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

ARTICLE INFO

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

ABSTRACT

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 non-verbal 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 principal component (1–6 Hz) occurring in the early part of N400 time window reflected detection of incoherence and a subsequent delta principal component (0–2 Hz) during the time window of N400 reflected integration of target words with the preceding text. Later, another delta principal component that was independent of the former occurred during the P600 time window, possibly mirroring further integrative, familiarity, or syntactically based processes (Steele et al., 2013). Although existing studies have verified that delta and theta band activity account for N400 and P600, it is still far from clear which frequency bands contribute to the N300. To our knowledge, using time-frequency analysis to study the multimodal metaphor comprehension has received few attention.

Although multimodal metaphor has been attracting more and more attention, there are still no neural experiments to uncover the specific cognitive mechanism underlying multimodal metaphor comprehension. Since multimodal metaphors are manifested in different modes, the comprehension process would recruit different sensory systems that make it more complex than verbal metaphor. In this regard, the present study intends to provide the electrophysiological evidence including event-related potential (ERP) and time-frequency measures for multimodal metaphor comprehension process. In order to effectively separate the underlying time-frequency components activities and reveal the contributions of TF components to the ERP components, we applied the TF-PCA method (Bernat et al., 2005). EEG data were recorded while participants viewed multimodal metaphors in which a vehicle served as the target domain in the pictorial mode and an animal served as source domain depicted by a written word. The task was to decide whether the metaphor was appropriate or not. Since previous ERP studies have indicated that the N300, N400, P600, delta and theta band index some aspects of semantic processing of both verbal and non-verbal stimuli, we would expect there to be these five neural components occurring in our study with larger N300, N400, P600, delta and theta activity in metaphorically inappropriate condition.

2. Material and methods

2.1. Participants

Eighteen undergraduates (9 females, aging from 19 to 24 years with mean age = 21.5, SD = 2.4) from Zhejiang university participated in this experiment and were paid for their participation. One participant's data was excluded for excessive recording artifacts. All participants were right-handed according to the handedness questionnaire adapted from Oldfield (1971) and had no history of current or past neurological or psychiatric illness. We conducted the experiment with the participants' written consent forms and approval of the Neuromanagement laboratory's ethics committee in Zhejiang University.

2.2. Experimental stimuli

Initial stimuli set included twenty black and white pictures of different kinds of vehicles that were fictitious and twelve Chinese characters depicting animals including tiger, lion, wolf, leopard, rabbit, cat, horse, cow, fox, sheep, dog and deer. Half of the vehicles were categorized into Sport Utility Vehicles (SUV) and the others were minicars, both of which were common in our daily life. Thus we had 240 pairs (20 vehicles × 12 animals) of vehicle-animal combinations (e.g., SUV-tiger, minicar-lion). Then we asked 156 student volunteers who did not participated in the later ERP experiment to rate the metaphorically appropriate degree between the vehicle and the animal in each pair (animal as the source domain and vehicle as the target domain) with the seven-point Likert scale ranging from 1 (the most inappropriate) to 7 (the most appropriate) in a preliminary test. All the 240 pairs were pseudo-randomly presented to the students on the computer. The result showed that the ten SUVs-lion pairs (mean = 5.791, SD = 1.297), SUVs-tiger pairs (mean = 5.744, SD = 1.258), ten minicars-rabbit pairs (mean = 5.651, SD = 1.498) and minicars-cat (mean = 5.116, SD = 1.647) had the highest appropriate degree. On the contrary, the ten SUVs-cat pairs (Mean = 1.496, SD = 0.911), SUVs-rabbit pairs (mean = 1.636, SD = 1.185), ten minicars-lion pairs (mean = 1.465, SD = 0.893) and minicars-tiger pairs (mean = 1.512, SD = 0.894) had the highest inappropriate degree. Thus the above eighty pairs of vehicle-animal were selected as the ERP experiment's stimuli. The appropriate group consisted of these forty pairs including SUVs-lion, SUVs-tiger, minicars-rabbit and minicars-cat, and the inappropriate group consisted of the other forty pairs including SUVs-cat, SUVs-rabbit, minicars-lion, and minicars-tiger. All the pictures of vehicles were processed by Adobe Photoshop 13.0 to have the same luminance, shade and size and the Chinese words of animals were Song typeface with the same font size.

2.3. Experimental procedure

Sentence-final word paradigm and stimulus1–stimulus2 (S1–S2) are the two paradigms commonly used in the metaphor studies (Yang et al., 2013). In the present study, we used the S1-S2 experiment paradigm (see Fig. 1) since it is also used in many non-verbal semantic processing studies (e.g., Federmeier and Kutas, 2002; Hamm et al., 2002), which may be suitable for multimodal metaphor study. S1-S2 experiment paradigm presented two stimuli in sequence. The first stimulus (S1) was the vehicle picture, and the second (S2) was the Chinese word of an animal. The picture-word pairs were randomly presented twice. The participants were asked to sit in the chair comfortably 70 cm away from computer screen in a sound-proof room with a visual angle of 2.58° × 2.4°. There were four blocks with 40 trials each. All the stimuli were presented in the center of screen with grey background. In each trial, a '+' was presented for 500 ms firstly, and the interval of the grey screen between the '+' and S1 was ranging from 500 to 700 ms (average interval was 600 ms). After the grey screen disappeared, a picture of a vehicle (S1) was presented for 1000 ms followed by a varied inter-stimuli interval from 200 to 400 ms (average interval was 300 ms), then a Chinese word of an animal (S2) with a presenting duration of 1000 ms. Once the word of an animal appeared, participants had 1500 ms to make a judgment whether the metaphor using the animal for the preceding vehicle was appropriate or not, and they were asked to response as soon as possible. Half of participants pressed 2 (appropriate) or 3 (inappropriate) and the other half of participants pressed 3 (appropriate) or 2 (inappropriate) for counterbalance. There was another grey screen occurring for 500 ms before the next trail. There was no repetition of either the picture or word stimuli on any two consecutive trials. The participants conducted a brief exercise before the experiment.

2.4. Electroencephalogram data recording

Electroencephalogram data was continuously recorded (band pass 0.05–100 Hz, sampling rate 1000 Hz) with Neuroscan Synamp2 Amplifier (Scan 4.3.1, Neurosoft labs, Inc.), adopting an electrode cap with 64 Ag/AgCl electrodes mounted according to the extended international 10–20 system and referenced to the left mastoid. The vertical electrooculogram was recorded from the infra-orbital and supra-orbital electrodes on the right eye, and the horizontal electroculogram was recorded from electrodes on the outer canthi of both eyes. Electrode impedance was maintained below 5 kΩ in the experiment.

2.5. Data analysis

2.5.1. ERP analysis

EEG data was preprocessed offline by software of Scan 4.5 (Compumedics NeuroScan Inc., Herndon, Virginia, USA). Electroencephalogram recordings were extracted from 200 ms before the onset of the Chinese words of animals to 800 ms after the onset, with the first

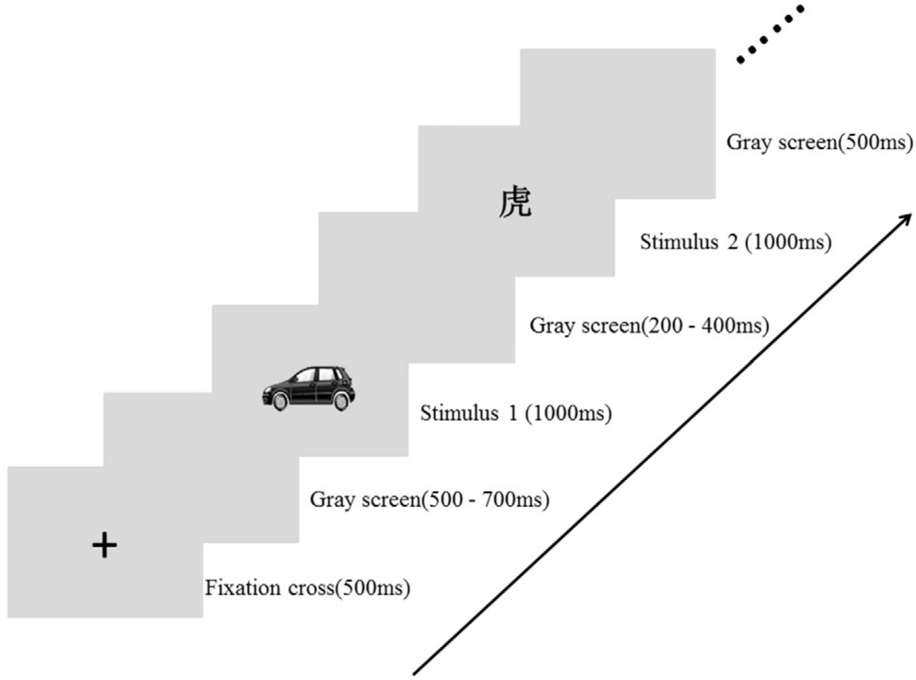

Fig. 1. Experimental procedure. In each trial, the participants viewed a picture of a vehicle (S1) and a Chinese word of an animal (S2) in sequence and judged whether the metaphor using an animal for the vehicle was appropriate or not when the word appeared. They were asked to make the decision as soon as possible by pressing a keypad.

