國立臺灣師範大學文學院英語學系

碩士論文

Department of English, College of Liberal Arts

National Taiwan Normal University

Master's Thesis

隱喻思維：漢英雙語能力的認知神經處理

Metaphorically Thinking: Insights from Neurocognitive

Processing in Mandarin-English Bilingualism

劉岩峻

Dylan Scott Low Yan Jun

指導教授：詹曉蕙 博士

Advisor：Shiaohui Chan, Ph.D.

中華民國 113 年 7 月

July 2024

Acknowledgements

care and support of these great people. And speaking of great people in my MA journey, to Stephanie (who was my dorm manager, Mandarin tutor, co-author, project partner, and general Taiwan guide): thank you for all the laughs and foodie trips around Gongguan we shared!

Finally, the family members (besides Michael, who already got his own spotlight paragraph). I wish to thank my parents for their nourishing support in all areas of my life outside academe. In particular, I want to thank my mother for coming all the way from Singapore to Taipei to help me move out of my dorm room at a critical time and staying by my side until the thesis project was completed. Without my mother caring for me when I was in poor health during data acquisition, this project might have not been finished on time either. I am also grateful to my brother, Dixon, for both the happiness he brings to my heart and the technical knowledge about all things computational which he has shared with me. And I thank Janiah Tillery, my sweetheart, for empowering me emotionally, caring for me when I caught COVID, and being a part of my life during these past four years in Taiwan. Janiah, you have made everything so much more beautiful.

One more acknowledgement is in order. I wish to acknowledge Maia the Cat (born 17 April, 2022 in Hualien, Taiwan), whose kitty cuteness was a source of joyful distraction during the much of the gruelling quantitative training. Maia, thanks for accompanying me on my lap while I processed data in R, and for sitting on my notes and playing with my pen when I was preparing for my statistics exam. I'll see you and Janiah in Baltimore.

Dylan Scott Low Yan Jun 劉岩峻

Taipei, Taiwan

15 July, 2024

摘要

　　學習新的語言也涉及學習表達意義的新方式與規範，而隱喻便是其中之一。儘管已有大量先前研究採用事件相關電位（Event-Related Potential；ERP）的方法對隱喻語言的神經認知處理進行瞭解，但很少有研究從雙語的角度進行探討。本論文探討第二語言英語使用者是否以不同於母語使用者的方式處理隱喻語言，並呈現 ERP 初步結果。實驗共招募兩組參與者：在台灣的英語母語者與非英語母語者的台灣人。此外，此研究亦使用文化智商量表（Cultural Intelligence Scale）來測量跨文化能力，以控制兩組參與者學習和適應新文化進而擴展到新意義表達方式的能力。實驗材料由語境-目標句對組成，其中目標句根據動詞分為隱喻或字面兩種情境。14 名非英語母語者和 6 名英語母語者的 ERP 平均數據的質性比較結果顯示，母語和非母語使用者在處理動詞隱喻時存在差異：非母語組在目標句的動詞上觀察到 N400，而母語組在賓語的名詞上可能存在 N400。此外，在目標句最後一個詞之後的 500-800 毫秒的時間窗裡，兩個語言組都觀察到持續的晚期負向波，這可能代表著隱喻與字面情境之間的記憶檢索差異。總而言之，本論文支持了不同語言背景的人以不同方式處理隱喻的觀點——具體而言，對於英語母語者和非英語母語者的句子理解歷程而言，隱喻處理在不同階段會產生不同的影響。

關鍵字：隱喻、雙語、第二語言習得、語言處理歷程、心理語言學、神經語言學、事件相關電位（ERP）

Abstract

Learning a new language involves learning new ways and norms to mean things, including metaphor. Despite much prior research on the neurocognitive processing of metaphorical language using the event-related potential (ERP) methodology, few studies have explored it from a bilingual perspective. This thesis presents preliminary results from an ERP study investigating whether L2 English speakers process metaphorical language differently from native speakers. To undertake this goal, two groups were recruited: native English speakers in Taiwan and Taiwanese non-native English speakers. Additionally, intercultural competence was measured using the Cultural Intelligence Scale (CQS) to control for participants' ability to learn and adapt to new cultures and, by extension, new ways of meaning. Materials consisted of context-target sentence pairs, where target sentences were metaphorical or literal depending on the verb. ERPs based on the grand averages of 14 non-native participants and 6 native English participants were analysed. The results indicated qualitative differences in how native and non-native speakers processed verbal metaphors, with an N400 observed at the verb for the non-native group and a possible N400 seen at the object head noun for the native group. Additionally, a late sustained negativity was observed for both language groups at the 500–800 ms time window following the target sentence final word, possibly indicating memory retrieval differences between metaphorical and literal conditions. In sum, this thesis lends support to the notion that persons of different language backgrounds process metaphors differently—specifically, that metaphors affect different points of the sentence comprehension process for non-native speakers and native-speakers of English.

Keywords: Metaphor, Bilingualism, Second Language Acquisition, Language Processing, Psycholinguistics, Neurolinguistics, Event-related potentials (ERP)

Table of Contents

List of Tables

Chapter 1   Introduction

Metaphors are ubiquitous in our everyday uses of language (Lakoff & Johnson, 1980). Often, metaphors are used to help us understand abstract concepts in terms of more concrete and accessible ones. For example, in the domain of economics and finance, we may speak of reducing expenditure with verbal expressions such as *cutting costs*, *slashing prices*, *axing budgets*, and, with the recent flamboyant 2023 election campaign of Javier Milei in Argentina, *taking a chainsaw to the economy*. Such metaphors allow the abstract concept of financial expenses, which is intangible, to be described as if it were a tangible physical object that can be cut and slashed with tools such as an axe or chainsaw, even though using such tools on budgets, prices, and economies is, literally speaking, impossible (Grice, 1975; Searle; 1993/1979). Yet, because such metaphorical readings draw upon more readily accessible concepts, often grounded in physical bodily experience that may be perceived with the five senses, metaphors facilitate understanding and communication. With metaphors, people can achieve a range of communicative goals not possible with mere literal language by shaping and influencing how others, or themselves, think about abstract ideas (Thibodeau et al., 2019).

In Cognitive Linguistics, the conceptual system is hypothesised to be organised as a collection of conceptual frames, as proposed by Fillmore (1985; see also Barsalou, 1992), which may be mapped in processes of metaphor (Kövecses, 2015; 2020). Yet, people who speak different languages use different metaphors to communicate. For example, Liu (2002) notes the prevalence of sports-related metaphors in American English, with idioms like *to get to first base* and *to play hard ball*, and food-related metaphors in Mandarin, such as 吃醋 *chīcù* "to be jealous" (lit. "eat vinegar"), 丢飯碗 *diū fànwǎn* "to lose one's job" (lit. "to lose rice bowl"), and 炒冷飯 *chǎo lěngfàn* "repeat something old with nothing new to add" (lit. "to fry cold rice"). Though Link (2013) observes in his comparison of Chinese and English metaphors that different cultures likely share more metaphors than they do not, a phenomenon Kövecses

(2010, p. 200) attributes to "universal bodily experience", the metaphorical mappings between abstract and concrete frames in idioms such as the above appear to be specific to either American or Chinese culture. Observations such as these have led to the idea that individual cultures exist as systems of conceptual frames, or "ways of meaning" (Agar, 1994; Halliday & Hasan, 1985; Hasan, 2015/1996).

In the domain of cognitive neuroscience, especially electroencephalography (EEG), metaphors have been approached as a productive topic for language processing research, particularly semantic processing. One main point at issue is whether metaphors are directly or indirectly accessed with respect to literal meanings. A second point of debate concerns the extent to which metaphorical language is processed differently from literal language during comprehension, and what psycholinguistic factors modulate such differences. However, whilst there exists a range of EEG studies investigating monolingual metaphor processing in the brain (e.g., Pynte et al., 1996; Coulson & Van Petten, 2002; 2007; Arzouan et al., 2007; Lai et al., 2009; 2019; De Grauwe et al., 2010; Bambini et al., 2016), neurocognitive experimental studies on bilingual metaphor processing are currently few in number (see Chapter 2).

Yet, bilingualism offers opportunities to investigate the role in which differences in language background, which can be seen as differences between cultural and conceptual systems, influence metaphor processing. Such opportunities are, hence, not being adequately capitalised upon in the EEG research field. Furthermore, as will be discussed in Chapter 2, much of the current bilingual metaphor processing research suffers from design problems which harm their claims to ecological validity. As shall be seen, the primary design problem with the previous bilingual studies stems from a fundamental failure to grant due attention to linguistic form, a point raised by Cameron (1999). In other words, the way a conceptual metaphor is put into words matters. For example, the finance-related metaphorical expressions mentioned earlier (*cutting costs*, *slashing prices*, *axing budgets*, and *taking a chainsaw to the*

economy) represent different linguistic realisations of a single conceptual metaphor, that MONEY IS A SOLID (Silaški & Kilyeni, 2014), at the level of form. By presenting source and domain concepts out of their realisational context (for example by presenting *money is a solid* as the stimulus instead of *axing the budget*), such experimental designs may fail to simulate metaphor comprehension processes that people's brains undertake in everyday language use outside the laboratory setting.

To contribute to the limited metaphor processing literature from the lens of bilingualism and culture, in this thesis I design and present an ERP experiment aimed at answering the following research question in the context of Mandarin-English bilingualism in Taiwan:

- Do L2 speakers of English process metaphorical language differently from native speakers of English?

At the same time, I attempt to maintain a claim to the experiment's ecological validity through the use of naturalistic language materials as stimuli, validated through an internet-based norming study conducted with native English speakers.

Chapter 2   Literature Review

2.1. Metaphors and Social-Semiotics

As cognising animals, a fundamental feature of our human cognitive system is the ability to create and learn systems of signs—semiotic systems. At its essential core, semiosis refers to the perceiving of one thing (a signifier), and from that, understanding another (a signified) (de Saussure, 1959/1916). Perhaps the most quintessential semiotic systems we have developed are those for language and for culture. In language, a certain word may carry a certain meaning, and in culture, a certain artefact or behaviour may carry certain significance (Low et al., 2022).

Equally interesting as the distinctions between linguistic semiotics and cultural semiotics are the intersections between them, specifically the ways in which language and culture are non-distinct. Much of the blurriness between language and culture lies at the stratum of meaning. So extensive are the overlaps between them, that the linguist M.A.K. Halliday proposed to consider "language as one among a number of systems of *meaning* [my emphasis] that, taken all together, constitute human culture" (Halliday & Hasan, 1989/1985, p. 2). It is important to note that here, Halliday intended the word "meaning" to be read as a verb, hence signifying a kind of semiotic activity or process. This is language as *social-semiotic* (Halliday & Hasan, 1989/1985, p. 1). The *social-semiotic* view is echoed by Agar (1994), who coined the term *languaculture* to refer to the unity of language and culture at the level of "meaning" in the sense of the "activity of meaning".

At the heart of languaculture is metaphor. Using Lakoff and Johnson's (1980) influential theory of *conceptual metaphor* as a starting point, I define metaphor as the conceptualisation of one thing in terms of another. For example, in sentences like *The cat warmed up to me* or *She gave me a warm smile*, the concept of AFFECTION is construed as WARMTH and expressed linguistically using lexical items primarily associated with WARMTH.

In this view, a metaphor is a semiotic link between concepts which serve as the meaning base for linguistic signs (Halliday & Matthiessen, 1999). That is, a concept signifies another concept, so strictly speaking, conceptual metaphor constitutes a pre-linguistic semiosis. From the viewpoint of linguistic form, then, conceptual metaphor is a kind of signified. Thus, if a naïve view of semiosis is adopted and it is assumed that for each signifier there is only one signified (hence ignoring polysemy), with conceptual metaphor at least two semiotic links are necessitated—one between the two concepts that comprise the semantics, and one between the linguistic form and the semantics. This is illustrated visually in Figure 2.1. Following Lakoff and Johnson (1980), the semiotic link between the two concepts may be referred to as *metaphorical mappings*, whereas the general linguistic term *realisation* may be used to refer to the semiotic link between a metaphor and its linguistic form.

Figure 2.1. Conceptual metaphor in a sign diagram in the style of de Saussure (1959/1916).
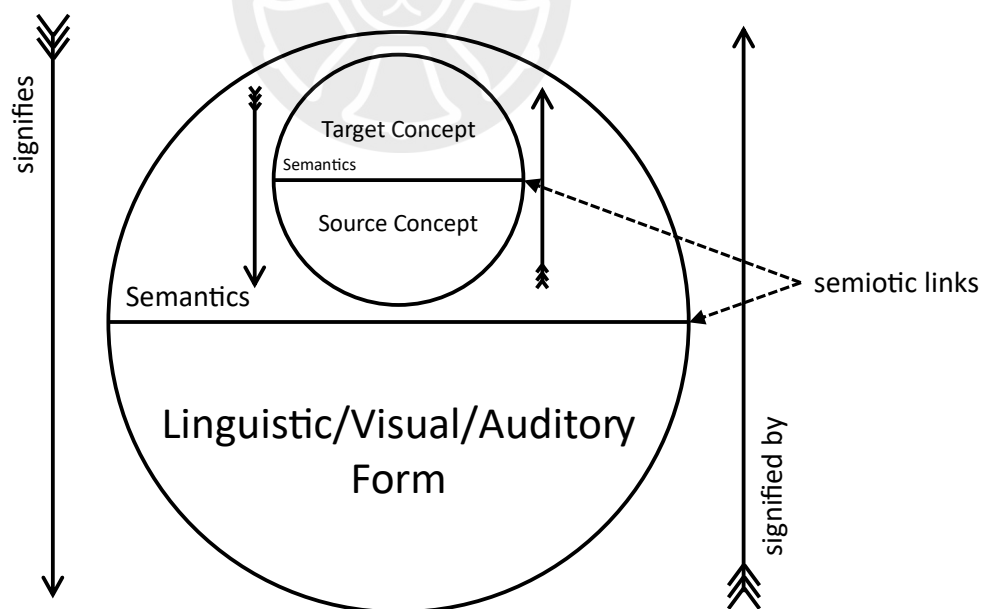
Distinguishing these two types of semiosis, metaphorical mapping and linguistic realisation, that is inherent in metaphorical language use is crucial, for they should not be conflated. As Cameron (1999, p. 8) notes, it is important for researchers to decide early if metaphor should be considered "a phenomenon of language, or of thought, or both," and to consider what such a decision implies. In *Metaphors We Live By*, Lakoff and Johnson (1980) focus exclusively on the latter, that is, metaphor as a phenomenon of thought (conceptual mappings), so even though the data on which they base their findings are linguistic expressions of English, it can be said that their Conceptual Metaphor Theory (CMT) does not actually include a theory of metaphor realisation. Rather, CMT is a theory of the *conceptual system*, and not a theory of the *linguistic system*. In fact, conceptual metaphors can also be realised in systems other than the linguistic, such as imagery, or in complex assemblages combining them, as is common in art and film.

But even when the scope is limited to the realm of language, there is immense complexity, which suggests that the link between concept and form is particularly important to consider in metaphor research (Cameron, 1999). For example, it has been observed that conceptual metaphors tend to be highly productive in that speakers can realise a metaphor in many different ways, such as single-word expressions, multi-word expressions, and even extended units of discourse (Shutova, 2015; Rai & Chakraverty, 2020). This means that while languacultures may share many conceptual metaphors cross-linguistically (Kövecses, 2010; Link, 2013), certain metaphor realisation patterns in form may be specific to a language. In other words, different languages may have different ways of expressing metaphorical meanings at the level of form. This leads to the question: do our brains access conceptual metaphors differently when reading in another language? Bilingualism offers an interesting theatre in which to attack this question.

6

Within psycholinguistics, a significant area of research has been whether knowledge of multiple languages is stored separately or not in memory, and whether different languages entail different conceptual systems (see Ansah, 2011; Athanasopoulos, 2015; Van Assche et al., 2020 for overviews). On one side are studies (e.g., Kolers, 1963; Gerard & Scarborough, 1989) that support a "separate-storage" view of concepts in relation to the bilingual lexicon. Others, meanwhile, have argued on the basis of experimental and linguistic evidence for a "shared-storage" view such that L2 forms access, at least largely, the same conceptual system as L1 forms (e.g., Kroll & Stewart, 1994; Kroll & de Groot, 1997; Lamb, 1999; Ansah, 2011). To take just one example of this "shared-storage" view for the purposes of illustration, Figure 2.2. represents Lamb's (1999) model of the cognitive relationship between form and meaning in a multilingual person, built upon his neurocognitive theory of language rooted in cortical neuroanatomy (see Lamb, 2004, Chapter 17; 1999, Chapter 17). Lamb's model in particular predicts that the acquisition of a new language entails learning new forms, but also new "ways of meaning" or "semiotic styles", to use terms from the work of Hasan (2015/1996, p. 191). This is a view echoed by psycholinguists such as Athanasopoulos (2015) and Pavlenko (1999)[1].

Figure 2.2. Lamb's cognitive model of multilingualism. Adapted from Lamb (1999, p. 42).

---
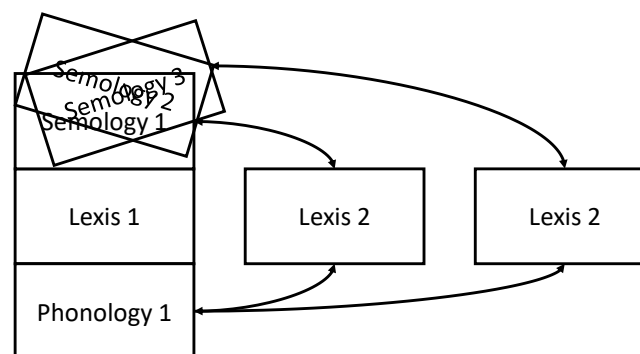
[1] Note however, that Lamb's diagram does not seem to distinguish between the conceptual and semantic systems above lexis, labelling both as "semology". The failure to distinguish the conceptual and semantic systems is a major point of criticism raised by Pavlenko (1999) towards many psycholinguistic models of bilingual lexical access. However, to be precise, Lamb uses "semology" as a cover term, and has referred to the conceptual system as "outer semantics" and the linguistic-semantic system as "inner semantics" (see Lamb, 1999, pp. 146–147).

The discussion thus far underscores the key themes of the present thesis. From the perspective of the Hallidayan and Agarian views of language as *social-semiotic* and language as *languaculture*, it may be predicted that aspects of the conceptual system which are developed from culture-specific experience, such as certain conceptual metaphors, to be strongly semiotically associated (a realisational relationship) with the linguistic forms from that languaculture.

2.2. Metaphor in the Theory of Norms and Exploitations

Throughout this thesis, metaphors are assumed to reflect patterns of thought, as per the CMT proposed by Lakoff and Johnson (1980). However, as mentioned earlier in section 2.1. at the beginning of this chapter, CMT is a theory of the conceptual system and does not itself offer an account of the realisation of conceptual metaphors at the level of linguistic form. In this section, then, I give a brief overview one account of linguistic metaphor realisation: the Theory of Norms and Exploitations (TNE) by Hanks (2004; 2006; 2013), a corpus-driven theory of language which situates the lexicon at its centre. The key question pertinent to the following discussion, then, is this: "How do we determine when a stretch of language is metaphorical or not?"

TNE proposes two main ideas: (1) that the meaning of a word is determined by its environment and, more generally, its context of use, and (2) that words have normal usages and meanings which can be distinguished from marginal usages and meanings. The first idea essentially echoes Firth's (1957/1968, p. 11) famous dictum that "You shall know a word by the company it keeps", which Hanks (2004; 2006) extends by suggesting that word meanings should be defined in terms of the *potential* of the word to contribute meaningfully to an utterance when inserted into some syntactic pattern. In other words, individual words do not

have meanings but rather *meaning potentials*[2], some or all of which may be activated when the word is used in specific syntactic constructions (Hanks, 2006; 2013).

The second idea establishes the notion that a language consists in a rule-governed system of norms, which Hanks (2013, p. 5) describes as *socially salient*, therefore "conventional and for that very reason unmemorable". According to Hanks, knowledge of norms in a language constitutes the *competence* that a learner of the language must acquire, in a suggestive mirroring of Chomsky's (1986) terminology. Breaking from the Chomskyan concept of competence, however, Hanks proposes a second rule-governed system which learners must acquire: knowledge of how to *exploit* the norms to express new meanings. Exploitations, in contrast to normal uses of language, tend to be more *cognitively salient* and memorable, and hence they are employed by language users to achieve rhetorical effects such as metaphor (Hanks, 2004; 2013).

In TNE, then, the view is that words in isolation have meaning potentials which can be realised by the word's usage within some range of contexts. These contexts vary along a cline of normalcy, such that more normal usages are more frequent and socially salient, whilst less normal usages are less frequent and more cognitively salient. With regards to metaphor, Hanks (2006) suggests that *metaphoricity* be similarly considered as a cline, challenging the truth-conditional semantics view (e.g., Davidson; Grice, 1975; Searle, 1993/1979) that a stretch of language is only either metaphorical or literal. In that view, metaphoricity is solely a binary variable. In contrast, Hanks (2006, p. 19) proposes that metaphoricity is instead gradable, where "literal meaning" refers to the parts of a word's meaning potential which are activated in its most normal contexts of use, making TNE compatible with connectionist theories of cognition (e.g., Lamb, 1999; Feldman, 2006).

---

[2] Hanks (2004) borrows the term *meaning potential* from Halliday's work in Systemic Functional Linguistics, where it instead refers to choices in meaning which are then realised in lexicogrammatical forms.

According to Lakoff and Johnson's (1980) CMT, a conceptual metaphor occurs when there is a mapping between two conceptual domains, which may be reflected in lexical realisation. By introducing the notions of norms and exploitations in language use, TNE arrives at similar conclusions to psycholinguistic models of conceptual metaphor such as the Career of Metaphor model (Bowdle & Gentner, 2005; see Section 2.3.1.), that metaphors become less effective, or "dead", the more conventionalised or normalised they become. Hanks argues that while metaphors may begin as exploitative uses of language, which are cognitively salient and thus have a strong rhetorical effect, conventionalisation can transform them from exploitations into norms, which are socially salient and less rhetorically outstanding. What makes some stretch of language metaphorical then, is a combination of whether the realised meaning potentials are the result of a conceptual mapping as per Lakoff and Johnson's account plus how exploitative the utterance is, relative to normal language use in that specific context.

