Journal of Neurolinguistics 56 (2020) 100915

Contents lists available at ScienceDirect

Journal of Neurolinguistics

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

The role of animacy in metaphor processing of Mandarin Chinese: An Event-Related Potential (ERP) study

Haojie Ji[a], Senqing Qi[a], Shiyang Xu[c], Jinxin Chen[a], David Yun Dai[d], Yadan Li[a], Weiping Hu[a,b,*]

[a] MOE Key Laboratory of Modern Teaching Technology, Shaanxi Normal University, China
[b] Collaborative Innovation Center of Assessment Toward Basic Education Quality at Beijing Normal University, China
[c] Faculty of Health Sciences University of Macau, China
[d] State University of New York at Albany, USA

ARTICLE INFO

Keywords:
Metaphor
Animacy
ERP
N400
P600
Conflict strength

ABSTRACT

Many ERP studies have highlighted processing difficulty for animacy violation in language, more so than non-violated sentences. However, all of these findings were for non-figurative language, and studies have rarely looked at the figurative language, such as metaphors involving animacy violation, which could be integrated to make novel and acceptable meaning. The present study aimed at assessing the role of animacy in metaphorical comprehension of Mandarin Chinese. ERPs were recorded as healthy participants read metaphor or literal sentences with animate or inanimate actors as sentence-initial noun, while the target words were measured at the verb and the second argument (object) of the sentences with subject-verb-object (SVO) structure. As expected, the animacy violations elicited a significant N400 effect by the target verb in the metaphor with inanimate initial-nouns. At the objects, the N400 amplitudes associated with metaphors were not regulated by animacy. Subsequently, the analysis revealed a significant difference in the P600 amplitudes between inanimate and animate metaphorical conditions. The metaphors with inanimate actor elicited an attenuated P600 as compared with the animate counterparts and converged to the same level as literal sentences, reflecting the less effortful metaphor-relevant mapping process. These results suggest that animacy violation may facilitate the integration of the reanalysis stage for metaphorical comprehension, and the conflict strength (animacy violation vs. no animacy violation) modulates the time course of metaphor processing in Mandarin Chinese. The present study yields new insights into the role of animacy in the online linguistic comprehension.

1. Introduction

Metaphor is recognized as being closely associated with creative cognitive processes (Benedek et al., 2014; Rutter et al., 2012; Semino & Steen, 2008), and has played an important role in poetry and other genres literature (Shen, 2008). In metaphoric expressions, the inanimate objects are specified as persons (and sometimes vice versa), establishing thereby animacy-related thematic role violations, such as "*Life has cheated me*" (Lakoff & Johnson, 2003). In this case, the *life* is like a bad guy who cheats us. The animacy violations between the inanimate actor (e.g., 'life') and a verb (e.g., 'cheat') constitute a strong conflict, producing a novel

* Corresponding author. MOE Key Laboratory of Modern Teaching Technology, Shaanxi Normal University, China.
E-mail address: huweiping@snnu.edu.cn (W. Hu).

https://doi.org/10.1016/j.jneuroling.2020.100915
Received 21 July 2019; Received in revised form 8 April 2020; Accepted 9 May 2020
Available online 04 August 2020
0911-6044/

expression. In their cognitive metaphor theory, Lakoff and Johnson (2003) propose that human experience of physical entities (including our own bodies) provides the foundation for a very wide range of ontological metaphors, which represent a way of viewing inanimate objects such as events, activities, emotions, thoughts, etc., as entities and substances. Ontological metaphors allow us to comprehend a wide variety of experiences with nonhuman objects, expressed in terms of human motivations, characteristics, and activities. Moreover, a metaphor (like personification or zoosemy) can be used to produce more novel and vivid expressions, reflecting the creativity in language (Shen, 2008).

However, animacy is so universal in human linguistic expression that it tends to be taken for granted, thus becoming inconspicuous (Dahl and Fraurud, 1996). This is especially true of metaphoric expression. In the range of metaphorical language, a considerable amount of electrophysiological literature has already been discussed theoretically and investigated empirically (See Bohrn, Altmann, & Jacobs, 2012 for a review). However, the question remains as to whether animacy plays an important role in metaphor processing. Research on cognitive linguistic has proceeded completely independently between the traditions of metaphor and animacy. Yet, increasing crosstalk between the two lines of research has enhanced mutual fertilization. The work reported in this study represents such a fusion.

1.1. Metaphor

Conventionally, studies of the steps in metaphoric comprehension have included two main models. The two models differ as to whether the processor always starts from the literal meaning (the indirect access account), or whether only the figurative/metaphorical meaning is processed (the direct access account). The indirect access model originates from approaches by Grice (1975) and Searle (1979). This approach would predict that metaphor requires a longer response time (Janus & Bever, 1985), or larger scalp-recorded potential than literal expression (Bambini, Bertini, Schaeken, Stella, & Di Russo, 2016; Lai, Curran, & Menn, 2009; Weiland, Bambini, & Schumacher, 2014). In contrast, the direct access view holds that metaphor processing does not involve more steps than the interpretation of literal expression, and a reader needs only to select the appropriate one. According to this reasoning, literal and metaphorical meanings should be processed in an equal number of steps (Gibbs, 1994; Glucksberg, 2008). It can also fall in-between; the *graded salience hypothesis* (GSH) posits that the comprehension process is influenced by the ‘salience’ of the stimuli (Giora, 2003). According to the hypothesis by Giora (1997), it is the salient information that determines the priority of access in that the salient meaning is always preferred. More specifically, conventionality can make a metaphor more salient. For conventional metaphors, the figurative meaning is commonly more salient than the literal one, and thus the figurative meaning is mostly accessed first. However, when a novel or unfamiliar metaphor is encountered, the salient meaning is the literal one, and the figurative meaning is inferred later. According to the prediction by GSH (Giora, 1997), novel or unfamiliar metaphors generally elicit larger ERPs amplitudes than conventional metaphors or literal expressions. For instance, Lai et al. (2009) manipulated the conventionality in metaphor and reported larger negativity for novel metaphor (e.g., *Every second of our time was attacked*) compared to conventional condition (e.g., *Every point in my argument was attacked*) during later time window 440–560 ms. Such results have been observed by a series of subsequent studies (e.g., Goldstein, Arzouan, & Faust, 2012; Tzuyin Lai & Curran., 2013).

A majority of Event-Related Potential (ERP) research on metaphor comprehension has focused on two particular ERP components: a centro-parietal N400, and a later parietal P600/LPC. Most researchers reported a biphasic N400–P600 effect that allegedly implicates different stages in conceptual mapping (Coulson & Van Petten, 2002; De Grauwe, Swain, Holcomb, Ditman, & Kuperberg, 2010). Other researchers reported a N400 response (Pynte, Besson, Robichon, & Poli, 1996; Tzuyin Lai & Curran., 2013), or a P600 response separately (Schmidt, Debuse, & Seger, 2007; Tang, Qi, Wang, Jia, & Ren, 2017). The former was initially related to retrieval of lexical-conceptual information (Kutas & Van Petten, 1988), and also reflects the integration of lexical meaning with context (Hagoort. et al., 2004). The latter is considered as a reanalysis of sentence-level processing conflicts (Frenzel, Schlesewsky, & Bornkessel-Schlesewsky, 2011; Yang, Bradley, Huq, Wu, & Krawczyk, 2013), and an interpretation in a wider context (Brouwer, Fitz, & Hoeks, 2012). Generally, ERP studies with a N400 effect favor the direct access model, while studies focusing on the biphasic pattern or a later effect argue against the direct view (Bambini et al., 2016).

The GSH has indicated explicitly that salient meanings were characterized by conventionality, prototypicality, familiarity or frequency. Although some researchers have explored the influence of conventionality (e.g., Giora et al., 2003), and familiarity or frequency (e.g., Pynte et al., 1996) on metaphor processing, the prototypicality is not studied. In linguistic theory, animacy determines the role prototypicality of noun arguments in sentences (Bornkessel-schlesewsky & Schlesewsky, 2010), and thus is a salient information factor that has potential impacts on metaphorical sentence comprehension.

1.2. Animacy

Animacy is regarded as a universal language phenomenon (Comrie, 1989). Psycholinguistic research has repeatedly confirmed that the human language comprehension system is sensitive to the animacy–inanimacy distinction of sentence arguments. The animacy hierarchy is the main constituent of semantic information (Caramazza & Mahon, 2003), and is related to distinct patterns of brain activation (Caramazza & Shelton, 1998; Gobbini et al., 2011). Since animacy indicates whether a noun is capable of being an ideal Actor (by virtue of the fact that it is sentient and capable of self-initiated, volitional action) or not, the fundamental distinction between inanimate and animate entities is considered a critical factor in language comprehension (Lowder, P. C. Gordon, 2015).

In the past decades, numerous studies have investigated the role of noun animacy during online sentence comprehension, and the results have repeatedly confirmed the animacy effect. Based on the literature's consensus on animacy, the ideal Actor is usually an animate subject. In contrast, an inanimate subject takes a role of the undergoer (e.g. Bock, 1986; Dahl, 2008; Malchukov, 2008;

Philipp, Bornkessel-Schlesewsky, Bisang, & Schlesewsky, 2008). If assignment of the argument role is inconsistent with expectations (such as animacy violation), it will evoke a strong animacy effect, and producing robust electrophysiological responses in the brain. Usually, inanimate initial noun-phrase (NP) argument(s) would lead to incremental processing costs, as inanimate entities are not ideal agents (Kuperberg, Sitnikova, Caplan, & Holcomb, 2003; Nakano, Saron, & Swaab, 2010; N. Bourguignon, Drury, Valois, & Steinhauer, 2012). The evidence from ERP studies using several languages (English, Chinese, etc.) revealed a centro-parietal N400 for inanimate (vs. animate) noun arguments as the Actor (Rosler, Pütz, Friederici, & Hahne, 1993; Weckerly & Kutas, 1999). However, other researchers reported P600/LPC (Kolk, Chwilla, Van, & Oor, 2003; Hoeks, Stowe, & Doedens, 2004; Kuperberg, Kreher, Sitnikova, Caplan, & Holcomb, 2007; Nakano et al., 2010; S.; Frenzel et al., 2011), or biphasic N400–P600 (LPC) (Szewczyk & Schriefers, 2011).