200 ms as baseline. Trials containing amplifier clipping, bursts of electromyography activity, or peak-to-peak deflection exceeding $\pm 80\ \mu V$ and incorrect appropriateness judgment were rejected before averaging. After rejecting all artifacts of EEG, at least 30 overlaps were available for each condition. EEGs were digitally re-referenced to the average of the left and the right mastoids and were filtered with a 30 Hz low-pass filter (24 dB/Octave). At last, the EEG data were averaged separately for appropriate and inappropriate conditions.

Based on the visual inspection of the ERP waveforms and topographic maps (see Fig. 2), two negative components within different time windows (latency at 280–360 and 380–460 ms respectively) in frontal and parietal regions were found. Similar to previous ERP studies (Federmeier and Kutas, 2002; Balconi and Amenta, 2010), these two negative components should be the N300 with distribution in frontal regions and N400 with distribution in parietal regions. Besides, a late positive component (latency at about 550–750 ms) was observed at posterior electrodes, which might be referred to as the P600 reported in previous metaphor studies (Coulson and Van Petten, 2002; Yang et al., 2013). Therefore, A $2 \times 2 \times 3$ within participants repeated-measures analysis of variance (ANOVA) on the mean amplitudes of N300 in the previously mentioned time windows of 280–360 ms was performed in the ERP statistical analysis, taking Type (appropriateness vs. inappropriateness), Localization (F: F1, Fz, F2; FC: FC1, FCz, FC2) and Lateralization (left: F1, FC1; middle: Fz, FCz; right: F2, FC2) as within-participant factors. Similarly, we conducted the 2 (Type: appropriateness vs. inappropriateness) $\times$ 2 (Localization: CP-CP3, CPz, CP4; P-P3, Pz, P4) $\times$ 3 (Lateralization: left-CP3, P3; middle-CPz, Pz; right-CP4, P4) repeated-measures ANOVA on the mean amplitudes of N400 and P600 in the time windows of 380–460 ms and 550–750 ms respectively. The Greenhouse-Geisser correction was used for sphericity departures and Bonferroni method was applied in multiple comparisons when appropriate.

2.5.2. Topographic analysis

To determine whether the N300 was different from the N400 in terms of scalp distribution, we performed a 2 (Type: appropriateness vs. inappropriateness) $\times$ 2 (Component: N300 (280–360 ms) vs. N400 (380–460 ms)) $\times$ 2 (Location: anterior-F3, Fz, F4; posterior-P3, Pz, P4) $\times$ 3 (Laterality: left-F3 & P3; middle-Fz & Pz; right-F4 & P4) repeated-measures ANOVA on the mean amplitudes that were normalized according to Mccarthy and Wood (1985). This topographic analysis method has been applied to determine whether scalp distributions differ across experimental conditions or components in many studies, e.g., (Mcpherson and Holcomb, 1999; Hamm et al., 2002; Murray et al., 2008). The significant interaction effects between Component and Location or Laterality were expected if N300 and N400 indeed showed different scalp distributions (Mcpherson and Holcomb, 1999). The Greenhouse-Geisser and Bonferroni correction were applied when appropriate.

2.5.3. Time-frequency decompositions

First, we extracted a longer EEG epoch window from 500 ms before and 1500 ms after the onset of the Chinese words to ensure that edge effects would not contaminate the time windows of interest (Cohen et al., 2007). As mentioned above, we mainly focused on the delta and theta activities. Therefore, a 3rd order high-pass Butterworth filter at 1 Hz and a low-pass at 7 Hz for theta activity and a 3rd order low-pass Butterworth filter at 3 Hz for delta activity were conducted to condition-averaged ERP waveforms (butter function, Matlab, Inc.). This overlapping filtering had the advantage of eschewing artificially isolating activities in different frequency-bands and showing where the overlapping activity was best characterized through PCA analysis (Steele et al., 2013). Next, the filtered signals were transformed into time-frequency energy distributions for the entire epoch using the binomial reduced interference distribution variant of Cohen's class of time-frequency transformations (Bernat et al., 2005). Finally, PCA was applied to each time-frequency distributions (TFDs) in the time range from 0 to 781.25 ms and frequency range from 0 to 10 Hz for delta and theta bands respectively. This TF-PCA approach is detailed in Bernat et al. (2005), which extracts the primary time-frequency components accounting for the greatest variances across all time-frequency points. The TF-PCA analyses were carried out by Psychophysiology Tool Box developed by Edward Bernat (download in the website of http://www.ccnlab.umd.edu/Psychophysiology_Toolbox).

Three components were extracted based on the scree plots (see Figs. 3 & 4) for both delta and theta bands, which accounted for a total of

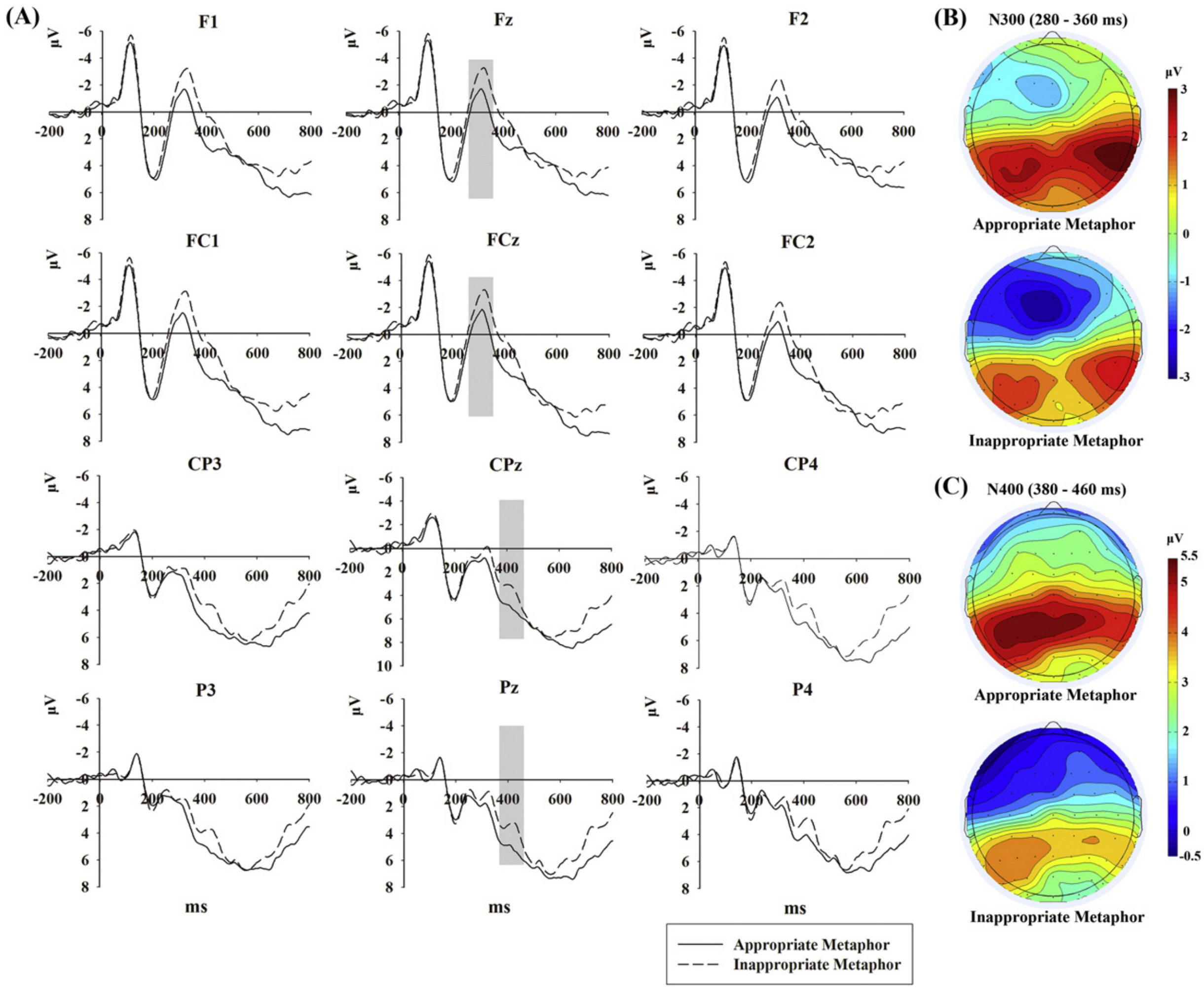

Fig. 2. (A) Grand averaged ERPs elicited by appropriate metaphors versus inappropriate metaphors at 12 electrodes at the frontal and parietal areas; (B) Topographic maps of N300 (280–360 ms) in appropriate condition (top) and inappropriate condition (bottom). (C) Topographic maps of N400 (380–460 ms) in appropriate condition (top) and inappropriate condition (bottom).