2.3. Neural Correlates of Metaphor

2.3.1. ERP research and metaphor

There is a fairly developed experimental literature on the neural correlates of metaphor processing, at least in monolingual contexts. Many such studies bear on the fundamental question of whether metaphors are directly or indirectly accessed in comprehension. Different theories of metaphor comprehension predict either direct or indirect access as a function of different factors. The classical pragmatic view (e.g., Grice, 1975; Searle, 1993/1979), based on truth-conditional semantics, takes the position of metaphorical meanings as deviations from the actual meanings associated with a stretch of language. This classical view is interpreted by cognitive neuroscientists to imply that literal meanings are accessed first and then followed by metaphorical meanings in the comprehension process (e.g., Coulson & Van Petten, 2002; Lai

et al., 2009; De Grauwe et al., 2010), a hypothesis that received early support from the behavioural results of Janus and Bever (1985).

In contrast, direct access models adopt a parallel-processing view wherein metaphoric and literal meanings are simultaneously processed and subjected to the same cognitive mechanisms (e.g., Coulson, 2008; Glucksberg, 2003; Gibbs, 1984), supporting Lakoff's (1987) idea that linguistic meanings are radially organised into core and extended meanings. Thus, indirect models treat metaphors as necessarily "special", whilst direct models do not presuppose any inherent processing "specialness" that distinguish them from literal language besides psychological factors known to affect radially organised conceptual categories (e.g., familiarity, salience, and frequency) (Coulson & Van Petten, 2002; Coulson, 2008; Giora, 2002).

Additionally, there are also models which escape simple classification into direct or indirect types, for example, Gentner and Wolff's (1997) Structure Mapping Model (SMM). Somewhat similar to the indirect access models, the SMM proposes at least two processing stages, but rather than assuming a literal processing and then metaphorical processing stage, it suggests that source and target concepts are first made to align in terms of their similarities, and then metaphorical inferences between the two are made. Related to the SMM, Bowdle and Gentner (2005) also propose the Career of Metaphor Model (CMM), which suggests that conventional and novel metaphors are processed differently. Therefore, in a similar vein to direct access models, an underlying assumption of both the SMM and CMM is that conventionality modulates whether and how "special" a metaphor is processed in comparison to literal language. Thus, the SMM and CMM occupy a complicated position between the direct and indirect models.

In ERP studies, the ERP component that has featured most prominently in metaphor processing research on the question of direct vs. indirect access is the N400. As its name

11

suggests, it is a negative-going wave that, after a meaningful stimulus is presented, onsets at about 250 ms, peaks at about 400 ms, and offsets around 550 ms (Baggio & Hagoort, 2011; Kutas & Federmeier, 2011). It is usually largest over the central and parietal electrode sites and is slightly larger over the right hemisphere than over the left (Luck, 2014). Historically, the N400 component was discovered when Kutas and Hillyard (1980) observed it when sentences ended with semantically incongruous words (as in *He took a sip from the transmitter*) versus congruous words (as in *He took a sip from the waterfall*). It is typically associated with violations of semantic predictions, though there are disagreements about whether such violations occur in semantic structure-building processes, i.e., a "combinatorial" account (Osterhout & Holcomb, 1992; Brown & Hagoort, 1993), or access/retrieval from semantic memory, i.e., a "pre-activation" account (Lau et al., 2016; 2008; Kutas & Federmeier, 2000), or a combination of both, i.e., a "multiple generators" account (Hagoort, 2013; Baggio & Hagoort, 2011).

A second ERP component that has been commonly observed in metaphor experiments is the P600. Osterhout and Holcomb (1992) notably observed the P600 when they compared well-formed sentences that contained temporary syntactic ambiguity (e.g., *The broker persuaded to sell the stock was sent to jail*) with sentences that did not (e.g., *The broker persuaded the man to sell the stock*). The P600 was originally thought to be a marker of syntactic reanalysis (Osterhout & Holcomb, 1992), so much so that it has alternatively been referred to as the Syntactic Positive Shift (SPS) (Hagoort et al., 1993). As this latter name suggests, it is a slow late positive shift in the ERP waveform that generally onsets at about 500 ms after the presentation of a stimulus (Swaab et al., 2012). The neurolinguistic literature also sometimes refers to the P600 as the Late Positive Component (LPC), a name which is shared

with similar wave deflections seen in other cognitive domains, such as the P3 or P300[3] in attention and memory ERP research (Polich, 2012). It lasts for several hundred milliseconds and is observed to peak at around 600 ms post-stimulus, though a clear peak is not always seen. It is also usually largest over the posterior electrode sites and sometimes the anterior sites.

However, despite its initial association with syntactic processes, later studies began to report P600-like observations in experiments involving semantic violations where an N400 might be expected instead (Swaab et al., 2012; Brouwer et al., 2012 Kuperberg et al., 2003). Kuperberg et al. (2003), for example, observed a "semantic P600" when they compared sentences containing thematic role animacy violations (e.g., *For breakfast the eggs would only eat toast and jam*) to semantically well-formed baseline sentences (e.g., *For breakfast the boys would only eat toast and jam*). Both sentence types were syntactically well-formed but elicited a P600 rather than N400. However, sentences containing non-thematic role pragmatic violations (e.g., *For breakfast the boys would bury toast and jam*) were found to elicit the N400 when compared with the baseline sentences. Observations such as these have motivated theories of what the P600 indexes, such as Kuperberg's (2007) proposal that the P600 is produced by conflicts between syntactic and semantic processing streams and Brouwer et al.'s (2012) suggestion that the P600 reflects the semantic structure-building processes previously associated with the N400, whilst the N400 itself reflects semantic memory retrieval. In any case, the idea that the P600 is merely an indicator of syntactic violations or reanalysis processes, akin somewhat to a syntactic mirror image of the N400, is no longer widely upheld and its involvement in semantic processes is now a fertile field of research and debate.

---

[3] In explaining his preference for the position-based name "P3" over the time-based name "P300", Luck (2014, p. 9) implies that the P3 and P600 refer to the same component when he states that "…in language experiments, the P3 wave generally follows the N400 wave, making the term P300 especially problematic." However, the term "P300" is also used to refer to a family of components rather than a single component, and there is reason to believe that the P600 is distinct from at least one of the other members of this family, the P3b (Swaab et al., 2012).

2.3.2. Monolingual studies of metaphors

It is against the backdrop of the above debates between direct and indirect models and about the natures of the N400 and P600 components that ERP results in metaphor research are often interpreted. Most studies reported a biphasic N400-P600 effect (that is, observations of both components) when comparing literal stimuli with metaphorical stimuli (Pynte et al., 1996; Coulson & Van Petten, 2002; 2007; Arzouan et al., 2007; De Grauwe et al., 2010), though some reported N400 only (Lai et al., 2019; 2009; Lai & Curran, 2013;) or P600 only (Yang et al., 2013). Bambini et al. (2016) reported one experiment in which the biphasic N400-P600 effect was replicated, and another experiment in which only a P600 was seen. Therefore, results seem inconsistent across different studies.

Pynte et al. (1996) sought to test three hypotheses related to metaphor access: the indirect access view, the direct access view, and the context-dependent view. In their initial experiment, all stimuli (French sentences of the form *Those X are Y*) were presented in isolation, without any preceding context. When they contrasted familiar metaphors with literal sentences, they observed a significant N400 effect equally distributed across scalp sites. They then conducted three follow-up experiments to investigate the possible effect of metaphoricity in the P600 time window. Their second experiment contrasted unfamiliar and familiar metaphors, but they found no significant effects in both the N400 and P600 time windows. Their third experiment further manipulated the presence of relevant or irrelevant preceding contexts. They found that by pairing relevant contexts with familiar metaphors and irrelevant contexts with unfamiliar metaphors, they could observe significant differences in both the N400 and P600 windows. However, in the P600 window, the unfamiliar metaphors displayed more negative-going waveforms than the familiar metaphors. The researchers interpreted these results as evidence against the indirect hypothesis since both the N400 and P600 windows were affected by a contrast between familiar and unfamiliar metaphors. Finally, they conducted a fourth

14

experiment that contrasted the inverse pairings of the third (familiar + irrelevant context vs. unfamiliar + relevant context), finding significant effects of contextuality in both N400 and P600 windows. The researchers observed a larger N400 amplitude for the familiar + irrelevant condition than the unfamiliar + relevant condition and a larger P600 amplitude for the unfamiliar + relevant condition than the familiar + irrelevant condition. They interpreted these results as meaning that relevant contexts can reduce N400 amplitude and elicit a late positivity relative to irrelevant contexts. They conclude that their results support a context-dependent view of metaphor access.

Subsequent to Pynte et al.'s study, Coulson and Van Petten (2002) conducted a partial replication and follow-up. Unlike the design of Pynte et al. (1996), Coulson and Van Petten (2002) allowed variation in syntactic structure across their stimuli, and their stimuli appeared to be more natural as a result, though they lacked preceding contexts. The researchers contrasted three conditions: metaphorical, literal, and literal mapping. The literal mapping condition represents situations where mappings are made between conceptual models, similar to metaphor, but still invoking the literal meanings of the terms involved (e.g., *The ring was made of tin, with a pebble instead of a gem*.) They reported that metaphors elicited larger N400s than literal sentences with literal mappings falling in between, and that metaphors also elicited a larger late positivity at parietal and occipital scalp sites. They interpreted the N400 observations as indicative of a gradient of comprehension difficulty across the three conditions and the P600 as an index of a conceptual retrieval process.

However, while the increased negativity in the N400 time window in metaphorical sentences relative to literal sentences is reliably replicated across most studies, the role and appearance of the P600 component is rather tenuous. In Lai et al. (2009), for example, the researchers do not observe a P600 component. This study investigated the contrast between literal, conventional metaphorical, novel metaphorical, and semantically anomalous sentences.

The researchers found a graded N400 effect across the four conditions, reproducing partially the observations of Pynte et al. (1996) and Coulson and Van Petten (2002). However, whilst the conventional metaphorical condition and literal condition converge in a later time window (440 ms to 560 ms), the negativity of the N400 in the novel metaphorical and anomalous conditions appeared to extend to the later time window as well. Another study by Yang et al. (2013) interestingly reported only the P600 and did not observe any N400 component. Here, the researchers presented probe words followed by a literal or metaphoric sentence to the participants, where the probe words were manipulated for conceptual domain which may or may not be congruent to the following sentence. Despite expecting to observe N400s, only the P600 component was seen in all conditions. The researchers suggested that this may index a process of reanalysis caused by the unexpected content of their metaphors, similar to how unexpected semantic anomalies elicited the P600 instead of N400 in Kuperberg et al. (2003). They also observe a larger P600 in incongruent conditions than congruent conditions, interpreting incongruent contexts as requiring more intensive reanalysis processes.

Bambini et al. (2016) also investigated the role of context in metaphor processing, and interestingly found that context could have an attenuating effect on the N400 commonly elicited by metaphors. They conducted two experiments. For both experiments, the stimuli consisted of two-sentence passages in Italian. Each passage consisted of a critical sentence and a preceding context sentence. The critical sentences contained nouns in a literal relation (e.g., *shark* and *fish*; in Italian: *squalo* and *pesce*) or a metaphorical relation (e.g., shark and lawyer; in Italian: *squalo* and *avvocato*. For their Experiment 1, the context sentences did not support any explicit relationship between the critical noun and other noun (e.g., literal condition: *Do you know what that fish is? A shark.*; metaphorical condition: *Do you know what that lawyer is? A shark.*). The results showed an N400 over the medial-central sites and a P600 over the medial-parietal sites. While a significant effect of metaphoricity was found for both the N400

and P600 time windows, they did not find significant effects of metaphor familiarity. The researchers interpreted the results as a successful replication of the biphasic N400-P600 effect seen in previous studies that compared metaphorical stimuli with literal stimuli. For their experiment 2, Bambini et al. (2016) constructed context sentences that made explicit a semantic link between the critical noun and its associated noun (e.g., literal condition: *That fish is really aggressive. It is a shark*.). For the metaphorical condition, this semantic link would correspond to the common ground between the metaphor's source and target, which facilitates the metaphorical mapping (e.g., *That lawyer is really aggressive. He is a shark*.). As the researchers expected, the results showed no significant effect of metaphoricity in the N400 time window, though a significant effect was seen in the P600 window over the frontal, central, and parietal electrodes, with the metaphorical condition eliciting the positivity.

Overall, Bambini et al. (2016) found that the N400 was only visible when there was minimal context, though the P600 was seen in both experiments. Consequently, their results suggested a particular sensitivity to pragmatic context for the N400 that is important to consider in metaphor experiments, though they remained agnostic as to whether their results support a direct or indirect access view. They suggested that the N400 may index processing efforts related to comprehension in the absence of supporting context, such as when participant's predictions about upcoming words are violated. Thus, in this view, the N400 has nothing to do with metaphorical processing per se, and more to do with lexical access. They interpreted the P600 as an index of pragmatic inference, in this case an inference from the literal meaning to the intended meaning of a metaphorical expression. Context is important for such inferential processes.

2.3.3. Bilingual studies of metaphors

The literature on ERP studies on metaphor processing in bilingual adults is still nascent, as few attempts have been made to use ERPs to study bilingual figurative language processing (García et al., 2015). The following discussion centres on one of the main questions addressed in this literature: when processing conceptual metaphors (hereafter, CM) in the L2 and in the L1, what are the differences between the ERP response? A summary of how each study addresses this question is provided in Table 2.1.

Thus far, the main L1-L2 language pairings investigated have been Mandarin-English and Polish-English. Chen et al. (2013) sought to investigate whether CMs are processed differently by upper-intermediate Chinese-English bilingual speakers when presented in L2 English versus L1 Chinese. The participants, who were "relatively proficient" in English, were shown sentences in four conditions: Chinese literal, Chinese CM, English literal, and English CM. Sentences were syntactically constrained into the simple copular verb construction of the form *X is Y* (e.g., "生命是唱片 *shēngmìng shì chàngpiàn*" or "Zeal is fire"). The participant's task was to decide if each sentence was metaphorical or not. The authors formulated their hypothesis based on the predictions of the Graded Salience Model (GSM; Giora, 2002), which states that meanings that are salient (conventionally lexicalised, context-independent, or prominent) are processed first regardless of literalness or contextual fit: L1 Chinese CMs should be easier to access than L2 English CMs because L1 Chinese CMs should be more salient. Hence it was expected that English CMs elicit larger N400 effects than Chinese CMs. The ERP results showed a graded N400 effect: the largest N400 was observed for the English CM condition, followed by English literal sentences and then Chinese CMs and Chinese literal sentences. The researchers did not find a significant difference in responses to Chinese CMs and Chinese literal sentences, suggesting that the participants processed both with similar ease. The researchers concluded that their results supported the Graded Salience Model; Chinese

18

CMs were so salient that they could be processed as easily as Chinese literal sentences. Correspondingly, English literal sentences were less salient than the Chinese literal sentences and CMs, and English CMs even less salient than English literal sentences, explaining the graded N400.

Table 2.1. Summary of bilingual studies and results on ERP amplitude differences for CMs in L1 and L2.

Author(s)	L1/L2	L2 Proficiency	Research Question	Results (N400)	Results (LPC)
Chen et al. (2013)	Chinese/English	Upper-intermediate	Do bilinguals recruit different neural pathways to process CMs in their L1 and L2?	Graded N400 effect. No difference between Chinese CM and Chinese literal. Larger N400 for English literal than Chinese literal. Larger N400 for English CM than Chinese CM and English literal.	None reported.
Jankowiak et al. (2017)	Polish/English	Late high proficiency	How do bilinguals process novel vs. conventional CMs in their L1 and L2?	Early window (300-400 ms): English verb+noun dyads in general elicited smaller N400 than Polish dyads. Late window (400-500 ms): Graded N400 effect independent of language. Anomalous dyads elicited the largest N400, followed by novel CMs, conventional CMs, and literal dyads.	LPC time window: Smaller LPC for Polish novel CMs than Polish conventional CMs. LPC for both English novel and conventional CMs was reduced relative to literal dyads.
Wang (2018)	Chinese/English	High vs. low proficiency	Do bilinguals recruit L1 neural pathways when processing L2 CMs?	No significant difference across both groups was found between Chinese CMs and Chinese literal sentences. Graded N400 effect: the N400 was larger for English CMs than English literal sentences across both groups. Larger N400 response to English CMs from the low group than from the high group. No significant difference for the high group was found between Chinese and English sentences in general.	LPC time window: CMs elicited a late negativity which partially overlapped with the LPC. Low group: A larger P600 response to English CMs than to Chinese CMs was found.
Jankowiak et al. (2021)	Polish/English	Late high proficiency	How do bilinguals process novel CMs vs. novel similes vs. literal sentences vs. anomalous sentences presented in their L1 and L2?	English: No significant differences between anomalous sentences, novel CMs, and novel similes. Polish: Graded N400 effect. Anomalous sentences elicited the largest N400, followed by novel CMs, novel similes, and then literal sentences.	Across both languages: Novel CMs and anomalous sentences elicited a larger late negativity than novel similes.
Tang et al. (2022)	Chinese/English	Late high proficiency	Would scientific CMs evoke larger N400 waves than conventional CMs in both L1 and L2?	Larger N400 for scientific CMs than conventional CMs. Larger N400 at the parietal region for English than for Chinese.	Smaller LPC for scientific CMs than conventional CMs. Smaller LPC for the English than for Chinese.

These results are partly contradicted by Jankowiak et al. (2017). Since monolingual studies had previously shown that novel CMs generally elicit a larger N400 than conventionalised CMs, the researchers sought to examine the brain responses of advanced Polish-English bilingual participants towards novel CMs, conventional CMs, literal utterances, and semantically anomalous utterances in both their L1 and L2. The stimuli were syntactically constrained into verb + noun dyads (e.g., *to exercise liberty*). General between-language differences were found in the early N400 time window (300-400 ms): L2 English dyads generally elicited a smaller N400 than L1 Polish dyads. The researchers suggested that this may be caused by weaker semantic links in the L2 as compared to the L1. However, a graded N400 effect independent of language was found in the later N400 time window (400-500 ms). Here, anomalous dyads elicited the largest N400, followed by novel CMs, conventional CMs, and then literal utterances. The authors interpreted this to suggest that lexico-semantic processes in non-literal meaning comprehension were not influenced by the language being L1 or L2.

Meanwhile, in the LPC time window (500-800 ms), novel CMs were found to evoke a smaller LPC than conventional CMs and literal utterances, but only in Polish. In English, a reduced LPC was observed for both novel and conventional CMs relative to literal utterances. This was contrary to the researchers' expectations that anomalous utterances and novel CMs would evoke larger LPC responses than conventional CMs and literal utterances, and they suggest that the reduced LPC responses may be caused by a sustained negative component that may have entered the LPC time window from the N400 window.

Wang (2018) also observed a late negativity, peaking during the LPC time window at around 780 ms, as well as a larger N400 for L2 English CMs than English literal sentences. In Wang's study, the aim was to investigate if bilingual speakers (L1 Chinese and L2 English) recruit different neural mechanisms to process CMs in their L1 and L2, and whether bilinguals

20

of different L2 proficiency levels do so differently. They recruited a high and a low L2 proficiency group. The stimuli consisted of Chinese and English sentences in four conditions: Chinese literal, Chinese CM, English literal, and English CM. The stimuli were constrained to the form *X is Y* (e.g., *Theories are fathers* and *Eyes are organ* [sic]). The ERP results showed no significant differences between the Chinese literal sentences and Chinese CMs across both proficiency groups. A graded N400 effect was also generally observed, with English CMs eliciting a larger N400 than English literal sentences, followed by the Chinese CMs/literal sentences. In the low proficiency group, a larger N400 response towards English CMs was observed than in the high proficiency group. Furthermore, the high proficiency group did not show significant differences between processing Chinese and English in the N400 window. For both proficiency groups, a late negativity was observed in the LPC window following the CM stimuli, peaking around 780 ms. This masked the LPC which was expected to be elicited by the CMs. Wang interpreted this late negativity to represent secondary semantic integration processes, following Arzouan et al. (2007). Yet, Wang also noted a "larger P600" was observed for English CMs in the low proficiency group in comparison to the high proficiency group, suggesting that the less proficient L2 learners exerted higher cognitive effort in the later time window when assessing the validity of their earlier analyses.

Jankowiak et al. (2021) focused on the processing of novel CMs in L1 Polish and L2 English and aimed to test the effect of comparison structure on ease of CM comprehension. Typically, the comparison structure of a CM takes the form *X is (a) Y*, but it is also possible to realise a CM as a simile with the comparison structure *X is like (a) Y*. The researchers recruited native speakers of Polish who were late proficient speakers of L2 English. The stimuli were 50% in Polish and 50% in English, divided into four conditions: novel nominal CMs (e.g., *Memory is a bag.*), novel similes (e.g., *Memory is like a bag.*), literal sentences, and anomalous sentences. In the ERP results, the researchers found a graded N400 effect in Polish but a

21

different graded pattern in English. In Polish, anomalous sentences evoked the largest N400, followed by novel CMs, then novel similes, and then literal sentences. In contrast, in English, the N400 effect for novel CMs and similes converged with the anomalous sentences. The researchers interpreted this to suggest that, during lexico-semantic access (indexed by the N400), comparison structures do not play a role in CM processing in the L2 and that both CMs and similes in the L2 are initially processed similarly to anomalous sentences. Yet, within the LPC time window (600-800 ms), the between-language differences disappeared and a sustained negativity was observed for both languages for novel CMs and anomalous sentences but not for novel similes and literal sentences. The researchers suggest this negativity might index ongoing difficulty in meaning integration, prolonged activation of semantic information, or activation of non-literal comprehension pathways.

