

Towards a metaphor-annotated corpus of Mandarin Chinese

Xiaofei Lu¹  · Ben Pin-Yun Wang¹

© Springer Science+Business Media Dordrecht 2017

Abstract Building on the success of the VU Amsterdam Metaphor Corpus, which comprises English texts annotated with metaphor following the Metaphor Identification Procedure Vrije Universiteit (MIPVU; Steen et al. in *Cogn Linguist* 21 (4):765–796, 2010a; Steen et al. in *A method for linguistic metaphor identification: from MIP to MIPVU*. John Benjamins, Amsterdam/Philadelphia, 2010b), this study has three aims: (1) to adapt and evaluate the transferability and reliability of MIPVU for Mandarin Chinese; (2) to construct a corpus of Chinese texts annotated for metaphor using the adapted procedure; and (3) to examine the distribution of metaphor-related words across Chinese texts in three different written registers: academic discourse, fiction, and news. The results of our inter-annotator reliability test show that MIPVU can be reliably applied to linguistic metaphor identification in Chinese texts. Our metaphor-annotated corpus consists of texts randomly sampled from the Lancaster Corpus of Mandarin Chinese, totaling 30,012 words (about 10,000 for each register). Data analysis reveals that approximately one out of every nine lexical units in our Chinese corpus is related to metaphor, that there is considerable variation in metaphor density across different registers and lexical categories, and that metaphor density is significantly lower in Chinese than in English texts. Our assessment of the replicability of MIPVU for Mandarin Chinese adds to the groundbreaking methodological contribution that Steen et al. (2010a, b) has made to metaphor research. The metaphor-annotated corpus of Mandarin Chinese contributes a valuable language resource for Chinese metaphor researchers,

✉ Xiaofei Lu
xxL13@psu.edu

Ben Pin-Yun Wang
benpywang@gmail.com

¹ Department of Applied Linguistics, The Pennsylvania State University, 304 Sparks Building, University Park, PA 16802, USA

and our analysis of the distribution of metaphor-related words in the corpus offers useful new insights into the extent and use of metaphor in Chinese discourse.

Keywords Corpus annotation · Cross-linguistic comparison · Metaphor · Metaphor density · Metaphor identification · Register variation

Abbreviations

1PL	First person plural pronoun
ADV	Adverbial marker (<i>de</i> 地)
ASSOC	Associative (<i>-de</i> 的)
COP	Copular
NOM	Nominalizer
3SG	Third person singular pronoun
ASP	Aspect marker
CL	Classifier
GEN	Genitive (<i>-de</i> 的)
NEG	Negator

1 Introduction

Metaphor was traditionally viewed as a rhetorical device used to achieve linguistic artistry (e.g. “My love is like a red, red rose.”). However, this view has been most radically challenged by Lakoff and Johnson (1980/2003), who argued that metaphor is pervasive in language, and more importantly, in thought as well. Under this cognitive view, the human conceptual system is held to be fundamentally metaphorical in nature, whereby one concept is often understood in terms of one or more other concepts (e.g. Kövecses 2010; Lakoff and Turner 1989). This can be illustrated by the classical ARGUMENT IS WAR conceptual metaphor, which is manifested in such wide-ranging everyday expressions as “Your claims are *indefensible*,” “He *attacked every weak point* in my argument,” and “His criticisms were right *on target*” (Lakoff and Johnson 1980/2003, p. 4, italics original). These examples show that we tend to use the source conceptual domain WAR to talk about the target conceptual domain ARGUMENT. Conceptual metaphors, or source-target domain mappings, as such are not simply a means for understanding metaphorical language use, but are seen as an organizing principle of the human conceptual system. Conceptual metaphors are revealing of the underlying beliefs and worldviews of members of different cultures (cf. Kövecses 2005; Lakoff and Johnson 1999; Yu 2008). The cognitive approach to metaphor has not only given rise to a booming research program within the cognitive linguistics community, but also attracted interest from many other disciplines in which metaphor is a relevant subject of inquiry, such as anthropology, philosophy, psychology, psycholinguistics, communication studies, translation studies, discourse analysis, natural language processing, and language learning and teaching (cf. Gibbs 2008).