Some researchers attempted to highlight the N400–P600 distinction as related to animacy violation. These factors include role prototypicality mismatch (Weckerly & Kutas, 1999), semantic attraction (Kim & Osterhout, 2005), and context (Nieuwland & van Berkum, 2005). However, there seems to be very little agreement on what factor(s) elicit the N400 component. Meanwhile, the P600 component is more closely associated with unresolved processing conflicts between the inanimate actors, and verbs requiring an animate actor. These findings were confirmed by several empirical investigations conducted in diverse languages (Kolk et al., 2003; Hoeks et al., 2004; Kuperberg et al., 2007).

So far, previous studies regarding the influence of animacy on language processing focused on non-figurative language (i.e., literal sentence). What is less clear is whether animacy influences metaphorical comprehension. For the literal meaning of the metaphor, animacy violations are semantically anomalous (Nieuwland & Van Berkum, 2006), and processed as unresolvable conflicts. However, animacy violations in poetry and other genres literature could usually be integrated into the novel and unconventional meanings, which is acceptable in metaphorical meaning. Therefore, the time-course dynamic processing mechanism of animacy violation in metaphor should differ from that for the non-metaphorical language reported in previous research.

1.3. Present study

In general, Chinese, as a language not susceptible to morphological syntax (Liu, Li, Shu, Zhang, & Chen, 2010; Zhang, Yu, & Boland, 2010), is easily affected by animacy information in language processing (Li, 1996; Mac Whinney, 2004). A metaphor contains diversified lexical collocations with different animacy hierarchies in constructing novel, creative, and imaginative figurative expressions (Shen, 2008). As mentioned above, animacy as salient information (Bornkessel-schlesewsky & Schlesewsky, 2010) would affect the metaphorical comprehension of Mandarin Chinese. The manner in which animacy could potentially impact the time course of key processes in metaphor is difficult to establish based on the preceding information. How do readers process these distinctive metaphors? ERPs provide an ideal method for studying sentence processing (Kazmerski, Blasko, & Dessalegn, 2003), and are very suitable for exploring processing conflicts between arguments for linguistic phenomena as they unfold in time. They allow for comparison of the distinctive processing patterns between metaphorical sentences with and without animacy violations.

Our aim in this study is to understand the role of animacy properties of sentence-initial nouns (NP1) as agents within SVO (Subject-Predicate-Object) sentences in the real-time comprehension of Mandarin Chinese metaphor. Using a metaphorically/literally plausible judgment task, we recorded ERPs elicited by the target words at the predicate verb and object of SVO sentences whose sentence-initial noun was an animate or inanimate actor in Mandarin Chinese. SVO structure was used for the following reasons: 1) to manipulate the animate variable, since most verbs limit the animacy hierarchy of arguments as to which entities can be actors at sentence-initial nouns (J.M. Szewczyk, 2011); 2) to limit thematic processing to a minimum; 3) to keep in line with the Chinese expression of linear order. A quadruplet shared the same predicate verb, which could be matched with animate and inanimate subjects arguments in order to make the literal sentences reasonable and interpretable under conditions of both animate and inanimate head-nouns.

Based on previous ERP studies on metaphor and animacy, our predictions were twofold: 1) Metaphors with inanimate initial NP were expected to evoke animacy violation N400 effects on the predicate verb (i.e., animacy violation condition); and 2) Relative to animate conditions, there would be a P600 effect on objects (NP2) for inanimate metaphors (i.e., animacy violation condition). First, the predicate verb was defined as the middle processing stage of the sentence. Once the verb was reached, the linguistic comprehensive system would monitor difficulties of integration arising from animacy violation conditions of metaphors. If the N400 amplitudes are shown sensitive to role prototypicality mismatches (Bornkessel-schlesewsky & Schlesewsky, 2010), there should be larger N400 elicited by inanimate metaphorical condition than the animate counterparts because the animacy violation between the inanimate head-nouns and the verb. Second, NP2 was defined as the final processing stage. Because animate and inanimate metaphors have differing conflict strengths, there are different P600 magnitudes between the two categories of metaphors. In metaphors, animacy initial-nouns are capable of being an ideal Actor, which were prominence information than inanimate ones and thus are computed first (Bornkessel-schlesewsky & Schlesewsky, 2010). However, the inanimate NP1 of metaphor would cause intensive unresolvable conflicts at the literal-level integration. If the P600s reflects the level of difficulty in integrating sentence-level interpretation conflicts (Frenzel et al., 2011), bigger P600s should be elicited by inanimate metaphors than animate ones. As far as we know, this is the first cognitive linguistic study to address the issue of animacy effects on metaphor processing.

Table 1A
Sample stimuli.

Condition	Examples	Explanation
High-animate Metaphor	老师点亮前途	The teachers light up the future.
Inanimate Metaphor	苦难点亮前途	The sufferings light up the future.
High-animate Literal	男人点亮木头	The man lit the firewood.
Inanimate Literal	火柴点亮木头	The match lit the firewood.

Note:The 'light up' and 'lit' were the same word '点亮' (*dianliang*) in Chinese.

2. Materials and methods

2.1. Participants

Twenty-five right-handed, native Chinese-speaking students (12 males, mean age 19.12 years, $SD = 0.95$, Range = 18–20 years) at the Shaanxi Normal University participated in the EEG experiment to receive volunteer service certification. All had normal or corrected-to-normal vision. None had neurological disorders or major head injuries. Three subjects were excluded because of excessive artifacts (> 10% of rejected trials) and excessive error rate in behaviour response (> 10%), leading to a final sample size of 22 subjects (10 males). The experimental standards of the study were approved by the local Review Board for Human Participant Research. Each subject provided written informed consent before participating.

2.2. Construction and classification of stimuli

The original materials were selected from contemporary Chinese poetry and other genres literature, and edited by researchers based on the SVO structure (As some original materials from poetry are not SVO structure sentences, we added the subjects or objects in order to complete the SVO sentences.). Then 124 quadruplets were formed by original materials and 40 quadruplets from them were selected based on pretest. Animacy (animate vs. inanimate) and Sentence Type (metaphoric vs. literal) formed a cross-factor design. Therefore each quadruplet contained four possible combinations of the two animate levels and two sentence types: High-animate Metaphor, Inanimate Metaphor, High-animate Literal, and Inanimate Literal (henceforth, Ha-M, Ina-M, Ha-L, Ina-L). Examples are shown in Table 1A. A half of the sentences were metaphors, which were adapted on the basis of initial materials collected from poetry and other genres literature. Their literal counterparts were constructed by replacing the subjects and objects of metaphorical sentences in terms of the verb's literal meaning. These sentences were matched in terms of five semantic features: familiarity, figurativeness, valence, interpretability, and cloze probability, in order to exclude possible confounding factors. Each sentence consisted of a transitive verb and two nouns, with an animate or inanimate concrete noun as the subject (NP1) and an inanimate noun as the object (NP2).

By and large, the concrete nouns were used as NP2 of literal sentences, and the abstract nouns were selected in metaphor sentences in order to construct a rational meaning accord with each sentence type. Hence the degrees of concreteness of the objects within the SVO sentence used within the main experiment are different in two sentence types. Given that the 'lexical concreteness' (Holcomb, Kounios, Anderson, & West, 1999; Zhang, Guo, Ding, & Wang, 2006) could potentially impact the time course of metaphorical comprehension (Forgács et al., 2015), an independent lexical decision task was designed to verify whether the lexical concreteness of object nouns with in two sentence types affects sentence comprehension in the main experiment. In the supplementary lexical decision task, the stimuli included 40 abstract nouns (the objects within metaphorical sentences), 40 concrete nouns (the objects within literal sentences), and 25 anomalous words. This proportion of meaningful words and non-words was selected according to previous research on Chinese concreteness effect by Zhang et al. (2006). Given the small number of anomalous words used in their study, we added some words for anomalous condition (total are 25 anomalous words in the present study).

2.3. Pre-test

First, the stimuli sentences were pretested for familiarity, figurativeness, valence, and interpretation. All raters (*n = 40, mean age = 21.5*) are native speakers of Chinese, selected from two undergraduate classes in the Department of Linguistics at the Xi'an Technological University. Each received cash for participation. 496 sentences (4 conditions × 124 stimuli sentences) were divided into four blocks using a Latin Square design, so that each stimuli sentence of one set appeared only once per block. As a result, each block contained 31 Ha-M sentences, 31 Ina-M sentences, 31Ha-L sentences, and 31 Ina-L sentences. Sentences in each block were then randomized. A given participant saw all four blocks during the course of the pre-test.

Forty students who did not participate in either the ERP experiment or the other pretest were asked to judge the overall familiarity, figurativeness, valence, and interpretability of each sentence on a 5-point Likert Scale. For the familiarity task, participants were instructed to rate their frequency of exposure to the sentence and its meaning, using a scale from 1 (*very unfamiliar*) to 5 (*very familiar*). For the figurativeness task, participants were instructed to rate how literal an interpretation each sentence suggested, using a scale from 1 (*very literal*) to 5 (*very figurative*). For the emotional valance task, participants were instructed to rate the emotional experience caused by a given stimuli sentence, using a scale from 1 *(negative emotion),* to 5 *(positive emotion),* with the middle point

Table 1B
Summary of *mean values* and *standard deviation* from pre-tests for selected experimental materials.