Tang et al. (2022) sought to test if scientific CMs would evoke a larger N400 response than conventional CMs in both L1 and L2 with two experiments, one with Chinese stimuli and one with English translations of the Chinese set. Both language stimuli were constrained in form into *X is Y* (e.g., 電路是階梯 *diànlù shì jiētī* or *A charge is flow*). The stimuli were split into 3 conditions: scientific CMs, conventional CMs, and literal sentences. According to the researchers, scientific CM are used by scientists to describe abstract concepts in terms of simpler ones. The target domains of the scientific CM are thus more complex than those of conventional CMs, which are used more frequently in daily life. The researchers recruited native speakers of Chinese with upper-intermediate proficiency, described as "late unbalanced". The ERP results showed that, in both languages, there was a graded N400 effect with scientific CMs evoking a larger N400 response than conventional CMs; in contrast, the LPC was smaller for scientific CMs than conventional CMs. The researchers suggest that the LPC reduction might be caused by a prolonged negativity which was greater for the scientific CMs. Between-language effects were also found wherein English scientific CMs evoke a larger N400 response

than Chinese scientific CMs at the parietal region. The authors suggested that the scientific terms used in the scientific CMs may be infrequent in both L1 and L2, though still more unfamiliar in L2 which led to greater difficulty in semantic access for the L2 scientific terms. Notably, the LPC was smaller in the English experiment than in the Chinese experiment.

Taken together, the five studies just presented represent a nascent ERP literature on the bilingual processing of CMs. Some consistencies, as well as inconsistencies, can be noted. Generally, L2 CMs can be expected to evoke larger N400 responses than L1 CMs (Chen et al., 2013; Wang, 2018; Tang et al., 2022). Notably, this expectation is contradicted by the results from Jankowiak et al. (2017), whereas Jankowiak et al. (2021) did not compare the amplitudes of the N400 in the L2 with the corresponding amplitudes in the L1. On the contrary, Jankowiak et al. (2017) reported that their L2 word dyads elicited smaller N400 responses than L1 word dyads, but this only applied to the early N400 time window (300-400 ms), following which a language-independent graded N400 effect emerges. It is not clear whether this observation is an effect of Jankowiak et al.'s (2017) more precise description of the N400 time window; the Chinese studies notably did not differentiate between the early and late windows of the N400.

Another general finding was that the LPC is attenuated by a prolonged negativity which was larger for less salient stimuli (Jankowiak et al., 2017; 2021; Wang, 2018; Tang et al., 2022). The only study which did not report the LPC and the negativity associated with CM processing was Chen et al. (2013), who did not discuss the late time window. Another outlier was Wang (2018). Even though Wang (2018) reported the LPC in a way consistent with the other studies, they also mentioned that they observed a "larger P600" in the L2 CM processing of their low proficiency group compared with the high proficiency group. Crucially, it should be noted that some of the English stimuli examples presented in Wang (2018) are ungrammatical (e.g., "Eyes are organ." and "Love is rose." [sic]), which may be a confounding factor.

Three general criticisms can be made about these five studies. The first relates to ecological validity. Nearly all the studies used the *X is Y* form to present their stimuli, except Jankowiak et al. (2017). However, the effectiveness of this construction, or the verb + noun dyad, to evoke CMs is questionable. For example, *Theories are fathers* (Wang, 2018) does not seem to effectively evoke the CM THEORIES ARE FATHERS, which may be realised more implicitly in natural language, as discussed in Chapter 1. Relatedly, the studies did not account for the cultural distribution of CMs nor their discourse context. Some CMs exist in one language but not others; for example, the conceptualisation of the HEART as the ORGAN OF THINKING exists in Chinese but not English (Yu, 2003). Finally, as Pynte et al. (1996) and Bambini et al. (2016) show, the role of context in modulating the CM processing has important consequences for ERP results. CMs naturally occur within supportive contexts. Unless such contexts are accounted for, it is hard to generalise any results about how CMs are processed in the brain's daily linguistic life.

2.4. The current study

In their essence, the three criticisms highlighted above are problems of uncontrolled confounding variables. The unnaturalness of the linguistic stimuli used in prior studies represents a problem with accounting for the role of syntax in comprehension. Such unnaturalness no doubt stems from the desire to eliminate syntactic variation as a possible confound (Keating & Jegerski, 2015), though it does so at the risk of introducing the precise lack of variation thereof as a confound in its steed. The second and third criticisms jointly target the lack of concern for the role of culture and discourse context in prior studies. Yet, metaphors are not produced in context-free vacuums, and brains belong to people who have lives of their own which are intricately wound up in specific contexts of situation (Agar, 1994). The bilingual

studies cited above do not adequately account for this crucial aspect of human language behaviour.

Consequently then, it is with a mind to address these shortfalls in factorial control that the present study is conceived. The present study aims to answer the following research question (repeated from Chapter 1):

- Do L2 speakers of English process metaphorical language differently from native speakers of English?

Due to the difficulty of recruiting native English speakers in Taiwan, the current thesis only presents preliminary results and discussion towards this end (see Chapter 4). Whilst recruitment and data collection are ongoing at the time of this writing, this thesis is able to report on the grand average ERPs of 14 participants in the non-native group (non-native speakers of English whose L1 is Taiwanese Mandarin) and to conduct a qualitative comparison with the grand average ERPs of 6 participants in the native English speaker group.

Chapter 3   Methodology

The study was approached as a 2×2 design (language background $\times$ metaphoricity). Language background is defined here as a between-subjects independent variable (IV) with two levels: native and non-native speakers of English. Metaphoricity is defined as a within-subjects IV with two levels: literal and metaphorical. More details on metaphoricity are explained below in Section 3.2.1.

3.1. Participants

For this study, two groups of participants were recruited: one group of non-Taiwanese native English speakers (hereafter the 'English group'; $n = 8$, 2 female, 6 male, mean age = 30.0), one group of Mandarin-English bilinguals from Taiwan (hereafter the 'Mandarin group'; $n = 15$, 8 female, 7 male, mean age = 23.5). During recruitment, English group participants were restricted to individuals from the United States, United Kingdom, Canada, Australia, Ireland, and New Zealand, while Mandarin group participants were restricted to individuals from Taiwan. During data analysis, 2 participants from the English group and 1 participant from the Mandarin group were excluded due to artefact contamination in the EEG data, so only 6 English group participants (1 female, 5 male, mean age = 28.83) and 14 Mandarin group participants (7 female, 7 male, mean age = 23.29) were included in the results (see Chapter 4). The English proficiency of the Mandarin group was determined based on standardised tests (e.g., TOEFL, TOEIC, IELTS, etc.) undertaken by the participants in the past. Additionally, participants' handedness was assessed using a simplified 10-item version of the Edinburgh Handedness Inventory (Oldfield, 1971). Prior to recruitment, potential participants were asked to self-report any history of neurological or psychiatric disorders or recent use of psychoactive medication in order to exclude any participants with such history.

The Cultural Intelligence Scale (CQS) questionnaire was used to quantify the intercultural competence of the participants. It is expected that intercultural competence may relate to the ability to learn new norms and ways of meaning that come with a second language or culture, within a dynamic view of cultural construction (Low et al., 2020; Waters, 2014). Hence, the variation in intercultural competence between groups should be controlled. The CQS was developed based on the cultural intelligence theory of Earley and Ang (2003) (see Nguyen-Phuong-Mai, 2017, for an overview). It has been validated for cross-cultural reliability by comparative studies in Singapore and the United States (Van Dyne et al., 2008) and in Taiwan-related contexts, such as for the measurement of the intercultural competence of Taiwanese living abroad (e.g., Wang et al., 2017; Uen et al., 2018) and of international students studying in Taiwan (e.g., Lin et al., 2012). For the purposes of this study, it has also been augmented with a modified version of the biographical information questionnaire from the Intercultural Competence Assessment project of the European Commission (INCA, 2004). The full questionnaire used in this study is available in Appendix A.

The CQS in this study is adapted from Ang et al. (2007) and Van Dyne et al. (2008), with 20 items being rated on a 5-point Likert-scale (1 = strongly disagree, 5 = strongly agree). The 20 items are categorised into four psychometric properties of culture intelligence: metacognitive (4 items), cognitive (6 items), motivational (5 items), and behavioural (5 items). Sub-scores are calculated for each psychometric property by averaging the responses for items that pertain to each property. An overall score is then calculated as the average of all the sub-scores. Therefore, sub-scores and overall scores can only range between 1 and 5. A higher overall score is correlated with better ability to adjust to and function effectively in new cultures or other culturally diverse settings.

3.2. Materials

Seventy-seven sentence pairs in English were initially created, either solely by the experimenter or by interaction with OpenAI ChatGPT 3.5 or Google Gemini 1.0. Sentence pairs consisted of a context sentence followed by a target sentence. Target sentences varied by syntactic structure and length (in terms of words), but all contained a transitive main verb and ended in either an adjunct prepositional phrase or a functionally peripheral embedded clause (see Figure 3.1.; cf. Van Valin & LaPolla, 1998; Fawcett, 2000; 2008).

Figure 3.1. Structural variation across target sentences. Bolded items (verb, object noun head, and sentence-final word) represent points marked for ERP analysis via event codes.
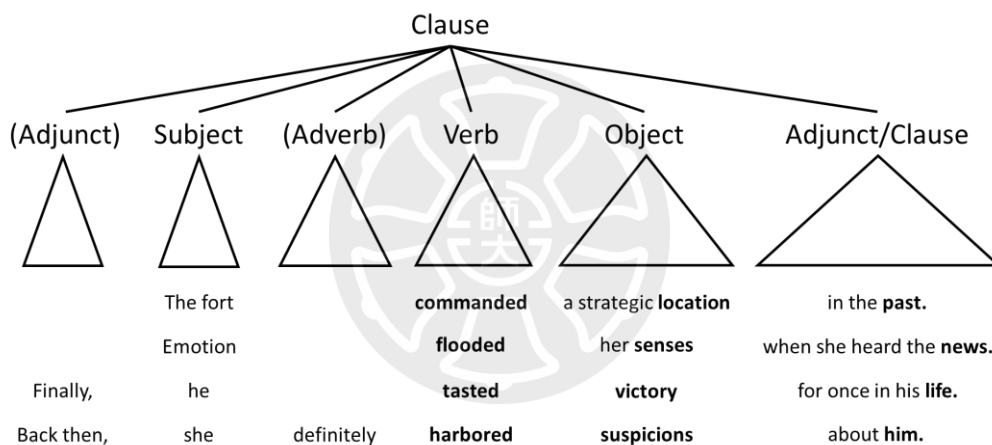
Two versions of each target sentence were created: one whose main verb is used in a literal way and one whose main verb is used in a metaphorical way (cf. Hanks, 2004). Verbs were chosen to contrast metaphorical and literal conditions, instead of nouns as in many previous studies (see Chapter 2), in order to increase the resemblance between the stimuli and natural metaphor use. Cameron (1999) notes that the *X is Y* notation as popularised by Lakoff and Johnson (1980) has misleadingly implied that metaphors are prototypically realised in nominal form (see also Deignan, 2008). But evidence exists suggesting otherwise, in particular that verbal forms are in fact the most common realisations for metaphor, at least in Western

Indo-European languacultures such as English. For example, Cameron (1997), working on U.K. English, reports that "verb metaphors…vastly outnumber nominal metaphors at both Single Word and Phrase levels". Similar preferences for verbal rather than nominal metaphors have been reported in Brazilian Portuguese (Sardinha, 2008) and Spanish EFL (Chapetón, 2010).

Both literal and metaphorical versions of a target sentence shared the same context sentence. Structurally, between 0 and 2 words may intervene between the main verb and the head noun of the object noun phrase, and between 1 and 6 words may intervene between the object head noun and the sentence-final word. Unlike target sentences, context sentences were not constrained for syntactic structure and were included to provide discursive context for the target sentences. Context sentences were designed to be unrelated to the metaphors used in the metaphorical target sentences, such that metaphors may not be predicted from the context sentences. Some example materials are shown in Table 3.1.

Table 3.1. Examples of the experimental materials. The main verb of the target sentence, which determines if the condition is metaphorical or literal, is marked in bold.

Item No.	Condition	Context Sentence	Target Sentence
1	Metaphorical	There are ruins of an old fort on that hill over there.	The fort **commanded** a strategic location in the past.
1	Literal	There are ruins of an old fort on that hill over there.	The fort **occupied** a strategic location in the past.
2	Metaphorical	Today, my mother that discovered she had won the lottery.	Emotion **flooded** her senses when she heard the news.
2	Literal	Today, my mother that discovered she had won the lottery.	Emotion **overwhelmed** her senses when she heard the news.
3	Metaphorical	Yesterday, a saleswoman sold my brother a broken computer.	He completely **consumed** her lies in a matter of seconds.
3	Literal	Yesterday, a saleswoman sold my brother a broken computer.	He completely **believed** her lies in a matter of seconds.
4	Metaphorical	Yesterday, my brother won his first competition.	Finally, he **tasted** victory for once in his life.
4	Literal	Yesterday, my brother won his first competition.	Finally, he **experienced** victory for once in his life.

The seventy-seven sentence pairs were then tested for naturalness, comprehension difficulty, and metaphoricity in an internet-based norming study with native English speakers who reside outside Taiwan (thus distinct from the main experimental English group). The sentence pairs were first split into two lists counterbalanced for condition (a sentence pair does not occur in more than one conditional variant in the same list). Four online surveys were created with Microsoft Forms; two for each list. In each survey, participants rated sentence pairs for naturalness and difficulty of comprehension, both on a five-point Likert scale (1 = very unnatural or very difficult, 5 = very natural or very easy). Additionally, they were also asked to judge if the main verb of the target sentence was used literally or metaphorically (three options: Metaphor, Literal, or "I don't know"). The surveys were distributed on websites such as Reddit, SurveyCircle, SurveySwap, Psychological Research on the Net, LinkedIn, and Facebook. Table 3.2. gives a summary of the four surveys in the norming study.

Table 3.2. Summary of the four internet surveys deployed in the norming study.

Survey	List	Part	Total Items	Total Participants	Rejected Participants	Retained Participants	Rejection Rate
1	1	1	39	39	0	39	0%
2	1	2	38	38	0	38	0%
3	2	1	39	79	27	52	34%
4	2	2	38	38	7	31	18%

Survey 3 (List 2 Part 1) and Survey 4 (List 2 Part 2) included attention checks which allowed the elimination of participants who were unlikely to have been paying attention to the stimuli whilst filling out the survey. No participants were rejected in Survey 1 (List 1 Part 1) and Survey 2 (List 1 Part 2).[4] After responses from participants who failed the attention checks were removed, participant ratings for naturalness and comprehension difficulty were averaged

---

[4] Due to a human error, the attention checks were not implemented in Survey 1 and Survey 2. Without being able to check if participants were paying attention, I do not reject any participant data in those surveys.

for each of the 77 sentence pairs in each list. The mean naturalness and difficulty ratings are summarised in Table 3.3. and plotted in Figure 3.2. using the *Flexplot* package in R (Fife, 2020; R Core Team, 2022). A criterion was set such that items which scored below 4.00 for naturalness or difficulty would be eliminated. Two items were eliminated by this criterion.

Table 3.3. Summary of the mean naturalness and comprehension difficulty (ease) ratings of the materials in the norming study.

List	Mean Naturalness (overall)	Mean Ease (overall)	Mean Naturalness (MET)	Mean Naturalness (LIT)	Mean Ease (MET)	Mean Ease (LIT)
1	4.55	4.67	4.51	4.59	4.62	4.72
2	4.62	4.72	4.58	4.66	4.69	4.75
Both	4.59	4.70	4.55	4.63	4.66	4.74

Figure 3.2. Bee-swarm plots of the naturalness ratings and difficulty (ease) ratings for 154 sentence-pairs across both lists 1 and 2. The ratings had been previously averaged across participants.
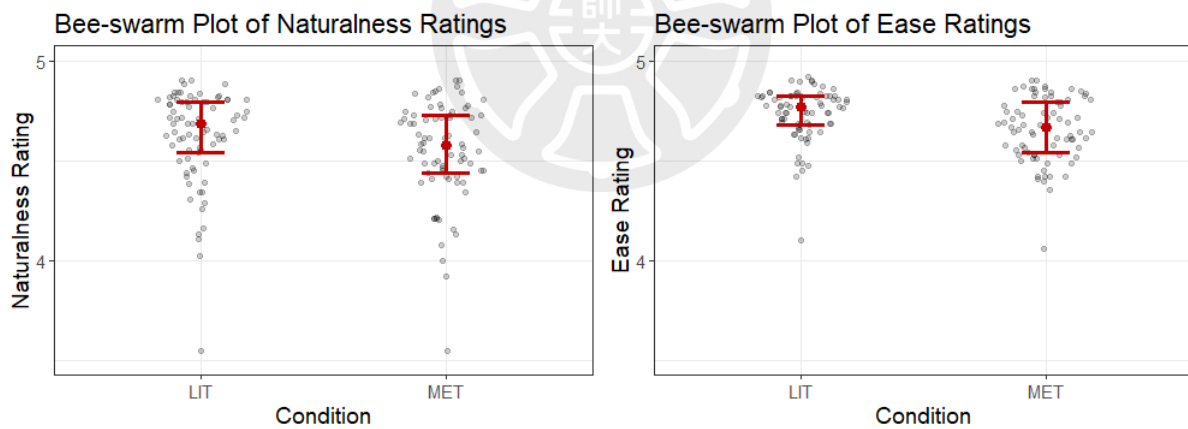
Paired samples *t*-tests were conducted to check for potential differences in the averaged naturalness ($n = 154$, M = 4.55) and difficulty ratings ($n = 154$, mean = 4.63) across both lists together by condition. The results of the *t*-tests showed no significant difference in average naturalness ratings between the metaphorical and literal conditions ($t(76) = 1.914$, $p = .059$), though a significant difference was found for the average difficulty ratings ($t(76) = 3.385$, p

< .01). To gain a better understanding of this difference, Cohen's *d* was calculated for the difficulty ratings (*d* = 0.524, 95% CI [0.200, 0.847]), indicating a moderate effect size.

A second criterion by metaphoricity judgements was set (partly inspired by the Theory of Norms and Exploitations; see Section 2.2). A sentence pair was classified as metaphorical if at least 20% more participants judged it as metaphorical than literal (e.g., 60% says sentence pair X is metaphorical vs. 30% who says it is literal). The same applies to literal classifications. If no such consensus emerges, then the item is eliminated. Items are also eliminated if the consensus among participants turns out to be the opposite of the intended condition (e.g., sentence X is designed to be metaphorical but consensus judges it as literal). Applying the metaphoricity criterion resulted in a further elimination of 34 items, bringing the total item count to 43. Finally, further inspection of the materials found that some literal verbs were duplicated across more than one item. To prevent repetition effects, 3 more items were removed from the materials, giving a final item count of 40 (20 sentence pairs per condition).

3.3. Procedure

The experiment was conducted in the Neurolinguistics Lab at National Taiwan Normal University, in a partially soundproof room. Participants were asked to sign a consent form and then to complete two questionnaires before the experiment: a questionnaire for health and demographic information and the Cultural Intelligence Scale questionnaire (see Appendix A). While the participant filled in both questionnaires, a 32-channel EEG electrode cap was placed on the participant's head and set up to begin recording the EEG.

After the electrode cap set-up was completed, the participant was guided to sit approximately 80 cm in front of the presentation computer and to hold a Logitech F310 game controller in a comfortable position, with their thumbs over the "back" (left) and "start" (right) buttons. They were told that those are the only two buttons they would need to use. The

participant was instructed to read the sentence pairs on the screen with utmost attention. Text was presented in white 18 pt Consolas font against a black background. Each sentence pair began with a 1000 ms fixation cross at the centre of the screen. Then, context sentences were presented on the screen in one slide for 3500 ms, followed target sentences presented word-by-word (see Figure 3.3.). Each word appeared for 300 ms and was then followed by a 200 ms blank, except for sentence-final words, which were followed by a 1200 ms blank. If the word was followed by any punctuation such as a comma or a full stop (period), the punctuation mark appeared together with it. A sentence pair constituted 1 trial and the main experiment consisted of 40 trials in total.

Figure 3.3. Stimulus presentation procedure for a single trial ending with a comprehension question.
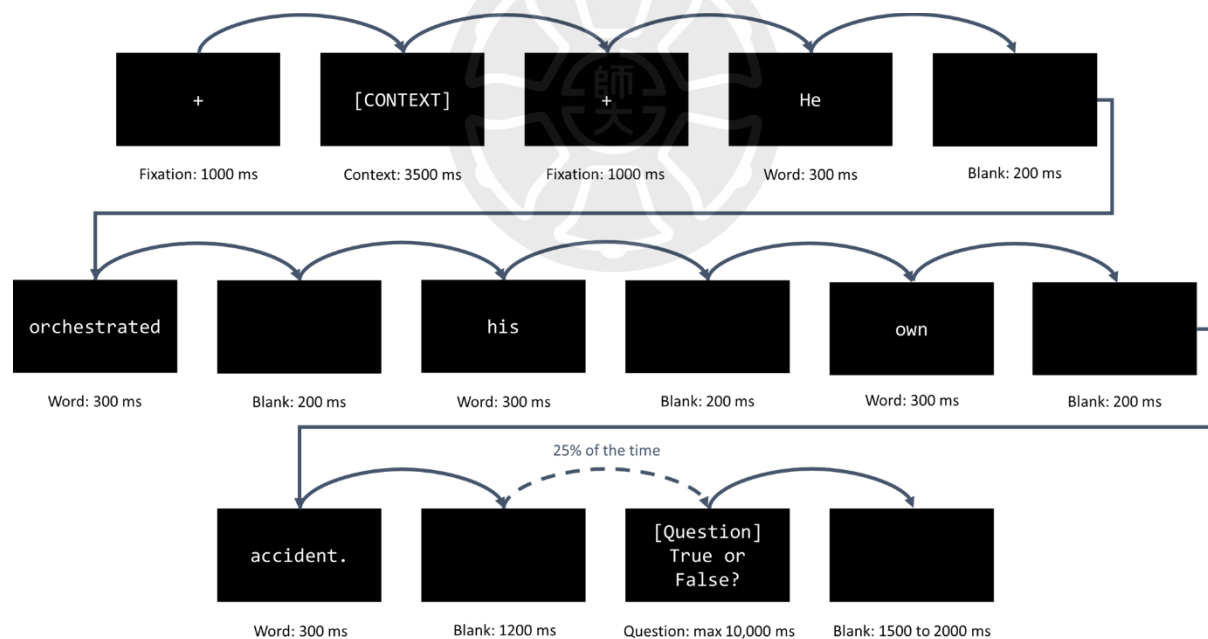
To ensure that participants paid attention to the stimuli, 25% of trials were followed by a true-or-false comprehension question about a statement related to the content of the sentence pair. Participants were told that comprehension questions appear randomly for 25% of trials, though the sequence of trials was actually predetermined in the same pseudorandomised order

for all participants. For the comprehension question, a statement was displayed on the screen for a maximum of 10,000 ms (10 sec), alongside a prompt reminding the participant which button on the controller signified "true" or "false". During this time, the participant pressed either the left or right button on the controller to indicate their response. The keys assigned to indicate "true" or "false" were counterbalanced across participants, such that half of the participants used their right thumb to respond "true" and the other half used their left thumb to respond "true". The opposite thumb was used to respond "false". After the participant made a response, a blank of duration between 1500 and 2000 ms appeared before the experiment proceeded to the next trial.