The potential value of large text corpora for metaphor research has long been recognized (e.g. Deignan 2005; Partington 2006; Stefanowitsch and Gries 2006). A

fundamental attraction of a corpus-assisted approach, as noted by Partington (2006), is that the large volume of naturally-occurring data allows metaphor researchers to identify recurring patterns of metaphorical language use that would otherwise be difficult to reveal. Such recurrent patterns provide useful insights into language users' systematic behavior and attitudes, the linguistic systems they appropriate, as well as the discourse contexts in which they operate (Partington 2006). A number of strategies have been employed by metaphor researchers to identify or extract instances of metaphorical language use from corpora that have not been annotated for metaphor. As summarized in Stefanowitsch (2006), such strategies include manual identification, searching for lexical items from the source or target domain, searching for sentences containing lexical items from both the source and target domains, searching for metaphors based on 'markers of metaphor', and extracting metaphors from corpora annotated for semantic fields or domains. While it is difficult to exhaustively extract all instances of metaphorical language use without corpora that have been fully annotated for metaphor, these strategies have facilitated the identification and analysis of linguistic expressions instantiating particular types of conceptual mappings or mappings between specific source and target domains.

Despite the usefulness of the corpus-assisted approach and the strategies for metaphor identification and extraction discussed above, researchers have reported several methodological challenges or limitations in their applications. Three of these issues are especially worth noting as they directly bear upon the goal and methodological choices of the current study. First, researchers have found it difficult to clearly differentiate between literal and metaphorical language use when identifying linguistic metaphors (e.g. Semino et al. 2004). Second, they have also reported difficulty in constructing the inventory of source and target domain categories and in determining the level of abstraction at which such categories should stand when extrapolating conceptual metaphors from linguistic metaphors (e.g. Semino et al. 2004; Shutova et al. 2013). Third, without fully metaphor-annotated corpora, the scope of metaphor research is limited to examining selected sets of metaphors and their related linguistic expressions and does not extend to systematic investigations of patterns of metaphor in usage as well as distributions of metaphor across different registers (e.g. Steen et al. 2010a, b).

In response to the first challenge, Steen and colleagues (Pragglejaz Group 2007; Steen et al. 2010a, b) argued for explicit and reliable procedures for identifying linguistic metaphors in discourse. The Pragglejaz Group (2007) developed the "Metaphor Identification Procedure" (MIP), which consists of a sequence of steps for determining whether each word conveys a metaphorical meaning in context. They also presented data from a case study that demonstrated both the reliability of the procedure and the complications involved in its application. Steen et al. (2010a, b) further proposed an extended and refined version of MIP, that is, the "Metaphor Identification Procedure Vrije Universiteit" (MIPVU). The new annotation protocol was also found to yield a high level of inter-annotator reliability. The highly explicit and comprehensive nature of MIPVU makes it an effective tool for minimizing the difficulty in differentiating literal and metaphorical language use when identifying linguistic metaphors. Rather than attempting to provide a solution for the difficulty in extrapolating conceptual metaphors from linguistic metaphors, Steen and colleagues

(Pragglejaz Group 2007; Steen et al. 2010a, b) considered the independence of MIP and MIPVU from conceptual analysis an important advantage. However, the reliable and systematic analyses of linguistic metaphor in usage enabled by MIP and MIPVU would undoubtedly facilitate later extrapolation of conceptual metaphors. Finally, Steen and colleagues argued that, with text corpora reliably annotated for metaphor following well-defined procedures such as MIPVU, researchers can systematically describe all metaphor in discourse. Indeed, Steen et al. (2010a, b) annotated a large number of texts representing four registers using MIPVU and reported substantial cross-register variation in the distribution of metaphor-related words. These annotated texts are now available online as the VU Amsterdam Metaphor Corpus (VUAMC).¹

To date, research on metaphor in Chinese has followed two main lines of investigation. The first was concerned with the examination of cross-domain mappings underlying specific metaphorical linguistic expressions. To this end, researchers largely adopted one or more of the strategies summarized in Stefanowitsch (2006) to extract potentially relevant linguistic expressions manifesting specific types of conceptual mappings from either general-purpose or self-compiled, specialized corpora (e.g. Chiang and Duann 2007; Chiu and Chiang 2011; Han 2014; Hsieh 2006; Jing-Schmidt and Peng 2017; Lu and Ahrens 2008; Tay 2015; Yu and Jia 2016). The second utilized lexical and ontological resources and corpus-driven methods to identify source domains or to determine mapping principles between source and target domain pairings (e.g. Ahrens et al. 2003, 2004; Chung et al. 2005; Chung and Huang 2010; Gong et al. 2008). For example, Ahrens et al. (2003, 2004) used both frequency information and information derived from WordNet and Suggested Upper Merged Ontology (SUMO) for mapping principle verification. Chung and Huang (2010) adopted a data-driven, bottom-up approach to source domain identification based on the frequency of collocational patterns. These lines of research have provided invaluable insights into source-target domain mappings and methods of their identification. However, in the absence of fully metaphor-annotated Chinese text corpora, researchers have had to focus on specific sets of metaphorical expressions or source and target domain pairings and have not been able to investigate the distribution and density of all metaphors in the corpus being analyzed.

