, China

A R T I C L E   I N F O

Article history:
Received 3 May 2016
Received in revised form 12 August 2016
Accepted 9 September 2016
Available online 10 September 2016

Keywords:
Multimodal metaphor
Event-related potentials
Time-frequency
N300
Delta
Theta

A B S T R A C T

The present study examined the event-related potential (ERP) and time-frequency components correlates with the comprehension process of multimodal metaphors that were represented by the combination of “a vehicle picture + a written word of an animal”. Electroencephalogram data were recorded when participants decided whether the metaphor using an animal word for the vehicle rendered by a picture was appropriate or not. There were two conditions: appropriateness (e.g., sport utility vehicles + tiger) vs. inappropriateness (e.g., sport utility vehicles + cat). The ERP results showed that inappropriate metaphor elicited larger N300 (280–360 ms) and N400 (380–460 ms) amplitude than appropriate one, which were different from previous exclusively verbal metaphor studies that rarely observed the N300 effect. A P600 (550–750 ms) was also observed and larger in appropriate metaphor condition. Besides, the time-frequency principal component analysis revealed that two independent theta activities indexed the separable processes (retrieval of semantic features and semantic integration) underlying the N300 and N400. Delta band was also induced within a later time window and best characterized the integration process underlying P600. These results indicate the specific cognitive mechanism of multimodal metaphor comprehension that is different from verbal metaphor processing, mirrored by several separable processes indexed by ERP components and time-frequency components. The present study extends the metaphor research by uncovering the functional roles of delta and theta as well as their unique contributions to the ERP components during multimodal metaphor comprehension.

1. Introduction

Metaphor is ubiquitous in our daily life, e.g., contained in a famous poem, or an encouraging speech, or a brand's logo. Compared to direct and literal claims, indirect metaphorical claims are often more impressive. In the view of conceptual metaphor theory (CMT), metaphor is more than a matter of language, but of thought and reason that involves complex cognitive processes (Lakoff, 1993). The essence of metaphor is the analogical mapping between two conceptual domains, from the source domain to the target domain (Lakoff, 1993; Coulson and Van Petten, 2002; Yang et al., 2013). This cross-domain mapping necessitates selection of relevant attributes of source domain and filter of irrelevant ones (Gernsbacher and Robertson, 1999). Comprehensions of metaphorical and literal expressions have been verified to involve the similar cognitive mechanism, but the metaphor processing demands more cognitive effort since the ambiguous and remote semantic relationship between source domain and target domain increases the semantic integration difficulty (Coulson and Van Petten, 2002; Yang et al., 2013). Following the CMT, Forceville suggested that verbal expression should not be the only external manifestation of conceptual metaphor, other non-verbal manifestations such as pictures, sounds, gestures, and smells can also convey metaphorical meanings (Forceville and Urios-Aparisi, 2009). Thus he extended the metaphor as monomodal metaphor and multimodal metaphor. Monomodal metaphors are defined as metaphors whose target and source domains are exclusively rendered in the same mode while multimodal metaphors are defined as those in which the target domain and source domain are manifested in different modes such as combinations of image + speech, image + written text, image + sound, image + sound + written text (Forceville, 2007; Ortiz, 2011). Even though the existence of multimodal metaphor has been admitted, most previous empirical studies regarding metaphor were carried out with the use of linguistic stimuli (words or sentences) only. Extending the studies of metaphor from the verbal form to full or partly (i.e., multimodal) non-verbal form is important to understand the mechanism of

* Corresponding author at: School of Management, Zhejiang University, Hangzhou, China.
E-mail address: wqz@zju.edu.cn (Q. Wang).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.ijpsycho.2016.09.007

metaphor comprehension completely and apply it effectively to more fields other than literature. Given the lack of experimental studies, the cognitive mechanism underlying the multimodal metaphor processing has not been clear yet. In present study, we intend to conduct an ERP experiment to explore the cognitive process of multimodal metaphor comprehension by using pictures and Chinese words to express concepts of target domain and source domain respectively.