The participants were told to sit as still as possible and to limit the urge to blink except on the context sentence or during the comprehension question, since these time periods would not be submitted to ERP analysis. To familiarise the participants with the stimulus presentation procedure, 8 practice trials were run before the main experimental session. In the main experimental session, the participants underwent the stimulus presentation procedure until all 40 trials were completed. Two breaks of indefinite duration were provided, during which participants could request eyedrops if their eyes felt tired or dry. The total time for the experiment (including the practice session and breaks) usually took 12–20 minutes to complete.

3.4. EEG recordings

This experiment used the E-Prime 3 software (Psychology Software Tools, Inc., 2016) in a Windows 7 environment to present the experimental stimuli and record participants' behavioural responses to the comprehension questions. In addition, E-Prime 3 was programmed to send event codes to the EEG digitisation computer to mark specific events on the recorded EEG. For the present study, ten kinds of events were marked: the onset of the context sentence, the onset of the first word of the target sentence (if not the subject noun

head/pronoun), the onset of the target sentence subject's noun head or pronoun, the onset of the target sentence main verb, the onset of the target sentence object's noun head or pronoun, the onset of the target sentence's final word, the onset of a question, the participant's behavioural response (correct and incorrect), and the onset of a break.

Electroencephalograms were recorded from a 32-channel electrode cap (Quik-Cap, Compumedics NeuroMedical Supplies, USA/Australia). The 32 channels were: FP1, FP2, F7, F3, FZ, F4, F8, FT7, FC3, FCZ, FC4, FT8, T3, C3, CZ, C4, T4, TP7, CP3, CPZ, CP4, TP8, A1, T5, P3, PZ, P4, T6, A2, O1, OZ, and O2. The average voltage between A1 and A2, placed at the left and right mastoids, was used as the reference voltage. Four other electrodes were placed on the participant's face to monitor vertical and horizontal eye movements—two on the outer canthus of each eye and two on the upper and lower ridges of the left eye. The impedance of all electrodes was kept below 5 kΩ.

3.5. Data analysis

3.5.1. Behavioural data analysis

The behavioural data in the main experiment consisted of true/false responses by the participant that could be correct or wrong. Accuracy was computed as the percentage of correct responses relative to the number of questions. To each of the 10 questions which participants encountered in the experiment, participants responded with "true", or "false", or no response at all. A correct response is one in which the participant's choice of "true" or "false" aligned with the predetermined correct answer for the question. Failure to respond to the question within the 10,000 ms (10 sec) timeframe was counted as a wrong response. A criterion was set such that participants were excluded from the ERP analysis if they provided more than 4 wrong responses. Additionally, E-Prime 3 collected the reaction times (in ms) of participant behavioural responses, which were submitted to inferential statistical analyses (Welch's two-

way *t*-test and Cohen's *d*) in order to investigate possible effects of language group on reaction time (see Section 4.1.1.).

A second set of behavioural data was obtained through the Cultural Intelligence Scale questionnaire administered to all participants prior to EEG collection. These questionnaire scores were computed and calculated according to the scoring procedure specified by Ang et al. (2007) and Van Dyne et al. (2008) (see Section 3.1.). Scores derived from the questionnaire were then submitted to inferential and descriptive statistical analyses (Welch's two-sample *t*-test, F-test for equality of variances, and Levene's test) to investigate potential effects of language group on the derived scores as well as to compare within-group variability (see Section 4.1.2.).

3.5.2. ERP data analysis

The EEG data was analysed using the EEGLAB (Delorme & Makeig, 2004) and ERPLAB Studio 11.02 (Lopez-Calderon & Luck, 2014; latest release: https://github.com/ucdavis/erplab) interactive toolboxes for MATLAB (The MathWorks Inc., 2022).

ERP analyses were time-locked with event codes (sent by E-Prime 3) to the onsets of the target verbs, noun head of the object of the verb, and sentence-final words. The EEG preprocessing steps are briefly described here. First, channel locations are specified for the raw continuous EEG upon importing into ERPLAB Studio. Then, the continuous EEG is submitted to a high-pass filter (IIR Butterworth, 0.1 Hz, db/oct = 12) prior to artefact correction via Independent Component Analysis (ICA; Jung et al., 2000). In order to accelerate the ICA, IC decomposition was performed on a duplicate EEG dataset that had been additionally submitted to a heavy bandpass filter (IIR Butterworth, 1.0 Hz, db/oct = 48) and downsampled from 1000 Hz to 100 Hz. Time segments without event codes (buffer before event code = 500 ms; buffer after event code = 1500 ms) were deleted from the duplicate dataset. The duplicate dataset was

36

then subjected to automatic artefactual time segment rejection (voltage threshold = 500 µV to 800 µV depending on participant; moving window width = 1000 ms; window step = 500 ms) before the IC decomposition algorithm is run. The decomposed IC weights are then transferred to the original dataset for artefactual components to be removed. If necessary, bad channels are then interpolated (ignored channels: HEOR, HEOL, VEOU, VEOL). An event list is then created to assign each event code to a bin to categorise event codes by condition.

The continuous EEG data were epoched 200 ms before the onset and 1000 ms after it. Channel operations were then performed to coalesce the four eye electrode channels into two bipolar channels, such that VEOU and VEOL were merged into VEOG, and HEOL and HEOR were merged into HEOG. Then, artefact detection was applied using the moving window peak-to-peak (threshold voltage for blinks = 100 µV; threshold voltage for scalp channels = 125 µV), step-like artefacts (threshold voltage for blinks = 50 µV; threshold voltage for horizontal eye movements = 32 µV), and simple voltage threshold (voltage limits for scalp channels = -150 to 150 µV) functions in ERPLAB. Average ERPs were then computed for each participant's data. Finally, grand average ERPs for all participants within an experimental group were computed.

Note that due to the difficulty of recruiting English speakers, the sample sizes for the two language groups were unequal; therefore, the between-subjects factor was not included in further statistical analysis. For the ERP data analysis, three within-subjects variables are specified: metaphoricity (metaphor, literal), anteriority (anterior, central, posterior), and hemisphere (left, midline, and right). Nine representative electrodes are selected (F3, FZ, F4, C3, CZ, C4, P3, PZ, and P4) as well as two time windows: 350–500 ms (for N400), 500–800 ms (for P600 or LPC). Because there is only 500 ms between the onset of a word and the onset of the next word, except for the sentence-final word, the P600/LPC time window can only be analysed for the sentence-final word. The mean amplitudes for each time window (N400 time

window for the verb, object noun, and sentence-final word and P600/LPC time window for the sentence-final word) were calculated and submitted to a 3-way repeated-measures ANOVA. Should Mauchley's test of sphericity show a violation of the sphericity assumption, the Greenhouse-Geisser correction is applied to the p-values. If interaction effects are found, post hoc pairwise comparisons are conducted with Bonferroni correction applied.

3.6. Expected findings

To recapitulate, the present study aims to answer the following research question:

- Do L2 speakers of English process metaphorical language differently from native speakers of English?

Based on the previous studies discussed in Chapter 2, there are some predictions that can be made. To start, it may be predicted that, for the native English speakers, an N400 effect will be observed for the metaphorical condition, followed by a P600. This prediction is supported by the extensive monolingual metaphor processing ERP literature which do not control for the role of discourse context. Yet, if it is expected that context plays a significant role, it may then be expected that there would be no visible N400 for the metaphorical condition, following from the results reported by Bambini et al.'s (2016) second experiment and Pynte et al.'s (1996) third and fourth experiments. However, it is an open question as to how these predictions regarding the role of context, based mainly on monolingual studies, might generalise to metaphor processing in non-native language speakers. The materials used in the bilingual studies surveyed in Section 2.3.3. did not provide context for the metaphors that were presented. While the current study did not seek to compare supportive context versus minimal

context conditions as Bambini et al. (2016) did, comparing how persons of different language backgrounds process metaphors in context may shed some light on this issue.

Chapter 4   Results

This thesis presents preliminary results from participant EEG data that has been collected so far. At the time of writing, 15 participants in the Mandarin group had been recruited and had their data analysed. Due to artefact contamination, 1 participant's data was excluded from further analysis, so only 14 Mandarin group participants (7 female, 7 male, mean age = 23.29) were included in the following statistical analyses and grand average ERPs. For the English group, 8 participants had been recruited, but 2 had their data excluded due to artefact contamination, so only ERPs from 6 English group participants (1 female, 5 male, mean age = 28.83) were grand averaged.

4.1. Behavioural results

4.1.1. Reaction time and comprehension question accuracy

Participants were instructed to respond to comprehension questions ($n = 10$) that appeared after 25% of trials ($n = 40$) by pressing a button to indicate "true" or "false". Reaction times and accuracy were recorded by E-Prime 3. A summary of the mean reaction times (Mandarin: 2946.31 ms; English: 2170.32 ms) and accuracy rate (Mandarin: 93.98%; English: 90.00%) for both groups is shown in Table 4.1. Welch's two-sample *t*-test indicated a significant difference in mean reaction time between the two language groups ($t(92.678) = -4.826$, $p < .001$). Cohen's *d* for the reaction times by language group ($d = -0.787$, 95% CI [-1.126, -0.449]) indicated a moderate effect size, which is supported by visual inspection of the corresponding bee-swarm plot in Figure 4.1. A chi-squared test of independence was then performed to examine the relationship between language group and question accuracy. The results showed no significant association between language group and question accuracy ($\chi^2(1, N = 180) = 0.144$, $p = .704$).

Table 4.1. Summary of the mean reaction times (in milliseconds) and accuracy rates to comprehension questions by language group.

Language Group	Mean RT (ms)	Mean Accuracy Rate
Mandarin (*n* = 14)	2946.308	93.98%
English (*n* = 5)	2170.320	90.00%

Figure 4.1. Bee-swarm plots of the reaction times of participants towards the comprehension questions by language group.
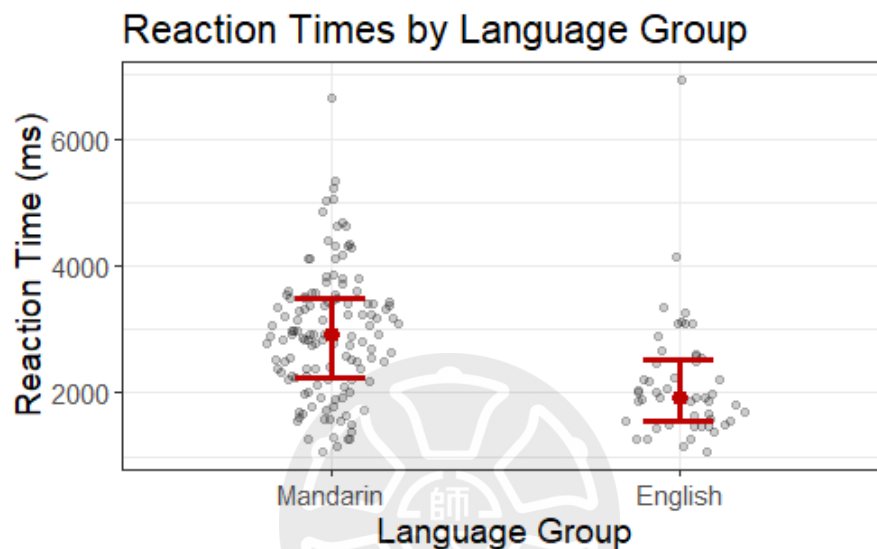
A criterion was set such that participants who answered more than 4 questions incorrectly (accuracy rate < 60%) were excluded from ERP analysis, though none of the participants in both the English group and Mandarin group met this criterion. One participant from the English group was excluded from the summary in Table 4.1. because they were only presented with 4 comprehension questions instead of 10, due to a technical error. However, because this participant answered 3 questions correctly (accuracy rate = 75.00%), it was decided that their data would be kept in the ERP analyses.

4.1.2. Intercultural competence scores

As mentioned in Section 3.1., to control for variation in the intercultural competence of participants between language groups, the Cultural Intelligence Scale (CQS; Ang et al., 2007;

Van Dyne et al., 2008) questionnaire was employed. The CQS in this study is adapted from Ang et al. (2007) and Van Dyne et al. (2008), The CQS is a questionnaire of 20 items categorised into four psychometric categories of cultural intelligence: metacognitive (4 items), cognitive (6 items), motivational (5 items), and behavioural (5 items). Participants respond to each item by giving a rating on a 5-point Likert-scale (1 = strongly disagree, 5 = strongly agree). Sub-scores are calculated for each psychometric property by averaging the ratings across items that relate to each category. The sub-scores are then averaged to produce an overall score for each participant. A summary of the sub-scores and overall scores on the CQS by language group is shown in Table 4.2. A bee-swarm plot for the overall CQS scores by language group is displayed in Figure 4.2.

Table 4.2. Summary of the mean Cultural Intelligence Scale sub-scores and overall scores by language group.

Language Group	Category				
	Metacognitive	Cognitive	Motivational	Behavioural	Overall
Mandarin (n = 15)	3.95	3.12	3.77	3.73	3.73
English (n = 8)	4.06	2.85	4.25	3.75	3.75

Five inferential statistical analyses using Welch's two-sample t-test were conducted on the CQS data to discern possible effects of language group on scores within the four categories and scores overall. The overall t-test found no significant effects of language group on the within-subjects averaged scores across categories ($t(69.833) = 0.460$, $p = .647$). The t-tests for the metacognitive ($t(19.855) = 0.401$, $p = .693$), cognitive ($t(13.593) = -0.989$, $p = .340$), motivational ($t(20.751) = 1.506$, $p = .147$), and behavioural ($t(16.734) = 0.044$, $p = .965$) categories similarly did not reveal any statistically significant effects of language group. These findings are supported by a visual inspection of the bee-swarm plot in Figure 4.2.

Figure 4.2. Bee-swarm plots of the CQS scores in the Mandarin group ($n = 15$) and English group ($n = 8$).
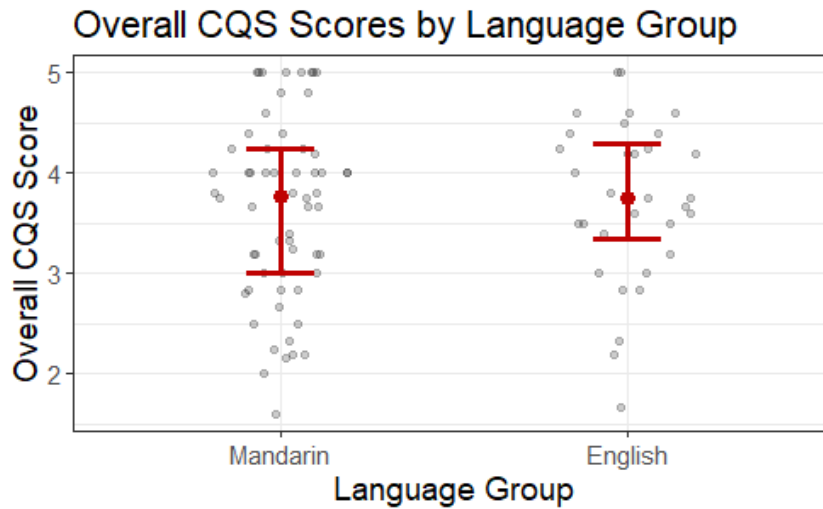
Next, the variability of scores between the two language groups was investigated. The Mandarin group showed slightly greater variability with a variance of 0.810 and standard deviation of 0.900, compared with a variance of 0.648 and standard deviation of 0.805 for the English group. To detect statistically significant differences, an *F*-test for equality of variances and Levene's test were conducted. The *F*-test results indicated no significant difference in variances by language group ($F(31, 59) = 0.799$, $p = .504$). Levene's test also showed no significant difference in variances between the groups ($F(1, 90) = 0.840$, $p = .362$). These findings are also supported by a visual inspection of the bee-swarm plot in Figure 4.2. Therefore, while the Mandarin group showed slightly more variability than the English group, this difference is not likely to be practically significant.

4.2. ERP results: Mandarin Group

4.2.1. Main verb event

The averaged waveforms of the Mandarin group at the target sentence main verb are plotted in Figure 4.3.

Figure 4.3. Mandarin group brainwaves at the main verb with the literal condition plotted in cyan and metaphorical condition in red.
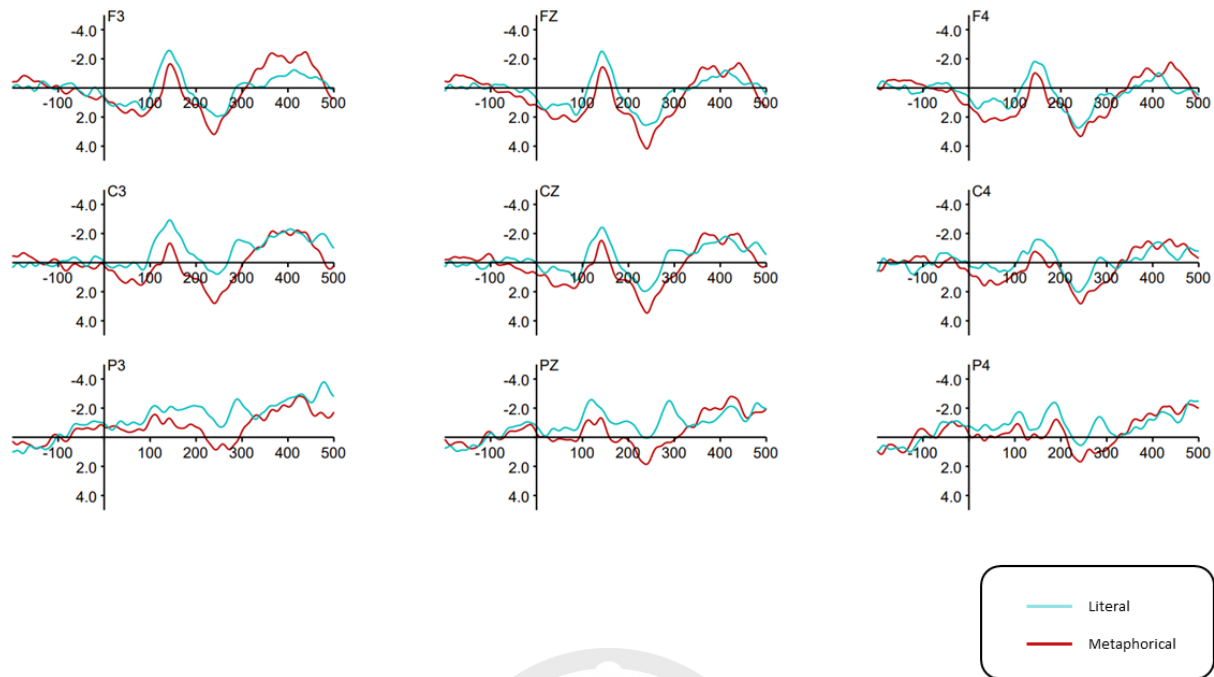
As can be seen in Figure 4.3., both metaphorical and literal conditions elicited N1 and P2 waves (relative to the reference line), most clearly at the frontal (F3, FZ, F4) and central (C3, CZ, C4) electrode sites. The N1 and P2 are visual processing related components known to be most prominent at these sites (Luck, 2014). While the brainwaves for both conditions were largely unified in the pre-stimulus timeframe (-200 ms to onset), they begin to diverge from the stimulus onset with the literal condition being slightly more negative throughout the N1 and P2 time windows. However, upon entering the 350–500 ms window, the brainwaves for the metaphorical condition appears to increase in negativity relative to the literal condition across most of the electrode sites, resulting in a reunification of the two waveforms in this period. The exception is in the graph for the F3 electrode site, where the brainwave for the metaphorical condition appears to be more clearly negative than the literal condition brainwave, though only slightly more so than at other frontal sites. This observation is congruent with the corresponding topographic scalp maps shown in Figure 4.4.

Figure 4.4. Mandarin group topographic maps showing the contrast between metaphorical and literal conditions at the main verb in the 350–500 ms time window.