Building on the success of VUAMC and complementing previous research on metaphor in Chinese, this study aims to adapt and evaluate the transferability and reliability of MIPVU for Mandarin Chinese, to construct a corpus of Chinese texts annotated for metaphor using the adapted procedure, and to examine the distribution of metaphor-related words across Chinese texts in three different written registers: academic discourse, fiction, and news. Following Steen et al. (2010a, b) and recognizing the lack of well-established procedures for extrapolating conceptual metaphors from linguistic metaphors, the present study focuses on the annotation of linguistic metaphors only and leaves the analysis of conceptual metaphors to future work. MIPVU was initially designed for identifying metaphor in English, but has been shown to be applicable to Dutch (Pasma 2012) and Russian (Badryzlova et al. 2013) with some necessary adjustments. Our assessment of the replicability of MIPVU for Mandarin Chinese will add to the groundbreaking methodological contribution that Steen et al. (2010a, b) has made to metaphor research. Furthermore, the construction

¹ The VU Amsterdam Metaphor Corpus is available at: <http://ota.ahds.ac.uk/headers/2541.xml>.

Table 1 The Metaphor Identification Procedure (Pragglejaz Group 2007)

-
- 1 Read the entire text/discourse to establish a general understanding of the meaning
 - 2 Determine the lexical units in the text-discourse
 - (a) For each lexical unit, establish its meaning in context, that is, how it applies to an entity, relation, or attribute in the situation evoked by the text (contextual meaning). Take into account what comes before and after the lexical unit
 - (b) For each lexical unit, determine if it has a more basic contemporary meaning in other contexts than the one in the given context. For our purposes, basic meanings tend to be
 - More concrete (what they evoke is easier to imagine, see, hear, feel, smell, and taste);
 - Related to bodily action;
 - More precise (as opposed to vague);
 - Historically older
 Basic meanings are not necessarily the most frequent meanings of the lexical unit
 - (c) If the lexical unit has a more basic current/contemporary meaning in other contexts than the given context, decide whether the contextual meaning contrasts with the basic meaning but can be understood in comparison with it
 4. If yes, mark the lexical unit as metaphorical
-

and analysis of the metaphor-annotated corpus of Mandarin Chinese will not only contribute a useful language resource for researchers but also offer new insights into the extent and use of metaphor in Chinese discourse.

2 From MIP to MIPVU and the VU Amsterdam Metaphor Corpus

The Pragglejaz Group, which consisted of 10 metaphor scholars from several different disciplines, made the first attempt “to create an explicit, reliable, and flexible method for identifying metaphorically used words in spoken and written language” (Pragglejaz Group 2007, p. 2). The outcome of their efforts was the Metaphor Identification Procedure (MIP), summarized in Table 1 (Pragglejaz Group 2007, p. 3). As can be seen, the procedure was designed to allow metaphor researchers to determine whether each lexical unit in a discourse conveys a metaphorical meaning based on its use in the specific context it appears in.

Steen and colleagues later noted several aspects of MIP that warranted refinement or extension (cf. Steen 2007; Steen et al. 2010a, b). Specifically, with respect to the types of metaphor-related words covered, they considered it useful to include not only indirect metaphors but also direct metaphors, such as simile and analogy, and implicit metaphors in such forms as substitutions and ellipses that can be potentially explained by cross-domain mappings. Furthermore, they argued that lexical signals of potential cross-domain mappings, such as markers of simile and analogy (e.g. *like*, *as*), should also be coded. Finally, they deemed it necessary to specify a mechanism for handling newly coined formations (e.g. *honey-hunting*) that may be potentially explained by cross-domain mappings. As a result of these considerations, Steen and colleagues developed an extended version of MIP, namely,

Table 2 The Metaphor Identification Procedure Vrije Universiteit (Steen et al. 2010b)

