

*Hongjun Chen, Xiaoshuang Peng, Qilin Lu, and Huili Wang
ERP Differences Between Literary and Non-literary Metaphors

Abstract: In this work the distinction between literary and non-literary metaphor comprehension is addressed by neurophysiological methodology to further explore the neural mechanisms of metaphor processing and promote the understanding of the differences between literary and non-literary metaphors. We have used event-related potentials (ERPs) in experiments to examine whether differences in neural mechanisms exist between the two. Amplitudes of the N400 and the Late Positive Component (LPC) ERP components (350–400ms, 750–800ms respectively) were more negative for literary metaphors when compared with literal sentences and non-literary metaphors. This study has produced empirical evidence for the differences between literary metaphor and non-literary metaphor comprehension, which expands the relevant studies and possibly helps to advance the understanding of how metaphor in literature is employed.

Keywords: ERP; LPC; literary metaphor; metaphor neural mechanism; N400

***Corresponding author, Hongjun Chen:** Language and Cognition Institute, School of Foreign Languages, Dalian University of Technology, Dalian, China, E-mail: chenhj@dlut.edu.cn

Xiaoshuang Peng: Foreign Language Department, Anhui Sanlian University, Hefei, China, E-mail: pengxiaoshuang@yahoo.com.cn

Qilin Lu: Institute of Neuroinformatics, Dalian University of Technology, Dalian, China, E-mail: qinli.lu@yahoo.com

Huili Wang: Language and Cognition Institute, School of Foreign Languages, Dalian University of Technology, Dalian, China, E-mail: huiliw@dlut.edu.cn

1 Introduction

Consider a sentence such as *My nurse is an angel*. In this sentence, some characteristics of an angel, like “beautiful” or “patient”, are transferred to that of a nurse. Such expressions are not supposed to be understood literally but convey connotative meanings metaphorically. As a matter of fact, metaphorical statements are omnipresent in people’s daily communication (Richards, 1936; Gibbs, 1994; Glucksberg, 2003) and are also prevalent in a

wide range of discourses when utilized in the description of scientific theories (Boyd, 1993) and in political discussions (Nelkin, 2001). Richards (1936) posited that beyond the sciences metaphoric language was seldom observed. Lakoff and Johnson (1980) proposed even more extremely that the human conceptual system is “fundamentally metaphorical in nature”.

How differently metaphor is processed from literal meanings is a problem that has puzzled researchers for many years. Numerous studies have been carried out to discover whether metaphorical expressions are processed directly or indirectly. According to their research, scholars have proposed different models of metaphor comprehension.

The conceptual metaphor view has suggested that the conceptual system used by people to think about things is fundamentally metaphorical in nature (Lakoff & Johnson, 1980). Lakoff and Johnson have claimed that a metaphor is a mapping of the structure of a source model onto a target model. In other words, people can make use of information from a concrete semantic domain to understand concepts in another, more abstract domain. In Lakoff’s opinion, metaphors should be automatically comprehended, without the effort of awareness.

A different model proposed by Gentner and Bowdle (2001) indicated that novel and conventional metaphors were processed in different ways. According to “the career of metaphor” theory, conceptual mappings between different domains can be completed by either comparison or categorization, and determined by the degree of conventionality of the applied metaphors. In other words, as metaphors become conventionalized, a shift in the mode of processing moves from comparison to categorization. Novel metaphors are comprehended as comparisons, in which the alignment of the target concept with the base concept is structurally accomplished. Conventional metaphors may be comprehended either as a process of matching the target domain with the base domain (a comparison process) or as an idea of regarding the target as a member of the superordinate metaphoric category belonging to the base domain (a categorization process). Various empirical studies support the claim that compared to novel metaphors, conventional metaphors are more quickly processed.

The graded salience hypothesis explains people’s understanding of novel and conventional metaphors in a new perspective. According to this hypothesis, the factor that determines precedence of access is salience (Giora, 1997). During the initial stages of metaphor comprehension, salient meanings, either metaphorical or literal, are always automatically processed. For example, in comprehending conventional metaphors (e.g., “life is a

journey”), both their literal (e.g., trip) and metaphoric (e.g., life) meanings will be activated. To be more specific, the metaphorical meaning in the conventional metaphor is more salient than the literal one, so in “dead” metaphors the metaphorical meanings are activated first. However, in comprehending novel metaphors, their literal meanings may be initially activated, while metaphorical meanings are slow in processing because of the rejection of literal meanings of the metaphorical expression in the first place. The graded salience hypothesis demonstrates that since the figurative meanings in conventional metaphors are salient, the speed for comprehending conventional metaphors and literal statements should be the same, while non-salient, novel metaphors should be processed by discarding the salient literal meaning at the initial stage. Therefore, compared to conventional metaphors, novel metaphors are less salient and slower to be processed.

Event-related potentials (ERPs) are very small voltages generated in the brain structure in response to specific events or stimuli (Blackwood & Muir, 1990). ERPs can be elicited by a wide variety of sensory, cognitive, or motor events. They are thought to reflect the summed activity of postsynaptic potentials produced when a large number of similarly oriented cortical pyramidal neurons (in the order of thousands or millions) fire in synchrony while processing information (Peterson, Schroeder, & Arezzo, 1995). With very high temporal resolution, ERPs are ideal for the research of the time course of sentence processing. Furthermore, ERPs allow random presentation of stimuli across conditions. The use of multiple scalp electrode sites makes it possible to infer the location of the source of activation. A good number of ERP experiments have been conducted to investigate the internal process of metaphor comprehension. Some of the experiment materials are from Indo-European languages with alphabetic orthographies, such as English (Tartter, Gomes, Dubrovsky, et al, 2002; Lai, Curran, & Menn, 2009), Spanish (Sotillo, Carretie, Hinojosa, et al., 2005), French (Pynte, Besson, Robichon, et al., 1996; Bonnaud, Gil, & Ingrand, 2002; Iakimova, Passerieux, Laurent, et al., 2005), and the Semitic language Hebrew (Arzouan, Goldstein, & Faust, 2007). This has led to rising curiosity about whether languages in the Chinese-Tibetan language family, which use pictographic writing systems such as Chinese, cause the same neural activation during metaphor processing. Some Chinese researchers have used the ERP paradigms to explore the unique features of neural mechanisms used to process written Chinese metaphors (Wang, 2007; Chen, Tang, Wang, et al., 2009).

In this study, the two ERP components involved are the N400 and the Late Positive Component (LPC). The N400 ERP component is one of the widely used language-related ones, a negative-going wave starting at around 200–250ms and peaking at around 400ms after stimulus onset. The modulation of N400 is thought to show the ease of connecting the meanings of the incoming words to semantic structure in memory with sentence and discourse level context (Federmeier, Wlotko, De Ochoa-Dewald, & Kutas, 2007). In recent years, the N400 has been widely reported to study metaphor comprehensions (Pynte et al., 1996; Tarter et al., 2002; Coulson & Van Petten, 2002; Iakimova, et al., 2005; Arzouan et al., 2007).