Main Verb, Metaphorical, 350 – 500 ms   Main Verb, Literal, 350 – 500 ms
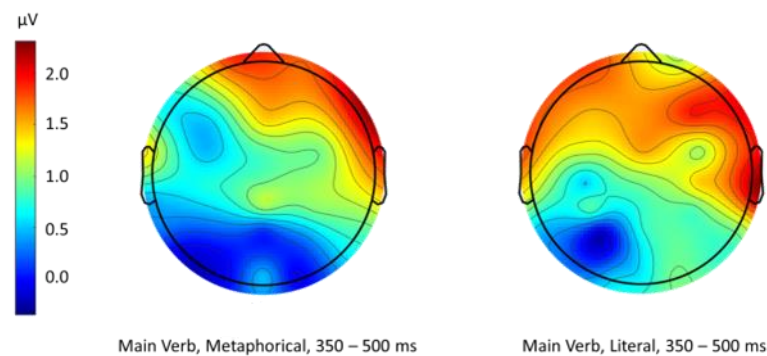
The topographic maps in Figure 4.4. display the mean voltage distributions across the scalp in the 350–500 ms time window in the metaphorical condition (left) and literal condition (right). As can be seen in the scalp map for the metaphorical condition, there is a negative region in the frontal-left region which is absent in the literal condition scalp map. This region corresponds to the location of the F3 electrode site.

The ERP data above were submitted to a three-way repeated measures ANOVA with the within-subjects factors of metaphoricity (literal, metaphorical), hemisphere (left, right), and anteriority (anterior, central, parietal). The ANOVA results are summarised in Table 4.3.

Table 4.3. Mandarin group verb ERP analysis: Summary of the degrees of freedom, F-values, and p-values of a repeated measures ANOVA for the 350–500 ms time window. Significance codes: *** $p < .001$, ** $p < .01$, * $p < .05$

Variables	*350–500 ms*		
	df	*F-value*	*p-value*
Main Effects			
Metaphoricity	1, 13	0.217	0.648
Hemisphere	2, 26	6.661	0.005**
Anteriority	2, 26	18.321	< 0.001***
Two-way Interactions			
Metaphoricity × Hemisphere	2, 26	1.35	0.277
Metaphoricity × Anteriority	2, 26	0.894	0.421
Hemisphere × Anteriority	4, 52	0.142	0.966
Three-way Interactions			
Metaphoricity × Hemisphere × Anteriority	4, 52	3.233	0.019*

Focusing on metaphoricity, the main effect of metaphoricity was found to be not significant ($F(1, 13) = 0.217$, $p = .648$). However, a significant three-way interaction was found between metaphoricity, hemisphere, and anteriority ($F(4, 52) = 3.233$, $p < .05$). Post-hoc pairwise comparisons of estimated marginal means (using Bonferroni correction) on the significant three-way interaction indicated by the ANOVA revealed that the contrast between metaphorical and literal conditions was not significant at any electrode site (Left × Anterior: t = 1.379, $p = .191$). A full summary of results for this post-hoc analysis is shown in Table 4.4.

Table 4.4. Mandarin group verb ERP analysis: Summary of post-hoc pairwise comparisons of the literal – metaphorical contrast for the 350–500 ms time window.

Hemisphere	Anteriority	df	t-ratio	p-value
Left	Anterior	13	1.379	0.191
Midline	Anterior	13	0.237	0.816
Right	Anterior	13	0.908	0.380
Left	Central	13	-0.166	0.871
Midline	Central	13	0.199	0.845
Right	Central	13	0.604	0.557
Left	Posterior	13	-0.588	0.567
Midline	Posterior	13	0.619	0.547
Right	Posterior	13	0.677	0.510

4.2.2. Object noun event

The second event of interest was the head noun of the nominal object following the verb. The averaged waveforms of the Mandarin group at the target sentence object noun are displayed in Figure 4.5.

Figure 4.5. Mandarin group brainwaves at the object noun with the literal condition plotted in cyan and metaphorical condition in red.
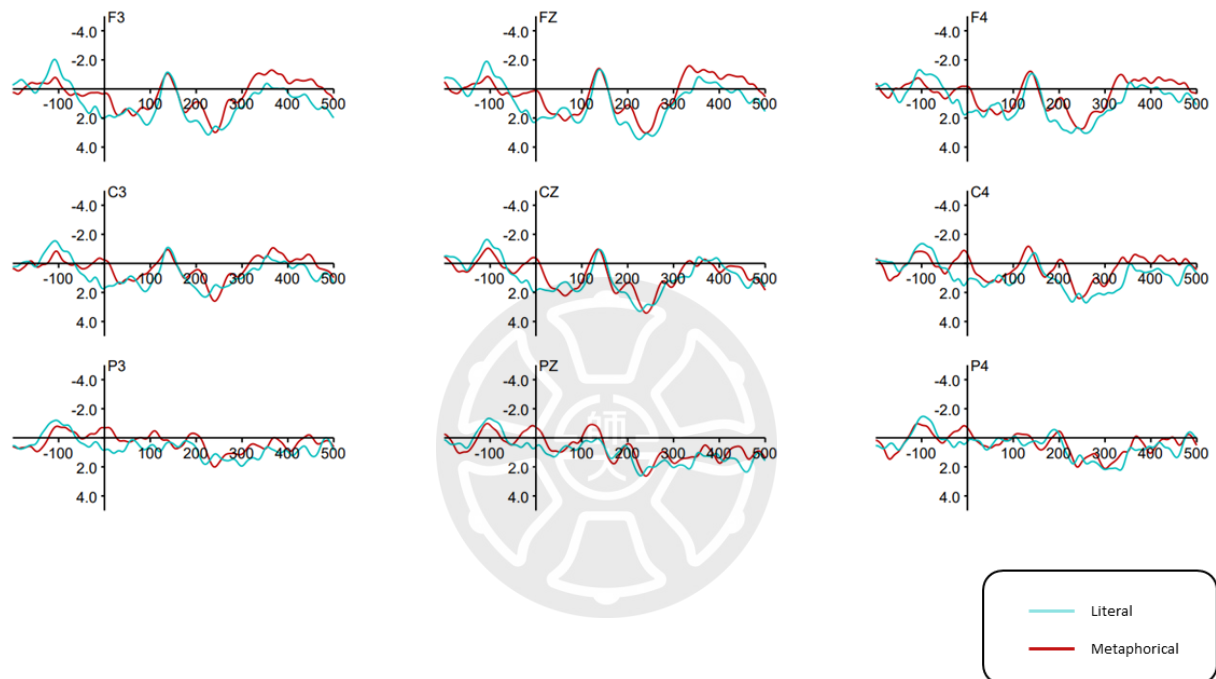
As Figure 4.5 shows, both metaphorical and literal conditions again elicited N1 and P2 waves most clearly at the frontal and central sites. The brainwaves diverge in the pre-stimulus timeframe (-200 ms to 0 ms), suggesting that a prior ERP component is still in action due to the presentation of non-identical stimuli before the noun (main verbs, determiners, or adjectives; see Section 3.2). The two brainwaves reunite during the N1 and P2 timeframes but diverge again slightly around 300 ms at the frontal, central-left, and central-right electrode sites, with the metaphorical condition showing more pronounced negativity. This pattern is captured in the topographic scalp map for the metaphorical condition in Figure 4.6.

Figure 4.6. Mandarin group topographic maps showing the contrast between metaphorical and literal conditions at the object noun in the 350–500 ms time window.

Object Noun, Metaphorical, 350 – 500 ms   Object Noun, Literal, 350 – 500 ms
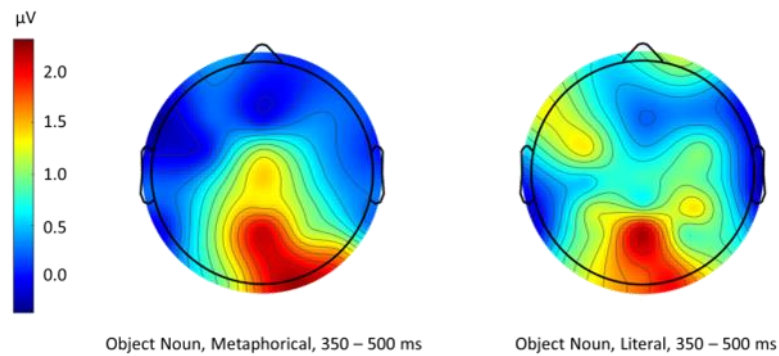
As before, the ERP data was submitted to a three-way repeated measures ANOVA. However, no significant effects or interactions involving the metaphoricity factor (main effect: $F(1, 13) = 0.444$, $p = .517$) were found (see Table 4.5.).

Table 4.5. Mandarin group object noun ERP analysis: Summary of the degrees of freedom, F-values, and p-values of a repeated measures ANOVA for the 350–500 ms time window. Significance codes: *** $p < .001$, ** $p < .01$, * $p < .05$

Variables	*350–500 ms*		
	df	*F-value*	*p-value*
Main Effects			
Metaphoricity	1, 13	0.444	0.517
Hemisphere	2, 26	1.458	0.251
Anteriority	2, 26	11.819	< 0.001***
Two-way Interactions			
Metaphoricity × Hemisphere	2, 26	0.576	0.569
Metaphoricity × Anteriority	2, 26	0.157	0.855
Hemisphere × Anteriority	4, 52	3.574	0.012*
Three-way Interactions			
Metaphoricity × Hemisphere × Anteriority	4, 52	0.980	0.427

4.2.3. Sentence-final word event

The third event of interest was the final word of the target sentence. The averaged waveforms of the Mandarin group at the sentence-final word are displayed in Figure 4.7., epoched from -200 ms to 1000 ms.

Figure 4.7. Mandarin group brainwaves at the sentence-final word with the literal condition plotted in cyan and metaphorical condition in red.
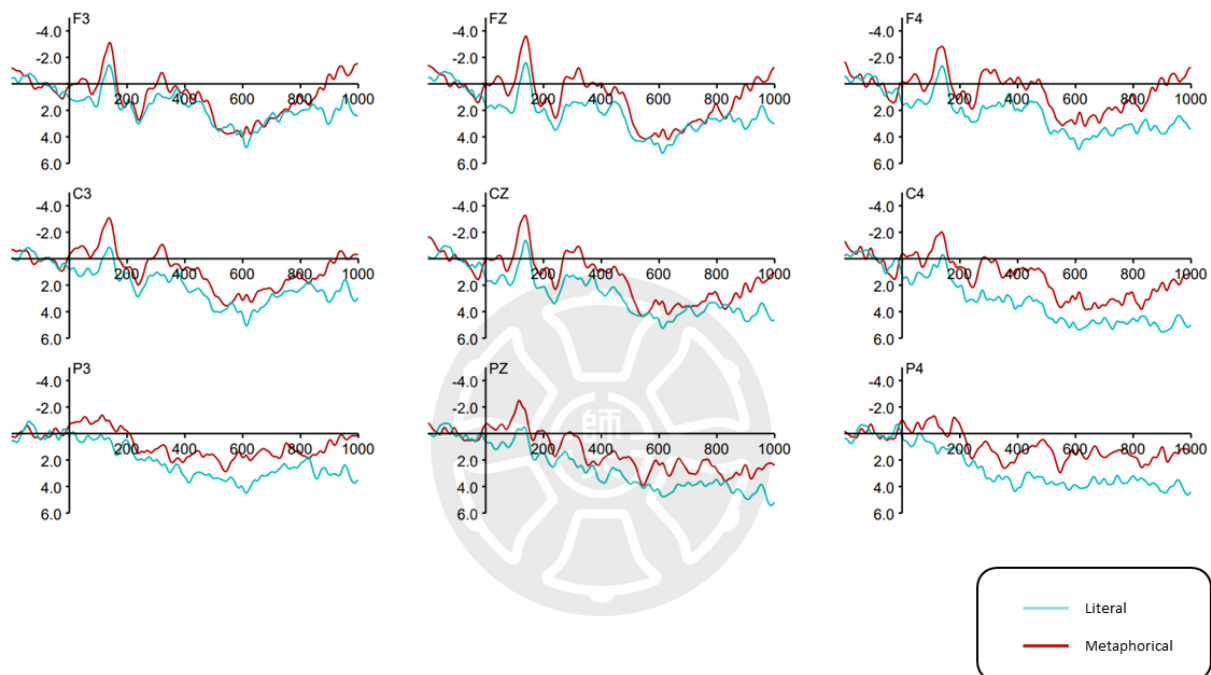
As with the previous waveforms, the peaks of the N1 and P2 components can be observed in both conditions over the frontal and central sites. The brainwaves for the two conditions are split at the onset, with the metaphorical condition remaining more negative than the literal condition throughout the 1000 ms epoch. Considering this initial split, the two conditions do not seem to deviate further in a clear manner during the 350–500 ms and 500–800 ms time windows. However, the metaphorical condition brainwave appears to be more negative-going at around 800 ms relative to the literal conditions across all the electrode sites. See Figure 4.8. for the topographic maps.

Figure 4.8. Mandarin group topographic maps showing the contrast between metaphorical and literal conditions at the sentence-final word in the 350–500 ms time window and 500–800 ms time window.

Sentence-Final Word, Metaphorical, 350 – 500 ms   Sentence-Final Word, Literal, 350 – 500 ms   Sentence-Final Word, Metaphorical, 500 – 800 ms   Sentence-Final Word, Literal, 500 – 800 ms
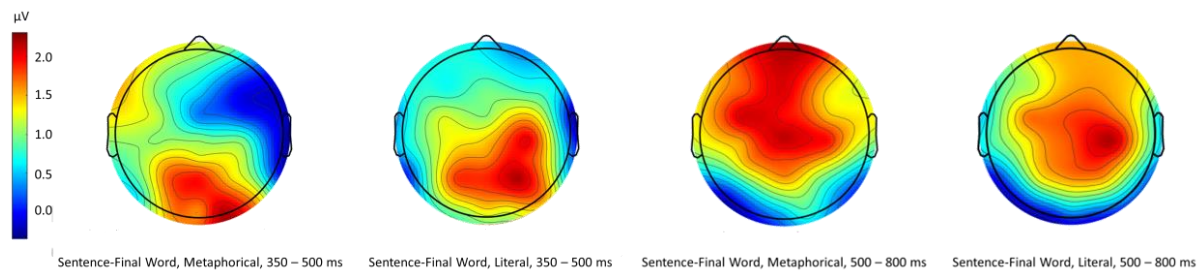
The ERP data for the sentence-final word was submitted to three-way repeated measures ANOVAs for two time windows: 350–500 ms and 500–800 ms. For the 350–500 ms time window, no main effect of metaphoricity was found ($F(1, 13) = 0.596$, $p = .454$). However, a significant two-way interaction was found between metaphoricity and hemisphere ($F(2, 26) = 4.296$, $p < .05$). The ANOVA summary for this time window can be found in Table 4.6.

Table 4.6. Mandarin group sentence-final word ERP analysis: Summary of the degrees of freedom, F-values, and p-values of a repeated measures ANOVA for the 350–500 ms time window. Significance codes: *** $p < .001$, ** $p < .01$, * $p < .05$

Variables	350–500 ms		
	df	F-value	p-value
Main Effects			
Metaphoricity	1, 13	0.596	0.454
Hemisphere	2, 26	0.054	0.947
Anteriority	2, 26	7.876	0.002**
Two-way Interactions			
Metaphoricity × Hemisphere	2, 26	4.296	0.024*
Metaphoricity × Anteriority	2, 26	0.855	0.437
Hemisphere × Anteriority	4, 52	1.298	0.283
Three-way Interactions			
Metaphoricity × Hemisphere × Anteriority	4, 52	0.673	0.614

To further investigate the significant interaction between metaphoricity and hemisphere in the 350–500 ms time window, post-hoc pairwise comparisons were conducted, the results

of which are summarised in Table 4.7. The pairwise comparisons found no significant contrasts between the literal and metaphorical condition in any of the scalp regions.

Table 4.7. Mandarin group sentence-final word ERP analysis: Summary of post-hoc pairwise comparisons of the literal – metaphorical contrast for the 350–500 ms time window.

Hemisphere	Anteriority	df	t-ratio	p-value
Left	Anterior	13	-0.115	0.910
Midline	Anterior	13	0.547	0.594
Right	Anterior	13	0.947	0.361
Left	Central	13	0.644	0.531
Midline	Central	13	0.713	0.489
Right	Central	13	1.341	0.203
Left	Posterior	13	0.667	0.516
Midline	Posterior	13	0.678	0.510
Right	Posterior	13	1.351	0.200

A second ANOVA analysis was conducted for the 500–800 ms time window following the sentence-final word, summarised in Table 4.8. Similar to the previous time window analysis, no significant main effect of metaphoricity was found ($F(1, 13) = 0.823$, $p = .381$) while a significant interaction between metaphoricity and hemisphere was found ($F(2, 26) = 3.696$, $p < .05$). As before the data was submitted to a post-hoc pairwise comparisons analysis. The post-hoc analysis revealed no significant contrasts between the literal and metaphorical conditions at any of the scalp regions. The post-hoc analysis for the 500–800 ms time window is summarised in Table 4.9.

Table 4.8. Mandarin group sentence-final word ERP analysis: Summary of the degrees of freedom, F-values, and p-values of a repeated measures ANOVA for the 500–800 ms time window. Significance codes: *** $p < .001$, ** $p < .01$, * $p < .05$

Variables	500–800 ms		
	df	F-value	p-value
Main Effects			
Metaphoricity	1, 13	0.823	0.381
Hemisphere	2, 26	1.746	0.194
Anteriority	2, 26	1.023	0.374
Two-way Interactions			
Metaphoricity × Hemisphere	2, 26	3.696	0.039*
Metaphoricity × Anteriority	2, 26	2.334	0.117
Hemisphere × Anteriority	4, 52	1.752	0.153
Three-way Interactions			
Metaphoricity × Hemisphere × Anteriority	4, 52	0.514	0.726

Table 4.9. Mandarin group sentence-final word ERP analysis: Summary of post-hoc pairwise comparisons of the literal – metaphorical contrast for the 500–800 ms time window.

Hemisphere	Anteriority	df	t-ratio	p-value
Left	Anterior	13	0.245	0.810
Midline	Anterior	13	0.275	0.787
Right	Anterior	13	0.917	0.376
Left	Central	13	0.613	0.551
Midline	Central	13	0.490	0.633
Right	Central	13	1.413	0.181
Left	Posterior	13	1.264	0.228
Midline	Posterior	13	1.134	0.277
Right	Posterior	13	1.991	0.068

4.3. ERP results: English Group

4.3.1. Main verb event

This section focuses on the brainwaves for the averaged ERPs of 6 participants in the English group. To begin, the grand average brainwaves of the English group for the target sentence main verb are plotted in Figure 4.9.

Figure 4.9. English group brainwaves at the main verb with the literal condition plotted in cyan and metaphorical condition in red.
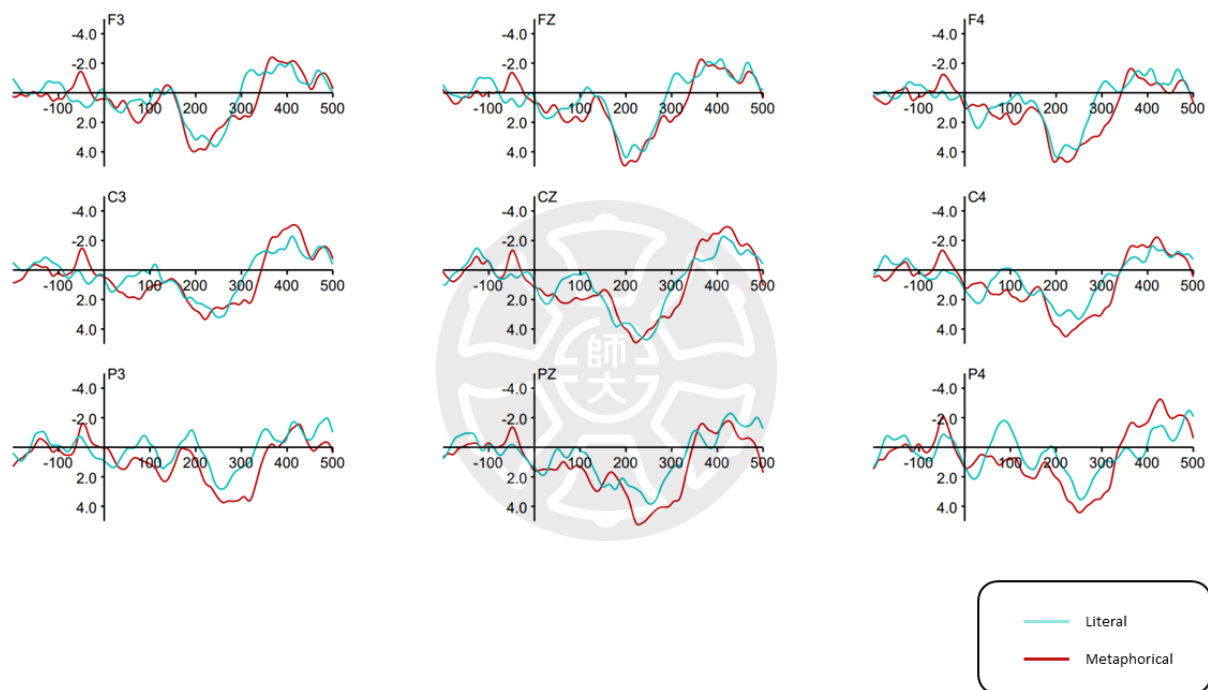
Figure 4.10. English group topographic maps showing the contrast between metaphorical and literal conditions at the main verb in the 350–500 ms time window.