Event-related potentials (ERP) method has been used to provide a high temporal resolution to uncover the time-course of cognitive mechanism underlying metaphor comprehension. The N400, a negative component of the ERPs occurring around 300–500 ms after onset of stimuli, has been observed in most metaphor studies (Pynte et al., 1996; Coulson and Van Petten, 2002; Arzouan et al., 2007; Cornejo et al., 2009; Ibáñez et al., 2011; Yang et al., 2013; Schneider et al., 2014). N400 can index the semantic processing and is sensitive to semantic incongruence (Kutas and Hillyard, 1980; Pynte et al., 1996; Balconi and Amenta, 2010; Kutas and Federmeier, 2011). An early metaphor study compared the ERPs when participants were reading sentences with literal, literal mapping and metaphorical endings respectively, and found that the N400 amplitude evoked by metaphorical condition was the largest, by literal mapping condition the intermediate and by literal condition the smallest (Coulson and Van Petten, 2002). This finding provides neural evidence for the argument that literal and metaphor comprehensions have the same cognitive mechanism, but the conceptual integration in metaphor comprehension is more difficult. N400 is also modulated by the familiarity of metaphorical expressions and meaningfulness or relatedness of the sentences (Paivio, 1991; Arzouan et al., 2007; Kutas and Federmeier, 2011; Schneider et al., 2014). For example, novel metaphors elicit larger N400 compared to conventional metaphors and literal expressions, whereas the unrelated ones elicit the largest, indicating that unrelated stimuli have the highest semantic incongruence and thus the most difficulty of semantic integration (Arzouan et al., 2007; Schneider et al., 2014). N400 is not specific to verbal stimuli, but also associated with semantic processing of many nonverbal stimuli such as pictures, videos, actions and mathematical symbols (see (Kutas and Federmeier, 2011) for review). For instance, the last picture that was semantically incongruent with preceding picture sequence conveying a story elicited a larger N400 than the semantically congruent picture did (West and Holcomb, 2002). Likewise, the contextually inappropriate movie endings also evoked a greater N400 (Sitnikova et al., 2008). In general, N400 amplitude can reflect the detection of semantic incongruence regardless of modality of stimuli. Another ERP component that occurs during the non-verbal stimuli (especially picture) processing is the N300 peaking at about 300 ms. N300 used to be considered as a sub-potential of N400, and few attention was paid to it until McPherson and Holcomb (1999) found that there were significant differences between N300 and N400 not only in time window but also in topography with N300 distributing more frontally and N400 in posterior areas. They also found N400 was more sensitive to semantic relatedness than N300. N300 has been considered as an image-specific ERP potential and can represent the semantic processing of pictures with the activation of image-based system (Federmeier and Kutas, 2002; Hamm et al., 2002; West and Holcomb, 2002; Franklin et al., 2007). Generally, larger N300 amplitude is elicited in semantically incongruent conditions such as between-category mismatches (Federmeier and Kutas, 2002; Hamm et al., 2002; Franklin et al., 2007) and discourse-level mismatches (West and Holcomb, 2002). A recent metaphor study also found larger N300 occurred in metaphorical sentences processing than literal ones, which implies the contribution of image-based system to metaphor comprehension (Balconi and Amenta, 2010).