Pynte et al. (1996) examined the differences in familiar vs. unfamiliar nominal metaphors in French. In their experiments, participants were asked to read familiar or unfamiliar metaphors presented alone or preceded by a relevant or irrelevant context. They found that in spite the fact that both familiar and unfamiliar metaphors elicited larger N400s than literal sentences, regardless of the metaphor familiarity, metaphors appropriate to the context elicited an N400 smaller than did metaphors inappropriate to the context.

Iakimova et al. (2005) investigated how both people with schizophrenia and non-patients comprehended highly conventionalized “dictionary metaphors” in French. Their findings suggested that in all participants, incongruous sentences elicited the largest N400, while literal sentences elicited larger N400 than metaphorical ones.

Arzouan et al. (2007) examined conventional metaphorical word pairs (e.g., “lucid mind”) and novel ones (e.g., “conscience storm”) in Hebrew. They found that novel metaphorical pairs were more difficult to process than conventional ones. They suggested that novel and conventional metaphors seemed to be accessed in a similar way, but differ in the degree of processing difficulty. Wang Xiaolu (2009) compared conventional vs. novel nominal metaphors (e.g., “护士是天使 Hu shi shi tian shi (‘nurses are angels’)” vs. “爱情是玫瑰 ai qing shi mei gui (‘love is a rose’)”) in Chinese. She found that novel metaphors elicited larger N400s than conventional metaphors and literal sentences.

However, inconsistent results have been produced by ERP studies on metaphor. Some studies suggested that metaphors are no more difficult to process than the literals (Pynte et al., 1996; Iakimova et al., 2005). Others suggested that metaphors elicited larger N400s than the literals (Arzouan et al., 2007; Wang, 2007). We suggest that the conflicting findings may have

arisen from the inconsistent linguistic forms of metaphor stimuli (e.g. nominal vs. sentential metaphors) and varying degrees of conventionality.

Another widely used ERP component in studies of language processing is a group of positive-going waves that usually begin 500–600 ms and extend until approximately 900ms after word onset. This is the P600 that derives from semantic implausibilities. Various studies have found that no single factor leads to this effect, but rather it is elicited by a set of factors: the degree of implausibility expressed in critical words, the degree of contextual constraint for words with an alternative interpretation, the ability of the participants in completing the tasks, and individual differences in memory retrieval capacity (Kuperberg, 2007). However, none of these factors alone is sufficient for eliciting a P600. A general consensus has been reached that additional analysis is helping to achieve the conviction that the participants interpret the input fully at the final stage (see Kuperberg, 2007; Van de Meerendonk, Kolk, Vissers, et al., 2010). Some ERP studies have been conducted to investigate LPC in relation to metaphor comprehension (Pynte et al., 1996; Coulson & Van Petten, 2002; Iakimova et al., 2005; De Grauwe, Swain, Holcomb, et al., 2010).

Though the studies of Pynte et al. (1996), as we have discussed in the previous section, found a larger-amplitude N400 elicited by metaphorical sentence endings than by literal sentence endings, they found no significant LPC differences between final words of the metaphorical (vs. literal) sentences.

Iakimova et al. (2005) found that anomalous words elicited larger LPC than metaphorical and literal words. However, metaphorical and literal words evoked similar LPC effect, which led them to the conclusion that metaphorical meaning was processed immediately during metaphorical sentence comprehension.

De Grauwe et al. (2010) examined the time-course of processing metaphorical and literal sentences using ERPs. Twenty-four normal participants performed the task of reading familiar nominal metaphors, literal, and semantically anomalous sentences and were judged by whether or not they were meaningful. They found that a significantly larger LPC was evoked by metaphorical than by the literal sentences, which is in line with the Pynte et al. (1996) findings.

The conflicting findings from these ERP studies may have resulted from some methodological and theoretical disparities. For instance, in some studies, critical words used across conditions may be different in the degree of frequency, concreteness and imageability (e.g., Laurent, Denhieres,

Passerieux, et al., 2006; Tartter et al., 2002). In most previous studies, whether syntactic structure or complexity was consistent between the literal and metaphorical sentences is unclear (Coulson & Van Petten, 2002, 2007; Iakimova et al., 2005; Laurent et al., 2006), while both of them are factors suggested to affect the amplitude of LPC (e.g., Friederici, Hahne, & Saddy, 2004; Kuperberg, 2007). The LPC's amplitude is also known to be affected by different tasks designed in different studies, such as reading for comprehension (Pynte et al., 1996), lexical decisions (Laurent et al., 2006), plausibility judgments (Arzouan et al., 2007; Iakimova et al., 2005) and sensibility ratings (Lai et al., 2009). Individual differences between participants may also lead to different ERP results (Katzmerski, Blasco, & Dessalegn, 2003). Furthermore, metaphors with different degrees of conventionality may have been used in different studies, and in some cases, more than one type of metaphor was used in the same study (Iakimova et al., 2005; Tartter et al., 2002), which may contribute the different ways of metaphor processing (Arzouan et al., 2007; Giora 2003).

We argue that these studies neglected a fundamental issue, that is, a strictly defined standard to distinguish novel metaphors from conventionalized ones. In some studies, metaphors selected from poems have been regarded as novel, for example, in the novel metaphor stimuli in Goldstein et al.'s experiment (Goldstein, Arzouan, & Faust, 2008). Whether or not metaphors in literature are more creative and novel than metaphors outside literature is an issue that is difficult to resolve empirically (Semino & Steen, 2008).

Metaphorical expressions drawn from literature are acknowledged to be more original, novel, complex, and difficult to understand than those found in non-literary discourses. Most scholars seem to agree that men of letters intentionally create novel metaphors to go beyond our ordinary conceptual resources and to arouse interesting and complex images. Therefore, such metaphors are assumed to be different from metaphors outside literature.

However, in which way metaphors in literature differ from metaphors elsewhere is a controversial issue. Some scholars put emphasis on the discontinuity between literary metaphors and non-literary metaphors, while others put emphasis on the continuity between literary metaphors and non-literary metaphors.

As a particularly important representative of modern linguistic approaches to literary texts, the Formalist view of literature is characterized by "the aesthetically intentional distortion of the linguistic components of the work, in other words the intentional violation of the norms of the

standard” (Mukařovský, 1970). In the Formalist view, metaphorical expressions are regarded as a particular type of linguistic deviation including the semantic level of language, for it is argued that metaphorical expressions are, if understood literally, absurd, illogical, or nonsensical (e.g., Short, 1996: 43). Though the studies mentioned above have realized the cognitive functions of metaphor (e.g., Leech, 1969: 158; Nowotny, 1965: 60), they put stress mainly on the linguistic level of metaphorical expressions. Addressing another perspective, Tsur’s (1987, 1992) cognitive poetics adopted cognitive theories to systematically explain “the relationship between the structure of literary texts and their perceived effects” (1992: 1). Tsur focused on explaining in which way the peculiar characteristics of individual novel metaphors in poetry contribute to unique effects. He suggested that logical contradictions found in metaphorical expressions are resolved by canceling irrelevant features of the vehicle and projecting the remaining features to the tenor (1987: 79; 1992: 209). Since the 1980s cognitive theorists have re-evaluated the function of metaphors in everyday, non-literary language, and explained metaphors in literature from a new perspective. Lakoff and Turner (1989) claimed that metaphors found in poetry were created by using conceptual metaphors that also underlie everyday metaphorical expressions. In other words, the seemingly novel metaphors in poetry were actually realized by using the same metaphorical concepts we all use in everyday language. Therefore, this approach attributed primacy to metaphor outside literature, and viewed metaphor in literature as the creative product of everyday, non-literary metaphors. Although the approaches discussed thus far differ from each other in some aspects, they share the same assumption, that is, that literary metaphors are more novel. However, whether metaphors in literature are more creative and novel than metaphors outside literature is an issue that is hard to prove empirically. Little work has been done in this area to date, for literary scholars generally have analyzed selected texts in testing their hypotheses. A notable exception is Goatly (1997), who compared metaphorical expressions extracted from six different genres (in English). He found that compared to other genres, modern lyric poetry has more novel and extended metaphors. However, many more factors need to be taken into consideration, for example, methodology (e.g., Crisp, Heywood, & Steen, 2002; Heywood, Semino, & Short, 2002), the relevant dimensions of metaphor (Steen, 1999) and the size of the data samples. Making use of informants is another approach to compare metaphor in literature and metaphor elsewhere.