Main Verb, Metaphorical, 350 – 500 ms   Main Verb, Literal, 350 – 500 ms
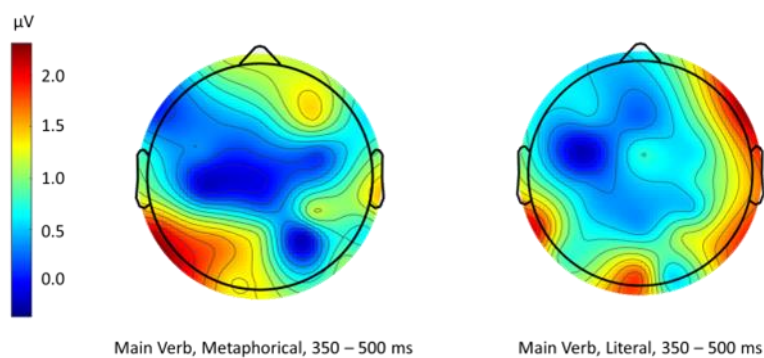
As with the Mandarin group brainwaves (in Figure 4.3.), the N1 and P2 visual processing components can be observed fairly clearly over the frontal and central electrode sites in Figure 4.9. Due to the smaller number of participants factored into the grand average, the brainwaves for both conditions appear to exhibit some influence of alpha waves which have not been fully reduced through the averaging process.[5] Taking the alpha wave influence into consideration, the brainwaves for the literal and metaphorical conditions do not appear to significantly deviate from one another. This observation is supported by the topographic maps in Figure 4.10., which do not show very different voltage distributions.

The ERP data for the English group brainwaves at the target sentence main verb were submitted to a three-way repeated measures ANOVA, summarised in Table 4.10. The analysis did not find any significant main effect of metaphoricity ($F(1, 5) = 0.070$, $p = .801$) nor any significant interactions between metaphoricity and hemisphere ($F(2, 10) = 0.466$, $p = .641$) or anteriority ($F(2, 26) = 0.520$, $p = .610$) or both hemisphere and anteriority ($F(4, 20) = 1.413$, p $= .266$).

Table 4.10. English group verb ERP analysis: Summary of the degrees of freedom, F-values, and p-values of a repeated measures ANOVA for the 350–500 ms time window. Significance codes: *** p < .001, ** p < .01, * p < .05

Variables	350–500 ms		
	df	F-value	p-value
Main Effects			
Metaphoricity	1, 5	0.070	0.801
Hemisphere	2, 10	0.577	0.579
Anteriority	2, 10	0.635	0.550
Two-way Interactions			
Metaphoricity × Hemisphere	2, 10	0.466	0.641
Metaphoricity × Anteriority	2, 10	0.520	0.610
Hemisphere × Anteriority	4, 20	1.708	0.188
Three-way Interactions			
Metaphoricity × Hemisphere × Anteriority	4, 20	1.413	0.266

---

[5] Alpha waves are EEG oscillations at 10 Hz which occur most prominently at posterior sites, usually when participants are tired, though some individuals exhibit prominent alpha waves even when fully rested and alert (Luck, 2014)

4.3.2. Object noun event

The English group ERPs for the second event of interest, the object noun, are displayed in Figure 4.11.

Figure 4.11. English group brainwaves at the object noun with the literal condition plotted in cyan and metaphorical condition in red.
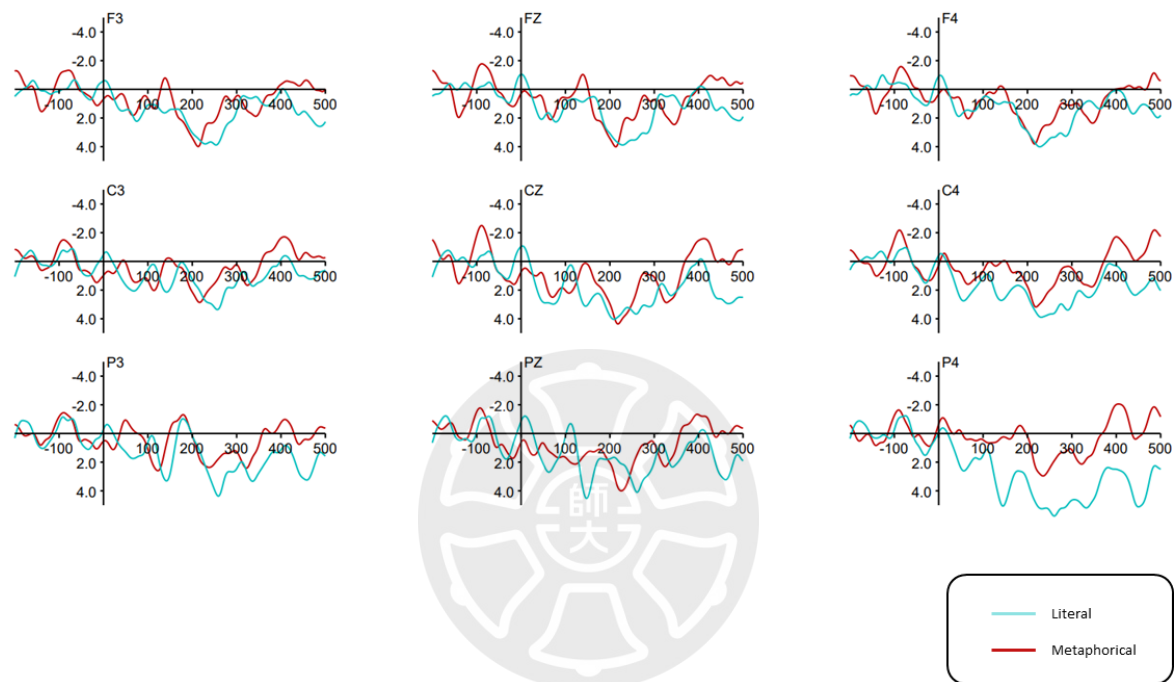
The brainwaves for both metaphorical and literal conditions appear to align well in the pre-stimulus timeframe (-200 ms to onset). The N1 and P2 components can be seen in the frontal and central sites. As before, the influence of alpha waves features quite strongly in the averaged ERPs, especially in the parietal electrode sites. Over frontal and central sites, the brainwaves for the metaphorical and literal conditions appear to part around 350 ms or 400 ms post-stimulus, with the metaphorical condition wave more negative-going relative to the literal brainwave. At the P4 site, the two conditional brainwaves appear to split right from the onset, with the literal condition deflecting positively throughout the whole 500 ms epoch relative to

the metaphorical condition. This positive deflection is captured in the topographic scalp map

for the literal condition (right) in Figure 4.12.

Figure 4.12. English group topographic maps showing the contrast between metaphorical and literal conditions at the object noun in the 350–500 ms time window.

Object Noun, Metaphorical, 350 – 500 ms   Object Noun, Literal, 350 – 500 ms
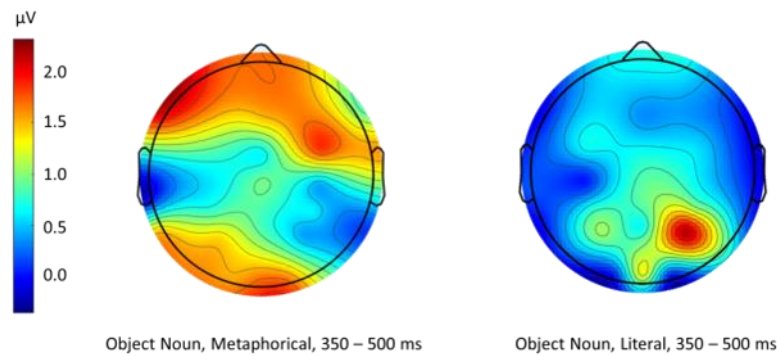
The ERP data for the English group at the object noun was submitted to a three-way

repeated measures ANOVA, summarised in Table 4.11.

Table 4.11. English group object noun ERP analysis: Summary of the degrees of freedom, F-values, and p-values of a repeated measures ANOVA for the 350–500 ms time window. Significance codes: *** $p < .001$, ** $p < .01$, * $p < .05$

Variables	*350–500 ms*		
	df	*F-value*	*p-value*
Main Effects			
Metaphoricity	1, 5	6.915	0.047*
Hemisphere	2, 10	0.255	0.781
Anteriority	2, 10	1.937	0.195
Two-way Interactions			
Metaphoricity × Hemisphere	2, 10	0.742	0.501
Metaphoricity × Anteriority	2, 10	1.511	0.267
Hemisphere × Anteriority	4, 20	1.648	0.201
Three-way Interactions			
Metaphoricity × Hemisphere × Anteriority	4, 20	0.875	0.496

The ANOVA found a significant main effect of metaphoricity ($F(1, 5) = 6.915$, $p < .05$) and no significant interactions with hemisphere or anteriority. To further investigate the main metaphoricity effect for the English group at the object noun, post-hoc pairwise comparisons were conducted, summarised in Table 4.12. Significant contrasts between the literal and metaphorical conditions were found at the right-central ($t = 2.628$, $p < .05$), left-posterior ($t = 2.757$, $p < .05$), and right-posterior ($t = 3.322$, $p < .05$) scalp regions.

Table 4.12. English group object noun ERP analysis: Summary of post-hoc pairwise comparisons of the literal – metaphorical contrast for the 350–500 ms time window. Significance codes: *** $p < .001$, ** $p < .01$, * $p < .05$

Hemisphere	Anteriority	df	t-ratio	p-value
Left	Anterior	5	1.177	0.292
Midline	Anterior	5	0.936	0.392
Right	Anterior	5	0.772	0.475
Left	Central	5	1.005	0.361
Midline	Central	5	1.928	0.112
Right	Central	5	2.628	0.047*
Left	Posterior	5	2.757	0.040*
Midline	Posterior	5	1.808	0.130
Right	Posterior	5	3.322	0.021*

Finally, the English group ERPs for the third event of interest, the sentence-final word of the target sentence, are shown in Figure 4.13.

Figure 4.13. English group brainwaves at the sentence-final word with the literal condition plotted in cyan and metaphorical condition in red.
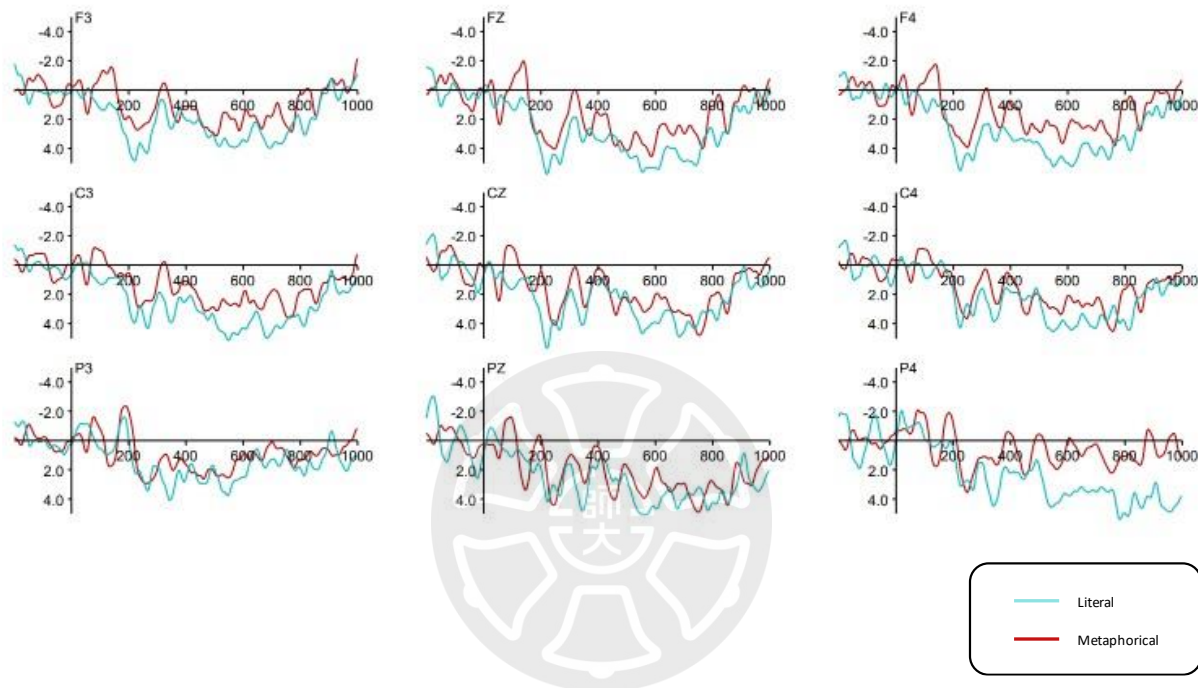

The brainwaves for the two conditions are mostly unified in the pre-stimulus timeframe (-200 ms to onset) and N1 and P2 components are visible. Alpha wave contamination is most severe at the parietal sites. For most electrode sites except P4, there does not seem to be any significant split between the metaphorical condition and literal condition brainwaves in the 350–500 ms time window. However, at the P4 electrode, the waveform for the metaphorical condition seems to experience a negative-going deflection relative to the literal condition wave. This is reflected in the topographical scalp maps plotted in Figure 4.14 (compare third and fourth maps from the left).

Figure 4.14. English group topographic maps showing the contrast between metaphorical and literal conditions at the sentence-final word in the 350–500 ms time window and 500–800 ms time window.

Sentence-Final Word, Metaphorical, 350 – 500 ms   Sentence-Final Word, Literal, 350 – 500 ms   Sentence-Final Word, Metaphorical, 500 – 800 ms   Sentence-Final Word, Literal, 500 – 800 ms
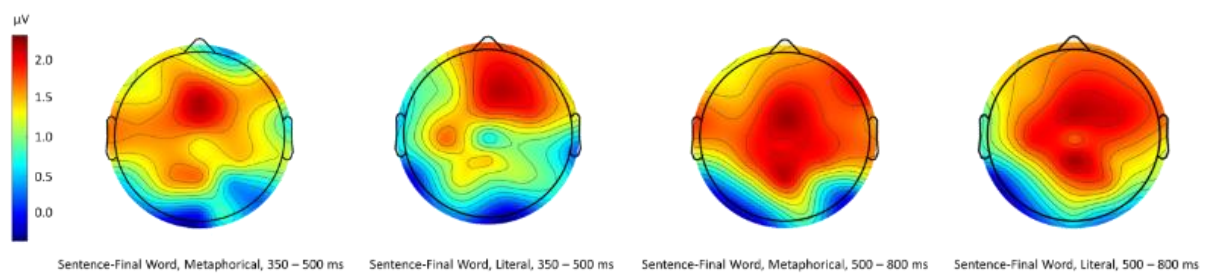
The ERP data of the English group at the sentence-final word were submitted to three-way repeated measures ANOVA for two time windows: 350–500 ms and 500–800 ms. In the 350–500 ms time window, the analysis did not detect any significant main effect of metaphoricity ($F(1, 5) = 1.831$, $p = .234$) nor any significant interactions with hemisphere or anteriority. The ANOVA results for the 350–500 ms window are summarised in Table 4.13.

Table 4.13. English group sentence-final word ERP analysis: Summary of the degrees of freedom, F-values, and p-values of a repeated measures ANOVA for the 350–500 ms time window.

Variables	350–500 ms		
	df	F-value	p-value
Main Effects			
Metaphoricity	1, 5	1.831	0.234
Hemisphere	2, 10	0.270	0.769
Anteriority	2, 10	0.706	0.517
Two-way Interactions			
Metaphoricity × Hemisphere	2, 10	0.826	0.466
Metaphoricity × Anteriority	2, 10	0.231	0.798
Hemisphere × Anteriority	4, 20	2.763	0.056
Three-way Interactions			
Metaphoricity × Hemisphere × Anteriority	4, 20	1.056	0.403

A second ANOVA was conducted for the 500–800 ms time window of the sentence-final word in the English group ERP data, summarised in Table 4.14. A significant main effect of metaphoricity was found ($F(1, 5) = 26.693$, $p < .05$) and no significant interactions between metaphoricity and hemisphere or anteriority.

Table 4.14. English group sentence-final word ERP analysis: Summary of the degrees of freedom, F-values, and p-values of a repeated measures ANOVA for the 500–800 ms time window.

Variables	500–800 ms		
	df	F-value	p-value
Main Effects			
Metaphoricity	1, 5	26.693	0.004*
Hemisphere	2, 10	2.975	0.097
Anteriority	2, 10	2.479	0.134
Two-way Interactions			
Metaphoricity × Hemisphere	2, 10	0.701	0.519
Metaphoricity × Anteriority	2, 10	0.158	0.856
Hemisphere × Anteriority	4, 20	2.456	0.079
Three-way Interactions			
Metaphoricity × Hemisphere × Anteriority	4, 20	2.248	0.100

To further investigate the significant main effect of metaphoricity in the 500–800 ms time window in the English group ERPs, post-hoc pairwise comparisons were conducted. The pairwise comparisons revealed significant contrasts between the two conditions at the right-anterior ($t = 3.492$, $p < .05$) and right-posterior ($t = 2.838$, $p < .05$) scalp regions, which aligns with the visual inspection of the P4 waveforms and topographical scalp maps. The post-hoc pairwise comparisons are summarised in Table 4.15.

Table 4.15. English group sentence-final word ERP analysis: Summary of post-hoc pairwise comparisons of the literal – metaphorical contrast for the 500–800 ms time window.

Significance codes: *** $p < .001$, ** $p < .01$, * $p < .05$

Hemisphere	Anteriority	df	t-ratio	p-value
Left	Anterior	5	2.054	0.095
Midline	Anterior	5	2.094	0.090
Right	Anterior	5	3.492	0.017*
Left	Central	5	2.112	0.088
Midline	Central	5	0.975	0.374
Right	Central	5	1.663	0.157
Left	Posterior	5	0.713	0.506
Midline	Posterior	5	1.657	0.158
Right	Posterior	5	2.838	0.036*

Chapter 5   Discussion

5.1. Interpretations of ERP results

To recap once more, the research question of this study is:

- Do L2 speakers of English process metaphorical language differently from native speakers of English?

It was predicted in Section 3.6. that there would be two possible scenarios for the 350–500 ms time window, particularly for the English group. Specifically, following results such as Coulson and Van Petten (2002) and Lai et al. (2009), an N400 effect should be observed for the metaphorical condition relative to the literal condition. Bilingual study results from Chen et al. (2013), Jankowiak et al. (2017) and Tang et al. (2022) additionally suggest that such an N400 should be observed for non-native speakers such as the Mandarin group. However, if context is expected to play a significant role in metaphor processing, as suggested by Pynte et al. (1996) and Bambini et al. (2016), then little to no N400 effect should be observed. In the present experiment, all target sentences were preceded by discursive context.

In the Mandarin group, the metaphorical condition at the target sentence verb induced an N400 in the frontal-left region (Figure 4.3), which was likely detected by the corresponding ANOVA as a significant three-way interaction between metaphoricity, hemisphere, and anteriority. This did not seem to emerge in the ERPs for the English group for the 350–500 ms time window at the verb (Figure 4.9). While the English group ERPs displayed alpha waves which were not fully reduced through the averaging process due to the relatively small number of participants, the brainwaves for the metaphorical and literal conditions did not appear to significantly deviate from one another throughout the epoch.

Since participants have only seen the subject and main verb of the target sentence at this point, it is possible that the non-native speakers may have experienced surprisal at some of the verbs in the metaphorical condition. For example, one of the trials contained the sequence *the interviewer grilled* and another had *the audience devoured*. In these cases, the subject noun phrases and preceding context give some indication that the main verb may be meant in a metaphorical sense. While the context sentences were not designed to allow prediction if a verb is used metaphorically or not prior to seeing the verb, they do reduce the probability that any of the metaphorical verbs are interpreted literally by native speakers. Native speakers of English are likely to know from convention that *grilled* in the context of an interview and *devoured* in the context of listening to a speech are meant metaphorically rather than literally, but these conventions may be less familiar to non-native speakers. To use the terminology of Hanks (2004; 2006; 2013; see Section 2.2.), the metaphorical verb choice may be perceived as a lexical norm by native English speakers ("direct access"; see Section 2.3.1.) and a lexical exploitation by the non-native speakers ("indirect access").

By the time the object head noun is perceived and processed, the effect of metaphoricity on the ERPs in the 350–500 ms window had disappeared for the Mandarin group (Figure 4.5.). At central and parietal sites, the brainwaves for both conditions were closely aligned with each other. At frontal sites, some slight negativity is seen in the metaphorical condition waveform relative to the literal condition, but no statistically significant effects or interactions involving metaphoricity were captured by the ANOVA (Table 4.5.). The results suggest that by the time of the object noun, any surprisal that had been experienced at the main verb had been resolved. Interestingly, this implies that the contrast between metaphorical and literal verbs had little effect on how the object noun is processed by the Mandarin group.

In contradistinction, some obvious divergences between the two conditions can be seen in the corresponding time window for the English group (Figure 4.11.). For the English group

ERPs at the object noun, the waveforms for the two conditions seem to split around 350 ms or 400 ms post-onset, with the metaphorical wave more negative. It is possible that this is an N400, and pairwise comparisons had furthermore indicated statistically significant conditional contrasts at the right-central, right-posterior, and left-posterior regions (Table 4.12). Many prior ERP studies on metaphor analysed epochs time-locked to the object noun, which also happened to be the sentence-final word, and observed N400 effects at that point (e.g., Pynte et al., 1996; Chen et al., 2013; Bambini et al., 2016; Wang, 2018; Tang et al., 2022), therefore lending some support to a possible N400. Since the object noun does not violate any English lexical norms, it is possible that the N400 here may index processes related to conceptual mappings evoked by metaphors. Such mappings may involve access or retrieval from semantic memory, as in "pre-activation" account of the N400 (Lau et al., 2016; 2008; Kutas & Federmeier, 2000). Since no N400 was found for the object noun in the Mandarin group, it is unlikely that the N400 here indexes semantic comprehension difficulties, as suggested by Coulson and Van Petten (2002).

However, there is also reason to doubt that the effect seen in the English group ERPs at the object noun is truly an N400. In those previous studies, the metaphorical contrast also tended to be dependent on the object noun, rather than the main verb as in the current experiment. Thus, it is challenging to explain why a metaphor-related N400 would emerge on the object noun when the contrast between metaphorical and literal meanings is on the main verb. Additionally, due to the low participant count for the English group, alpha waves had survived the averaging process into the grand average ERPs, and in fact seem even more pronounced at the object noun than in the English group ERPs for the main verb. Therefore, it is possible that the negative-going deflection in the metaphorical condition waveforms at the noun is simply artefactual. Reanalysis with a larger participant pool in the English group is required to confirm this.