Condition	Familiarity	Figurativeness	Emotional Valence	Interpretability	Cloze Probability (target 1)	Cloze Probability (target 2)
HaM	2.93(0.52)	3.58(0.36)	3.23(0.76)	3.28(0.40)	0.01(0.05)	0.01(0.04)
InaM	2.79(0.47)	3.67(0.34)	3.17(0.69)	3.29(0.43)	0.01(0.03)	0.01(0.04)
HaL	3.41(0.58)	1.72(0.38)	3.12(0.55)	3.90(0.38)	0.21(0.29)	0.08(0.19)
InaL	3.49(0.45)	1.82(0.36)	3.13(0.29)	3.82(0.36)	0.20(0.29)	0.12(0.48)

representing neutral emotion. For the interpretability task, participants were given the following instruction: "If you feel that the sentence is very easily interpretable, give it a 5. If the sentence takes you a very long while to come up with an interpretation, give it a 1."

Second, two other pretests were conducted separately on cloze probability for predicate verbs and objects of stimuli sentences, as these two positions are interesting targets (Predicate verb was target 1, and NP2 was target 2). 496 sentences were divided into 4 versions using a Latin-square procedure. Twenty-eight students who did not participate in either the ERP experiment or the other pretest were randomly assigned to one of the four versions of each Cloze Probability Rating. They were then asked to record the first word that came to mind, and to make the completed sentence as natural as possible. Their mean familiarity, figurativeness, valence, and interpretability of whole sentences, and the cloze probability of the two target words are listed in Table 1B.

A repeated-measures ANOVA were computed with the factors TYPE and the factor ANIMACY and revealed no significant Animacy main effect, or any interactive effect, for 1) rate of familiarity; 2) figurativeness; 3) interpretability of the whole sentence; or 4) cloze probability ($p > .1$), and significant Sentence Type main effect for 1) rate of familiarity ($F_{(1, 39)} = 37.30, p < .001, \eta_P^2 = .49$); 2) figurativeness ($F_{(1, 39)} = 1055.57, p < .001, \eta_P^2 = .96$); 3) interpretability of the whole sentence ($F_{(1, 39)} = 69.68, p < .001, \eta_P^2 = .64$); or 4) cloze probability ($F_{(1, 6)} = 31.49, p < .001, \eta_P^2 = .45$ for target 1; $F_{(1, 6)} = 4.66, p < .05, \eta_P^2 = .11$ for target 2). Furthermore, there was no significant Animate main effect, Sentence Type main effect, or any interactive effect for rate of Emotional Valence ($p > .1$). Pairwise comparison' results indicated that high- and low-animacy conditions within the same sentence type did not differ significantly from each other ($p > .1$).

2.4. Procedure

2.4.1. Sentence comprehension (metaphorically/literally plausible) task

Participants were seated in a comfortable chair in a sound-attenuated, electrically-shielded room and instructed to read each sentence carefully. Sentences were then presented word-by-word in white color on a black background. Each trial began with a fixation point ('+') displayed at the center of the screen for 500 ms. After this, an experimental sentence was presented in the following time sequence: blank (200–500 ms randomly), NP1 (1000 ms), blank (200–500 ms randomly), predicate verb (1000 ms), blank (200–500 ms randomly), and NP2 (1000 ms). After the presentation of final segment of each sentence, there was a question mark until participants responded with a button press. Participants were asked to judge whether the SVO sentence conveys a literal meaning or metaphoric meaning, and press a corresponding key. Response time was within 3000 ms after the target onset. They were aware of the presence of the two types of expressions and were given examples of each. The assignment for the index finger of the left or right hand was counterbalanced across participants. Following the judgment task there is a 1000 ms intertrial interval of a blank screen. The overall sequence of events for each trial is illustrated in Fig. 1.

Using a Latin Square design, the 160 sentences (4 conditions/sentence frames × 40 target words) were divided into 4 blocks such

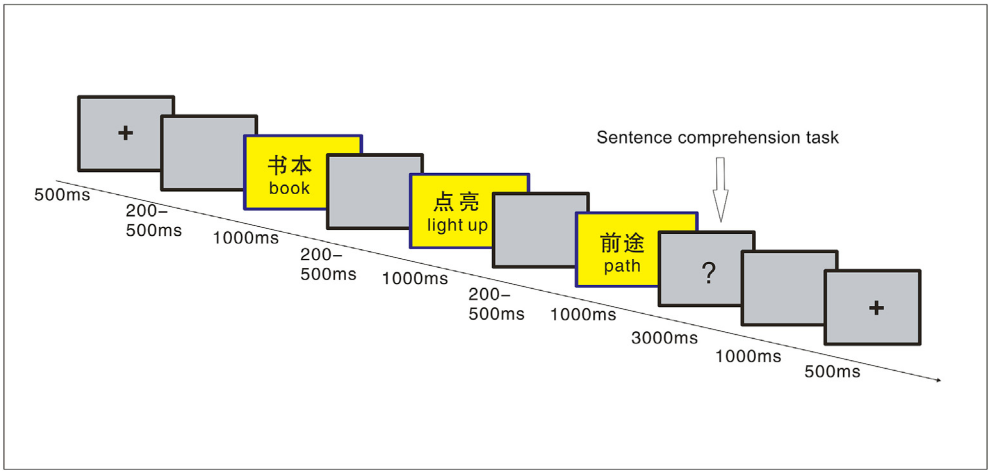

Fig. 1. Experimental paradigm.

that each condition from each quadruplet appeared only once per block. As a result, each block contained 10 Ha-M, 10 Ina-M, 10 Ha-L, and 10 Ina-L. There was a short break between the two consecutive blocks. Before the main experiment session, participants performed a practice block of 16 sentences, which had constructions similar to the stimuli in the formal experiment. Including electrode application, an experimental session lasted approximately 2 h.

2.4.2. Lexical decision task

A lexical decision task was conducted to examine the degree to which the concreteness of object nouns influences sentence comprehension by recording ERP response elicited by NP2 nouns in an independent experiment. Each trial began with a cross point that was displayed for 1600 ms–1800 ms randomly. Following the cross, a stimulus was presented on the screen for 800 ms, then replaced by the cross point for the next trial. Participants were asked to press a button as accurately and quickly as possible if the presented stimulus was not a meaningful Chinese word. This procedure was based on the experiment of Zhang et al. (2006), and was modified for stimuli presentation time to 800 ms, which was sufficient to recognize and understand the single word. Before the formal experiment, practice trials were administered to the participants to ensure that the testing procedure was well understood. The presentation order of the sentence comprehension task and lexical decision task was counterbalanced as ABBA sequence.

2.5. Electrophysiological recording

Scalp voltages were recorded at 64 scalp sites using Ag/AgCl electrodes with the CURRY 7 system (Compumedics Neuroscan, Texas, USA), monitored by the CURRY recorder software and connected to a SynAmp amplifier (Compumedics Neuroscan, Texas, USA). EEG recording was referenced to the left and right mastoid, with a ground electrode on the medial frontal aspect. The vertical electrooculogram (EOG) was recorded supra- and in fraorbitally at the left eye; the horizontal EOG was recorded from the left versus the right orbital rim. All interelectrode impedances were maintained below5 kΩ. Amplified analog voltages were digitized at 500 Hz/channel with a band-pass filter of DC to 100 Hz.

Offline processing of EEG signal data was performed in MATLAB using ERPLAB toolbox (Lopez-Calderon & Luck, 2014). Portions of EEG containing large muscle artifacts or extreme offsets (identified by visual inspection) were removed. EEG data was down-sampled to 250 Hz, and re-referenced to the average of the left and right mastoids. A digital band-pass filter of 0.01–30 Hz was applied to the EEG recordings. Independent component analysis (ICA) was performed on continuous data for each participant to remove components relevant to eye movements and eye blinks. The ICA-corrected EEG data were segmented into epochs using a baseline of −200 to 0 ms for stimulus-locked averages. Epochs were 1200 ms long, with a 200 ms pre-stimulus baseline. Any epochs with EEG voltages exceeding a threshold of ± 75 μV were excluded from the average.

For the sentence comprehension task, the overall valid rate was 96.11%, equal for all conditions. The mean rate of valid trials for the target 1 (predicate verb) was 96.48% for the HaM condition, 95.80% fort he InaM condition, 96.93% for the HaL condition and 95.45% for the InaL condition. The mean rate of valid trials for the target 2 (NP2) was 96.59% for the HaM condition, 95.57% for the InaM condition, 96.25% for the HaL condition and 95.80% for the InaL condition. For the lexical decision task, the overall valid rate was 95.99%, equal for three conditions (Abstract: 96.59%; Concrete: 95.57%; Anomalous: 95.82%).

2.6. Data analysis

In accordance with previous literature on the N400 and P600 for metaphor (Bambini et al., 2016; Lai et al., 2009) and visual inspection based on the grand averaged data, we defined two different time windows (320–450 ms for the N400 and 500–700 ms for the P600) of predicate verb and two different time windows (300–420 ms for the N400 and 460–640 ms for the P600) of NP2 (see also Table 2A) that were submitted to the analyses for sentence comprehension task. The selection of time windows of lexical decision task was same as the NP2 in sentence stimuli.

The SPSS statistical software package (Version 20.0, SPSS) was used for the statistical analyses. Only significant results are presented (alpha level was set at 0.05). For the ERPs, grand averages were computed per participant and condition, from 200 ms prior to critical constituent onset of target words (Predicate verb and NP2) until 1000 ms post-onset. The ERP data were analyzed using

Table 2A
Summary of statistical analysis results for the target words of verb and NP2 in the selected time-windows.