In their study, Katz et al. (1988) presented 204 literary and 260 non-literary metaphors to assess 10 psychological dimensions, and asked 634 informants to judge their degree of ease in comprehension, metaphoricity, imageability, and so on. The result did not show substantial differences between literary and non-literary metaphors. However, even if the data used in this experiment were reliable, other methodological problems need to be taken into consideration. Steen (1994) reported that literary metaphors did differ from journalistic metaphors when measuring various dimensions for two languages: Dutch and English.

Although scholars have taken various approaches to establish differences between literary and non-literary metaphors, few studies have been conducted by means of electrophysiological methods. The present study is an attempt to overcome this deficit by making use of event-related potentials (ERPs) to shed light on whether there exist differences between literary and non-literary metaphor comprehension.

The research foci of the present study are as follows: to determine (1) whether there exists neural mechanism differences between comprehension of Chinese metaphor in literature and non-literary metaphors, and (2) whether the amplitudes of N400 and LPC provide sensitive indices of differences between Chinese literary metaphors and non-literary metaphors.

2 The Experiment

In the present study we used ERPs to examine the processing differences between metaphor in literature and metaphor outside literature, and directly compared them to the processing of literal sentences. Manipulating the effects of familiarity should contribute to characterizing the processing of metaphor in literature. Similar to Pynte et al. (1996), participants were asked to perform the task of judging whether the expressions are metaphorical or not, and in this way metaphorical expressions that were comprehended could be separated from those that were not.

2.1 Subjects

Twenty-four (fourteen men and ten women) undergraduates and postgraduates at Dalian University of Technology participated in the study

for monetary compensation. These participants were between 22 and 27 years in age (mean: 24; S.D.: 1.38). All subjects were right-handed, native Chinese speakers, and had normal or corrected to normal vision. None had any neurological disorder or major head injury that was diagnosed as having a long-term side effect. Data were discarded from 3 subjects due to excessive ocular artifacts from EEG recordings.

2.2 Stimuli

In order to eliminate confounding factors that would undermine the experimental results, the experimental stimuli had to be created according to certain requirements: first, the number of target words (Chinese characters) should be the same; second, the syntactic structure for each type should be identical. In this case noun-noun constructions with a literal subject and a literal/metaphoric predicate (e.g., literal expression: “芭蕾舞是舞蹈” Ba li shi wu dao (‘ballet is a kind of dance’); literary metaphor: “金柳是新娘” Jin liu shi xin niang (‘golden willow is a bride’); non-literary metaphor: “水利是命脉” Shui li is ming mai (‘hydraulic projects are lifelines’) were chosen.

A total of 240 sentences in 3 groups (80 literal sentences, 80 metaphors in literature, and 80 metaphors outside literature) were selected as the stimuli. Following Katz and his colleagues’ studies in 1988, our literary metaphors were obtained in the following manner: First, this author consulted a professor of Chinese, who recommended anthologies of modern and contemporary poems written in Chinese. The sources are mainly the poems composed by well-known Chinese poets 徐志摩 (Xu Zhi Mo), 朱自清 (Zhu Zi Qing), 冰心 (Bing Xin), 余光中 (Yu Guang Zhong), 汪国真 (Wang Guo Zhen), 舒婷 (Shu Ting), 海子 (Hai Zi), and so on. This author then searched these works for specific figurative expressions, which were rewritten, when necessary, to adapt to the “A is B” structure. During this process, an attempt was made to maintain the original words and elaborations. Finally, a professor of Chinese fiction was asked to remove the inappropriate selections from the pool. The final sample of metaphors consisted of 80 clauses, representing the works of 11 different poets. The non-literary metaphors and literal sentences were obtained from economic, political, and various discourses while other literary discourses were excluded.

One pretest was conducted in order to check whether the stimuli were familiar. 100 native Chinese speakers from Dalian University of Technology volunteered for participation. The participants were instructed to rate each

sentence on a scale from 1 to 5. The instructions for the familiarity scale were: “If you have heard similar expressions frequently before and feel that the meaning is highly familiar, give it a 5. If you have heard similar expressions relatively frequently before and feel that the meaning is familiar, give it a 4. If you have heard similar expressions occasionally before and feel that the meaning is somewhat familiar, give it a 3. If you have heard similar expressions once or twice before and feel the meaning is somewhat unfamiliar, give it a 2. If you have never heard it before and feel that the meaning is unfamiliar, give it a 1”.

After the familiarity test, the familiarity degree of each sentence was analyzed statistically. Literary metaphors/non-literary metaphors/literal sentences that were rated as familiar and highly familiar were chosen to be the stimuli in the formal experiment. In the ERP experiment, each subject saw 40 literary metaphors, 40 non-literary metaphors and 40 literal sentences. The sentence length in Chinese was five or six characters. Using E-prime software, each sentence was represented in three pictures, that is, the subject, the linking verb, and the target word were each given a representative picture; 120 sentences were split into 4 blocks. As a result, each block contained 10 literary metaphors, 10 non-literary metaphors and 10 literal sentences. Sentences in each block were then randomized.

2.3 Procedure

Participants first completed a consent form and a subject information form including gender, major, language proficiency, etc., followed by Sensor Net setup and a brief practice session. The main experiment was conducted in a quiet room with white noise in the background and dim light. The participants were asked to horizontally gaze at a black dot in the middle of the screen, which was 40–60cm away from the subjects. The experimental paradigm was partially replicated from the paradigm of Wang (2009) in stimulus presentation. Each picture in each sentence was presented for 500ms. At the offset of the sentence-final target word, a dark screen was presented for 2500ms. Upon seeing the dark screen, participants were asked to judge whether the sentence bore a metaphorical meaning or not. When the subject decided that the sentence had no metaphorical meaning, he/she had to press the button “f”; when the subject decided that the sentence had metaphorical meaning, he/she had to press the button “j”. Once a response

was made, the program moved on to the next trial. Subjects were instructed to refrain from facial, eye, and bodily movement during the task.