75.44% (delta) and 63.96% (theta) of variance in the decompositions respectively. These six principal components scores served as dependent measures in the following analysis. According to topographic maps of delta PCs (see Fig. 3), the obvious differences between the appropriate and inappropriate conditions mainly distributed in the fronto-central and centro-parietal regions. Thus, the nine electrodes (FC3, FCz, FC4, C3, Cz, C4, CP3, CPz, CP4) were selected for delta analysis. A $2 \times 3 \times 3$ within participants repeated-measures analysis of variance (ANOVA) on the three PCs was performed separately, taking Type (appropriateness vs. inappropriateness), Localization (FC: FC3, FCz, FC4; C: C3, Cz, C4; CP: CP3, CPz, CP4) and Lateralization (left: FC3, C3, CP3; middle: FCz, Cz, CPz; right: FC4, C4, CP4) as within-participant factors. Based on the visual inspection of topographic maps of theta PCs (see Fig. 4), a pronounced frontally-distributed inappropriate-appropriate difference was identified. Therefore, we chose the six electrodes including F3, Fz, F4, FC3, FCz, FC4 for theta analysis. Similarly, we performed a $2 \times 2 \times 3$ within participants repeated-measures analysis of variance (ANOVA) on three theta PCs respectively, taking Type (appropriateness vs. inappropriateness), Localization (F: F3, Fz, F4; FC: FC3, FCz, FC4) and Lateralization (left: F3, FC3; middle: Fz, FCz; right: F4, FC4) as within-participant factors. The Greenhouse-Geisser and Bonferroni correction were applied when appropriate.

2.5.4. Regression analyses

We performed the regression analyses to assess the relationships between TF components (i.e., delta and theta) and time-domain components (N300, N400, P600). In each regression, TF components served as the predictor variables while TD components served as the dependent variables. It's worth noting that only TF components that exhibited significant experimental condition differences were used for regression analyses. Besides, the delta-PC2 was not selected for the regression analysis of N300 or N400 because this component occurred later than N400 time window. Delta-PC3 was also excluded for N300 analysis since it was less activated and seemed not different between the appropriate and inappropriate conditions within N300 time window. Similarly, the theta-PC2 was also not chosen for the regression analysis of P600 because this component occurred in the time window about 200–400 ms that was earlier than P600. The averaged inappropriate minus appropriate difference scores were used for both predictor and dependent variables. The regression models were established with the backward method in SPSS statistical software (SPSS Inc., Chicago, Illinois, USA) and only the results of final models were reported. Additional Pearson correlation analysis based on the grand averages of TF measures was conducted to further dissociate the theta and delta components.

3. Results

3.1. Behavioral results

Behavioral result showed that participants had longer response time (700.394 ± 202.336 ms) in appropriate condition than inappropriate

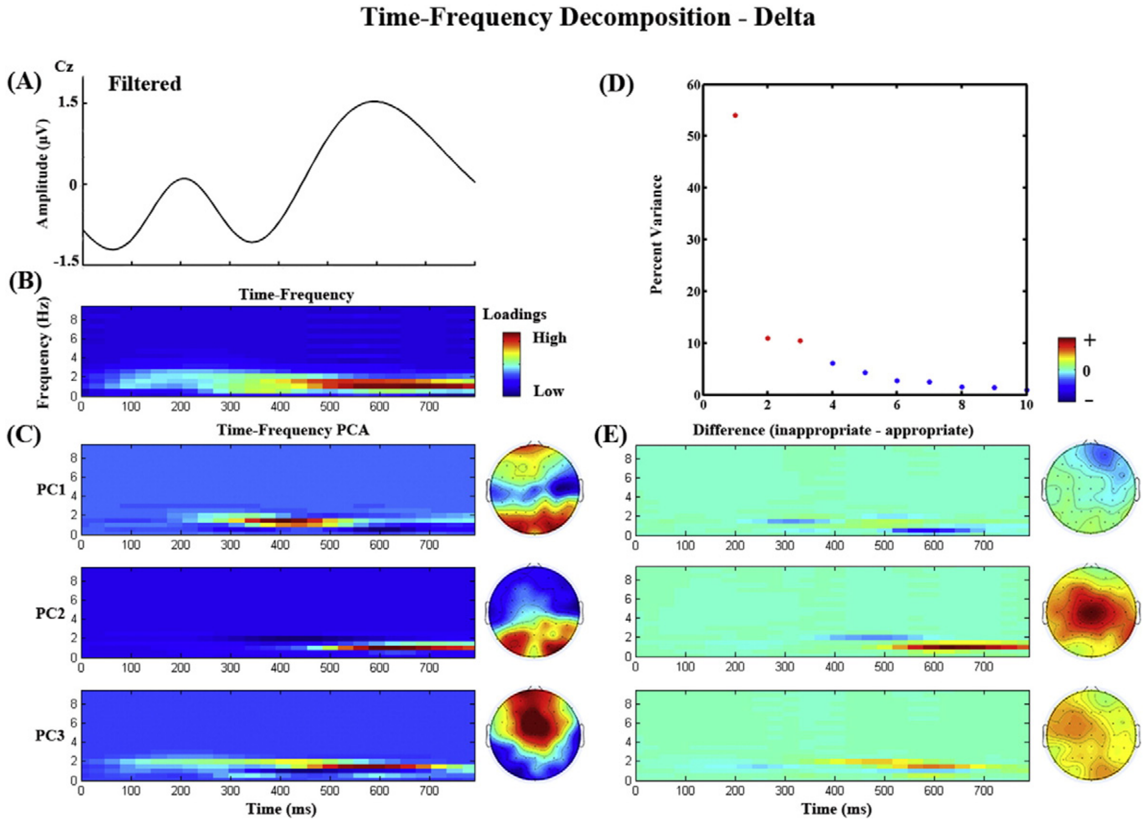

Fig. 3. Time-frequency (TF) decomposition of delta. (A) Grand averaged delta (<3 Hz) filtered ERP at Cz during the time window of 0–800 ms. (B) Grand averaged TF representation of delta filtered ERP in the time range from 0 to 781.25 ms and frequency range from 0 to 10 Hz. (C) Grand averaged TF representations of delta principal component 1–3 in the time range from 0 to 781.25 ms and frequency range from 0 to 10 Hz accompanied by their topographic maps. (D) Scree plot of percent variance accounted for by the first 10 components (the three extracted components were highlighted in red). (E) Grand averaged inappropriate-appropriate difference TF representations of delta principal component 1–3 in the time range from 0 to 781.25 ms and frequency range from 0 to 10 Hz accompanied by their topographic maps.

condition (663.019 ± 187.51 ms), and the Paired *t*-test result indicated that there existed significant difference between the two conditions ($t = 2.405$, $P = 0.029$).

3.2. ERP results

ANOVA results of N300 indicated the significant main effects of Type ($F_{(1, 16)} = 7.966$, $P = 0.012$, $\eta^2 = 0.332$) and Lateralization ($F_{(2, 32)} = 10.281$, $P < 0.001$, $\eta^2 = 0.391$). But no significant main effect of Localization was observed ($F_{(1, 16)} < 1$, $P = 0.652$, $\eta^2 = 0.013$). The average amplitude of N300 in inappropriate condition (mean = −2.226, SE = 0.922) was more negative than that in appropriate condition (mean = −0.643, SE = 1.108). For lateralization effect, the pairwise comparison showed significant difference between left and right lines ($P_{left, right} = 0.011$, $M_{left} = -1.646$, $SE_{left} = 0.934$, $M_{right} = -0.941$, $SE_{right} = 0.997$) and middle and right lines ($P_{middle, right} = 0.005$, $M_{middle} = -1.716$, $SE_{middle} = 1.024$), but not between left and middle lines ($P_{left, middle} > 0.5$). There were not significant interaction effects of factors.