		Predicate Verb						Object (NP2)					
		N320-450			P500-700			N300-420			N460-640		
Midline		*F*	*p*	η^2	*F*	*p*	η^2	*F*	*p*	η^2	*F*	*p*	η^2
	Type	7.37	.013*	.26	2.86	.11	.12	3.71	.07	.15	5.73	.03*	.21
	Animacy	.081	.78	.00	.29	.60	.01	2.74	.11	.12	6.81	.02*	.25
	Type × Animacy	7.05	.02*	.25	3.68	.07	.15	1.07	.31	.05	4.31	.05*	.17
Lateral		*F*	*p*	η^2	*F*	*p*	η^2	*F*	*p*	η^2	*F*	*p*	η^2
	Type	6.75	.02*	.24	2.15	.16	.09	2.21	.15	.10	4.96	.04*	.19
	Animacy	.512	.48	.02	.055	.82	.00	3.73	.07	.15	5.80	.03*	.22
	Type × Animacy	3.25	.09	.13	1.13	.30	.05	.86	.36	.04	3.36	.08	.138

*:p value reached a significant level at least 0.05.

repeated-measures ANOVAs, with the following condition factors: 1) Sentence type (metaphor vs. literal); 2) Animacy (Animate NP1 vs. Inanimate NP1) and 3) the topographical factor region of interest (ROI) over mean amplitudes per condition and time window. All three are within-subjects factors. The four lateral ROIs were grouped by location as follows: left anterior (F7/F3/FC5/FC1), left posterior (CP5/CP1/P7/P3), right anterior (F4/F8/FC2/FC6), and right posterior (CP2/CP6/P4/P8). The six electrodes for the midline were grouped pairwise: frontal (Fz/FCz), central (Cz/CPz), and parietal (Pz/POz). Lateral and midline electrodes were analyzed separately.

All repeated-measures analyses of variance (ANOVA) results received Greenhouse-Geisser correction if the sphericity assumption was violated. In these cases, we reported the uncorrected degrees of freedom and the corrected p values. Post-hoc multiple comparisons were carried out using Bonferroni-adjusted corrections. Effect sizes are presented as partial eta-squared (η_P^2) for F tests.

3. Results

3.1. Behavioral results

For the metaphorically/literally plausible judgment task, we calculated mean accuracy rates and reaction times for each sentence type. Overall, task response accuracy ($M = 95.1\%$, $SD = 0.03$) for button presses suggested that the sentences were read well for comprehension. One-way repeated-measures ANOVA revealed no differences between all four conditions (HaM, InaM, HaL, and InaL) as far as accuracy rates ($F_{(1,\ 21)} = 2.03$, $p = .16$, $\eta_P^2 = .09$) and reaction times ($F_{(1,\ 21)} = 1.42$, $p = .27$, $\eta_P^2 = .18$).

3.2. Event-related potentials

For the electrophysiological scalp data, the time-windows of interest were predefined by visual inspection. Amplitude measurements were based on the average amplitude within the specified time windows (see the results of statistical analysis in Table 2A). As shown in Table 2A, mean amplitudes of the predicate verb were extracted from the 320–450 ms time window (N400) and the NP2 were extracted from the 300–420 ms (N400) and 460–640 ms time windows (P600). In cases where main or interaction effects could be observed, additional post hoc ANOVA were carried out to compare each metaphorical category with the others. We focused on the main effects and interaction effects involving the factors of ANIMACY and SENTENCE TYPE (The following displays "TYPE"). Grand averaged waveforms for the Verb and the NP2 (sentence-final words) are displayed in Fig. 2 and Fig. 3.

3.2.1. Predicate verb —— middle processing stage

3.2.1.1. N320-450 ms. A repeated-measures ANOVA revealed significant main effects of TYPE for both the midline ($F_{(1,\ 21)} = 7.38$, $p < .05$, $\eta_P^2 = .26$), and the lateral ($F_{(1,\ 21)} = 6.75$, $p < .05$, $\eta_p^2 = .24$), with metaphor receiving a significantly more negative mean amplitude (−1.81 μV for midline, −0.92 μV for lateral) over literal (−1.16 μV for midline, −0.47 μV for lateral). ANIMACY manipulation did not yield a significant effect. There were significant interaction effects between TYPE and ANIMACY for midline ($F_{(1,\ 21)} = 7.05$, $p < .05$, $\eta_p^2 = .25$). We followed this interaction by computing separate repeated-measures ANOVA to tested the effects of TYPE identity for each condition of animacy. Ina-M elicited a greater negativity potential (−2.11 μV) than Ina-L (−0.93 μV)($F_{(1,21)} = 10.84$, $p < .01$, $\eta_p^2 = .34$). There was no significant difference between Ha-M (−1.50 μV) and Ha-L (−1.39 μV) conditions ($F_{(1,21)} = 0.20$, $p = .66$, $\eta_p^2 = .01$).

3.2.1.2. N500-700 ms. Statistical analyses showed no significant effect for the 500–700 ms time-window (see in Table 2A).

3.2.2. NP2 —— final processing stage

3.2.2.1. N300-420 ms. For the time window from 300 to 420 ms of NP2, the analysis only showed a significant interaction of the factors ROI × TYPE for lateral ($F_{(3,\ 63)} = 3.48$, $p < .05$, $\eta_p^2 = .14$). Further analysis indicated that the mean amplitudes for metaphors ($M = 0.06$ μV) were more positive than for the literal ones ($M = -0.55$ μV) on the left anterior (LA) region ($F_{(1,21)} = 5.80$, $p < .05$, $\eta_p^2 = .22$), while there was no significant difference between the metaphor and literal for the other three regions ($F_{(1,21)} = 0.40$, $p = .53$, $\eta_p^2 = .02$ for LP, $F_{(1,21)} = 2.60$, $p = .12$, $\eta_p^2 = .11$ for RA, $F_{(1,21)} = 0.04$, $p = .84$, $\eta_p^2 = .00$ for RP).

3.2.2.2. P460-640 ms. A statistical analysis revealed a main effect of TYPE for both the midline ($F_{(1,21)} = 5.73$, $p < .05$, $\eta_p^2 = .21$), and the lateral ($F_{(1,21)} = 4.96$, $p < .05$, $\eta_p^2 = .19$), and a main effect of ANIMACY for both midline ($_{F(1,21)} = 6.82$, $p < .05$, $\eta_p^2 = .25$), and the lateral ($F_{(1,21)} = 5.81$, $p < .05$, $\eta_p^2 = .22$). Further analysis indicated that the waveforms for the metaphors ($M = 4.13$ μV for midline; $M = 2.54$ μV for lateral) were much more positive than for the literals ($M = 3.14$ μV for midline; $M = 1.92$ μV for lateral), both for midline and lateral. Pairwise comparisons showed that the difference between Metaphor and Literal was statistically significant ($p < 0.05$ for both). Meanwhile, the waveforms for the Ha-M conditions ($M = 3.98$ μV for midline; $M = 3.28$ μV for lateral) were more positive than for the Inanimate ones ($M = 2.47$ μV for midline; $M = 1.99$ μV for lateral) both for midline and lateral. There was also a significant interaction between ANIMACY and TYPE for midline ($F_{(1,\ 21)} = 4.32$, $p = .05$, $\eta_p^2 = .17$). Significant interactions were followed by separate repeated-measures ANOVA for metaphorical and literal sentences, with the within-subjects factors of ROI. The statistical analysis showed that Ha-M (4.64 μV) elicited a stronger positive amplitude than Ina-M (3.61 μV) ($F_{(1,\ 21)} = 11.28$ $p < .01$, $\eta_p^2 = .35$), while there was no significant effect ($F_{(1,\ 21)} = 1.31$, $p = .27$, $\eta_p^2 = .058$) between the Ha-L (3.32 μV) and Ina-L (2.96 μV).