3 ERP Recording and Data Analysis

3.1 ERP Recording

The electroencephalogram (EEG) was continuously recorded from a 32 channel-Quick cap (Neuroscan Inc.). Electrode impedances were kept below 5 K Ω . The EEG signals were continuously recorded with a band-pass from 0.05 to 100 Hz with a sampling rate of 500Hz. Off-line ERPs were averaged, within each type for each subject, and time-locked after the onset of the final words. Average waveforms were digitally low pass filtered (20 Hz cut-off frequency). Average ERPs, from -200ms to 1000ms after the onset of the target words, were computed as a function of sentence types with 200ms before the onset of the target words as the baseline. Data of three participants were excluded due to heavy ocular artifacts, as noted previously.

ERPs were derived by averaging correctly classified trials on each condition for each participant, that is, literal sentences judged as not conveying a metaphorical meaning, non-literary metaphors and literary metaphors judged as conveying a metaphorical meaning. The mean number of trials averaged for the three conditions was 30.

After visually checking each waveform, we chose two different time windows (i.e., 350–400ms, 750–800ms) to capture the N400 and the P600. ERPs were analyzed separately for 4 groups of electrode sites (i.e. anterior, midline, central, and posterior).

3.2 Statistical Analysis

Only RTs and ERPs for correct responses were considered in the analysis. A two-way ANOVA was performed for 4 groups of electrode sites (i.e. anterior, midline, central, and posterior) and three levels of condition factors. When appropriate, the Greenhouse-Geisser correction was applied (Greenhouse & Geisser, 1959).

4 Results

4.1 Behavior Results

Behavior data consisted of two factors: participants' judgment accuracy and reaction times for the stimuli. According to the purpose of this study, the means of judgment accuracy and reaction time were analyzed using repeated measures of ANOVA. For the reaction times, there was a main effect between conditions [$F(3, 57) = 6.944, p = 0.003 < 0.05$]. Pair-wise comparisons confirmed that except for no obvious differences between literal conditions and non-literary conditions, each condition differed from all others. For the judgment accuracy, a main effect between conditions was found [$F(3, 57) = 20.321, p < 0.05$]. Pair-wise comparisons indicated that each condition differed from all others.

For rate of errors, pair-wise comparisons revealed that comparison between non-literary metaphors and literary metaphors was significant ($p = 0.016$); comparison between non-literary metaphors and literal sentences was significant ($p = 0.016$); and a significant difference was also found between literary metaphors and literal sentences ($p < 0.05$). The result showed that literary metaphors were the most difficult to be processed, next were the non-literary metaphors, while literal sentences were processed most easily.

In reaction time, pair-wise comparisons confirmed that participants took a longer time to decide the meaningfulness of literary metaphors than that of non-literary metaphors ($p < 0.001$). No significant difference was found between non-literary metaphors and literal sentences ($p = 0.956$), while there exists a significant difference between literary metaphors and literal sentences ($p < 0.05$). This, to some extent, indicates that people can directly reach the metaphorical meanings of the non-literary metaphors without the first stage of discarding their literal meanings, but more solid evidence needs to be explored in ERP data.

4.2 Event-related Potentials

ERPs were derived by averaging correctly classified trials on three conditions for each participant. The waveforms in the present study were characterized by N400 and a later positive deflection, which appeared more negative for

the literary metaphor condition. Grand average waveforms of three conditions are presented at selected electrode sites in Figure 1. A negative-positive N2/P3 complex can be seen in the first 300ms after onset of the target words. The N2/P3 was followed by a negative-going N400. The N400 was investigated across the time-window: 350ms–400ms, and LPC were also investigated across the time windows: 750ms–800ms.

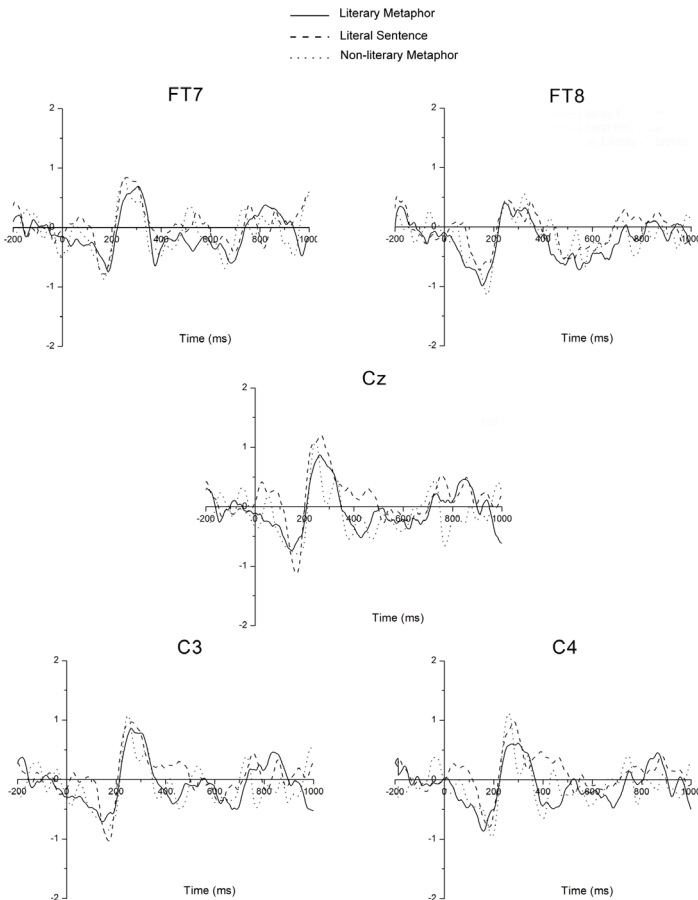


Figure 1: Grand average ERP waveforms recorded at selected electrode sites for the literary metaphor (solid lines), non-literary metaphor (dotted lines), and literal sentence (dashed lines) conditions

For all these time windows, ANOVAs, including three levels of Condition (literary metaphors, non-literary metaphors, and literal sentences), indicated significant main effects and/or interactions between condition and electrode sites ($p < 0.05$). Therefore, we report the results of planned pair-wise ANOVAs that compared each condition with one another. The focus was on the main effects and interactions concerning Condition, which were of greatest theoretical interest. The interactions between Condition, Hemisphere, and/or electrode sites (ES) not noted in the following section were not significant ($p > 0.05$). Near-significant main effects and interactions ($p < 0.1$) are discussed only when they were accompanied by at least one other significant main effect or interaction at another column.

Table 1: Pair-wise ANOVAs comparing ERPs to each condition in the N400 and LPC time windows (correct responses)

	Effect	N400	LPC
		F value	F value
Literary vs. literal			
Midline	C	10.589***	7.898**
	C X ES	1.828	4.217**
Anterior	C	3.204*	5.489**
	C X ES	1.871	1.601
Central	C	3.190*	4.271*
	C X ES	4.564**	9.557**
Posterior	C	5.087**	1.084
	C X ES	2.312	2.398
Non-literary vs. literal			
Midline	C	2.018	1.646
	C X ES	0.112	0.693
Anterior	C	0.797	0.541
	C X ES	0.542	0.423
Central	C	1.944	1.506
	C X ES	2.483	0.029
Posterior	C	3.758*	0.852
	C X ES	0.717	0.008

	Effect	N400	LPC
Literary vs. Non-literary			
Midline	C	7.143**	4.916**
	C × ES	2.509**	2.439**
Anterior	C	5.781**	3.581*
	C × ES	1.900	1.744
Central	C	4.345**	4.398**
	C × ES	6.841**	0.442
Posterior	C	0.010	0.462
	C × ES	3.869**	4.224**

C: main effect of Condition, degrees of freedom 1, 20. C × ES: interaction between Condition and electrode sites, degrees of freedom 5,100 (midline), 3, 60 (anterior), 1, 20(central), 2, 40(posterior). C × H: interaction between Condition and Hemisphere, degrees of freedom 1, 20. C × ES × H: interaction between Condition, electrode sites and Hemisphere, degrees of freedom 5,100 (midline), 3, 60 (anterior), 1, 20(central), 2, 40 (posterior).