For N400 analysis, ANOVA results showed that only main effect of Type was significant ($F_{(1, 16)} = 9.760$, $P = 0.007$, $\eta^2 = 0.379$). The effects of Localization ($F_{(1, 16)} < 1$, $P = 0.717$, $\eta^2 = 0.008$) and Lateralization ($F_{(2, 32)} < 1$, $P = 0.562$, $\eta^2 = 0.035$) were not significant. More negative N400 was observed in inappropriate condition than that in appropriate condition ($M_{inappropriate} = 3.694$, $SE_{inappropriate} = 0.958$; $M_{appropriate} = 5.087$, $SE_{appropriate} = 0.949$). No significant interaction effects of factors were observed.

ANOVA results of P600 revealed the significant main effects of Type ($F_{(1, 16)} = 6.885$, $P = 0.018$, $\eta^2 = 0.301$) and Lateralization ($F_{(2, 32)} = 3.977$, $P = 0.029$, $\eta^2 = 0.199$). But no significant main effect of Localization was observed ($F_{(1, 16)} = 2.508$, $P = 0.133$, $\eta^2 = 0.136$). The average amplitude of P600 in appropriate condition (mean = 6.671, SE = 0.845) was more positive than that in inappropriate condition (mean = 5.473, SE = 0.633). For Lateralization effect, the pairwise comparison showed significant difference only between left and middle lines ($P_{left, middle} = 0.017$).

Topographic analysis results showed that the interaction effects of Component × Location ($F_{(1, 16)} = 5.490$, $P = 0.032$, $\eta^2 = 0.255$) and Component × Laterality ($F_{(2, 32)} = 12.845$, $P = 0.001$, $\varepsilon = 0.717$, $\eta^2 = 0.445$) were both significant, which illustrated the scalp distributions of N300 and N400 were different. A further analysis of Component × Location indicated that Location simple effect was significant for the N300 ($F_{(1, 16)} = 5.725$, $P = 0.029$, $\eta^2 = 0.264$), but not for the N400 ($F_{(1, 16)} = 0.805$, $P = 0.383$, $\eta^2 = 0.048$). The distribution of N300 showed a more anterior negative than posterior one, whereas the distributions of the N400 were similar across anterior-posterior regions. Besides, Laterality simple effects were significant for both N300 and N400 ($F_{(2, 32)} = 9.242$, $P = 0.002$, $\varepsilon = 0.735$, $\eta^2 = 0.366$ for N300; $F_{(2, 32)} = 4.401$, $P = 0.035$, $\varepsilon = 0.717$, $\eta^2 = 0.216$ for N400). With respect to the pairwise comparisons in N300, significant differences between left and right lines ($P_{left, right} = 0.01$) and middle and right lines ($P_{middle, right} = 0.02$) were observed, whereas no such difference between left and middle lines ($P_{left, middle} = 0.152$) was found. In contrast, only difference between left and middle lines reached significant ($P_{left, middle} = 0.028$) in N400 pairwise comparisons.

3.3. TF-PCA results

In delta activity decomposition, only delta-PC2 and delta-PC3 exhibited significant main effects of Type ($F_{(1, 16)} = 5.168$, $P = 0.037$, $\eta^2 =$

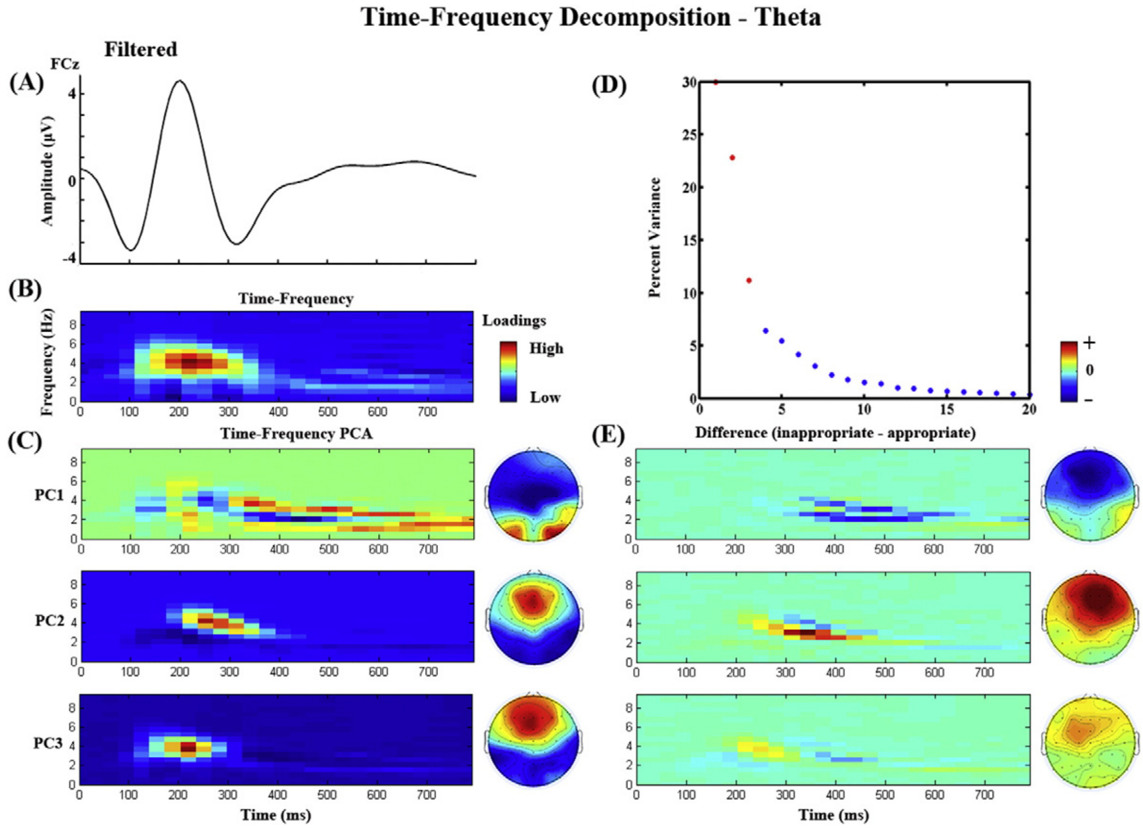

Fig. 4. Time-frequency (TF) decomposition of theta. (A) Grand averaged theta (1–7 Hz) filtered ERP at FCz during the time window of 0–800 ms. (B) Grand averaged TF representation of theta filtered ERP in the time range from 0 to 781.25 ms and frequency range from 0 to 10 Hz. (C) Grand averaged TF representations of theta principal component 1–3 in the time range from 0 to 781.25 ms and frequency range from 0 to 10 Hz accompanied by their topographic maps. (D) Scree plot of percent variance accounted for by the first 20 components (the three extracted components were highlighted in red). (E) Grand averaged inappropriate-appropriate difference TF representations of theta principal component 1–3 in the time range from 0 to 781.25 ms and frequency range from 0 to 10 Hz accompanied by their topographic maps.

0.244 for PC2; $F_{(1, 16)} = 4.959$, $P = 0.041$, $\eta^2 = 0.237$ for PC3). For PC2, delta activity in inappropriate condition was significantly greater than that in appropriate condition. The Localization ($P = 0.129$) and Lateralization ($P = 0.548$) effects of PC2 were not significant. For PC3, delta activity in inappropriate condition was significantly greater than that in appropriate condition. The Localization effect ($P = 0.129$) of PC3 was not significant while the Lateralization effect was significant ($F_{(2, 32)} = 4.140$, $P = 0.025$, $\eta^2 = 0.206$). No significant interaction effects of factors were found for both PC2 and PC3.