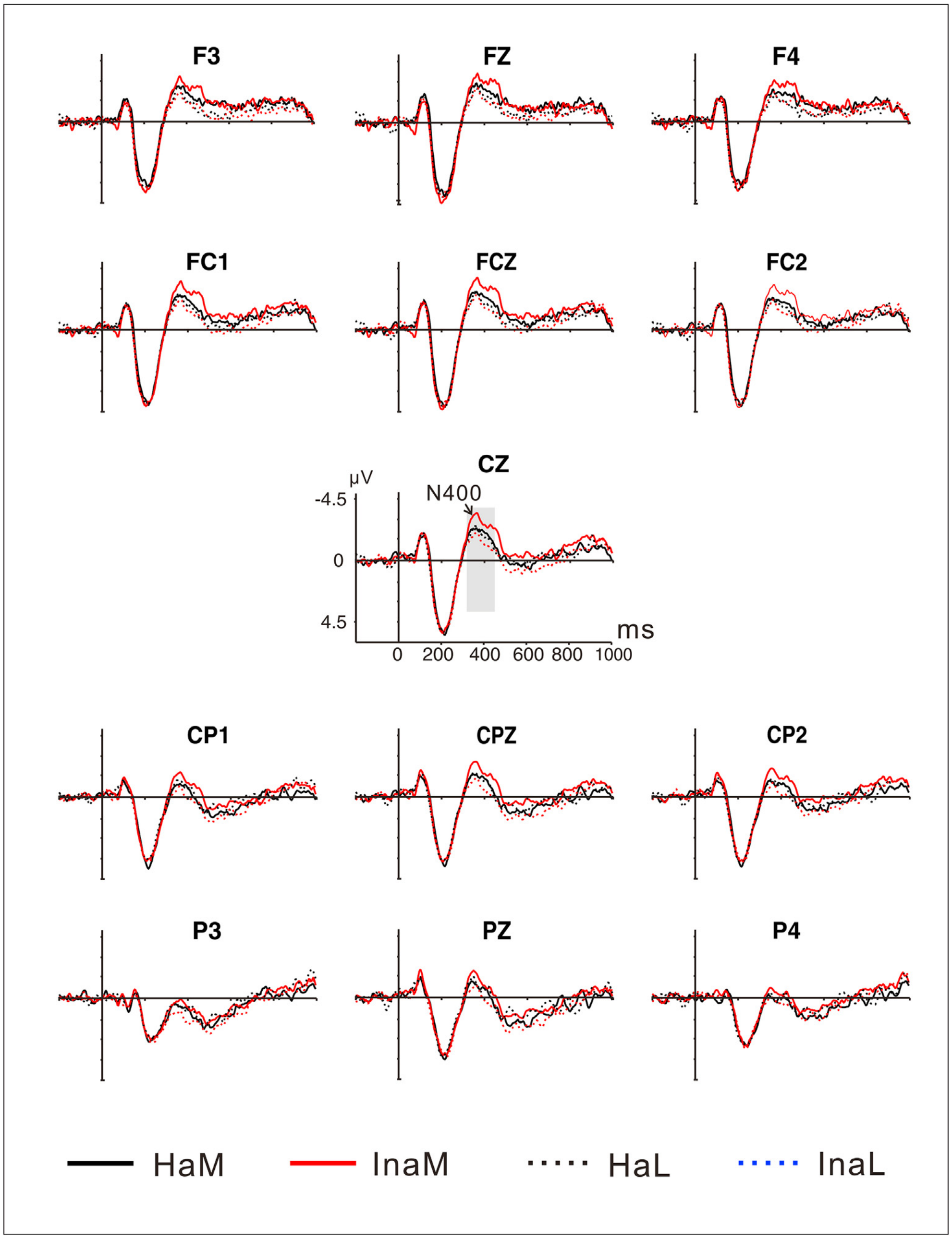

Fig. 2. Grand average ERP waveforms recorded at predicate verb for the chosen electrodes.

3.2.3. Lexical decision task

The same time-windows as the NP2 were adopted in the present analysis (see Table 2B for details regarding the statistical analysis). The grand averaged waveforms for the object nouns in Lexical Decision Task is displayed in Fig. 4.

3.2.3.1. N300-420 ms. For the time window from 300 to 420 ms of object noun, the analysis showed a significant interaction between TYPE and ROI for midline ($F\,(2, 44) = 3.20, p < .05, \eta_p^2 = .42$), and for lateral ($F\,(3, 66) = 8.18, p < .001, \eta_p^2 = .28$). The repeated-measures ANOVA was followed for anterior, central, and posterior separately, with the within-subjects factors of Word Type. The pairwise comparisons showed that there were significant differences between Abstract and Concrete nouns for Anterior (Abstract: $M = 0.027$ μV; Concrete: $M = -1.16$ μV; $p < .05$) and Central (Abstract: $M = 2.10$ μV; Concrete: $M = 0.87$ μV; $p < .05$), with no significant difference for Posterior (Abstract: $M = 3.46$ μV; Concrete: $M = 0.2.62$ μV; $p = .12$).

The repeated-measures ANOVA was followed for each region of lateral (LA, LP, RA, and RP) separately, with the within-subjects factors of Word Type. The pairwise comparisons showed that significant differences between Abstract and Concrete nouns were found

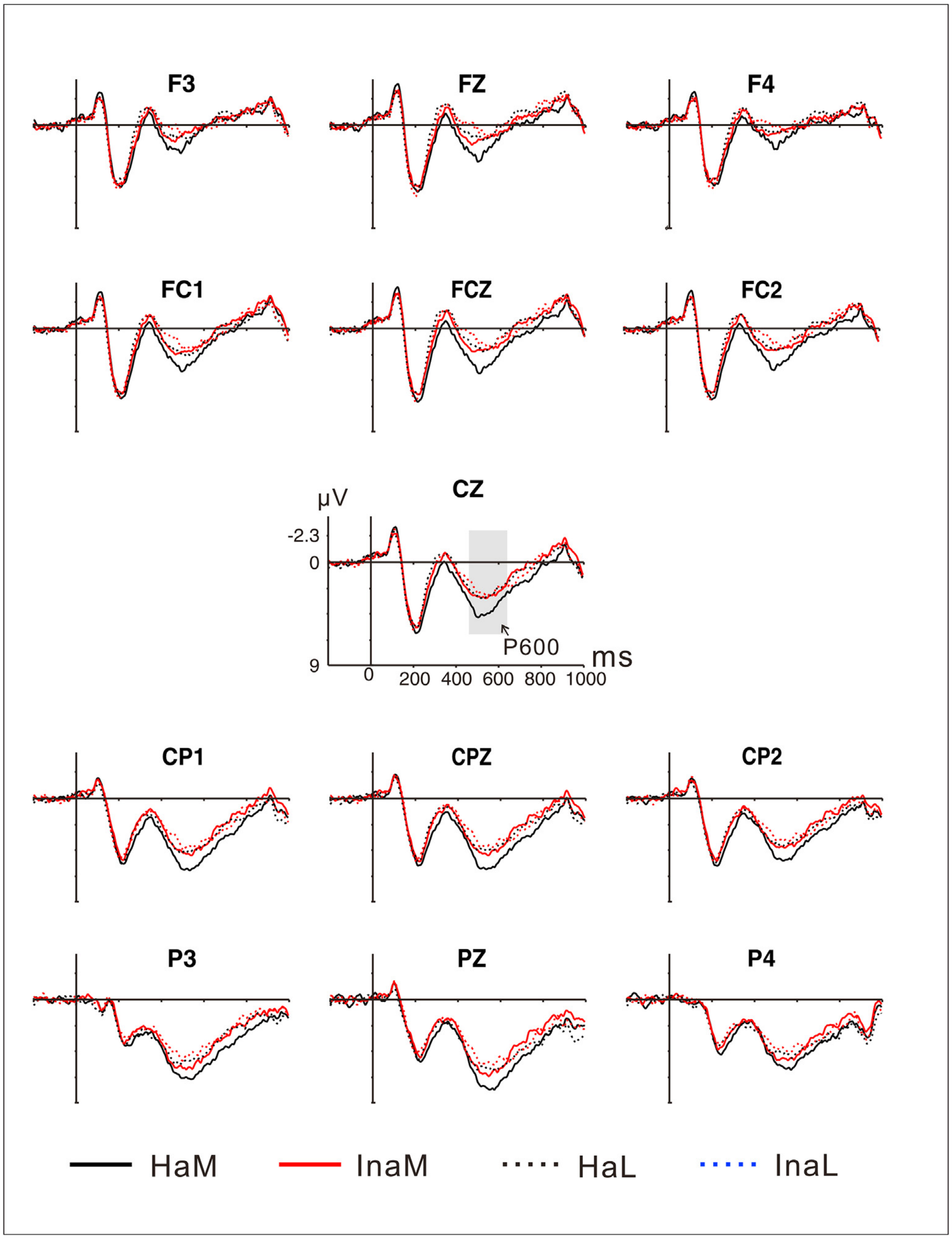

Fig. 3. Grand average ERP waveforms recorded at NP2 for the chosen electrodes.

on the Left Anterior (Abstract: $M = .52$ μV; Concrete: $M = -0.74$ μV; $p < .01$) and Right Anterior regions (Abstract: $M = 0.83$ μV; Concrete: $M = -0.08$ μV; $p < .05$). These differences were not found on the Left Posterior (Abstract: $M = 1.99$ μV; Concrete: $M = 1.43$ μV; $p = .12$) or the Right Posterior (Abstract: $M = 2{,}37$ μV; Concrete: $M = 1{,}71$ μV; $p = .06$).

3.2.3.2. P460-640 ms. The analysis for the 460–640 ms time window yielded significant main effects of TYPE for the both midline ($F_{(2,42)} = 11.27, p = .01, \eta_p^2 = .35$) and the lateral ($F_{(2,42)} = 7.11, p = .01, \eta_p^2 = .25$) electrodes. A significant interaction was also found between TYPE and ROI for midline ($F_{(4,\ 84)} = 11.93, p < .001, \eta_p^2 = .36$), and for lateral ($F_{(4,\ 84)} = 3.83, p < .05, \eta_p^2 = .15$). Pairwise comparison showed that only the Anomalous words category (5.30 μV for midline; 3.44 μV for lateral) was significantly more positive than abstract (3.30 μV, $p < .01$ for midline; 2.23 μV, $p < .05$ for lateral) and concrete (2.88 μV, $p < .00$ for midline; 1.96 μV, $p < .01$ for lateral) nouns. The most important result, however, is that the abstract and concrete words did not yield significant difference ($F_{(1,21)} = 1.89$, $p = .18$, $\eta_p^2 = .08$) for midline ($F_{(1,21)} = 2.46$, $p = .13$, $\eta_p^2 = .11$) and for lateral. This is suggested that the P600 effects of NP2 can be completely attributed to experimental manipulation.

Table 2B
Summary of results for target words in the selected time-windows.

		NP2 noun					
		N300-420			N460-640		
Midline		*F*	*p*	η^2	*F*	*p*	η^2
	Word Type	2.78	.14	.18	12.86	.00**	.38
	Word Type × ROI	3.20	.04*	.42	11.93	.00**	.36
Lateral		*F*	*p*	η^2	*F*	*p*	η^2
	Word Type	3.04	.07	.23	7.11	.01*	.25
	Word Type × ROI	8.18	.00**	.28	3.83	.02*	.15

*:p value reached a significant level at least 0.05.

Although the significant interactions between TYPE and ROI were also found for midline, there was no statistical difference between Abstract and Concrete ($F_{(1,21)} = 1.75$, $p = .68$, $\eta_p^2 = .01$ for anterior; $F_{(1,21)} = 4.00$, $p = .06$, $\eta_p^2 = .16$ for central; $F_{(1,21)} = 0.95$, $p = .34$, $\eta_p^2 = .04$ for posterior).