*p<0.1

**p<0.05

***P<0.01

****p<0.001

4.2.1 N400: 350–400ms

Literary vs. literal. Figure 2(a) shows the average ERP waveforms at selected electrodes for the two conditions. According to the output of the ANOVAs, in the time window of 350–400ms, the waveform to the literary metaphor was more negative than that elicited by literal sentence, reflected by significant main effects of Condition at midline and posterior sites and by a significant Condition by ES interactions at central sites.

Non-literary vs. literal. Figure 2(b) shows the average ERP waveforms at selected electrodes for these two conditions. According to the output of the ANOVAs, in the time window of 350–400ms, the direct contrast between the non-literary metaphor and literal sentence appeared to show a more negative late N400 to the non-literary metaphor than to the literal sentence, reflected by Condition by Electrode sites interactions that reached significance at

central and posterior sites, but the main effects of Condition only approached significance at posterior sites.

Literary vs. non-literary. Figure 2(c) on next 3 pages shows the average ERP waveforms at selected electrodes for the two conditions. According to the output of the ANOVAs, in the time window of 350–400ms, the waveform to the literary metaphor was more negative than that elicited by the non-literary metaphor, reflected by significant main effects of Condition at midline, anterior and central sites and by a significant Condition by ES interaction at posterior sites.

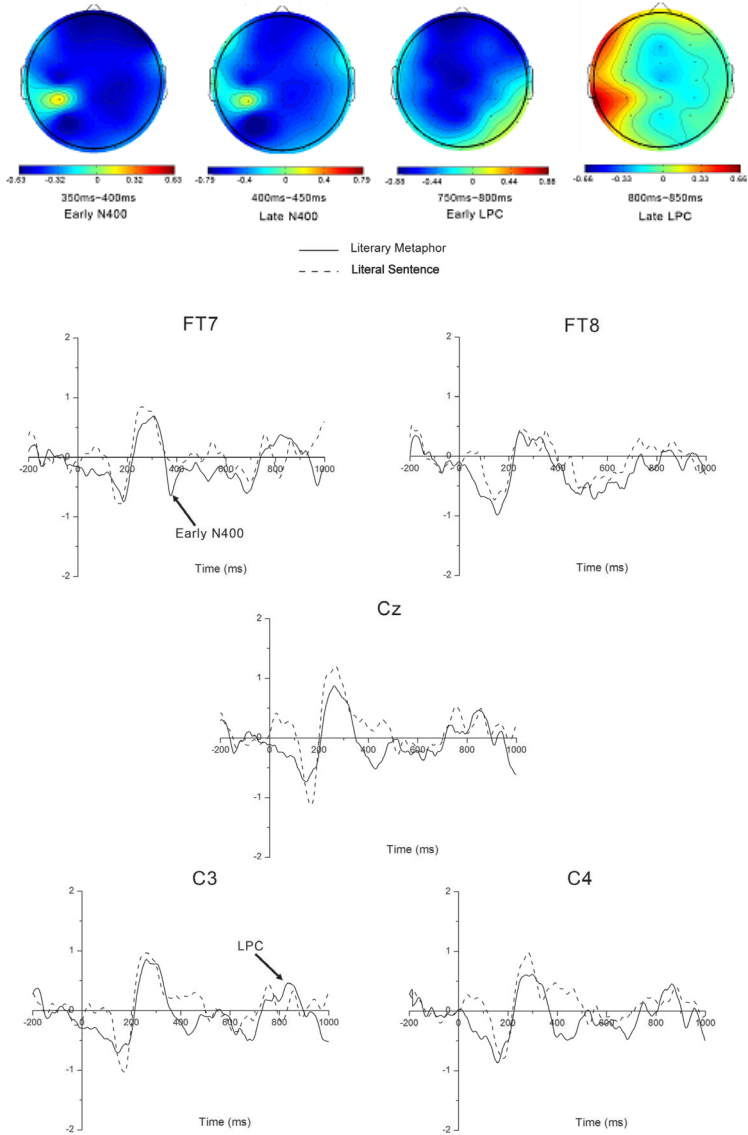
4.2.2 LPC: 750–800ms

Literary vs. literal. Figure 2(a) shows the average ERP waveforms at selected electrodes for these two conditions. According to the output of the ANOVAs, in the time window of 750–800ms, the waveform to the literary metaphors continued to be more negative than that elicited by literal sentence, reflected by significant main effects of Condition at midline and anterior sites and by a significant Condition by ES interactions at central sites.

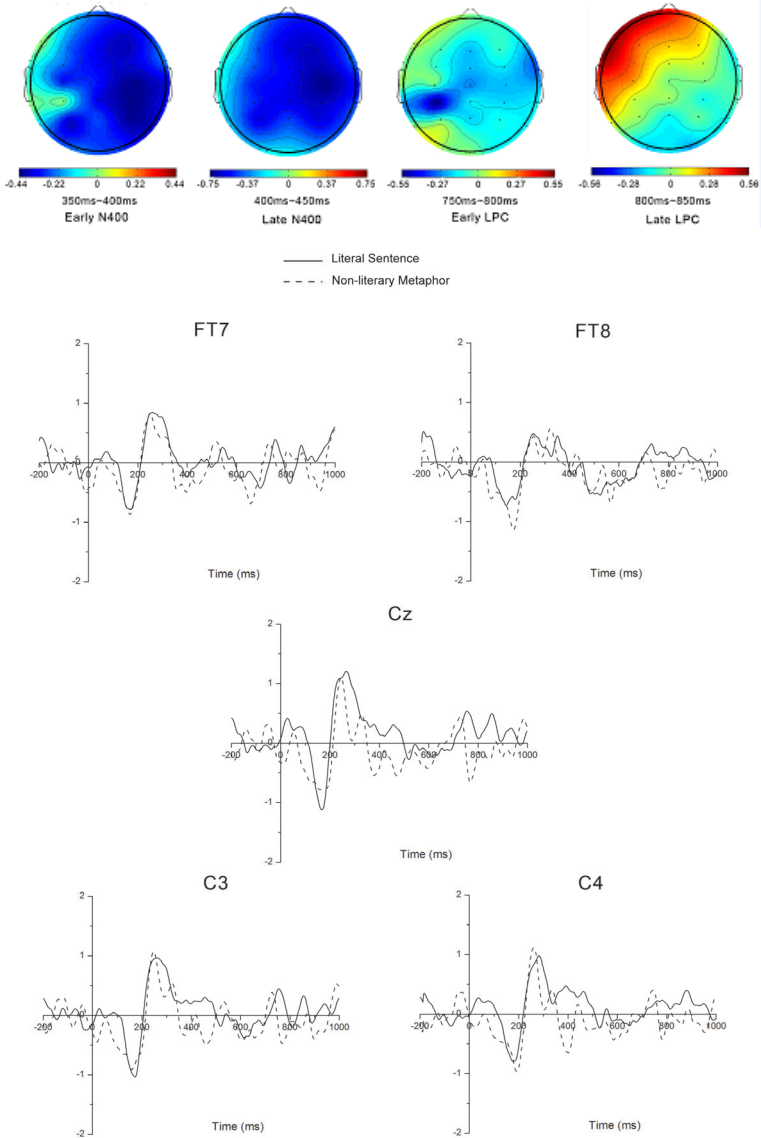
Non-literary vs. literal. Figure 2(b) shows the average ERP waveforms at selected electrodes for these two conditions. According to the output of the ANOVAs, in the time window of 750–800ms, the main effect of the condition was not significant at all electrode columns.