1. Find metaphor-related words (MRWs) by examining the text on a word-by-word basis
2. When a word is used indirectly and that use may potentially be explained by some form of cross-domain mapping from a more basic meaning of that word, mark the word as metaphorically used (MRW)
3. When a word is used directly and its use may potentially be explained by some form of cross-domain mapping to a more basic referent or topic in the text, mark the word as direct metaphor (MRW, direct)
4. When words are used for the purpose of lexico-grammatical substitution, such as third person personal pronouns, or when ellipsis occurs where words may be seen as missing, as in some forms of co-ordination, and when a direct or indirect meaning is conveyed by those substitutions or ellipses that may potentially be explained by some form of cross-domain mapping from a more basic meaning, referent, or topic, insert a code for implicit metaphor (MRW, implicit)
5. When a word functions as a signal that a cross-domain mapping may be at play, mark it as a metaphor flag (MFlag)
6. When a word is a new-formation coined, examine the distinct words that are its independent parts according to steps 2 through 5

MIPVU, the basic procedure of which is presented in Table 2 (Steen et al. 2010b, pp. 25–26).

In MIPVU, as is the case with MIP, the purpose of the annotation is to identify the linguistic manifestations of metaphor **rather than its underlying conceptual structures (i.e. mappings between source and target domains)**. Linguistic metaphor identification is interpreted as **producing data about “the semiotic structure of language in usage events”** (Steen et al. 2010b, p. 21), and **is not intended to make direct claims about cognitive processes**.

In both MIP and MIPVU, the unit of analysis examined is the ‘lexical unit’, which is generally a word, with some exceptions.² Metaphorical interpretations on the morphological, phraseological, and syntactical levels are not considered. Recognizing the importance of a systematic and explicit approach to the unit of analysis for a reliable quantitative analysis of the data, Steen et al. (2010b) provided clear guidelines for determining the boundaries of lexical units. Different from MIP, MIPVU relies on the distinction between part-of-speech (POS) categories in defining such units. The data used in developing MIPVU came from the British National Corpus (BNC), and all words annotated with an independent POS tag in the BNC were considered lexical units. The same lemma with different parts of speech (e.g. *work* as a noun and *work* as a verb) is analyzed as distinct lexical units. A number of fixed multi-word expressions or polywords (e.g. *of course*) are analyzed as one lexical unit in the BNC and therefore also in MIPVU. Phrasal verbs (e.g. *look up*), although segmented as two words, are treated as single lexical units because they “function as linguistic units designating one action, process, state or relation in the referential dimension of the discourse” (Steen et al. 2010b, p. 28). In addition, compounds spelt as one word (e.g. *underpass*) or two hyphenated words (e.g. *pitter-patter*) that can be found in the dictionary as such are also treated as

² In this paper, the terms ‘lexical unit’ and ‘word’ are used interchangeably in contexts where the differentiation is not important.

single lexical units, so are conventionalized compound nouns spelt as two words (e.g. *power plant*, with primary stress on the first word), again because they designate a single referent in the discourse. Finally, proper names and titles that appear in the dictionary and that have the same stress pattern as conventionalized compound nouns (e.g. *Pulitzer Prize*) are also treated as single lexical units.

In both MIP and MIPVU, indirect metaphors are operationalized by similarity or comparison. Specifically, a lexical unit is considered an indirect expression of metaphor if its contextual meaning is related to a more basic meaning by some form of similarity. For example, the contextual meaning of the word *valuable* in (1), namely, ‘very useful and important’, can be understood in comparison with its more basic meaning of ‘worth a lot of money’. In this case, the use of *valuable* counts as an instance of indirect metaphor.

- (1) Professional religious education teachers like Marjorie B Clark (Points of View, today) are doing *valuable* work in many secondary schools (...) (Steen et al. 2010b, p. 45, italics original)

As mentioned above, apart from indirect metaphors, MIPVU also codes direct and implicit metaphors. Direct metaphors are defined as lexical units that are used directly but that can be potentially explained by some form of cross-domain mapping to the local or main referent or topic of the text. Consider the use of *dragonfly* in (2): the lexical unit is used directly here since there is no contrast between its contextual and basic meanings. However, the direct use of the word introduces an incongruous local referent into the discourse, setting up a cross-domain comparison, specifically one between a person and a dragonfly. Therefore, *dragonfly* in this case is considered a metaphor-related word.

- (2) Sara was undressed and ready for bed but Jenny was fully clothed, moving about the room in her harlequin dress *like some angry restless dragonfly*. (Steen et al. 2010b, p. 93)

Different from indirect and direct metaphors, implicit metaphors do not have words that come from another semantic or conceptual domain. Rather, these are metaphors by substitution, which work through pro-forms, or by ellipsis, which work through absent words that could be inserted into grammatical gaps. Specifically, implicit metaphors include cases of substitution and ellipsis in which the notions that are substituted or the lexical units that are absent but understood convey direct or indirect meanings that can be explained by cross-domain mapping. To illustrate, the use of *it* in (3) constitutes an instance of implicit metaphor in that the pronoun here substitutes for the metaphorically used word, *step*.