4. Discussion

This study used ERPs to examine the role of animacy in the metaphorical comprehension of Mandarin Chinese, as well as its spatio-temporal processing dynamics. The metaphorical sentences with SVO structure containing the relationship between actor and undergoer arguments can be used as an innovative approach in metaphor studies. Behaviorally, the participants correctly comprehended the sentences, and showed similar response times across conditions. Based on Coulson and Van Petten's (2002) claim, equivalent processing time in the task does not imply equivalent effort. Indeed, differences in the ease of access, at the neural level, were reflected in fronto-central N400 and centro-parietal P600 amplitudes. The analysis of ERPs data showed that, compared with the literal sentence and metaphorical sentence with the high-animacy NP1, the inanimate metaphor elicited an increased N400 animacy effect on middle processing stage (predicate verb), which is in line with previous studies (Frenzel et al., 2011; Kim & Osterhout, 2005; Weckerly & Kutas, 1999). Contrary to the prior hypothesis, this study found an attenuated P600 effect on the final stage (NP2) for the metaphor with inanimate NP1, compared to their high-animacy counterparts. The graded P600 effect between animate and inanimate sentence suggests that animacy violation may facilitate the integration of the reanalysis stage for metaphor with inanimate NP1; consequently, animacy may modulate the metaphor processing in Mandarin Chinese. Taken together, the animacy effect observed in current metaphorical study suggested that the conflict strength (animacy violation vs. no animacy violation) modulates the time course of metaphor processing in Mandarin Chinese.

4.1. Middle processing stage

As hypothesized, the integration between inanimate actor and verb in the inanimate metaphor condition led to enhanced processing effort and was reflected in the increased N400 amplitude. The animacy violation effects observed in the present study are consistent with a wide range of psycholinguistics studies on diverse languages, for instance, in Dutch (Hoeks et al., 2004; Kim & Osterhout, 2005) and English (Nakano et al., 2010; Weckerly & Kutas, 1999). The increment of the N400 amplitude might index the increased cost in integrating information from various semantic regions (Chwilla D. J. et al., 1995; Federmeier & Kutas, 1999; Hagoort. et al., 2004; Rutter et al., 2012). Several researchers have suggested that once a current item is entered as input, the language processing system will immediately identify the verb's selection limitation and recognize its argument's coarse word category such as animacy (Kuperberg, Choi, Cohn, Paczynski, & Jackendoff, 2010; Paczynski & Kuperberg, 2011). This identification ensures that subject-verb integration generate an intended meaning. Obviously, the inanimate NP1 in metaphor condition (an abstract noun, e.g., *suffering*) is not an ideal Actor for the verb (e.g., *light up*), which semantically requires an animate subject (e.g., *teachers*). The verb's semantic information retrieval from the long-term memory fails to meet predictions by means of pragmatic inference and world-knowledge (Kutas, Van Petten, & Kluender, 2006). Such a failure will cause difficulties in integrating the verb with inanimate head nouns, and then will be reflected in the increased N400. The data suggest that language comprehension system has captured the conflicts of animacy violation at the literal-level integration.

4.2. Final processing stage

With respect to the N400, metaphor processing was not modulated by animacy, which may be due to the fact that the noun concreteness effect counteracts the metaphorical effect. On the left anterior region (LA), however, there was a more negative amplitude for NP2 with literal conditions (concrete nouns) than with the metaphorical ones. This pattern of results (more details will be provided later) is similar to that of the lexical decision task. This anomalous result could also be attributed to the influence of the noun concreteness effect on lexical access. In fact, the effect was rather weak and thereby lateralized (perhaps reflecting the activity of bilateral temporal sources) in the study.

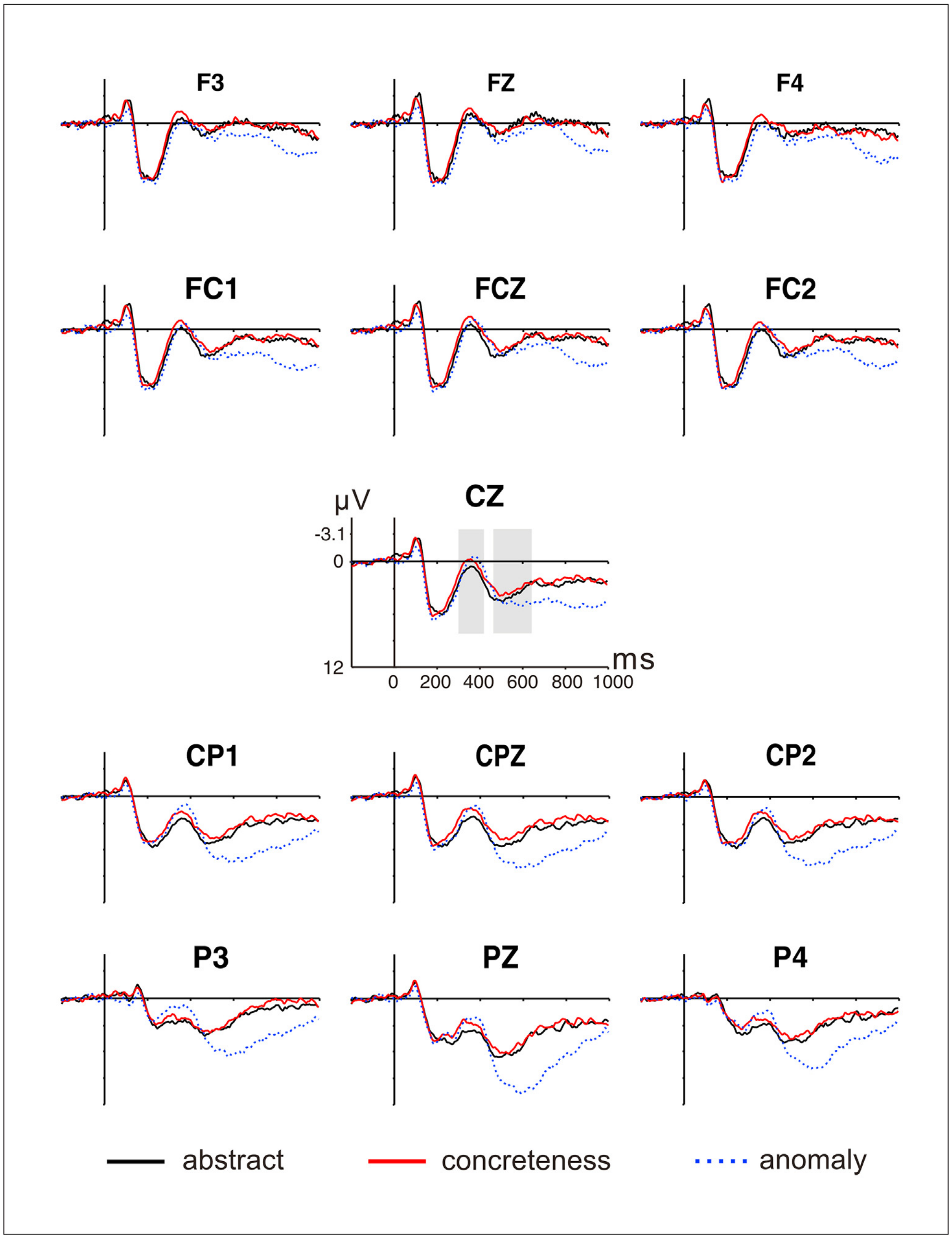

Fig. 4. Grand average ERP waveforms of object nouns in Lexical Decision Task for the chosen electrodes.

As for the P600, the results in the present study found that metaphors with high-animacy head nouns elicited a larger P600 than did literal expressions, which is consistent with previous studies of metaphor (DeGrauwe et al., 2010; Yang et al., 2013). The increment of the P600 in metaphor comprehension is considered to reflect either the extended processing cost of reanalysis in resolving inconsistencies between the literal meaning and the intended one, or the need for integration with the context and input stimuli of multiple semantic regions (Yang et al., 2013). In addition, van Herten, Kolk, and Chwilla (2005) proposed that P600 might reflect the involvement of cognitive or executive control processes in resolving response uncertainty in language processing. Within the framework of the pragmatic model, metaphor processing can be seen as the consequence of a pragmatic inference integrating the context, world knowledge, and lexical meaning of an utterance (Carston, 2010). In other words, understanding metaphor thus involves an inferential move from the literal meaning to the intended one. This process begins with the preconditions in the decoded meaning, connects contextual assumptions, and concludes with a set of conclusions warranted by the preconditions (Wilson & Carston, 2007). Taken together, the larger P600 amplitudes for metaphorical conditions with animate NP1 observed in the current study might reflect the enhanced efforts of reanalysis arising from pragmatically or inferentially-driven relevant mapping processes, including two remote domains.

The interesting finding was the graded effects on P600 amplitude between animate vs. inanimate metaphors, reflecting the different cognitive efforts in metaphor processing. With the same verbs and NP2, two varying metaphor conditions with animate and inanimate NP1 were rated similarly in terms of familiarity, metaphorical tendency, emotional valence and interpretability, even in cloze probability of NP2. However, the high-animacy metaphors (e.g., *The teachers light up the future.*) required greater processing effort than inanimate metaphor sentences did (e.g., *The sufferings light up the future.*). The findings for Ha-M conditions are in line with the indirect process hypothesis (Grice, 1975; Searle, 1979), which posits more effort necessary to process metaphorical as compared to literal expressions. Meanwhile, the results for Ina-M condition seemed in keeping with the directed process hypothesis (e.g., Gibbs, 1994), which suggests that the processing of metaphor does not involve more steps compared with the interpretation of literal utterances (Glucksberg, 2008). Globally, the graded effect of two categories of metaphors demonstrates that animacy affected the metaphorical comprehension, indicating that animacy plays the important role in processing metaphor. Similarly, previous empirical studies have indicated numerous factors modulating the processing costs in metaphor, for example, familiarity, appropriateness, context (Gibbs, 1994; Pynte et al., 1996; Bambini, Gentili, Ricciardi, Bertinetto, & Pietrini, 2011; Weiland et al., 2014). Also, this pattern of results partly supports the Graded Salience Hypothesis (GSH; Giora, 1997), suggesting that prominent meanings modulate metaphor processing (Giora, 2003). In the present study, salient information is not an animate argument in metaphor, but an inanimate argument. Animacy violation triggered by inanimate actor arguments captured the language processing system, and was a preferred mode of processing over non-violated sentences (with animate actors). Thereby, these findings supported that the conflict strength controlled by animacy modulated the metaphorical comprehension.