Literary vs. non-literary. Figure 2(c) shows the average ERP waveforms at selected electrodes for these two conditions. According to the output of the ANOVAs, in the time window of 750–800ms, the direct contrast between the literary metaphor and the anomalous sentence appeared to continue to show a more negative late N400 to the literary metaphor than to the literal sentence, reflected by a/the significant main effect of Condition at midline and central sites, Condition by Electrode sites interactions that reached significance at midline and posterior sites and Condition by Hemisphere interactions that reached significance at the posterior sites.

(a)



(b)



(c)

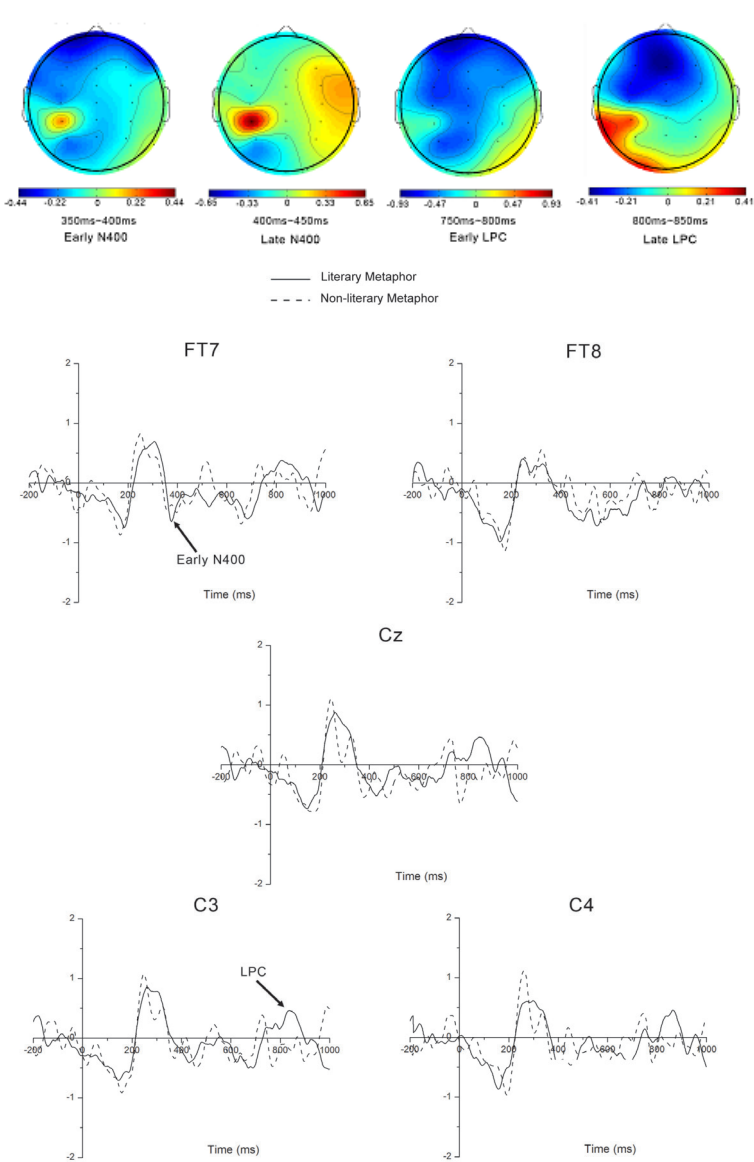


Figure 2: (a) literary vs. literal; (b) non-literary vs. literal; (c) literary vs. non-literary

5 Discussion

The behavioral results of this experiment show that people made more errors when deciding that a literary metaphorical expression had meaning than when deciding that a non-literary metaphor had metaphorical meaning and that a literal sentence did not have metaphorical meaning. This indicates that it was easier for them to understand non-literary metaphors and literal sentences than literary metaphors. For reaction time, people took longer to understand literary metaphor than non-literary metaphor and literal sentences.

5.1 N400

We examined the neural correlates of processing literary metaphors in comparison with literal sentences and non-literary metaphors. ERPs were examined after the onset of the final target words.

A negativity effect was observed only at some posterior sites to non-literary metaphors relative to literal sentences. It is possible that this reflected a transient, localized N400 effect resulting from the problems of semantically mapping the meaning of the non-literary metaphors onto their metaphorical context brought about by the early access to the literal meaning of the non-literary metaphors. For this reason, the effect was transient because the metaphorical meanings of the non-literary metaphors were accessed quickly afterwards, fitting with the metaphorical context and leading to N400 attenuation. Literary metaphors elicited a widespread N400, which was more negative than that elicited by literal sentences and non-literary metaphors throughout the N400 time-window. This indicates that it is more difficult for the literary metaphors to be semantically mapped onto their preceding context than literal sentences and non-literary metaphors.

Pynte et al. (1996), together with Coulson and Van Petten (2002) found that N400 evoked by metaphors (novel or conventional) was more negative than literal sentences. We partially replicated these previous studies in that in our experiment, the effect of the literary metaphors, which were assumed to be novel, was observed throughout the N400 time-window, while the effect of the non-literary metaphors, which were assumed to be less novel than literary metaphors, was localized to only a few electrode sites. Although participants may have had difficulty in accessing the metaphorical meaning

of the non-literary metaphors, contributing to the first access of the literal meaning, they reached the non-literary metaphors' metaphorical meaning quickly afterwards and successfully mapped this onto the preceding context; that is to say, both literal and non-literary metaphors were understood within the N400 time window. To be more specific, accessing the metaphorical meaning of the non-literary metaphors was delayed for less than 50ms and the N400 effect was not as widespread as that elicited by the literary metaphors. This indicates that in contrast to literary metaphors, participants need not reject the literal meaning of the non-literary metaphors before their metaphorical meaning may be accessed, which supports the assumption that literary metaphors are more novel than non-literary metaphors.