- (3) Naturally, to embark on such a step is not necessarily to succeed immediately in realizing *it*. (Steen et al. 2010b, p. 39, italics original)

In addition to different types of metaphors, MIPVU also annotates metaphor flags. These are lexical signals of potential cross-domain mappings that alert the language user of some form of contrast or comparison (cf. Goatly 1997). The types of metaphor flags include markers of simile and analogy such as *like* and *as*, more substantial lexical markers such as *compare*, *similar*, and *analogy*, and complex mental conception markers such as *regard as* and *imagine*.

Despite these explicit instructions, ambiguous cases still remain. MIPVU further adds to its binary scale (i.e. clear MRWs and non-MRWs) a third category of WIDLII, ‘When In Doubt, Leave It In’, for those borderline cases that cannot be resolved after group discussion among the annotators. When a metaphorical and a literal interpretation are equally plausible for the contextual use of a lexical unit, the unit is also coded with the tag WIDLII to signal its ambiguity.

It is important to note that neither MIP nor MIPVU codes historical metaphors, as historically older senses of lexical units are generally considered to be unavailable to typical contemporary language users. This consideration is aligned with the assumption that a metaphor is shorthand for “metaphorical to some language user” (Steen 2007). In MIPVU, ‘language user’ is operationalized as “the idealized native speaker of English as represented in the description of English by the dictionary of the particular period” (Steen et al. 2010b, p. 7). Both the Praggeljaz Group (2007) and Steen et al. (2010a, b) used the *Macmillan English Dictionary for Advanced Learners* as a lexical tool to assist decisions regarding the status of lexical units, the contextual and basic meanings of the lexical units, and the degree of distinctness between the two meanings. Steen et al. (2010a, b) also used the *Longman Dictionary of Contemporary English* to obtain a second opinion in cases of doubt.

The VU Amsterdam Metaphor Corpus (Steen et al. 2010a, b) represents the most systematic effort in metaphor annotation to date. This corpus consists of nearly 190,000 lexical units annotated following MIPVU. These metaphor-annotated texts cover four English-language registers: news, conversation, fiction, and academic discourse, with approximately 50,000 words for each register. The texts were sampled from a subset of BNC-Baby, a four-million word subset of the British National Corpus. Steen et al. (2010b) documented the challenges encountered in annotating texts from each register and reported substantial cross-register variation in the distribution of metaphor-related words.

Pasma (2012) and Steen et al. (2010b) further examined the applicability of MIPVU in annotating Dutch texts. Steen et al. (2010b) reported that in general, MIPVU could be directly applied to Dutch, although some language-specific procedural problems need to be addressed. For example, one class of lexical units that required their own treatment in the procedure was separable complex verbs, which consist of a particle and a verb that may be separated from each other in some contexts. It was decided that these components would be consistently analyzed as one lexical unit. Another linguistic issue that arose had to do with polywords. Unlike in English, recognition of two or more words as one polyword proved difficult in Dutch, given the Dutch dictionary and the POS tags in the Dutch corpus employed. As such, the separate parts of potential polywords were analyzed as separate lexical units.

Badryzlova et al. (2013) discussed issues related to the applicability of MIPVU in annotating Russian texts. First, in cases where it was difficult to single out one basic meaning, they defined a group of basic meanings that share “the feature of concreteness, body-relatedness and preciseness” (p. 79). Second, they added special tags for idioms and proper names. Third, some words were defined in the Russian dictionaries by words with the same root but different parts of speech, and some imperfective and passive verbs by their perfective or active counterparts. Moreover,

the dictionaries also listed homonymous word forms with different parts of speech in the same entry. The authors thus needed to introduce a procedure for determining the basic meanings of such words. Finally, they noted that agglutinative and abbreviated compound words required separate analysis of their stems, and that in some cases the basic meaning of a word could be a stylistically marked meaning of the word.