Most importantly, the result of reduced P600 amplitudes under Ina-M condition is slightly different from those in previous studies on the animacy effect (usually non-figurative language), suggesting that animacy violations hamper processing in the late stage and thus increase processing costs (Weckerly & Kutas, 1999; G. R. Kuperberg et al., 2003; Hoeks et al., 2004; Nakano et al., 2010; Kuperberg et al., 2007). Conversely, the animacy violation here seems to facilitate metaphorical comprehension. One possible interpretation for the graded P600 difference is that this reduced processing cost may arise from a belief revision triggered by the animacy violation prior to the final stage. In the present metaphor case, animacy violation (between the inanimate NP1 and verb) means novelty and breaking conventional expressions. The inanimate actors as sentence-initial nouns seem to be a highly ideal Actor role for verb within metaphors compared with their animate counterparts (usually a person). This function makes it easier for inanimate actors to construct actor-undergoer relationships with inanimate NP2. Hence, the beliefs needed for an integration of the argument relationship were modulated in metaphor processing, which is in complete opposition to the processing of literal expressions mentioned in the introduction. The modulated beliefs may give rise to the predictive function to language processing systems, which allow arguments to be integrated easily when the input meaning conforms to certain cognitive principles. The facilitating effect may be the results of top-down processing, which makes language comprehension proceed speedily (van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005; J.M. Szewczyk, H. Schriefers, 2013), as reflected in the P600 amplitude (Davenport & Coulson, 2011; Van Petten & Luka, 2012).

An alternative explanation is that the reduced processing cost might reflect diffuse activation prepared for mapping processing, which provides the bridge of the two distant domains and allows the appropriate associations to be constructed (Coulson & Van Petten, 2002). Novel metaphor processing is involved in two different and often semantically remote domains requiring mental integration via inferencing or mapping processes (Schmidt, Debuse, & Seger, 2007; Weiland et al., 2014). The semantic distance between the relevant words involved in animacy violation (e.g., *sufferings* and *future.*) is rather remote compared to those of no animacy violations (e.g., *teachers* and *future.*). Accordingly, the comprehension system could activate an extensive semantic network prior to the final stage, so as to prepare for further metaphorical mapping processing. When the final words appear, the system only needs to make use of the activated concepts and inferences to output interpretations. Spreading activation is presumably a parallel process in the brain that makes only small demands on its resources (Kintsch, 2008), which is reflected in the decreased P600 amplitude. The processing pattern of the spreading activation may be a bottom-up mechanism prioritized by the prominent meaning (animacy violation).

Following the modular view (Jerry, 1983), the present evidence alludes to the possibility that metaphor processing is not only a bottom-up mechanism classified according to the reading information or prioritized by the attraction of highlighted meaning; it may also be a top-down mechanism that judges input information according to the established beliefs.

4.3. The noun concreteness effect

The lexical decision task showed that the classic noun concreteness effect existed on the distribution of the fronto-central region and left hemispheres, but only in the time window 300–420 ms, while there is no effect in the time window 460–640 ms. These results partially replicated the earlier findings reporting the only fronto-central distribution (e.g. Holcomb et al., 1999; Kanske & Kotz, 2007; H.A. Barber, Otten, Kousta, & Vigliocco, 2013). This may affect the comparison of brain responses between metaphorical and literal language in the time-window from 300 to 420 ms, but it does not affect the subsequent window from 460 to 640 ms. The most important point is that comparisons within each sentence type were not affected by the noun concreteness effect. Even so, this remains an unavoidable limitation for the present experiment, and improvements are warranted in further studies. Based on the current results, the noun concreteness effect had a limited impact on the sentence comprehension task.

In conclusion, the present study provides the first piece of evidence that animacy affects metaphorical comprehension, and the specific modulation is reflected in sentence initial arguments' animacy. The animacy violations in metaphor may facilitate the integration at the reanalysis stage for sentence comprehension and make the metaphor-relevant mapping process less effortful. We conclude that the conflict strength (animacy violation vs. no animacy-violation) modulates the time course of metaphor processing in Mandarin Chinese. These results yield new insights into the role of animacy in online linguistic comprehension and serve to fill a gap in the ERP literature concerning an interactive perspective on metaphor-animacy in Mandarin Chinese. Future studies on metaphor comprehension should take into account the mechanism of animacy violation that facilitates the metaphor processing, especially with respect to whether it works in a top-down or bottom-up fashion, or both. Doing so will help better characterize metaphor processing and comprehension and inform theoretical accounts of animacy violation.

5. Statement of significance

This study explores the role of animacy in the metaphorical comprehension of Mandarin Chinese. It demonstrates, by applying ERP measurements and observations, that the animacy of sentence initial-arguments (animacy vs. inanimate) specifically affects the integration at the reanalysis stage of metaphor-relevant mapping processes. These results provide preliminary evidence that conflict strength (animacy violation vs. no animacy-violation) contributes to metaphor processing. This is the first study of cognitive linguistics that introduces the issue of animacy effects from the non-figurative language into the field of metaphor, and thus has expanded the existing theoretical accounts of metaphor.

Funding

This work was Supported by the National Natural Science Foundation of China [grant numbers 31871118], by the Fundamental Research Funds For the Central Universities [grant numbers 2016CBZ009], and by the Academic Bookbones Fund of Shaanxi Normal University [grant numbers 18QNGG014].

Author contributions

Haojie Ji and Weiping Hu conceived and designed the study. Haojie Ji, Jinxin Chen, and Shiyang Xu performed the experiments and the statistical analyses. Haojie Ji conducted the interpretation of the data as well as wrote the first draft of the manuscript. Senqing Qi, Yadan Li, Yun Dai and Weiping Hu revised the manuscript critically for important content and provided general advice. All authors have approved the final manuscript.

Declaration of competing interest

The author declare no conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jneuroling.2020.100915.