Exactly what it is that the N400 indexes is a controversial issue, one that has been explained by different approaches (Federmeier & Kutas, 1999; Lau et al., 2009). Some studies assumed that N400 reflects the integration of the target words' meaning into its context (Hagoort et al., 2004), while others assumed that N400 reflects a dynamic process of retrieving the meaning of the target word from the semantic memory (Van Petten et al., 1999). The current study assumes at the retrieval phase, both literary and non-literary metaphorical meanings are required to be retrieved. At the integration phase, only literary metaphorical meanings were difficult to be integrated into the rest of the context.

5.2 LPC

A widespread early LPC effect was observed in/with literary metaphors, relative to literal sentences and non-literary metaphors (from 750 to 800ms). Coulson and Van Petten (2002), and De Grauwe et al. (2010) found that relative to the literally-interpreted words, the metaphorically-interpreted words also triggered a larger LPC effect. We partially replicated these previous studies in that in our experiment, the effect of the literary metaphors, which were assumed to be novel, was observed throughout the early LPC time-window, while the effect of the non-literary metaphors, which were assumed to be less novel than literary metaphors, was not significant throughout LPC time-window. In agreement with Coulson and Van Petten (2002), we suggested that this effect reflected a later attempt to process the metaphorical meaning of the literary metaphors by continuously retrieving the metaphorical meaning of the literary metaphors from the semantic memory and integrating it with the context.

The present study seems to be in agreement with the indirect access view, at first sight. The literary metaphors and non-literary metaphors differ in their N400 and LPC amplitude, suggesting a difference in processing effort, which can be interpreted in two ways: One possibility is before reaching the metaphorical meaning of the literary metaphors, participants needed to reject the firstly-accessed literal meaning. In this case, when processing literary metaphors, the literal meaning was accessed first, and thus would support the indirect view. The other possibility is participants were selecting among multiple meanings that appeared in the mind at the same time, which is consistent with the classic exhaustive access view (Onifer & Swinney, 1981, Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982), and in agreement with the direct access view.

Our data do not support cognitive theorists' Conceptual Metaphor Theory of literary metaphor comprehension, for the N400 and LPC effect in the literary metaphor condition indicated the conceptual mappings for processing literary metaphors were more cognitively taxing than literal sentences and non-literary metaphors. The Gradient Salience hypothesis is supported, for non-literary metaphors, which were assumed to be less novel than literary metaphors, and which were therefore supposed to be "salient", evoked a transient N400, which was not as widespread as that elicited by the literary metaphors. Literary metaphors, which were assumed to be novel and were therefore supposed to be "salient", elicited a widespread N400 and LPC, which was more negative than that elicited by literal sentences and non-literary metaphors throughout the N400 time-window and early LPC time-window.

The "career of metaphor" theory proposed by Gentner and Bowdle (2001) indicated that novel and conventional metaphors were processed in different ways. Conceptual mappings between different domains can be completed by comparison or categorization, which is determined by the degree of conventionality of metaphors. In other words, as metaphors become conventionalized, a shift in the mode of processing happens from comparison to categorization. The current finding established that there exist differences between non-literary and literary metaphor processing, and between literary metaphor and literal sentence processing; in other words, understanding literary metaphors is harder than understanding non-literary metaphors because comparing the concepts and creating new conceptual mappings are required for processing literary metaphors.

6 Conclusion and Implications

The current study has aimed to explore whether different neuro-mechanisms exist in processing Chinese metaphors in literature and Chinese metaphors outside literature. Researchers have tried different methods to explain the varying degrees of conventionality between metaphor in literature and metaphor outside literature, but few studies have been conducted by means of electrophysiological methods. In this study, ERP has been adopted to investigate the neural mechanisms involved in processing three different types of sentences: literary metaphors, non-literary metaphors, and literal sentences. The current study yielded empirical evidence to suggest that metaphors in literature are more creative and novel than non-literary metaphors.

In the present study, most of the participants were science students who, to some extent, are not expert in literature, especially poetry, which may lead to the prolonged activity in processing literary metaphors. Therefore, in a future study, the relative degree of literary attainment of the participants should be taken into consideration. In other words, differences among participants who are expert in literature and who are not expert in literature should be a factor worthy of research in processing literary metaphors. The selection of stimuli should be gained more finely. It is possible that there exist differences in the level of plausibility among the metaphors in different literature genres written by different writers. Moreover, the familiarity and conventionality of the stimuli, effect of word frequency, and effect of context are also important factors that influence metaphor processing. Furthermore, stimuli in future research can be in English or other languages. Also, differences that may exist between L2 learners' processing literary metaphors in the mother tongue and in a second language would be worthy of investigation as well.

As for the ERP data, a more accurate statistical methodology should be adopted in future research. In addition, the analysis of N400 and LPC, more components, such as N1 and N2, should be analyzed in order to study the characteristics of processing phonological and calligraphic Chinese characters. Lastly, low-resolution brain electro-magnetic tomography (LORETA) should be applied to establish a dynamic model of literary metaphor comprehension.

Acknowledgements: This work is funded by the research project for Humanities and Social Science, granted by the Ministry of Education, PRC, 2012. The project order: 12YJ CZH010. Thanks are due to two anonymous reviewers whose constructive criticisms substantially improved this manuscript.

References

- Arzouan, Y., Goldstein, A., Faust, M. (2007). Brainwaves are stethoscopes: ERP correlates of novel metaphor comprehension. *Brain Research*, 1160, 69–81.
- Blackwood, D. H. R., & Muir, W. J. (1990). Cognitive brain potentials and their application. *The British Journal of Psychiatry*, 157(9), 96–101.
- Bonnaud, V., Gil, R., & Ingrand, P. (2002). Metaphorical and non-metaphorical links: A behavioral and ERP study in young and elderly adults. *Clinical Neurophysiology*, 32(4), 258–268.
- Boyd, R. (1993). Metaphor and theory change. In A. Ortony (Ed.), *Metaphor and thought*, 2nd ed., (pp. 401–424). Cambridge, MA: Cambridge University Press.
- Chen, H., Tang, Y., Wang, H., Liu, W., Wang, Y, Ha, S., et al. (2009). Significant differences between metaphoric idioms and literal expressions in gradients and spatial involvement Chinese comprehension: An ERP proof supporting left-hemisphere activation gradients and spatial involvement. *Computing, Communication, Control, and Management(CCCM). ISECS International Colloquium* (vol. 4), 541–545.
- Crisp, P., Heywood, J., & Steen, G. J. (2002). Identification and analysis, classification and quantification. *Language and Literature*, 11(1), 55–69.
- De Grauwe, S., Swain, A., Holcomb, P. J., Ditman, T., & Kuperberg, G.R. (2010). Electrophysiological insights into the processing of nominal metaphors. *Neuropsychologia*, 48(7), 1965–1984. doi: 10.1109/CCCM.2009.5267520ICCSNA 2009.
- Federmeier, K. D., Wlotko, E. W., De Ochoa-Dewald, E., & Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain Research*, 1146, 75–84.
- Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language*, 41(1), 469–495.
- Friederici, A., Hahne, A., & Saddy, D. (2004). Distinct neurophysiological patterns reflecting aspects of syntactic complexity and syntactic repair. *Journal of Psycholinguistic Research*, 31(1), 45–63.
- Gentner, D. & Bowdle, B. F.(2001). Convention, form, and figurative language processing. *Metaphor and Symbol*, 16(3–4), 223–247.
- Gibbs, R. (1994). *The poetics of mind*. Cambridge, MA: Cambridge University Press.
- Giora, R. (2003). *On our mind: Salience, context and figurative language*. Oxford: Oxford University Press.
- Giora, R. (1997). Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics*, 8(3), 183–206.
- Glucksberg, S. (2003). The psycholinguistics of metaphor. *Trends in Cognitive Sciences*, 7(2), 92–96.