Several neuroimaging studies have found the more activation of brain during the metaphor processing in contrast to the literal and meaningless stimuli (Bohrn et al., 2012; Forgács et al., 2012; Rapp et al., 2012; Schneider et al., 2014), reflecting the recruitment of more cognitive resources in comprehending metaphor. Right hemisphere is considered to play a special role in metaphor comprehension. According to the coarse semantic coding theory, each stimulus will elicit a semantic field referring to the information activated in response to the stimulus, and left hemisphere is associated with the narrow semantic field while the right hemisphere for the broad semantic field (Beeman and Chiarello, 1998; Jung-Beeman, 2005; Forgács et al., 2012). Accordingly, right hemisphere is activated more in novel metaphor compared to conventional metaphor and literal languages because of the broad semantic relationship between concepts in novel metaphors (Stringaris et al., 2006; Mashal et al., 2007; Mashal et al., 2009; Diaz et al., 2011). In contrast to predominantly right hemisphere involvement, some metaphor studies also report bilateral hemispheric network (Coulson and Van Petten, 2007; Rapp et al., 2007; Schmidt and Seger, 2009; Balconi and Amenta, 2010) and meta-analyses indicate a predominantly left hemisphere lateralized network (Bohrn et al., 2012; Rapp et al., 2012). Schmidt and Seger (2009) suggested it is the factors including figurativeness, familiarity and difficulty that determine different hemispheric recruitment in semantic processing. Therefore, which hemisphere involvement is not due to metaphor per se, that is to say, metaphor comprehension is supposed to benefit from both hemispheres.

Different frequency bands activity underlying ERP components can index some separable cognitive processes that the ERP may not characterize clearly. A recently proposed Time-Frequency PCA (TF-PCA) approach (Bernat et al., 2005) has been applied to effectively separate activities in different frequency ranges that overlap in time and reveal the relationships between several time-frequency components (e.g., delta, theta) and ERP components (e.g., ERN, N2, FN, P300, N400, P600) in various tasks, for instance, Eriksen flanker task (Bernat et al., 2005; Hall et al., 2007), go/no-go task (Harper et al., 2014), oddball task (Bernat et al., 2007), gambling task (Bernat et al., 2011; Bernat et al., 2015) and lexical decision task (Steele et al., 2013). The time-frequency (TF) analysis method has been used in substantial studies regarding the language processing (Bastiaansen et al., 2002a, 2002b; Hald et al., 2006; Davidson and Indefrey, 2007). Power changes in theta frequency band are most often found sensitive to semantic violation. Generally, theta band activity increases more in response to semantic or syntactic violations and mainly distributes in frontal areas (Bastiaansen et al., 2002a, 2002b; Hald et al., 2006; Davidson and Indefrey, 2007). As theta activity is related with the degree of working memory load, larger activity of theta in the semantic violation condition indicates more difficulty in semantic integration (Bastiaansen et al., 2002a; Sauseng et al., 2010). A number of studies have also suggested that theta activity can indicate the access to memory system to store and retrieve semantic information during language processing (Hagoort et al., 2004; Bastiaansen et al., 2005; Bastiaansen et al., 2008; Maguire et al., 2010). For example, Bastiaansen et al. (2005) found increased theta activity when processing open-class words (e.g., nouns, verb) compared to closed-class words (e.g., conjunctions, prepositions), suggesting retrieval of lexical-semantic properties of the open-class words. Another study discovered that theta was involved in the retrieval of visual semantic properties or auditory semantic properties of a word during lexical decision (Bastiaansen et al., 2008). Delta is another TF component engaged in language processing (Carrus et al., 2011; Giraud and Poeppel, 2012; Harmony, 2013; Mai et al., 2016). Two recent studies have revealed the unique contribution of delta to the N400 and P600 when processing sentences (Roehm et al., 2004; Bernat et al., 2005). Roehm et al. (2004) found a theta (~3.5–7.5 Hz) activity appearing in the N400 time window and a delta (1–3 Hz) activity occurring within the P600 time window during reading grammatical and ungrammatical sentences, suggesting that the theta was associated with the linguistic problem detection and the delta reflected the conflict resolution process. Another study applied the TF-PCA approach mentioned above to decompose the ERP signal and successfully uncover the separable and sequential time-frequency components before, during and after the N400 elicited by target words that were unrelated or inference-related to the preceding text (Steele et al., 2013). Specifically, a theta