Satisfactory inter-rater reliability has been reported for MIP and MIPVU. For MIP, the Pragglejaz Group (2007) reported inter-annotator agreement among six raters using Cohen's Kappa. Computed using the means of agreement between each pair of raters, Kappa was .70 for news text and .56 for conversation. Computed using agreement across all six raters, Kappa was .72 for news texts and .62 for conversation. For the application of MIPVU to English texts, Steen et al. (2010a, b) reported consistent inter-rater agreement across four raters in a series of reliability tests. In the final test of the series, Fleiss's Kappa was .79 for academic text, .78 for conversation, .85 for fiction, .96 for news text, and .85 across all four registers. For the application of MIPVU to Dutch texts, Steen et al. (2010b) conducted three reliability tests. Inter-rater agreement among three independent raters, measured using Fleiss's Kappa, ranged from .74 to .80 for conversation, .77 to .86 for news text, and .79 to .82 across the two registers. Finally, for the application of MIPVU to Russian texts, Badryzlova et al. (2013) reported a Fleiss's Kappa of .65 in an initial reliability test (following the original protocol of MIPVU without adjustments) and of .90 in a second test (after necessary adaptations) among three annotators on four texts representing four registers (fiction, transcribed spoken, popular science/academic, and news texts). These results suggest that MIPVU can be reliably applied to texts of diverse registers and possibly in different languages.

3 Method

This section explains how we applied MIPVU to Mandarin Chinese texts and the necessary adjustments made with regard to the tools available (including the corpus and dictionaries used) and language-specific issues in the unit of analysis.

3.1 Materials

The texts annotated for this study were randomly sampled from the **Lancaster Corpus of Mandarin Chinese (LCMC)** (McEnery and Xiao 2004; Xiao 2017). LCMC contains five hundred 2000-word samples of written Chinese texts published in Mainland China around 1991. The one-million-word balanced corpus was designed to be a Chinese counterpart of the FLOB and Frown corpora of British and American English. LCMC was word segmented and **POS-tagged using the Chinese Lexical Analysis System developed by the Institute of Computing Technology, Chinese Academy of Sciences (ICTCLAS)**. The core lexicon of ICTCLAS incorporates a frequency dictionary of 80,000 words with POS information. For word segmentation, the system achieves a precision of 97.58% and a recall of 99.94%; for POS tagging, the system achieves an accuracy of 97.16% (Zhang and

Table 3 Information on the annotated texts

	Academic	Fiction	News
Text category	SCIENCE	GENERAL FICTION	PRESS REPORTAGE
Number of files	30	20	25
Total number of words	10,002	10,004	10,006

Liu 2002). LCMC was chosen for two main reasons. First, the corpus is balanced within the written register and relatively contemporary compared to other Chinese corpora. Second, and more importantly, the complete set of raw data in LCMC is licensed free of charge for use in non-profit-making research. It is therefore possible for researchers to gain free access to both the raw LCMC data and our annotation by obtaining two separate, free licenses.

To examine cross-register variations in metaphor density, this study set out to annotate files taken from three text categories in LCMC, as shown in Table 3. These categories are representative of three distinct registers, namely, academic, fiction and news texts. In other words, this study covers three of the four registers in VUAMC, with the exception of conversation. For academic and fiction texts, separate fragments (delimited by segments of paragraphs) were randomly taken from the beginning, middle, and ending sections of 30 files in the SCIENCE and 20 in the GENERAL FICTION categories.³ As for news texts, a total of 25 complete files from the category of PRESS REPORTAGE were included. The goal of this project was to annotate approximately 10,000 words for each register.

3.2 Tools

Following previous studies applying MIP(VU), we also relied on external lexical tools for making decisions regarding basic and contextual meanings of lexical units. Unfortunately, there are no large-scale, corpus-based dictionaries, like the *Macmillan English Dictionary for Advanced Learners* adopted by the Pragglejaz Group (2007) and Steen et al. (2010a, b), in Chinese. To make the most informed decisions, we decided to use two types of lexical tools: a lexical database, i.e. the Chinese WordNet (CWN), and two reference dictionaries. CWN, a language resource developed by Academia Sinica, subsumes more than 8700 lemmas (mostly medium-frequency words) and 23,000 senses (Huang et al. 2010). The information for each lexical entry includes part-of-speech, sense description, illustrative examples, corresponding English synset(s) from the Princeton WordNet, lexical semantic relations, etc. (Lee et al. 2009). CWN is well-suited for the purpose of the present project because of its clear divisions and elaborate definitions of word senses, accompanied with detailed lexical category information and corpus-attested examples.

³ The files in the SCIENCE category cover a broad range of academic disciplines, including humanities, social sciences, natural sciences, engineering, etc.