- Goatly, A. (1997). *The language of metaphors*. London: Routledge.
- Goldstein, A., Arzouan, Y., & Faust, M. (2008). Timing the metaphoric brain. *Brain Research in Language. Literary Studies*, vol.1, 205–223.
- Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. (2004). Integration of word meaning and world knowledge in language comprehension. *Science*, 304, 438–441.
- Heywood, J., Semino, E., & Short, M. H. (2002). Linguistic metaphor identification in two extracts from novels. *Language and Literature*, 11(1), 35–54.
- Iakimova, G., Passerieux, C., Laurent, J. P., & Hardy-Bayle, M. C. (2005). ERPs of metaphoric, literal, and incongruous semantic processing in schizophrenia. *Psychophysiology*, 42(4), 380–390.
- Katz, A., Marshark, M., & Clark, J. (1988). Norms for 204 literary and 260 non-literary metaphors on psychological dimensions. *Metaphor and Symbolic Activity*, 3(4), 191–214.
- Katzmarsi, V. A., Blasco, D. G., & Dessalegn, B. G. (2003). ERP and behavioral evidence of individual differences in metaphor comprehension. *Memory & Cognition*, 31(5), 673–689.
- Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research*, 1146, 23–49.
- Lai, V. T., Curran, T., & Menn, L. (2009). Comprehending conventional and novel metaphors: An ERP study. *Brain Research*, 1284(11), 145–155.
- Lakoff, G., & Turner, M. (1989). *More than cool reason: A field guide to poetic metaphor*. Chicago: The University of Chicago Press.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: The University of Chicago Press.
- Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (De)constructing the N400. *Nature Reviews Neuroscience*, 9, 920–933.
- Laurent, J. P., Denhieres, G., Passerieux, C., Iakimova, G. & Hardy-Bayle, M. C. (2006). On understanding idiomatic language: The salience hypothesis assessed by ERPs. *Brain Research*, 1068(1), 151–160.
- Leech, G. N. (1969). *A linguistic guide to English poetry*. London: Longman.
- Mukařovský, J. (1970). Standard language and poetic language. In D. C. Freeman (Ed.), *Linguistics and literary style*. New York: Holt, Rinehart and Winston, Inc.
- Nelkin, D. (2001). *Molecular metaphors: The gene in popular discourse*. *Nature Reviews/Genetics*, 2(July), 555–559.
- Nowotny, W. (1965). *The language poets use*. London: The Athlone Press.
- Onifer, W., & Swinney, D. (1981). Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias. *Memory and Cognition*, 9(3), 225–236.
- Peterson, N. N., Schroeder, C. E., & Arezzo, J. C. (1995). Neural generators of early cortical somatosensory evoked potentials in the awake monkey. *Electroencephalography and Clinical Neurophysiology*, 96(3), 248–260.
- Pynte, J., Besson, M., Robichon, F. H., & Poli, J. (1996). The time-course of metaphor comprehension: An event-related potential study. *Brain and Language*, 55(3), 293–316.
- Richards, I. (1936). *The philosophy of rhetoric*. London: Oxford University Press.

- Seidenberg, M. S., Tanenhaus, M. K., Leiman, J. M., & Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology*, 14(4), 489–537.
- Semino, E., & Steen, G. (2008). Metaphor in literature. In R.W. Gibbs, Jr. (Ed.), *The Cambridge handbook of metaphor and thought*, 232–246. Cambridge: Cambridge University Press.
- Short, M. (1996). *Exploring the language of pomes, plays and prose*. London: Longman.
- Sotillo, M., Carretie, L., Hinojosa, J. A., Tapia, M., Mercado, F., Martin, S.L., et al. (2004). Neural activity associated with metaphor comprehension: Spatial analysis. *Neuroscience Letters*, 373(1), 5–9.
- Steen, G. J.(1994). *Understanding metaphor in literature: An empirical approach*. London: Longman.
- Steen, G. J.(1999). Metaphor and discourse: Towards a linguistic checklist for metaphor analysis. In L. Cameron & G. Low (Eds.), *Researching and applying metaphor* (pp. 81–104). Cambridge: Cambridge University Press.
- Tartter, V. C., Gomes, H., Dubrovsky, B., Molholm, S., & Stewart, R. V. (2002). Novel metaphors appear anomalous at least momentarily: Evidence from N400. *Brain and Language*, 80(3), 488–509.
- Tsur, R. (1987). *On metaphorizing*. Jerusalem: Israel Science Publishers.
- Tsur, R. (1992). *Toward a theory of cognitive poetics*. Amsterdam: Elsevier (North Holland) Science Publishers.
- Van de Meerendonk, N., Kolk, H. H. J., Vissers, C. T. W. M., & Chwilla, D. J. (2010). Monitoring language perception: Mild and strong conflicts elicit different ERP patterns. *Journal of Cognitive Neuroscience*, 22(1), 67–82.
- Van Petten, C., Coulson, S., Rubin, S., Plante, E., & Parks, M. (1999). Time course of word identification and semantic integration in spoken language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(2), 394–417.
- Wang, X. (2009). *Chinese metaphorical cognition and its ERP imaging*. Beijing: Higher Education Press.

Bionotes

Hong-Jun Chen

Hong-Jun Chen (b. 1972) is a professor at Dalian University of Technology. His research interests include cognitive linguistics, neurolinguistics, SLA, and bilingualism. His publication include “Neural correlates of metaphor comprehension: An ERP study” (2010), “Neural substrates of Chinese and English metaphor processing” (2013), “The brain mechanisms in metaphor processing from the holistic perspective of neurolinguistics” (2013), and “Analysis of the cognitive mechanism and the pragmatic relevance of guessing riddles” (2014).

Xiao-Shuang Peng

Xiao-Shuang Peng (b.1987) is an assistant professor in the Foreign Language Department at Anhui Sanlian University. Her research interests include cognitive linguistics, cognitive pragmatics, and SLA. Her publication include “Relevance theory and metaphor: A case study in spouse-seeking advertisements” (2016).

Qi-Lin Lu

Qi-Lin Lu (b. 1982) is a postdoctoral researcher at University of Science and Technology of China. His research interests include neurolinguistics and neuroplasticity. Publications include “The different time courses of reading different levels of Chinese characters: An ERP study” (2011), “Short-term meditation induces white matter changes in the anterior cingulate” (2010), and Mechanisms of white matter changes induced by meditation” (2012).

Huili Wang

Huili Wang (b. 1966) is professor and director of the Institute for Language and Cognition in Dalian University of Technology. Her research interests include psycholinguistics, cognitive neurolinguistics, and experimental philosophy. Publications include “Economy is an organism: A comparative study of metaphor in English and Russian economic discourse” (2013) and “The role of preparation time length in asymmetrical switch cost: An ERP study based on overt picture naming” (2013).