In theta activity decomposition, only theta-PC1 and theta-PC2 showed significant main effects of Type ($F_{(1, 16)} = 6.918$, $P = 0.018$, $\eta^2 = 0.302$ for PC1; $F_{(1, 16)} = 5.404$, $P = 0.034$, $\eta^2 = 0.252$ for PC2). For PC1, theta activity was less activated in inappropriate condition than that in appropriate condition. On the contrary, theta activity showed more activated in inappropriate condition than that in appropriate condition for PC2. Besides, the Localization ($P = 0.176$) and Lateralization effects ($P = 0.473$) of PC1 were not significant, whereas both Localization ($F_{(2, 32)} = 11.828$, $P = 0.003$, $\eta^2 = 0.425$) and Lateralization effects ($F_{(2, 32)} = 6.871$, $P = 0.003$, $\eta^2 = 0.3$) were significant for PC2. Theta activity was greater in more anterior sites and significantly different only between middle and right lines ($P = 0.004$). There were not significant interaction effects of factors for PC1 and PC2.

3.4. Regression analysis results

According to the results in Table 1, the three regression models that predicted difference-scores between appropriate and inappropriate conditions of TD components from those of TF components were all significant. In model 1, the theta-PC1 and theta-PC2 difference-scores accounted for 66.1% of the variance in N300 score and both of theta-PC1 and theta-PC2 were unique predictors. In model 2, the theta-PC2 was the only TF component predicting difference-scores of N400 and explaining 23.0% of the variance. In model 3, delta-PC2 and delta-PC3 were unique predictors and accounted for 35.0% of the variance in P600 difference-scores. Pearson correlation analysis showed only delta-PC2 was significantly correlated with theta-PC1 ($r = 0.730$, $P = 0.001$). No other significant correlated relationship between delta and theta components were observed (all $P > 0.1$).

4. Discussion

The present study used the combination of "a vehicle picture + a written word of an animal" to explore the cognitive process of multimodal metaphor comprehension by event-related potential and time-frequency measurements. Besides the typical N400 effect, a N300

Table 1
Results of regression analysis in which time-frequency components served as the predictor variables while ERP components served as the dependent variables.

	Dependent variable	Predictors	Beta	F	Adjust R^2
Model 1	N300	Theta-PC1	0.340*	16.599***	0.661
		Theta-PC2	−0.651**		
Model 2	N400	Theta-PC1	0.527*	7.607*	0.230
Model 3	P600	Delta-PC2	−0.545*	5.779*	0.350
		Delta-PC3	0.577*		

* $P < 0.05$.
** $P < 0.01$.
*** $P < 0.001$.

component was also found. Compared to appropriate metaphor condition, larger N300 and N400 were observed in inappropriate metaphor condition. These results suggest multimodal metaphor comprehension shares similar cognitive mechanism regarding cross-domain mapping with verbal metaphor, but has an additional semantic categorization processing earlier. Participants in the present study might comprehend the multimodal metaphor by recruiting both image-based and language-based processing systems. A late positive component (i.e., P600) was also observed as previous metaphor studies (Coulson and Van Petten, 2002; Yang et al., 2013). Specifically, larger P600 was elicited in appropriate condition than that in inappropriate condition. Furthermore, we observed delta and theta activity as expected and discovered their unique contributions to the N300, N400 and P600 separately. In detail, theta-PC2 activity underlying N300 and delta-PCs activities underlying P600 were enhanced in inappropriate condition, whereas theta-PC1 underlying N400 showed the opposite effect. The present study implies the specific cognitive mechanism of multimodal metaphor comprehension that is different from verbal metaphor processing by effectively revealing dissociable processes underlying the ERP components.

Multimodal metaphor is constructed by target domain and source domain that are manifested in different representations (Forceville, 2007). In our study, the source domains were the animals depicted by written Chinese words while the target domains were the vehicles descripted using pictures. As we know, metaphor comprehension is an analogical mapping process that involves selection of specific features of the source domain (Gernsbacher and Robertson, 1999). Therefore, the way we design the representation formats of conceptual domains will influence the metaphor processing. For example, in a pictorial metaphor, some features need not to be retrieved from memory before mapping since they have been given in the picture (Forceville, 2007). Thus, pictorial metaphors are easier and less effortful to comprehend than verbal metaphors. In the condition of pictorial metaphor, human firstly need to extract information in the pictures. It is inferable that multisensory perceptions may be necessary to figure out the resemblances between target and source domains in multimodal metaphor comprehension. As for the present experiment design, pictorial information needed to be preprocessed and integrated into the context which would recruit image-based system.

We found a frontally-distributed negativity N300 occurred earlier than N400 in both appropriate and inappropriate conditions. In the light of previous studies, N300 is a specific ERP marker of the semantic processing of pictorial stimuli (Mcpherson and Holcomb, 1999; Federmeier and Kutas, 2002; Hamm et al., 2002; West and Holcomb, 2002). The occurrence of N300 in the present study suggests that multimodal metaphor comprehension involves an additional image-based semantic processing. In contrast to previous studies, the stimuli eliciting the N300 were the written words depicting animals rather than the pictures. N300 was not reported in most previous verbal metaphor studies except a recent work in which metaphor sentence elicited larger N300 than literal ones, suggesting the cross-domain mapping in metaphor comprehension may perform through a mental imagery (Balconi and Amenta, 2010). According to the dual-coding theory, concrete word is processed simultaneously by the image processing system and the semantic processing system (Paivio, 1991). Therefore, concrete word can also activate image-based processing like pictures do, which can be indexed by N300. Our results provide evidence for this theory and metaphor indeed can be processed both by language and image processing systems (Holcomb et al., 1999; West and Holcomb, 2002). Several studies find more evident N300 in between-category mismatches than within-category mismatches (Federmeier and Kutas, 2002; Hamm et al., 2002; Franklin et al., 2007). For example, in Hamm et al.'s (2002) study, word "magpie" followed by the picture of collie would elicited the N300 whereas huskie-collie pair would not. Thus, N300 is influenced by categorical relatedness (Hamm et al., 2002; Franklin et al., 2007). Similar to these studies, we found inappropriate metaphors elicited larger N300 amplitude than that the appropriate metaphors did. Since the animal and vehicle were not within the same category, this between-category mismatch would account for elicitation of N300 effect in our study. Besides, this categorization process might be partly based on the visual similarity comparison through mental image (Federmeier and Kutas, 2002; Maguire et al., 2010). There were larger mismatches for the inappropriate pairs because of clearer distinctions between the vehicles and animals in visual features, leading to greater N300 in inappropriate condition than appropriate condition. These results imply that the semantic categorization of objects proceed at early stage. With respect to the scalp distribution, N300 was more obvious in the frontal area, which was in agreement with previous studies that suggest frontal brain area is involved in picture processing (Mcpherson and Holcomb, 1999; West and Holcomb, 2002). Besides, N300 effect also showed a left-lateralized distribution. This result was consistent with Federmeier and Kutas (2002), who suggested left hemisphere was more sensitive to process between-category mismatch. Moreover, the categorization process involved extraction of visual features of vehicles and animals. The left hemisphere is more likely to be recruited since it is quick to activate contextually relevant semantic features of the upcoming stimuli (Federmeier and Kutas, 1999, 2002; Jung-Beeman, 2005).

In line with previous ERP studies about the metaphor (Arzouan et al., 2007; Balconi and Amenta, 2010; Schneider et al., 2014), a robust N400 effect was observed in current study. Inappropriate metaphors elicited greater N400 amplitude than appropriate ones, supporting the argument that the semantic incongruence between source and target domains is responsible for the larger N400 (Arzouan et al., 2007; Yang et al., 2013). This N400 effect indicates the similar cognitive mechanism between exclusively verbal metaphor and multimodal metaphor regarding the semantic processing at later stage. As for the scalp distribution, the N400 mainly distributed in parietal regions, which was the same as previous finding (Kutas and Federmeier, 2011). Contrary to the dominance of either left or right hemisphere in previous metaphor studies, we did not observe a lateralized distribution of N400. Based on coarse semantic coding theory, it is postulated that left hemisphere predominantly processes close semantic relationships and right hemisphere predominantly processes distant semantic relationships (Schmidt et al., 2007). Thus, left hemisphere is more activated for familiar metaphors while right hemisphere for novel ones (Forgács et al., 2012). In our study, the multimodal metaphor might be neither novel nor familiar to the participants. It is because the metaphors with representations of "vehicles + animals" have been used in many vehicle advertisements in daily life and some logo of vehicles are designed using animals such as Porsche with horse and Peugeot with lion. However, the stimuli of vehicles in our study were fictitious and thus not identical with vehicles already existing in reality. Hence, the homogeneous contribution between the two hemispheres was possible.