However, due to the limited coverage of CWN, we also used the following two reference dictionaries: *Xiandai Hanyu Guifan Cidian* 现代汉语规范词典 [A Standard Dictionary of Modern Chinese] (XHGC; Li 2014) and *Xiandai Hanyu Xuexi Cidian* 现代汉语学习词典 [A Learner's Dictionary of Modern Chinese] (XHXC; Commercial Press Dictionary Research Centre 2010). Both are authoritative, updated dictionaries of contemporary Chinese with a wide coverage of vocabulary and helpful notes on word usage (e.g. distinctions among near-synonyms or confusable words). The two dictionaries have different strengths in terms of metaphor identification. The former dictionary, XHGC, adds useful notations on the relations among different senses of a word. According to the preface, the editors strived to list the distinct senses of a word in the order of semantic extension. Such information is valuable in that historically older senses are more likely—though not necessarily—to be more basic, that is, more concrete, specific and human-oriented (Pragglejaz 2007; Steen et al. 2010b). The second dictionary, XHXC, presumably because it is specifically written for first language learners of Chinese, generally has more elaborate definitions, especially for multi-functional words such as prepositions and localizers. The illustrative examples in this dictionary also tend to be longer (thus with more contextual information) and more varied (including both concrete and abstract examples). For a given word not found in CWN, its meaning comparisons were performed using the dictionary that provides more fine-grained sense descriptions for the word between the two.

Still, the two reference dictionaries adopted have a number of limitations with regard to metaphor identification in natural discourse. For a few lexical units, their basic and extended meanings are conflated under one sense description in the lexical tools. To illustrate, according to XHGC, the noun *chū-lù* 出路 out-path has two senses, the first of which is “通向外面的路; 比喻生存的机会或发展的前途” ‘a way out; figuratively meaning chances of survival or future development.’ This sense description encompasses not only the literal meaning of the noun but also its metaphorical extension, delimiting the two uses by a semicolon. We therefore deemed the noun to have three separate senses. For another example, consider the sense descriptions for the verb *shàngzhǎng* 上涨 ‘to rise’: “(水位、价格等) 升高” in XHGC and “水位或价格等升高” in XHXC, both of which can be glossed as ‘(water level, price, etc.) to rise.’ The definitions usefully enumerate the two typical domains to which the verbal sense can be applied. The two listed domains are separated either by “、,” a Chinese slight-pause mark used to set off items in a series, or by the conjunction *huò* 或 ‘or.’ Apparently, the first domain is conceptually more basic whereas the other is a derived use from the meaning of change of spatial location. For cases in which sense conflation is suspected, like the two examples discussed here, the uses of punctuation marks or conjunctions by the Chinese dictionary definitions serve to delimit the distinct senses, or uses in different conceptual domains, of a given word.

Another limitation of the lexical tools is that their sense descriptions may use synonyms of the words being defined (cf. Pasma 2012; Veale et al. 2016). A case in point is the verb *shīqù* 失去 ‘to lose’: CWN uses the same word in its definition (“失去後述物件” ‘to lose the object(s) referred to hereinafter’), and the two reference dictionaries define the verb with its synonyms, *diūdiào* 丟掉 and *shīdiào* 失掉.

While *diūdiào* 丢掉 has clear concrete and abstract uses (e.g. to lose a cell phone vs. to lose a job), the other two words are only illustrated with abstract examples (e.g. to lose confidence/effect) in the dictionaries. After consulting different balanced Chinese corpora, we concluded that, like its synonym, *shīqù* 失去 does have a less frequent, but more concrete sense (e.g. to lose a body part).⁴ In cases of doubt like this, concordance analyses were performed to ascertain that the synonyms used to define one another have comparable uses.

A final issue is when the words segmented in LCMC cannot be found in CWN or either of the reference dictionaries. The coverage problem is expected to exist regardless of the dictionaries used, as has been shown in previous research on unknown word resolution in Chinese language processing (e.g. Lu 2007, 2008). The unfound words represent less than 4% of the tokens, as estimated using one-fifth of the overall data. More details about the lexical gaps and how we addressed them will be provided in Sect. 4.

3.3 Language-specific issues in the unit of analysis

This study followed the basic protocol of MIPVU, presented in Table 2, when applying the procedure (steps 1–5) to Chinese texts. In so doing, this study also took the word, or lexical unit, as the unit of analysis, excluding the considerations of metaphorical interpretations below (orthographic or morphemic) or above (phrasal or syntactic) the lexical level.⁵ However, since MIPVU was first devised for English, explanations for how the lexical types specific to Chinese were catered to with regard to the relevant unit of analysis are in order. Our general strategy for determining the boundaries of lexical units was to reduce the number of exceptions to the POS tagging provided by LCMC, as the corpus provides an objective, frequency-based word segmentation independent of the study of metaphorical use.