At the sentence-final word, the ERPs for the Mandarin group display statistically significant interactions between metaphoricity and hemisphere in both the 350–500 ms and 500–800 ms time windows (Table 4.6. and Table 4.8.). Visual inspection of the ERP graphs (Figure 4.7.) suggests that the brainwaves for the two conditions do not clearly depart from one another in those time windows over the left and midline sites. However, the metaphorical condition waveform appears to be more negative-going than the literal condition waveform at the right hemisphere sites, especially around 560 ms post-onset, which may underlie the significant metaphoricity × hemisphere interactions detected by the ANOVAs.

In previous studies, Wang (2018) and Jankowiak et al. (2021) also observed negativities in late time windows at the sentence-final word. For Wang (2018), a late negativity emerging at around 600 ms and peaking around 780 ms could be seen in their ERP graphs, which they interpreted as an index of secondary semantic integration processes. However, their graphs contrasted waveforms for low and high proficiency L2 speakers of English instead of metaphorical and literal conditions, which reduces the comparability of their results to the present observations. For Jankowiak et al. (2021), novel nominal metaphors elicited an N400 relative to literal sentences in the 350–450 ms time window, and this negativity was then sustained throughout the rest of their 900 ms epoch. The researchers interpreted the sustained negativity as an indication that novel metaphor comprehension is "cognitively taxing". If this reasoning is borrowed into the present study, it implies that the non-native speakers found that processing a metaphorical sentence required more time for semantic processing than a literal sentence.

However, in the context of the present study, this reasoning seems to be undermined by the fact that no effect of metaphoricity was found at the object noun in the Mandarin group ERPs. Additionally, Jankowiak et al.'s reasoning can only be successfully borrowed if it is assumed that the Mandarin group interpreted the metaphorical stimuli, which were designed to

be conventional in English, as novel and unfamiliar. While it was argued earlier that the N400 seen at the verb in the Mandarin group ERPs may have emerged due to unfamiliarity and surprisal at the metaphorical lexical choice, a crucial factor to consider is that by the time the sentence-final word is reached, much time had already passed to resolve any reactions to the unfamiliar metaphorical meaning, especially since all stimuli ended in a functionally peripheral adjunct or clause between 1 and 7 words in length following the object noun (including the sentence-final word itself; see Section 3.2.). This is unlike Jankowiak et al.'s design, which used nominal metaphors wherein the literal – metaphorical contrast is dependent on the sentence-final word (e.g., *Love is a <u>monastery</u>*). Therefore, it is likely that the late negativity observed at the sentence-final word in the current study's Mandarin group ERPs is qualitatively different from that observed by Jankowiak et al. (2021).

Interestingly, visual inspection of the sentence-final word ERPs for the English group appears to indicate slightly more pronounced negativity for the metaphorical condition in the 500–800 ms time window than in the Mandarin group (Figure 4.13). This is supported by the ANOVA and post-hoc results, which detected a main effect of metaphoricity and significant conditional contrasts at the right-anterior and right-posterior regions—the late negativity in the Mandarin group ERPs also seemed most prominent in the right hemisphere. The observation of late negativity in both the English and Mandarin groups suggest that it indexes a process independent of native language.

With this in mind, it is possible that the sentence-final word may have elicited post-reading memory or comprehension processes, possibly in anticipation of a potential upcoming comprehension question. The late negativity observed for both Mandarin and English groups bears some similarities to the Late Posterior Negativity (LPN) component seen in episodic memory research on retrieval (Johansson & Mecklinger, 2003; Wolk et al., 2007). According to Johansson and Mecklinger (2003), the LPN is an ERP component that onsets before or at

around the time when participants respond to a cue for memory retrieval. They describe its length as lasting several hundred milliseconds and its distribution as parietal and bilateral, centred around the PZ electrode site. Similarly, the late negativity observed here is at a point shortly before participants have to respond to a comprehension question, which might involve similar cognitive processes to episodic memory retrieval, though the distribution seems somewhat biased towards the right hemisphere for the present results. If the late negativity is related to the LPN, then it implies that the retrieval process for metaphorical sentences, in the context of preparing to respond to a comprehension check, differed from that for literal sentences in participants in both groups.

In conclusion, the results reported in this thesis, based on preliminary data and analysis, suggested that significant differences may hold in how native speakers of English and non-native speakers process metaphorical language use in English. Unlike previous studies, the present experiment manipulated metaphoricity at the main verb (see Section 3.2). Qualitative comparisons between the ERPs of the Mandarin and English group suggested that metaphor-related N400 effects may have emerged at the main verb for the Mandarin group and object noun for the English group. However, some caveats should be noted. First, the distribution of the possible N400 at the verb for the Mandarin group seemed isolated to the frontal-left region, which is different from the typical right-biased central-parietal distribution of the N400 component. Secondly, the ERPs for the English group contained the influence of alpha waves due to the small participant pool, which may have been a confounding factor. The potential confounding role of alpha waves entail that interpretations of the English group results need to be taken cautiously. Reanalysis with a larger participant pool is needed to support the interpretations regarding the English group, should the observations reported here continue to hold. Finally, the observed late negativity for the metaphorical condition in the results of both

language groups at the sentence-final word suggests a possible relationship between metaphoricity and episodic memory retrieval which could be investigated in a future study.

5.2. Limitations and Future Work

There are some limitations to the present thesis. The first is that the sample size of the English group is both small and unequal to the size of the Mandarin group. This weakness has several implications. For example, as has already been mentioned, the small sample size in the English group influences the robustness of the averaging process in deriving ERPs from EEG data. Visual inspection of the ERP grand average plots for the English group indicate that alpha waves were not fully reduced by the averaging process, introducing a confounding factor. As a result, any significance detected by statistical analyses performed such as ANOVAs and pairwise comparisons may not be reliable. Interpretations based on visual inspection of plots and statistical tests need to be checked against future analyses using a larger sample size. The small sample size of the English group also means that it is unequal with that of the Mandarin group. As a result, between-groups differences were not quantitatively investigated using inferential statistical tests. Therefore, this study only investigated between-group differences based on qualitative comparisons between ERP graphs, topographical maps, and within-group statistics. This limitation will be addressed by re-running analyses on a larger participant pool for the English group.

A second important limitation exists regarding the design of the experiment. Whilst previous studies reported a biphasic N400-P600 effect for metaphors, the 500–800 ms time window, when the P600 would typically appear, could not be investigated at the main verb and object noun. The reason is that the interstimulus intervals between the onsets of words in the target sentence are too short, since each word is presented for 300 ms and then followed by a blank of only 200 ms in length. A future version of this experiment with longer interstimulus

intervals could be considered to allow analysis in the 500–800 ms time window. A longer interstimulus interval would also mitigate occurrences where the onset of a critical word occurs while a previous ERP component is still active, as may have happened at the object noun in the Mandarin group ERPs (see Section 4.2.2.). Another approach to this problem might be to re-assign the event codes at the object noun to new event codes distinguishing different types of preceding stimuli (e.g., noun preceded immediately by verb, noun preceded immediately by adjective, and noun preceded immediately by determiner). Furthermore, more experimental materials could be designed to create an experiment with more trials per condition and the materials could be better controlled for the predictability of metaphors, imageability/concreteness of verb meanings, collocational or collostructional strengths and frequencies, and syntactic structure variability (cf. Trevisan & García, 2019). A norming study similar to the one reported in Section 3.2. could also be conducted with non-native speakers instead to control for the difficulty and naturalness of the materials from their perspective. Lastly, validated L2 English vocabulary difficulty lists may also be consulted to construct materials.

Two interesting lines of research may be pursued following the results of this study. One follow-up study might involve investigating the possible role of intercultural competence in language learning, focusing further on analysing the Mandarin group and building on themes on language and culture developed in Chapter 1 and Sections 2.1. and 2.2. As mentioned in Chapter 3, participants were administered the CQS (Cultural Intelligence Scale) questionnaire to assess their intercultural competence. The preliminary results of this questionnaire were presented and statistically analysed in Section 4.1.2. Participant data from the Mandarin group can be reorganised according to the results of the CQS questionnaire, adding intercultural competence as a new independent variable for statistical analysis alongside L2 proficiency. More specifically, multiple regression analysis could be conducted in R (R Core Team, 2022)

to compare whether intercultural competence or L2 proficiency serve as better predictors of the possible N400 observed in the 350–500 ms time window at the main verb. A second interesting line of research might involve investigating the relationship between metaphoricity and memory retrieval. In the present study, comprehension questions were included in the experiment as a form of attention check. A future design might manipulate the comprehension questions to probe for different aspects of the target sentence's meaning, such as the event described or the participants semantically involved in the event. Analysis could be conducted to explore the relationship between metaphorically or literally expressed events and memory retrieval. Such a future study would help inform whether the late negativity observed presently at the sentence-final word is indeed related to memory retrieval processes.

Chapter 6   Conclusion

This thesis aimed to explore the neurocognitive processing of metaphorical language in a bilingual context using the event-related potential (ERP) methodology. Specifically, it sought to investigate if L2 speakers of English process metaphorical language differently from native speakers of English. The preliminary findings suggested notable differences in how native English speakers and Mandarin-English bilinguals process metaphorical language.

The experiment materials consisted of context-target sentence pairs which could be either metaphorical or literal depending on the verb. For the Mandarin group, an N400 was observed at the main verb in the metaphorical condition, possibly indicating processing difficulties due to unfamiliarity or surprisal at English metaphor norms. Conversely, for the English group, no N400 emerged at the main verb, but a possible N400 was observed at the object noun, potentially indicating semantic retrieval processes related to conceptual mappings. However, the small sample size and the presence of alpha waves in the English group's ERPs make the N400 observation unreliable and necessitate reanalysis with a larger sample. The results also indicated a late negativity for both groups in the 500–800 ms time window at the sentence-final word, possibly related to the LNP component in episodic memory research.

In sum, this thesis contributes to understanding how bilinguals process metaphors in contexts resembling natural language use. Unlike previous studies, innovations in the experiment design presented here allowed for the analysis of metaphor processing at different points in the sentence comprehension process. Thus, this thesis highlights the importance of designing naturalistic materials in order to maximise generalisability of experimental results. It also encourages considering language in terms of "ways of meaning". We learn new ways of meaning when we learn a new languaculture, and conventional metaphors occupy a key aspect of languacultural competence. Ultimately, this work underscores the need for a nuanced approach to studying metaphor usage that incorporates cultural and contextual factors.

References

Agar, M. (1994). *Language shock: The culture of conversation*. New York: HarperCollins.

Ang, S., Van Dyne, L., Koh, C., Ng, K.Y., Templer, K.J., Tay, C., and Chandrasekar, N.A.
(2007). Cultural intelligence: Its measurement and effects on cultural judgement and
decision making, cultural adaptation and task performance. *Management and
Organization Review*, *3*(3), pp. 335–371. http://doi.org/10.1111/j.l 740-
8784.2007.00082.x

Ansah, G. (2011). *Metaphor and bilingual cognition: The case of Akan and English in Ghana*
[Unpublished PhD thesis]. Lancaster University.

Arzouan, Y., Goldstein, A., and Faust, M. (2007). Brainwaves are stethoscopes: ERP
correlates of novel metaphor comprehension. *Brain Research*, *1160*, pp. 69–81.
https://doi.org/10.1016/j.brainres.2007.05.034

Athanasopoulos, P. (2015). Conceptual representation in bilinguals: The role of language
specificity and conceptual change. In J.W. Schwieter (Eds.), *The Cambridge
handbook of bilingual processing*, pp. 275–292. Cambridge, UK: Cambridge
University Press. https://doi.org/10.1017/cbo9781107447257.012

Baggio, G., and Hagoort, P. (2011). The balance between memory and unification in
semantics: A dynamic account of the N400. *Language and Cognitive Processes*,
26(9), pp. 1338–1367. http://dx.doi.org/10.1080/01690965.2010.542671

Bambini, V., Bertini, C., Schaeken, W., Stella, A., and Di Russo, F. (2016). Disentangling
metaphor from context: An ERP study. *Frontiers in Psychology*, *7*(559).
https://doi.org/10.3389/fpsyg.2016.00559

Barsalou, L.W. (1992). Frames, concepts, and conceptual fields. In A. Lehrer and E.F. Kittay
(Eds.), *Frames, fields, and contrasts: New essays in semantic and lexical
organization*, pp. 21–74. Hillsdale, NJ: Lawrence Erlbaum Associates.

Brouwer, H., Fitz, H., and Hoeks, J. (2012). Getting real about semantic illusions: Rethinking the functional role of the P600 in language comprehension. *Brain Research*, *1446*, pp. 127–143. https://doi.org/10.1016/j.brainres.2012.01.055

Brown, C., and Hagoort, P. (1993). The processing nature of the N400: Evidence from masked priming. *Journal of Cognitive Neuroscience*, *5*(1), pp. 34–44. https://doi.org/10.1162/jocn.1993.5.1.34

Cameron, L. (1997). *Metaphorical use of language in educational discourse: A theoretical and empirical investigation* [Unpublished PhD thesis]. University of London, Institute of Education. Retrieved November 20, 2023, from: https://discovery.ucl.ac.uk/id/eprint/10020284/

Cameron, L. (1999). Operationalising 'metaphor' for applied linguistic research. In L. Cameron and G. Low (Eds.), *Researching and applying metaphor*, pp. 3–28. Cambridge, UK: Cambridge University Press.

Chapetón, C.M. (2010). Metaphor identification in EFL argumentative writing: A corpus-driven study. *Folios*, *2*(2), pp. 125–140. Retrieved November 20, 2023, from: https://www.researchgate.net/publication/262427021_Metaphor_Identification_in_EFL_Argumentative_Writing_A_Corpus-driven_Study

Chen, H., Peng, X., and Zhao, Y. (2013). An ERP study on metaphor comprehension in the bilingual brain. *Chinese Journal of Applied Linguistics*, *36*(4), pp. 505–519. https://doi.org/10.1515/cjal-2013-0034

Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use*. New York, NY: Praeger.

Coulson, S. (2008). Metaphor comprehension and the brain. In R.W. Gibbs (ed.), *The Cambridge handbook of metaphor and thought* (pp. 177-196). Cambridge: Cambridge University Press. Retrieved December 23, 2022, from: https://pages.ucsd.edu/~scoulson/Papers/coulson-handmet.pdf9

Coulson, S., and Van Petten, C. (2007). A special role for the right hemisphere in metaphor comprehension?: ERP evidence from hemifield presentation. *Brain Research*, *1146*, pp. 128–145. https://doi.org/10.1016/j.brainres.2007.03.008

Coulson, S., and Van Petten, C. (2002). Conceptual integration and metaphor: An event-related potential study. *Memory & Cognition*, *30*(6), pp. 958–968. https://doi.org/10.3758/bf03195780

De Grauwe, S., Swain, A, Holcomb, P.J., Ditman, T., Kuperberg, G. (2010). Electrophysiological insights into the processing of nominal metaphors. *Neuropsychologia*, *48*(7), pp. 1965–1984. https://doi.org/10.1016/j.neuropsychologia.2010.03.017

Deignan, A. (2008). Corpus linguistic data and conceptual metaphor theory. In M.S. Zanotto, L. Cameron, and M.C. Cavalcanti (Eds.), *Confronting metaphor in use: An applied linguistic approach*, pp. 149–162. Amsterdam/Philadelphia, PA: John Benjamins.

Delorme, A., and Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG. *Journal of Neuroscience Methods*, *134*, pp. 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

de Saussure, F. (1959). *Course in general linguistics* (W. Baskin, Trans.). New York: Philosophical Library. (Original work published 1916)

Earley, C., and Ang, S. (2003). *Cultural intelligence: Individual interactions across cultures*. Stanford, CA: Stanford University Press.

Fawcett, R.P. (2000). *A theory of syntax for systemic functional linguistics*.

Amsterdam/Philadelphia, PA: Johns Benjamins.

Fawcett, R.P. (2008). *Invitation to systemic functional linguistics trough the Cardiff*

grammar: An extension and simplification of Halliday's systemic functional grammar.

London: Equinox.

Feldman, J.A. (2006). *From molecule to metaphor: A neural theory of language*. Cambridge,

MA: The MIT Press.

Fife, D. (2020). *Flexplot: Graphically-based data analysis*. PsyArXiv.

https://doi.org/10.31234/osf.io/kh9c3

Fillmore, C.J. (1985). Frames and the semantics of understanding. *Quaderni di Semantica*, *6*,

pp. 222–255.

Firth, J. (1968). A Synopsis of Linguistic Theory, 1930-55. In F. R. Palmer (Ed.), *Selected*

Papers of J. R. Firth (1952–59) (pp. 168–205). London: Longmans. (Original work

published 1957)

García, O., Cieślicka, A.B., and Heredia, R.R. (2015). Nonliteral language processing and

methodological considerations. In R.R. Heredia and A.B. Cieślicka (Eds.)., *Bilingual*

figurative language processing, pp. 117–168. Cambridge, UK: Cambridge University

Press.

Gentner, D., and Wolff, P. (1997). Alignment in the processing of metaphor. *Journal of*

Memory and Language, *37*(3), pp. 331–355. https://doi.org/10.1006/jmla.1997.2527

Gibbs, R.W. (1984). Literal meaning and psychological theory. *Cognitive Science*, *8*(3), pp.

274–304. https://doi.org/10.1016/S0364-0213(84)80004-X

Giora, R. (2002). Literal vs. figurative language: Different or equal? *Journal of Pragmatics*,

34(4), pp. 487–506. https://doi.org/10.1016/S0378-2166(01)00045-5

Glucksberg, S. (2003). The psycholinguistics of metaphor. *Trends in Cognitive Sciences*, *7*(2), pp. 92–96. https://doi.org/10.1016/S1364-6613(02)00040-2

Grice, H.P. (1975). Logic and conversation. In P. Cole and J.L. Morgan (Eds.), *Syntax and semantics* (vol. 3; pp. 41–58). New York: Academic Press.

Hagoort, P. (2013). MUC (memory, unification, control) and beyond. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00416

Hagoort, P., Brown, C., and Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes*, *8*(4), pp. 439–483. https://doi.org/10.1080/01690969308407585

Halliday, M.A.K., and Hasan, R. (1989). *Language, context, and text: Aspects of language in a social-semiotic perspective* (2nd ed.). Oxford: Oxford University Press.

Halliday, M.A.K., and Matthiessen, C.M.I.M. (1999). *Construing experience through meaning: A language-based approach to cognition*. Continuum Press.

Hanks, P. (2004). The syntagmatics of metaphor and idiom. *International Journal of Lexicography*, *17*(3), pp. 245–274. https://doi.org/10.1093/ijl/17.3.245

Hanks, P. (2006). Metaphoricity is gradable. In A. Stefanowitsch and S.Th. Gries (Eds.), *Corpus-based approaches to metaphor and metonymy*, pp. 17–35. Berlin: Mouton de Gruyter.

Hanks, P. (2013). *Lexical analysis: Norms and exploitations*. Cambridge, MA: The MIT Press.

Hasan, R. (2015). *Ways of saying, ways of meaning: Selected papers of Ruqaiya Hasan* (C. Cloran, D. Butt, and G. Williams, eds.). London: Bloomsbury Academic. (Original work published 1996)

INCA. (2004). *Assessor manual*. European Commission. Retrieved November 19, 2023,

from: https://ec.europa.eu/migrant-integration/library-document/inca-project-

intercultural-competence-assessment_en

Jankowiak, K., Rataj, K., and Naskręcki, R. (2017). To electrify bilingualism:

Electrophysiological insights into bilingual metaphor comprehension. PLoS One,

12(4). https://doi.org/10.1371/journal.pone.0175578

Jankowiak, K., Naranowicz, M., and Rataj, K. (2021). Metaphors are like lenses:

Electrophysiological correlates of novel meaning processing in bilingualism.

International Journal of Bilingualism, *25*(3), pp. 668–686.

https://doi.org/10.1177/1367006921996820

Janus, R.A., and Bever, T.G. (1985). Processing of metaphoric language: An investigation of

the three-stage model of metaphor comprehension. *Journal of Psycholinguistic

Research*, *14*(5), pp. 473–487. https://doi.org/10.1007/BF01666722

Johansson, M., and Mecklinger, A. (2003). The late posterior negativity in ERP studies of

episodic memory: Action monitoring and retrieval of attribute conjunctions.

Biological Psychology, *64*, pp. 91–117. https://doi.org/10.1016/S0301-

0511(03)00104-2

Jung, T.P., Makeig, S., Humphries, C., Lee, T.W., McKeown, M.J., Iragui, V., Sejnowski,

T.J. (2000). Removing electroencephalographic artifacts by blind source separation.

Psychophysiology, *37*(2), pp. 163–178. https://doi.org/10.1111/1469-8986.3720163

Keating, G.D., and Jegerski, J. (2015). Experimental designs in sentence processing research:

A methodological review and user's guide. *Studies in Second Language Acquisition*,

37, pp. 1–32. https://doi.org/10.1017/S0272263114000187

Kolers, P.A. (1963). Interlingual word associations. *Journal of Verbal Learning and Verbal

Behavior*, *2*(4), 291–300. https://doi.org/10.1016/S0022-5371(63)80097-3

Kövecses, Z. (2010). Metaphor and culture. *Acta Universitatis Sapientiae, Philologica*, *2*(2), pp. 197–220. http://doi.org/10.1093/acprof:oso/9780190224868.003.0005

Kövecses, Z. (2015). *Where metaphors come from: Reconsidering context in metaphor*. Oxford: Oxford University Press.

Kövecses, Z. (2020). *Extended conceptual metaphor theory*. Cambridge, UK: Cambridge University Press.

Kroll, J.F., & de Groot, A.M.B. (1997). Lexical and conceptual memory in the bilingual: Mapping form to meaning in two languages. In A.M.B. de Groot & J.F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic* perspectives (pp. 169–199). Lawrence Erlbaum Associates Publishers.