By applying TF-PCA approach, we extracted two meaningful delta and theta principal components respectively, which exhibited significant condition effect (see Figs. 3 & 4). The regression analyses indicated that the theta-PC1 and theta-PC2 accounted for the N300 while only theta-PC1 accounted for the N400. In addition, P600 was characterized best by delta-PC2 and delta-PC3. As shown in the Fig. 4, theta-PC2 occurred within the time window corresponding best to N300 and had a frontal distribution. Previous studies have found frontal theta activity is modulated during processing of taxonomic semantic relationships and category judgment (Brickman et al., 2005; Maguire et al., 2010). As discussed above, semantic categorization processing proceeds in the N300 time window. Taken together, the N300-theta here exactly reflects the categorization process. According to previous studies, theta is implicated with storage and retrieval of semantic information from memory during language processing (Hagoort et al., 2004; Bastiaansen et al., 2005; Bastiaansen et al., 2008; Maguire et al., 2010). We know that metaphor comprehension is considered as an analogical mapping process that involves selection of specific features of the source domain (animal words in the current study) (Gernsbacher and Robertson, 1999) and semantic categorization processing is suggested to rely on perceptual features and require additional attention processes to identify the category relationship between objects (Maguire et al., 2010). Therefore, the N300-theta may characterize the retrieval of semantic properties (e.g., in terms of visual features) related to the vehicle and animal from semantic memory for category relationship identification (Maguire et al., 2010) and analogical mapping (Gernsbacher and Robertson, 1999). Due to the distant relationship between source domain and target domain in inappropriate metaphor, the participants needed to search more semantic information to compare, resulting in larger theta activity. Difference in theta-PC1 activity between appropriate and inappropriate conditions begun at the late part of N300 time window and reached largest during the N400 time window. Theta underlying N400 has been suggested related to working memory process that is involved in the semantic integration (Bastiaansen et al., 2002a, 2002b; Hald et al., 2006; Davidson and Indefrey, 2007). We speculated that the semantic integration might occur soon after the semantic categorization process. An early study proffered that the retrieval and integration of word meanings proceeded at the same time when reading sentences (Hagoort et al., 2004). However, our results indicate that these two processes could be dissociable (at least in the task of the current study) and the retrieval process precedes integration process. Within the N300 time window, the retrieval of semantic information indexed by theta-PC2 is the predominant process, whereas integration process indexed by the theta-PC1 occupies dominance during N400 time window.

N400 was found best accounted for by theta-PC1, which is in agreement with existing studies (Bastiaansen et al., 2002a, 2002b; Hald et al., 2006; Davidson and Indefrey, 2007). Previous language studies with time-frequency analysis suggested that theta activity increased in semantic violation conditions because there were more difficulty in semantic integration, which demanded more effort (Hald et al., 2006; Davidson and Indefrey, 2007). In contrast with these studies, our study found larger activity of theta band induced in appropriate condition than inappropriate condition. This inverse result suggests that the later semantic integration may not take place under inappropriate condition, which is consistent with a recent near-infrared spectroscopy (NIRS) study, in which metaphors elicited higher O_2HB than meaningless sentences (Schneider et al., 2014). The authors attribute this finding to the stronger ambiguity for metaphorical sentences regarding their meaningfulness. The ambiguity makes it more difficult to judge the metaphor as meaningful/appropriate or not compared to literal and meaningless sentences. In our study, therefore, judging inappropriate metaphors would be easier and once the animal was successfully detected as metaphorically inappropriate for the preceding vehicle, no further semantic integration was needed. However, as long as the ambiguity that whether the animal was metaphorically appropriate for the preceding vehicle or not still existed, the participant had to make further semantic integration, i.e. invest more effort to make the semantic integration, leading to a relative increase in theta activity. This speculation was also supported by our behavioral result of significantly longer response time in appropriate condition compared to inappropriate condition. Another potential factor might be the experimental stimuli of multimodal metaphors, in which the source domain (animal) and the target domain (vehicle) were in different categories. This difference in category made the judgment of “inappropriateness” under the inappropriate condition easier, whereas the judgment of “appropriateness” under the appropriate condition needed to semantically integrate across different categories, which required more effort and thus resulted in apparently greater activity of theta band.

In order to determine that N300 was different from N400, we conducted a topographic analysis as previous studies (Mcpherson and Holcomb, 1999; Hamm et al., 2002; Murray et al., 2008). The results verified the difference between N300 and N400 in terms of different scalp distributions. Specifically, N300 mainly distributed in the frontal areas while N400 had a widespread distribution. The N400 amplitude from anterior to posterior areas were not different. Frontally-distributed N400 is found when semantic processing pictorial stimuli (Willems et al., 2008; Kutas and Federmeier, 2011). As discussed above, the processing of animal word would recruit both image and semantic processing system in the current study (Paivio, 1991), which resulted in both anterior and posterior distributed N400. Our TF-PCA analysis results provided additional evidence to bear out this argument that N300 and N400 are two distinct components. As mentioned above, N300 was best characterized by theta-PC2 that reflected a retrieval of semantic information process while N400 was best characterized by theta-PC1 that mirrored the integration process. It indicates separable processes underlying N300 and N400.

P600 has been suggested sensitive to the syntactic violations (Bastiaansen et al., 2002b; Roehm et al., 2004; Davidson and Indefrey, 2007) as well as semantic anomalies (Kolk et al., 2003; Kuperberg et al., 2003; Van Herten et al., 2005). Recent metaphor studies also have observed the P600 effect following the N400 and considered it as a reflection of reanalyzing the information to ensure the accurate comprehension of the metaphor sentences (Coulson and Van Petten, 2002; De Grauwe et al., 2010; Yang et al., 2013). For example, participants would reanalyze the semantic features of the words in the metaphorical sentence to make sure the metaphorical meanings, which resulted in larger amplitude of P600 (Yang et al., 2013). Coulson and Van Petten (2002) suggested that the participant would retrieve more information from semantic memory for blending to ensure the metaphorical interpretation made sense. Thus, P600 may serve as an indicator of checking veridicality of one's analysis (Kolk et al., 2003; Van Herten et al., 2005). Delta has been found associated with the integration of information and attenuated when the integration is difficult (Roehm et al., 2004; Steele et al., 2013). Due to the ambiguity of appropriate metaphor, the participants were more likely to reanalyze appropriate metaphor to ensure the metaphorical meaning, reflected by larger P600. Moreover, delta activity underlying P600 may reflect an additional process of semantic integration to make sure whether the multimodal metaphor was appropriate or not. Smaller delta activity in appropriate condition revealed more effort to make the semantic integration.

5. Conclusion

In sum, the present study, to our knowledge, is the first time to use electrophysiological measures combining event-related potential (ERP) and time-frequency measures to examine multimodal metaphor comprehension process. The results suggest that comprehension of multimodal metaphor represented by combinations of pictorial and verbal modalities experiences several separable processes that are successfully characterized by ERP components (i.e., N300, N400 and P600) and the underlying TF components (i.e., delta and theta). Specifically, an early semantic categorization by activating image-based processing system was reflected by N300 and retrieval of semantic feature process was indexed by underlying theta activity in N300 time window. Subsequently, larger N400 was elicited by detecting semantic incongruence and another theta activity underlying N400 indexed the semantic integration process. At later stage, a reanalysis process was reflected by P600 and the additional integration process was indexed by delta activity. The combination of ERP and time-frequency analysis serves to yield a more complete picture of cognitive mechanism underlying the multimodal metaphor comprehension. It will be interesting to explore the functional roles of more frequency bands activity (alpha, beta and gamma) and their contributions to the ERP components during the multimodal metaphor comprehension. Different cognitive processes may be expected when applying other non-verbal stimuli (such as sound, video) to present metaphor concepts in multimodal metaphor research.

Acknowledgments

This work was supported by grants No.71371167 and No.71272167 from the National Natural Science Foundation of China, No. 09JZD0006 from the State Education Ministry of China as a key project and No. 2015N093 from the Social Science Association of Zhejiang Province as an annual project.

References

Arzouan, Y., Goldstein, A., Faust, M., 2007. Brainwaves are stethoscopes: ERP correlates of novel metaphor comprehension. Brain Res. 1160, 69–81. http://dx.doi.org/10.1016/j.brainres.2007.05.034.

Balconi, M., Amenta, S., 2010. "A fighter is a lion". Neuropsychological indexes in comprehending frozen metaphors. J. Pragmat. 42, 3246–3257. http://dx.doi.org/10.1016/j.pragma.2010.06.016.