References

Bambini, V., Bertini, C., Schaeken, W., Stella, A., & Di Russo, F. (2016). Disentangling metaphor from context: An ERP study. *Frontiers in Psychology, 7*, 559.
Bambini, V., Gentili, C., Ricciardi, E., Bertinetto, P. M., & Pietrini, P. (2011). Decomposing metaphor processing at the cognitive and neural level through functional magnetic resonance imaging. *Brain Research Bulletin, 86*(3–4) 0-216.
Barber, H. A., Otten, L. J., Kousta, S. T., & Vigliocco, G. (2013). Concreteness in word processing: Erp and behavioral effects in a lexical decision task. *Brain and Language, 125*(1), 47–53.
Benedek, M., Beaty, R., Jauk, E., Koschutnig, K., Fink, A., Silvia, P. J., et al. (2014). Creating metaphors: The neural basis of figurative language production. *NeuroImage, 90*, 99–106.
Berkum, J. J. A. V., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from erps and reading times. *Journal of Experimental Psychology Learning Memory and Cognition, 31*(3), 443–467.
Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology, 18*(3), 355–387.
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*(11), 2669–2683.
Bornkessel-schlesewsky, I. D., & Schlesewsky, M. (2010). The role of prominence information in the real time comprehension of transitive constructions: A cross-linguistic approach. *Language & Linguistics Compass, 3*(1), 19–58.
Bourguignon, N., Drury, J. E., Valois, D., & Steinhauer, K. (2012). Decomposing animacy reversals between agents and experiencers: An ERP study. *Brain and Language, 122*, 179–189.
Brouwer, H., Fitz, H., & Hoeks, J. (2012). Getting real about semantic illusions: Rethinking the functional role of the p600 in language comprehension. *Brain Research, 1446*(none), 127–143.
Caramazza, A., & Mahon, B. Z. (2003). The organization of conceptual knowledge: The evidence from category-specific semantic deficits. *Trends in Cognitive Sciences, 7*, 354–361.
Caramazza, A., & Shelton, J. R. (1998). Domain-specific knowledge systems in the brain: The animate-inanimate distinction. *Journal of Cognitive Neuroscience, 10*, 1–34.
Carston, R. (2010). Lexical pragmatics, *ad hoc* concepts and metaphor: A relevance theory perspective. *Journal of Italian Linguistics, 22*, 153–180.
Chwilla, D. J., Brown, C. M., & Hagoort, P. (1995). The N400 as a function of the level of processing. *Psychophysiology, 30*, 47–61.
Comrie, B. (1989). *Language universals and linguistic typology.* Chicago: University of Chicago Press.
Coulson, S., & Petten, C. V. (2002). Conceptual integration and metaphor: An event-related potential study. *Memory & Cognition, 30*(6), 958–968.
Dahl, Ö. (2008). Animacy and egophoricity: Grammar, ontology and phylogeny. *Lingua, 118*(2) 0-150.
Dahl, Ö., & Fraurud, K. (1996). Animacy in grammar and discourse. In D. Fretheim, & J. Gundel (Eds.). *Reference and referent accessibility* (pp. 47–64). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Davenport, T., & Coulson, S. (2011). Predictability and novelty in literal language comprehension: An erp study. *Brain Research, 1418*(none), 70–82.
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.
Federmeier, R., & Kutas, M. (1999). Right words and left words: Electrophysiological evidence for hemisphere differences in meaning processing. *Cognitive Brain Research, 8*, 373–392.
Frenzel, S., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2011). Conflicts in language processing: A new perspective on the n400–p600 distinction. *Neuropsychologia, 49*(3), 574–579.
Gibbs, R. W. (1994). *The poetics of mind:figurative Thought,Language,and understanding.* Cambridge,UK: Cambridge University Press.
Giora, R. (1997). Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics, 8*, 183–206.
Giora, R. (2003). *On our mind:Salience,Context and figurative language.* New York, NY: Cambridge University Press.
Glucksberg, S. (2008). How metaphors create categories – quickly. In R. W. Gibbs (Ed.). *The cambridge handbook of metaphor and thought* (pp. 67–83). Cambridge: Cambridge University Press.
Gobbini, M. I., Gentili, C., Ricciardi, E., Bellucci, C., Salvini, P., Laschi, C., et al. (2011). Distinct neural systems involved in agency and animacy detection. *Journal of Cognitive Neuroscience, 23*, 1911–1920.
Goldstein, A., Arzouan, Y., & Faust, M. (2012). Killing a novel metaphor and reviving a dead one: Erp correlates of metaphor conventionalization. *Brain and Language, 123*(2), 137–142.
Grice, P. H. (1975). Logic and conversation. In P. Cole, & J. Morgan (Vol. Eds.), *Speech acts. Syntax and semantics: Vol. 3*, (pp. 41–58). New York: Academic Press.
Hagoort, P. (2004). Integration of word meaning and world knowledge in language comprehension. *Science, 304*(5669), 438–441.
van Herten, M., Kolk, H. H., & Chwilla, D. J. (2005). An ERP study of P600 effects elicited by semantic anomalies. *Cognitive Brain Research, 22*(2), 241–255.
Hoeks, J. C., Stowe, L. A., & Doedens, G. (2004). Seeing words in context: The interaction of lexical and sentence level information during reading. *Cognitive Brain Research, 19*(1), 59–73.
Holcomb, P. J., Kounios, J., Anderson, J. E., & West, W. C. (1999). Dual-coding, context vailability, and concreteness effects in sentence comprehension: An electrophysiological investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 721–742.
Janus, R. A., & Bever, T. G. (1985). Processing of metaphoric language: An investigation of the three-stage model of metaphor comprehension. *Journal of Psycholinguistic Research, 14*, 473–487.
Jerry, A. Fodor. (1983). The modularity of mind. *Philosophical Review, 8*(1), 73–77.
Kanske, P., & Kotz, S. A. (2007). Concreteness in emotional words: Erp evidence from a hemifield study. *Brain Research, 1148*(none), 138–148.
Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language, 52*, 205–225.
Kintsch, W. (2008). How the mind computes the meaning of metaphor. In R. W. GibbsJr. (Ed.). *The cambridge handbook of metaphor and thought* (pp. 295–307). Cambridge: Cambridge University Press.
Kolk, H. H., Chwilla, D. J., Van, H. M., & Oor, P. J. (2003). Structure and limited capacity in verbal working memory: A study with event-related potentials. *Brain and Language, 85*(1), 1–36.
Kuperberg, G. R., Choi, A., Cohn, N., Paczynski, M., & Jackendoff, R. (2010). Electrophysiological correlates of complement coercion. *Journal of Cognitive Neuroscience, 22*(12), 2685–2701.
Kuperberg, G. R., Kreher, D. A., Sitnikova, T., Caplan, D. N., & Holcomb, P. J. (2007). The role of animacy and thematic relationships in processing active English sentences: Evidence from event-related potentials. *Brain and Language, 100*(3), 223–237.
Kuperberg, G. R., Sitnikova, T., Caplan, D., & Holcomb, P. J. (2003). Electrophysiological distinction in processing conceptual relationships within simple sentences. *Cognitive Brain Research, 22*, 117–129.
Kutas, M., & Van Petten, C. (1988). *Event-related potential studies of language. Advances in psychophysiology*. Greenwich, CT: JAI Press139–197.
Kutas, M., Van Petten, C., & Kluender, R. (2006). Psycholinguistics electrified II. In M. J. Traxler, & M. A. Gernsbacher (Eds.). *Handbook of psycholinguistics* (pp. 659–724). (2nd ed.). New York: Elsevier Press.
Lai, V. T., Curran, T., & Menn, L. (2009). Comprehending conventional and novel metaphors: An erp study. *Brain Research, 1284*(none), 145–155.
Lakoff, G., & Johnson, M. (2003). *Metaphors we live by.* Chicago: University of Chicago Press.
Li, P. (1996). The temporal structure of spoken sentence comprehension in Chinese. *Perception & Psychophysics, 58*(4), 571.
Liu, Y., Li, P., Shu, H., Zhang, Q., & Chen, L. (2010). Structure and meaning in Chinese: An erp study of idioms. *Journal of Neurolinguistics, 23*(6) 0-630.
Lopez-Calderon, J., & Luck, S. J. (2014). Erplab: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience, 8*, 213. https://doi.org/10.3389/fnhum.2014.00213.
Lowder, M. W., & Gordon, P. C. (2015). Natural forces as agents: Reconceptualizing the animate-inanimate distinction. *Cognition, 136*, 85–90.
Mac Whinney, B. (2004). A unified model of language acquisition. In J. Kroll, & A. De Groo (Eds.). *Handbook of bilingualism: Psycholinguistic approaches* (pp. 49–67). Oxford: OUP.
Malchukov, A. L. (2008). Animacy and asymmetries in differential case marking. *Lingua, 118*(2) 0-221.
Nakano, H., Saron, C., & Swaab, T. Y. (2010). Speech and span: Working memory capacity impacts the use of animacy but not of world knowledge during spoken sentence comprehension. *Journal of Cognitive Neuroscience, 22*, 2886–2898.
Nieuwland, M. S., & Berkum, J. J. A. V. (2005). Testing the limits of the semantic illusion phenomenon: ERPs reveal temporary semantic change deafness in discourse comprehension. *Brain Res Cogn Brain Res, 24*(3), 691–701.
Nieuwland, M. S., & Van Berkum, J. J. A. (2006). When peanuts fall in love: n400 evidence for the power of discourse. *Journal of Cognitive Neuroscience, 18*(7), 1098–1111.
Paczynski, M., & Kuperberg, G. R. (2011). Electrophysiological evidence for use of the animacy hierarchy, but not thematic role assignment, during verbargument processing. *Language & Cognitive Processes, 26*(9), 1402–1456.
Philipp, M., Bornkessel-Schlesewsky, I., Bisang, W., & Schlesewsky, M. (2008). The role of animacy in the real time comprehension of Mandarin Chinese: Evidence from auditory event-related brain potentials. *Brain and Language, 105*, 112–133.
Pynte, J., Besson, M., Robichon, F., & Poli, J. (1996). The timecourse of metaphor comprehension: An event-related potential study. *Brain and Language, 55*, 293–316.
Rösler, F., Pütz, P., Friederici, A., & Hahne, A. (1993). Event-related brain potentials while encountering semantic and syntactic constraint violations. *Journal of Cognitive Neuroscience, 5*(3), 345–362.
Rutter, B., KröGer, S., Hill, H., Windmann, S., Hermann, C., & Abraham, A. (2012). Can clouds dance? Part 2: An erp investigation of passive conceptual expansion. *Brain and Cognition, 80*(3).
Schmidt, G. L., Debuse, C. J., & Seger, C. A. (2007). Right hemisphere metaphor processing? Characterizing the lateralization of semantic processes. *Brain and Language, 100*(2), 127–141.
Searle, J. R. (1979). Metaphor. In A. Ortony (Ed.). *Metaphor and thought* (pp. 83–111). Cambridge: Cambridge University.
Semino, E., & Steen, G. J. (2008). Metaphor in literature. In R. W. GibbsJr. (Ed.). *The cambridge handbook of metaphor and tought* (pp. 232–245). Cambridge: Cambridge University Press.
Shen, Y. (2008). Metaphor and poetic figures. In R. W. GibbsJr. (Ed.). *The cambridge handbook of metaphor and thought* (pp. 295–307). Cambridge: Cambridge University Press.
Szewczyk, J. M., & Schriefers, H. (2011). Is animacy special?: Erp correlates of semantic violations and animacy violations in sentence processing. *Brain Research, 1368*(none), 208–221.
Szewczyk, J. M., & Schriefers, H. (2013). Prediction in language comprehension beyond specific words: An erp study on sentence comprehension in polish. *Journal of Memory and Language, 68*(4), 297–314.
Tang, X., Qi, S., Wang, B., Jia, X., & Ren, W. (2017). The temporal dynamics underlying the comprehension of scientific metaphors and poetic metaphors. *Brain*

Research, 1655, 33–40.
Tzuyin Lai, V., & Curran, T. (2013). ERP evidence for conceptual mappings and comparison processes during the comprehension of conventional and novel metaphors. *Brain and Language, 127*(3), 484–496.
Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology, 83*, 176–190.
Weckerly, J., & Kutas, M. (1999). An electrophysiological analysis of animacy effects in the processing of object relative sentences. *Psychophysiology, 36*(5), 559–570.
Weiland, H., Bambini, V., & Schumacher, P. (2014). The role of literal meaning in figurative language comprehension: Evidence from masked priming ERP. *Front. Hum. Neurosci. 9*, 583. https://doi.org/10.3389/fnhum.2014.00583.
Wilson, D., & Carston, R. (2007). A unitary approach to lexical pragmatics: Relevance, inference and ad hoc concepts. In N. Burton-Roberts (Ed.). *Advances in pragmatics* (pp. 230–260). Palgrave, 2007.
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. *Journal of Neurolinguistics, 26*(2).
Zhang, Q., Guo, C. Y., Ding, J. H., & Wang, Z. Y. (2006). Concreteness effects in the processing of Chinese words. *Brain and Language, 96*(1), 59–68.
Zhang, Y. X., Yu, J., & Boland, J. E. (2010). Semantics does not need a processing license from syntax in reading Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(3), 765–781.