Kroll, J.F., and Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, *33*(2), pp. 149–174. https://doi.org/10.1006/jmla.1994.1008

Kuperberg. G.R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research*, *1146*, pp. 23–49. http://dx.doi.org/10.1016/j.brainres.2006.12.063

Kuperberg, G.R., Caplan, D., Sitnikova, T., Eddy, M., and Holcomb, P.J. (2003). Neural correlates of processing syntactic, semantic, and thematic relationships in sentences. *Language and Cognitive Processes*, *21*(5), pp. 489–530. https:/doi.org/10.1080/01690960500094279

Kutas, M., and Federmeier, K.D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*(12), pp. 463–470. https://doi.org/10.1016/S1364-6613(00)01560-6

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, *62*, 621–647. https://doi.org/10.1146/annurev.psych.093008.131123

Kutas, M., and Hillyard, S.A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), pp. 203–205. https://doi.org/10.1126/science.7350657

Lai, V.T., and Curran, T. (2013). ERP evidence for conceptual mappings and comparison processes during the comprehension of conventional and novel metaphors. *Brain & Language*, *127*(3), pp. 484–496. https://doi.org/10.1016/j.bandl.2013.09.010

Lai, V.T., Curran, T. & Menn, L. (2009). Comprehending conventional and novel metaphors: An ERP study. *Brain Research*, *1281*, pp. 145–155. https://doi.org/10.1016/j.brainres.2009.05.088

Lai, V.T., Howerton, O., and Desai, R.H. (2019). Concrete processing of action metaphors: Evidence from ERP. *Brain Research*, *1714*, pp. 202–209. https://doi.org/10.1016/j.brainres.2019.03.005

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago, IL: The University of Chicago Press.

Lakoff, G. and Johnson, M. (1980). *Metaphors we live by*. Chicago, IL: The University of Chicago Press.

Lamb, S.M. (1999). *Pathways of the brain: The neurocognitive basis of language*. Amsterdam: Johns Benjamins.

Lamb, S.M. (2004). *Language and reality* (J. Webster, Ed.). London/New York: Continuum.

Lau, E.F., Namyst, A., Fogel, A., and Delgado, T. (2016). A direct comparison of N400 effects of predictability and incongruity in adjective-noun combination. *Collabra*, *2*(1), pp. 1–19. https://doi.org/10.1525/collabra.40

Lau, E.F., Phillips, C., and Poeppel, D. (2008). A cortical network for semantics: (De)constructing the N400. *Nature Reviews Neuroscience*, *9*, pp. pages 920–933. https://doi.org/10.1038/nrn2532

Lin, Y.-c., Chen, A.S.-y., and Song, Y.-c. (2012). Does your intelligence help to survive in a foreign jungle? The effects of cultural intelligence and emotional intelligence on cross-cultural adjustment. *International Journal of Intercultural Relations*, *36*(4), pp. 541–552. https://doi.org/10.1016/j.ijintrel.2012.03.001

Link, P. (2013). *An anatomy of Chinese*. Cambridge, MA: Harvard University Press.

Liu, D. (2002). *Metaphor, culture, and worldview: The case of American English and the Chinese language*. Lanham, MD: University Press of America.

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

Low, D.S., Aung, M.T., and Day, M.J. (2020). Cognitive sociology: Developing the 'diversity pathways' model in cultural neuroscience. *Human Behavior, Development and Society*, *21*(4), pp. 66–77. https://so01.tci-thaijo.org/index.php/hbds/article/view/243179

Low, D.S., Mcneill, I., and Day, M.J. (2022). Endangered languages: A sociocognitive approach to language death, identity loss, and preservation in the age of artificial intelligence. *Sustainable Multilingualism*, *21*(1), pp. 1–25. http://doi.org/10.2478/sm-2022-0011

Luck, S.J. (2014). *An introduction to the event-related potential technique* (2nd ed.). Cambridge, MA: The MIT Press.

Nguyen-Phuong-Mai, M. (2017). *Intercultural communication: An interdisciplinary approach: When neurons, genes, and evolution joined the discourse*. Amsterdam: Amsterdam University Press.

Oldfield, R.C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, *9*(1), pp. 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Osterhout, L., and Holcomb, P.J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*(6), pp. 785–806. https://doi.org/10.1016/0749-596X(92)90039-Z

Pavlenko, A. (1999). New approaches to concepts in bilingual memory. *Bilingualism: Language and Cognition*, *2*(3), pp. 209–230. https://doi.org/10.1017/S1366728999000322

Polich, J. (2012). Neuropsychology of P300. In E.S. Kappenman and S.J. Luck (Eds.), *The Oxford handbook of event-related potential components*. Oxford: Oxford University Press.

Psychology Software Tools, Inc. (2016). *E-Prime 3.0*. Pittsburgh, PA: Psychology Software Tools, Inc. https://pstnet.com/products/e-prime/

Pynte, J., Besson, M., Robichon, F.-H., and Poli, J. (1996). The time-course of metaphor comprehension: An event-related potential study. *Brain and Language*, *55*(3), pp. 293–316. https://doi.org/10.1006/brln.1996.0107

R Core Team. (2022). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. http://www.r-project.org/

Rai, S., and Chakraverty, S. (2020). A survey on computational metaphor processing. *ACM Computing Surveys*, *53*(2). https://doi.org/10.1145/3373265

Sardinha, T.B. (2008). Metaphor probabilities in corpora. In M.S. Zanotto, L. Cameron, and

   M.C. Cavalcanti (Eds.), *Confronting metaphor in use: An applied linguistic approach*,

   pp. 127–148. Amsterdam/Philadelphia, PA: John Benjamins.

Searle, J.R. (1993). Metaphor. In A. Ortony (Ed.), *Metaphor and thought* (2nd ed.; pp. 83–

   111). Cambridge, UK: Cambridge University Press. (Original work published 1979)

Silaški, N., and Kilyeni, A. (2014). The money is solid metaphor in economic and business

   terminology in English. *Professional Communication and Translation Studies*, *7*(1),

   pp. 73–79. Retrieved January 10, 2024, from:

   https://www.researchgate.net/publication/264929828_THE_MONEY_IS_SOLID_M

   ETAPHOR_IN_ECONOMIC_AND_BUSINESS_TERMINOLOGY_IN_ENGLISH

Swaab, T.Y., Ledoux, K., Camblin, C.C., and Boudewyn, M.A. (2012). Language-related

   ERP components. In E.S. Kappenman and S.J. Luck (Eds.), *The Oxford handbook of

   event-related potential components*. Oxford: Oxford University Press.

Tang, X., Shen, L., Yang, P., Huang, Y., Huang, S., Huang, M., and Ren, W. (2022).

   Bilingual processing mechanisms of scientific metaphors and conventional

   metaphors: Evidence via a contrastive event-related potentials study. *Frontiers in

   Psychology*, *13*(894114). https://doi.org/10.3389/fpsyg.2022.894114

The MathWorks Inc. (2022). *MATLAB version: 9.13.0 (R2022b)*. Natick: The MathWorks

   Inc. https://www.mathworks.com

Thibodeau, P.H., Matlock, T., and Flusberg, S.J. (2019). The role of metaphor in

   communication and thought. *Language and Linguistics Compass*, *13*(5).

   https://doi.org/10.1111/lnc3.12327

Trevisan, P., and García, A.M. (2019). Systemic functional grammar as a tool for

   experimental stimulus design: New appliable horizons in psycholinguistics and

neurolinguistics. *Language Sciences*, *75*, pp. 35–46.

https://doi.org/10.1016/j.langsci.2019.101237

Uen, J.-F., Teng, S.-K., Wu, L-C., and Tsao, S.-A. (2018, May). *The antecedents and*

consequences of cultural intelligence: An exploratory study of Taiwanese expatriates

[Paper presentation]. 5th International Conference on Business and Industrial

Research (ICBIR), Bangkok, Thailand. Retrieved, November 18, 2023, from:

https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8391245

Van Assche, E., Brysbaert, M., and Duyck, W. (2020). Bilingual lexical access. In R.R.

Heredia and A.B. Cieślicka (Eds.), *Bilingual lexical ambiguity resolution*, pp. 42–68.

https://doi.org/10.1017/9781316535967.004

Van Dyne, L., Ang, S., and Koh, C. (2008). Development and validation of the CQS: The

cultural intelligence scale. In S. Ang and L. Van Dyne (Eds.), *Handbook f cultural*

intelligence: Theory, measurement, and applications, pp.16–38. Routledge.

Van Valin, R.D., and LaPolla, R.J. (1997). *Syntax, structure, meaning and function*.

Cambridge: Cambridge University Press.

Wang, L., Wang, K.T., Heppner, P.P., and Chuang, C.-C. (2017). Cross-national cultural

competency among Taiwanese international students. *Journal of Diversity in Higher*

Education, *10*(3), pp. 271–287. http://doi.org/10.1037/dhe0000020

Wang, Q. (2018). Neural mechanism and representation of English and Chinese metaphors of

bilinguals with different second language proficiency: An ERP study. *Chinese*

Journal of Applied Linguistics, *41*(1), pp. 67-83. https://doi.org/10.1515/cjal-2018-

0004

Waters, T. (2014). Of looking glasses, mirror neurons, culture, and meaning. *Perspectives on*

Science, *22*(4), pp. 616–649. https://doi.org/10.1162/POSC_a_00152

Wolk, D.A., Schacter, D.L., Lygizos, M., Mandu Sen, N., Chong, H., Holcomb, P.J., Daffner, K.R., and Budson, A.E. (2007). ERP correlates of remember/know decisions: Association with the late posterior negativity. *Biological Psychology*, *75*(2), pp. 131–135. https://doi.org/10.1016/j.biopsycho.2007.01.005

Yang, F.-P.G., Bradley, K., Huq, M., Wu, D.-L., and Krawczyk, D.C. (2013). Contextual effects on conceptual blending in metaphors: An event-related potential study. *Journal of Neurolinguistics*, *26*(2), pp. 312–326. https://doi.org/10.1016/j.jneuroling.2012.10.004

Yu, N. (2003). Chinese metaphors of thinking. *Cognitive Linguistics*, *14*(2-3). https://doi.org/10.1515/cogl.2003.006

APPENDIX A: Cultural Intelligence Scale Questionnaire

Section 1: Biographical Information (adapted from INCA, 2004)

Name:
Age:
Profession/Course of Study:
Company/University:
Citizenship:

How many friends from abroad do you have? Please give an estimate.
How many languages do you speak well?
How often have you dealt with people from other countries in your professional/school life?
Have you ever worked in a work group with members from various cultures?

How many times have you read books that are written in a foreign language?
☐ 0–2 times ☐ 3–5 times ☐ 6–10 times ☐ more than 10 times

How many times have you been abroad?
☐ 0–2 times ☐ 3–5 times ☐ 6–10 times ☐ more than 10 times

How many different countries have you visited already?
☐ 0–1 ☐ 2–3 ☐ 4–6 ☐ more than 7

Which countries have you been to?

How long did your longest stay abroad last?
☐ one or two days ☐ 2 days–1 week ☐ 1 week–1 month O 1–5 months
☐ more than half a year

Section 2: Self-Report

Read each statement and select the response that best describes your capabilities. Select the answer that BEST describes you AS YOU REALLY ARE.
(1 = strongly disagree; 5 = strongly agree)

Q1. I am conscious of the cultural knowledge I use when interacting with people with different cultural backgrounds. (metacognitive)

Q2. I adjust my cultural knowledge as I interact with people from a culture that is unfamiliar to me. (metacognitive)

Q3. I am conscious of the cultural knowledge I apply to cross-cultural interactions. (metacognitive)

Q4. I check the accuracy of my cultural knowledge as I interact with people from different cultures. (metacognitive)

Q5. I know the legal and economic systems of other cultures. (cognitive)

Q6. I know the rules (e.g., vocabulary, grammar) of other languages. (cognitive)

Q7. I know the cultural values and religious beliefs of other cultures. (cognitive)

Q8. I know the marriage system of other cultures. (cognitive)

Q9. I know the arts and crafts of other cultures. (cognitive)

Q10. I know the rules for expressing nonverbal behaviors in other cultures. (cognitive)

Q11. I enjoy interacting with people from other cultures. (motivational)

Q12. I am confident that I can socialize with locals in a culture that is unfamiliar to me. (motivational)

Q13. I am sure I can deal with the stresses of adjusting to a culture that is new to me. (motivational)

Q14. I enjoy living in cultures that are unfamiliar to me. (motivational)

Q15. I am confident that I can get accustomed to the shopping conditions in a different culture. (motivational)

Q16. I change my verbal behavior (e.g., accent, tone) when a cross-cultural interaction requires it. (behavioural)

Q17. I use pause and silence differently to suit different cross-cultural situations. (behavioural)

Q18. I vary the rate of my speaking when a cross-cultural situation requires it. (behavioural)

Q19. I change my nonverbal behavior when a cross-cultural situation requires it. (behavioural)

Q20. I alter my facial expressions when a cross-cultural interaction requires it. (behavioural)

APPENDIX B: Full Experimental Materials List

Item No.	Condition	Context Sentence	Target Sentence
1	MET	There are ruins of an old fort on that hill over there.	The fort **commanded** a strategic location in the past.
1	LIT	There are ruins of an old fort on that hill over there.	The fort **occupied** a strategic location in the past.
2	MET	Today, my mother that discovered she had won the lottery.	Emotion **flooded** her senses when she heard the news.
2	LIT	Today, my mother that discovered she had won the lottery.	Emotion **overwhelmed** her senses when she heard the news.
3	MET	Yesterday, a saleswoman sold my brother a broken computer.	He completely **consumed** her lies in a matter of seconds.
3	LIT	Yesterday, a saleswoman sold my brother a broken computer.	He completely **believed** her lies in a matter of seconds.
4	MET	Yesterday, my brother won his first competition.	Finally, he **tasted** victory for once in his life.
4	LIT	Yesterday, my brother won his first competition.	Finally, he **experienced** victory for once In his life.
5	MET	Before the summer break, he committed himself to a strict workout routine.	He **sculpted** his body through hours of exercise.
5	LIT	Before the summer break, he committed himself to a strict workout routine.	He **developed** his body through hours of exercise.
6	MET	At the annual meeting, tensions were high as the board met with the CEO.	The board **slammed** the CEO's proposal as an impractical idea.
6	LIT	At the annual meeting, tensions were high as the board met with the CEO.	The board **criticized** the CEO's proposal as an impractical idea.
7	MET	In the quiet gallery, she stood before an intricate masterpiece.	She **absorbed** the beauty of the painting.
7	LIT	In the quiet gallery, she stood before an intricate masterpiece.	She **appreciated** the beauty of the painting.
8	MET	Today, the mayor told a completely different story about the scandal.	He totally **flipped** the narrative to his advantage.
8	LIT	Today, the mayor told a completely different story about the scandal.	He totally **changed** the narrative to his advantage.
9	MET	At school, two classmates bonded over shared interests and experiences.	They **forged** a strong friendship that lasted for years.
9	LIT	At school, two classmates bonded over shared interests and experiences.	They **established** a strong friendship that lasted for years.
10	MET	During the lecture, he carefully explained the difficult subject.	His explanation **illuminated** the topic without adding complexity.
10	LIT	During the lecture, he carefully explained the difficult subject.	His explanation **clarified** the topic without adding complexity.

Item No.	Condition	Context Sentence	Target Sentence
11	MET	The politician told her tale with conviction, making it almost believable.	Her fake story **planted** many falsehoods into the public mind.
11	LIT	The politician told her tale with conviction, making it almost believable.	Her fake story **introduced** many falsehoods into the public mind.
12	MET	After weeks, the detective found the missing evidence for a cold case.	The investigator **unearthed** a hidden truth about the crime.
12	LIT	After weeks, the detective found the missing evidence for a cold case.	The investigator **discovered** a hidden truth about the crime.
13	MET	Despite his efforts, my son couldn't meet his teacher's high expectations.	Her criticism **eroded** his confidence with its harshness.
13	LIT	Despite his efforts, my son couldn't meet his teacher's high expectations.	Her criticism **weakened** his confidence with its harshness.
14	MET	After a long interrogation, the enemy soldier finally gave in.	He **spilled** all the secrets he knew, hoping for his release.
14	LIT	After a long interrogation, the enemy soldier finally gave in.	He **revealed** all the secrets he knew, hoping for his release.
15	MET	After submitting my essay, I awaited feedback from my teacher.	My teacher **dissected** my essay in fine detail.
15	LIT	After submitting my essay, I awaited feedback from my teacher.	My teacher **analyzed** my essay in fine detail.
16	MET	A misunderstanding escalated tensions between my parents.	It further **fueled** the conflict in our family.
16	LIT	A misunderstanding escalated tensions between my parents.	It further **worsened** the conflict in our family.
17	MET	My boss arrived early at the office, ready to tackle the day's challenges.	He **juggled** many important issues as the company's director.
17	LIT	My boss arrived early at the office, ready to tackle the day's challenges.	He **managed** many important issues as the company's director.
18	MET	During our session, the therapist listened attentively to our concerns.	The therapist skillfully **steered** us towards a resolution.
18	LIT	During our session, the therapist listened attentively to our concerns.	The therapist skillfully **advised** us towards a resolution.

Item No.	Condition	Context Sentence	Target Sentence
19	MET	The speaker captivated the audience with her passionate speech.	The audience **devoured** every word of her story.
19	LIT	The speaker captivated the audience with her passionate speech.	The audience **accepted** every word of her story.
20	MET	My sister used to feel uneasy and unsure about her new co-worker.	Back then, she definitely **harbored** suspicions about him.
20	LIT	My sister used to feel uneasy and unsure about her new co-worker.	Back then, she definitely **had** suspicions about him.
21	MET	Before he left the farm, he woke up early each morning to tend to the crops.	He **shouldered** many hardships back in those days.
21	LIT	Before he left the farm, he woke up early each morning to tend to the crops.	He **endured** many hardships back in those days.
22	MET	After years of hard work, the writer finally finished writing her novel.	Eagerly, she **courted** many book publishers in her city.
22	LIT	After years of hard work, the writer finally finished writing her novel.	Eagerly, she **contacted** many book publishers in her city.
23	MET	The government implemented new regulations to stabilize prices.	The new policies successfully **cooled** the market by the end of the year.
23	LIT	The government implemented new regulations to stabilize prices.	The new policies successfully **moderated** the market by the end of the year.
24	MET	We listened intently as our grandfather recounted his adventures.	His words **painted** a vivid tale of the past.
24	LIT	We listened intently as our grandfather recounted his adventures.	His words **described** a vivid tale of the past.
25	MET	My rich cousin lived in a mansion with servants attending to her every need.	Her wealth **shielded** her from any hardship.
25	LIT	My rich cousin lived in a mansion with servants attending to her every need.	Her wealth **protected** her from any hardship.
26	MET	After a long day at work, my father came home to rest and relax.	A full night's sleep **recharged** him with energy for the next day.
26	LIT	After a long day at work, my father came home to rest and relax.	A full night's sleep **refreshed** him with energy for the next day.

Item No.	Condition	Context Sentence	Target Sentence
27	MET	We rushed through our project and barely stopped to think.	The short deadline **smothered** our creativity for good quality work.
27	LIT	We rushed through our project and barely stopped to think.	The short deadline **suppressed** our creativity for good quality work.
28	MET	During his performance, the magician held everyone's attention.	He skillfully **enchanted** the audience with his tricks and jokes.
28	LIT	During his performance, the magician held everyone's attention.	He skillfully **entertained** the audience with his tricks and jokes.
29	MET	The teacher explained the math problem to my sister in simple terms.	She **grasped** the concept quickly and with ease.
29	LIT	The teacher explained the math problem to my sister in simple terms.	She **understood** the concept quickly and with ease.
30	MET	We listened as the young refugee shared his heartfelt story of loss and escape.	His words **struck** us deeply with sadness.
30	LIT	We listened as the young refugee shared his heartfelt story of loss and escape.	His words **affected** us deeply with sadness.
31	MET	Seeking a new job, I arrived at the interview nervous but determined.	Still, the interviewer **grilled** me with many tough questions.
31	LIT	Seeking a new job, I arrived at the interview nervous but determined.	Still, the interviewer **questioned** me with many tough questions.
32	MET	My father once committed insurance fraud.	He **orchestrated** his own accident to claim a payout.
32	LIT	My father once committed insurance fraud.	He **planned** his own accident to claim a payout.
34	MET	My friends and I gathered to read a storybook together.	The book's author **weaved** a beautiful tale with vivid language.
34	LIT	My friends and I gathered to read a storybook together.	The book's author **wrote** a beautiful tale with vivid language.

90

Item No.	Condition	Context Sentence	Target Sentence
35	MET	The young influencer worked hard to promote his content on social media.	The ad campaign **amplified** his reach across many users.
35	LIT	The young influencer worked hard to promote his content on social media.	The ad campaign **increased** his reach across many users.
36	MET	My mother struggled with impulsiveness, often acting without thinking.	With meditation, she **tamed** her impulsive personality over time.
36	LIT	My mother struggled with impulsiveness, often acting without thinking.	With meditation, she **controlled** her impulsive personality over time.
37	MET	My billionnaire uncle avoided paying taxes by using loopholes and offshore accounts.	For years, he **shrouded** his sources of wealth behind networks of shell companies.
37	LIT	My billionnaire uncle avoided paying taxes by using loopholes and offshore accounts.	For years, he **concealed** his sources of wealth behind networks of shell companies.
38	MET	My co-worker noticed a substantial problem in our project and told our manager.	The manager even **echoed** her concerns to the director.
38	LIT	My co-worker noticed a substantial problem in our project and told our manager.	The manager even **repeated** her concerns to the director.
39	MET	As a journalist for a tabloid, I befriended the wife of a famous celebrity.	Still, she always **stonewalled** my requests for information.
39	LIT	As a journalist for a tabloid, I befriended the wife of a famous celebrity.	Still, she always **refused** my requests for information.
40	MET	My mother came from a poor working class family.	Her experiences **molded** her character as a kind and resourceful person.
40	LIT	My mother came from a poor working class family.	Her experiences **determined** her character as a kind and resourceful person.