Bastiaansen, M.C., Oostenveld, R., Jensen, O., Hagoort, P., 2008. I see what you mean: theta power increases are involved in the retrieval of lexical semantic information. Brain Lang. 106, 15–28. http://dx.doi.org/10.1016/j.bandl.2007.10.006.

Bastiaansen, M.C., Van Berkum, J.J., Hagoort, P., 2002a. Event-related theta power increases in the human EEG during online sentence processing. Neurosci. Lett. 323, 13–16. http://dx.doi.org/10.1016/S0304-3940(01)02535-6.

Bastiaansen, M.C., Van Berkum, J.J., Hagoort, P., 2002b. Syntactic processing modulates the θ rhythm of the human EEG. NeuroImage 17, 1479–1492. http://dx.doi.org/10.1006/nimg.2002.1275.

Bastiaansen, M.C., Van, D.L.M., Ter, K.M., Dijkstra, T., Hagoort, P., 2005. Theta responses are involved in lexical-semantic retrieval during language processing. J. Cogn. Neurosci. 17, 530–541. http://dx.doi.org/10.1162/0898929053279469.

Beeman, M., Chiarello, C., 1998. Right Hemisphere Language Comprehension: Perspectives from Cognitive Neuroscience. Psychology Press.

Bernat, E.M., Malone, S.M., Williams, W.J., Patrick, C.J., Iacono, W.G., 2007. Decomposing delta, theta, and alpha time-frequency ERP activity from a visual oddball task using PCA. Int. J. Psychophysiol. 64, 62–74. http://dx.doi.org/10.1016/j.ijpsycho.2006.07.015.

Bernat, E.M., Nelson, L.D., Baskin-Sommers, A.R., 2015. Time-frequency theta and delta measures index separable components of feedback processing in a gambling task. Psychophysiology 52, 626–637. http://dx.doi.org/10.1111/psyp.12390.

Bernat, E.M., Nelson, L.D., Steele, V.R., Gehring, W.J., Patrick, C.J., 2011. Externalizing psychopathology and gain-loss feedback in a simulated gambling task: dissociable components of brain response revealed by time-frequency analysis. J. Abnorm. Psychol. 120, 352–364. http://dx.doi.org/10.1037/a0022124.

Bernat, E.M., Williams, W.J., Gehring, W.J., 2005. Decomposing ERP time–frequency energy using PCA. Clin. Neurophysiol. 116, 1314–1334. http://dx.doi.org/10.1016/j.clinph.2005.01.019.

Bohrn, I.C., Altmann, U., Jacobs, A.M., 2012. Looking at the brains behind figurative language—a quantitative meta-analysis of neuroimaging studies on metaphor, idiom, and irony processing. Neuropsychologia 50, 2669–2683. http://dx.doi.org/10.1016/j.neuropsychologia.2012.07.021.

Brickman, A.M., Paul, R.H., Cohen, R.A., Williams, L.M., Macgregor, K.L., Jefferson, A.L., et al., 2005. Category and letter verbal fluency across the adult lifespan: relationship to EEG theta power. Arch. Clin. Neuropsychol. 20, 561–573. http://dx.doi.org/10.1016/j.acn.2004.12.006.

Carrus, E., Koelsch, S., Bhattacharya, J., 2011. Shadows of music-language interaction on low frequency brain oscillatory patterns. Brain Lang. 119, 50–57. http://dx.doi.org/10.1016/j.bandl.2011.05.009.

Cohen, M.X., Elger, C.E., Ranganath, C., 2007. Reward expectation modulates feedback-related negativity and EEG spectra. NeuroImage 35, 968–978. http://dx.doi.org/10.1016/j.neuroimage.2006.11.056.

Cornejo, C., Simonetti, F., Ibáñez, A., Aldunate, N., Ceric, F., López, V., et al., 2009. Gesture and metaphor comprehension: electrophysiological evidence of cross-modal coordination by audiovisual stimulation. Brain Cogn. 70, 42–52. http://dx.doi.org/10.1016/j.bandc.2008.12.005.

Coulson, S., Van Petten, C., 2002. Conceptual integration and metaphor: an event-related potential study. M&C. 30, 958–968. http://dx.doi.org/10.3758/BF03195780.

Coulson, S., Van Petten, C., 2007. A special role for the right hemisphere in metaphor comprehension? ERP evidence from hemifield presentation. Brain Res. 1146, 128–145. http://dx.doi.org/10.1016/j.brainres.2007.03.008.

Davidson, D.J., Indefrey, P., 2007. An inverse relation between event-related and time–frequency violation responses in sentence processing. Brain Res. 1158, 81–92. http://dx.doi.org/10.1016/j.brainres.2007.04.082.

De Grauwe, S., Swain, A., Holcomb, P.J., Ditman, T., Kuperberg, G.R., 2010. Electrophysiological insights into the processing of nominal metaphors. Neuropsychologia 48, 1965–1984. http://dx.doi.org/10.1016/j.neuropsychologia.2010.03.017.

Diaz, M.T., Barrett, K.T., Hogstrom, L.J., 2011. The influence of sentence novelty and figurativeness on brain activity. Neuropsychologia 49, 320–330. http://dx.doi.org/10.1016/j.neuropsychologia.2010.12.004.

Federmeier, K.D., Kutas, M., 1999. Right words and left words: electrophysiological evidence for hemispheric differences in meaning processing. Cogn. Brain Res. 8, 373–392. http://dx.doi.org/10.1016/S0926-6410(99)00036-1.

Federmeier, K.D., Kutas, M., 2002. Picture the difference: electrophysiological investigations of picture processing in the two cerebral hemispheres. Neuropsychologia 40, 730–747. http://dx.doi.org/10.1016/S0028-3932(01)00193-2.

Forceville, C., 2007. Multimodal metaphor in ten Dutch TV commercials. Public Journal of Semiotics. 1, 15–34.

Forceville, C., Urios-Aparisi, E., 2009. Multimodal Metaphor. Walter de Gruyter.

Forgács, B., Bohrn, I., Baudewig, J., Hofmann, M.J., Pléh, C., Jacobs, A.M., 2012. Neural correlates of combinatorial semantic processing of literal and figurative noun noun compound words. NeuroImage 63, 1432–1442. http://dx.doi.org/10.1016/j.neuroimage.2012.07.029.

Franklin, M.S., Dien, J., Neely, J.H., Huber, E., Waterson, L.D., 2007. Semantic priming modulates the N400, N300, and N400RP. Clin. Neurophysiol. 118, 1053–1068. http://dx.doi.org/10.1016/j.clinph.2007.01.012.

Gernsbacher, M.A., Robertson, R.R., 1999. The role of suppression in figurative language comprehension. J. Pragmat. 31, 1619–1630. http://dx.doi.org/10.1016/S0378-2166(99)00007-7.

Giraud, A.L., Poeppel, D., 2012. Cortical oscillations and speech processing: emerging computational principles and operations. Nat. Neurosci. 15, 511–517. http://dx.doi.org/10.1038/nn.3063.

Hagoort, P., Hald, L., Bastiaansen, M., Petersson, K.M., 2004. Integration of word meaning and world knowledge in language comprehension. Science 304, 438–441. http://dx.doi.org/10.1126/science.1095455.

Hald, L.A., Bastiaansen, M.C., Hagoort, P., 2006. EEG theta and gamma responses to semantic violations in online sentence processing. Brain Lang. 96, 90–105. http://dx.doi.org/10.1016/j.bandl.2005.06.007.

Hall, J.R., Bernat, E.M., Patrick, C.J., 2007. Externalizing psychopathology and the error-related negativity. Psychol. Sci. 18, 326–333. http://dx.doi.org/10.1111/j.1467-9280.2007.01899.x.

Hamm, J.P., Johnson, B.W., Kirk, I.J., 2002. Comparison of the N300 and N400 ERPs to picture stimuli in congruent and incongruent contexts. Clin. Neurophysiol. 113, 1339–1350. http://dx.doi.org/10.1016/S1388-2457(02)00161-X.

Harmony, T., 2013. The functional significance of delta oscillations in cognitive processing. Front. Integr. Neurosci. 7, 83. http://dx.doi.org/10.3389/fnint.2013.00083.

Harper, J., Malone, S.M., Bernat, E.M., 2014. Theta and delta band activity explain N2 and P3 ERP component activity in a go/no-go task. Clin. Neurophysiol. 125, 124–132. http://dx.doi.org/10.1016/j.clinph.2013.06.025.

Holcomb, P.J., Kounios, J., Anderson, J.E., West, W.C., 1999. Dual-coding, context-availability, and concreteness effects in sentence comprehension: an electrophysiological investigation. J. Exp. Psychol. Learn. Mem. Cogn. 25, 721. http://dx.doi.org/10.1037/0278-7393.25.3.721.

Ibáñez, A., Toro, P., Cornejo, C., Hurquina, H., Manes, F., Weisbrod, M., et al., 2011. High contextual sensitivity of metaphorical expressions and gesture blending: a video event-related potential design. Psychiatry Res. Neuroimaging 191, 68–75. http://dx.doi.org/10.1016/j.pscychresns.2010.08.008.

Jung-Beeman, M., 2005. Bilateral brain processes for comprehending natural language. Trends Cogn. Sci. 9, 512–518. http://dx.doi.org/10.1016/j.tics.2005.09.009.

Kolk, H.H.J., Chwilla, D.J., Van Herten, M., Oor, P.J.W., 2003. Structure and limited capacity in verbal working memory: a study with event-related potentials. Brain Lang. 85, 1–36. http://dx.doi.org/10.1016/s0093-934x(02)00548-5.

Kuperberg, G.R., Sitnikova, T., Caplan, D., Holcomb, P.J., 2003. Electrophysiological distinctions in processing conceptual relationships within simple sentences. Cogn. Brain Res. 17, 117–129. http://dx.doi.org/10.1016/S0926-6410(03)00086-7.

Kutas, M., Federmeier, K.D., 2011. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annu. Rev. Psychol. 62, 621–647. http://dx.doi.org/10.1146/annurev.psych.093008.131123.

Kutas, M., Hillyard, S.A., 1980. Event-related brain potentials to semantically inappropriate and surprisingly large words. Biol. Psychol. 11, 99–116. http://dx.doi.org/10.1016/0301-0511(80)90046-0.

Lakoff, G., 1993. The contemporary theory of metaphor. Metaphor and thought. 2, 202–251.

Maguire, M.J., Brier, M.R., Ferree, T.C., 2010. EEG theta and alpha responses reveal qualitative differences in processing taxonomic versus thematic semantic relationships. Brain Lang. 114, 16–25. http://dx.doi.org/10.1016/j.bandl.2010.03.005.

Mai, G., Minett, J.W., Wang, W.S., 2016. Delta, theta, beta, and gamma brain oscillations index levels of auditory sentence processing. NeuroImage 133, 516–528. http://dx.doi.org/10.1016/j.neuroimage.2016.02.064.

Mashal, N., Faust, M., Hendler, T., Jung-Beeman, M., 2007. An fMRI investigation of the neural correlates underlying the processing of novel metaphoric expressions. Brain Lang. 100, 115–126. http://dx.doi.org/10.1016/j.bandl.2005.10.005.

Mashal, N., Faust, M., Hendler, T., Jung-Beeman, M., 2009. An fMRI study of processing novel metaphoric sentences. Laterality 14, 30–54. http://dx.doi.org/10.1080/13576500802049433.

Mccarthy, G., Wood, C.C., 1985. Scalp distributions of event-related potentials: an ambiguity associated with analysis of variance models ☆. Electroencephalogr. Clin. Neurophysiol. 62, 203–208. http://dx.doi.org/10.1016/0168-5597(85)90015-2.

Mcpherson, W.B., Holcomb, P.J., 1999. An electrophysiological investigation of semantic priming with pictures of real objects. Psychophysiology 36, 53–65.

Murray, M.M., Brunet, D., Michel, C.M., 2008. Topographic ERP analyses: a step-by-step tutorial review. Brain Topogr. 20, 249–264. http://dx.doi.org/10.1007/s10548-008-0054-5.

Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9, 97–113. http://dx.doi.org/10.1016/0028-3932(71)90067-4.

Ortiz, M.J., 2011. Primary metaphors and monomodal visual metaphors. J. Pragmat. 43, 1568–1580. http://dx.doi.org/10.1016/j.pragma.2010.12.003.

Paivio, A., 1991. Dual coding theory: retrospect and current status. Canadian Journal of Psychology/Revue canadienne de psychologie. 45, 255. http://dx.doi.org/10.1037/h0084295.

Pynte, J., Besson, M., Robichon, F.-H., Poli, J., 1996. The time-course of metaphor comprehension: an event-related potential study. Brain Lang. 55, 293–316. http://dx.doi.org/10.1006/brln.1996.0107.

Rapp, A.M., Leube, D.T., Erb, M., Grodd, W., Kircher, T.T., 2007. Laterality in metaphor processing: lack of evidence from functional magnetic resonance imaging for the right hemisphere theory. Brain Lang. 100, 142–149. http://dx.doi.org/10.1016/j.bandl.2006.04.004.

Rapp, A.M., Mutschler, D.E., Erb, M., 2012. Where in the brain is nonliteral language? A coordinate-based meta-analysis of functional magnetic resonance imaging studies. NeuroImage 63, 600–610. http://dx.doi.org/10.1016/j.neuroimage.2012.06.022.

Roehm, D., Schlesewsky, M., Bornkessel, I., Frisch, S., Haider, H., 2004. Fractionating language comprehension via frequency characteristics of the human EEG. Neuroreport 15, 409–412. http://dx.doi.org/10.1097/00001756-200403010-00005.

Sauseng, P., Griesmayr, B., Freunberger, R., Klimesch, W., 2010. Control mechanisms in working memory: a possible function of EEG theta oscillations. Neurosci. Biobehav. Rev. 34, 1015–1022. http://dx.doi.org/10.1016/j.neubiorev.2009.12.006.

Schmidt, G.L., Seger, C.A., 2009. Neural correlates of metaphor processing: the roles of figurativeness, familiarity and difficulty. Brain Cogn. 71, 375–386. http://dx.doi.org/10.1016/j.bandc.2009.06.001.

Schmidt, G.L., Debuse, C.J., Seger, C.A., 2007. Right hemisphere metaphor processing? Characterizing the lateralization of semantic processes. Brain Lang. 100, 127–141. http://dx.doi.org/10.1016/j.bandl.2005.03.002.

Schneider, S., Rapp, A.M., Haeußinger, F.B., Ernst, L.H., Hamm, F., Fallgatter, A.J., et al., 2014. Beyond the N400: complementary access to early neural correlates of novel metaphor comprehension using combined electrophysiological and haemodynamic measurements. Cortex 53, 45–59. http://dx.doi.org/10.1016/j.cortex.2014.01.008.

Sitnikova, T., Holcomb, P.J., Kiyonaga, K.A., Kuperberg, G.R., 2008. Two neurocognitive mechanisms of semantic integration during the comprehension of visual real-world events. J. Cogn. Neurosci. 20, 2037–2057. http://dx.doi.org/10.1162/jocn.2008.20143.

Steele, V.R., Bernat, E.M., Van Den Broek, P., Collins, P.F., Patrick, C.J., Marsolek, C.J., 2013. Separable processes before, during, and after the N400 elicited by previously inferred and new information: evidence from time-frequency decompositions. Brain Res. 1492, 92–107. http://dx.doi.org/10.1016/j.brainres.2012.11.016.

Stringaris, A.K., Medford, N., Giora, R., Giampietro, V.C., Brammer, M.J., David, A.S., 2006. How metaphors influence semantic relatedness judgments: the role of the right frontal cortex. NeuroImage 33, 784–793. http://dx.doi.org/10.1016/j.neuroimage.2006.06.057.

Van Herten, M., Kolk, H.H., Chwilla, D.J., 2005. An ERP study of P600 effects elicited by semantic anomalies. Brain Res. Cogn. Brain Res. 22, 241–255. http://dx.doi.org/10.1016/j.cogbrainres.2004.09.002.

West, W.C., Holcomb, P.J., 2002. Event-related potentials during discourse-level semantic integration of complex pictures. Cogn. Brain Res. 13, 363–375. http://dx.doi.org/10.1016/S0926-6410(01)00129-X.

Willems, R.M., Zyürek, A., Hagoort, P., 2008. Seeing and hearing meaning: ERP and fMRI evidence of word versus picture integration into a sentence context. J. Cogn. Neurosci. 20, 1235–1249. http://dx.doi.org/10.1162/jocn.2008.20085.

Yang, F.-P.G., Bradley, K., Huq, M., Wu, D.-L., Krawczyk, D.C., 2013. Contextual effects on conceptual blending in metaphors: an event-related potential study. J. Neurolinguistics 26, 312–326. http://dx.doi.org/10.1016/j.jneuroling.2012.10.004.