Before we discuss Chinese-specific lexical types that require additional attention in applying MIPVU, a few words about the notion of ‘word’ in Chinese are needed. While words in English texts are generally delimited by white spaces, Chinese texts are made up of strings of characters without similar natural word delimiters. Therefore, the concept of “character” (*zì* 字) appears to be more intuitive to Chinese speakers than that of “word” (*cí* 词) (Packard 2000; Sun 2006). Each Chinese character represents a phonological syllable and, in most cases, corresponds to a morpheme, free or bound. Accordingly, a word in Chinese can be composed of more than one character. The absence of conventionalized word boundaries also makes the task of automatic word segmentation in Chinese more challenging, particularly with respect to ambiguity resolution and identification of unknown words (cf. Chen and Bai 1998; Chen and Liu 1992; Huang et al. 1997; Huang and Xue 2015; Lu 2007; Sun and Zou 2001).

⁴ The corpora consulted, besides LCMC, included the Chinese National Corpus and the Sinica Corpus.

⁵ Among previous studies that applied the earlier version of the annotation protocol (MIP) to Chinese, some (e.g. Duann and Huang 2015; Han 2014) also adapted the procedure to include metaphorical interpretations involved in the internal structures of Chinese characters and compound words.

Despite the lack of transparent word-boundary indicators in Chinese texts, the linguistic construct of ‘word’ can still be usefully characterized in Chinese. Li and Thompson (1981) define a Chinese word according to its semantic and syntactic independence and integrity, a practice, they argue, more in line with how ‘word’ is viewed in other languages. Not surprisingly, such characterization is also compatible with the definition of ‘lexical unit’ by Steen et al. (2010b, p. 27), namely, “grammatical units which designate one specific referent in the discourse.” The comparable operationalization of ‘lexical unit’ or ‘word’ constitutes a promising first step to assess the transferability of MIPVU to Chinese.

The issue of lexical units is further complicated by the productive morphological process of compounding in Chinese. Given the fuzzy boundary of words in Chinese, we followed the practice of LCMC in deciding what counts as a compound. The original treatment of (conventional) English compounds in MIPVU as single lexical units, which consist of distinct parts and designate one referent or process in discourse (Steen et al. 2010b), can also be applied to Chinese compounds. However, two types of Chinese compounds may present operational challenges to the demarcation of lexical units in Chinese texts: verb-object compounds (VOCs) and resultative verb compounds (RVCs), for the morphemes in these compounds can be separated. When used intact, VOCs and RVCs are segmented in LCMC as individual words with independent tags (mostly verbs), and therefore they are taken as single lexical units. When separated, the parts of a VOC or a RVC are segmented as isolated words with their own POS tags; in this case, each part of the compounds is analyzed as a distinct lexical unit.⁶

Chinese VOCs, which are composed of two morphemes with a syntactic relation of a verb and its direct object, do not form a uniform group in that they vary in their idiomaticity and separability of their parts (Chao 1968; Li and Thompson 1981). This latter dimension of VOCs is of particular relevance to the delimitation of lexical units. While some compounds of this type are entirely or almost inseparable (e.g. *dān-xīn* 担心 bear-heart ‘to worry’), and thus their wordhood is less controversial, others allow insertion of constituents of various grammatical complexity between the two morphemes. Take the verb-object compound *shēng-qì* 生气 produce-gas ‘angry’ as an example of how its separability can affect the decisions regarding lexical units as well as basic and contextual meanings. When used without any intervening element, *shēng-qì* 生气, taken as a single lexical unit in the word class of verbs, has a monosemous, emotion-related meaning, derived via the conceptual metaphor ANGER IS THE HOT GAS IN A CONTAINER (Yu 1998). In such cases, the compound, despite the metaphorical nature of its composition, is always treated as a non-MRW, for “MIPVU focuses on word use in

⁶ We also recognize that, by staying as close to the word segmentation results as provided by LCMC as possible with respect to the delimitation of lexical units, this study also inherited the constraints shared by Chinese tokenizers and their practical solutions to some thorny cases of discontinuous compounds, in particular, the VOCs split by aspect morphemes and the RVCs in the potential form (see the discussion in this section below). An alternative solution is to manually re-segment these split compounds, based on the more widely held view in the Chinese linguistics literature (e.g., Chao 1968; Li and Thompson 1981) that treats the aspects markers within VOCs as suffixes as well as the potential markers *de* 得 and *bù* 不 between the two parts of RVCs as infixes.

context, not on the result of metaphoric word formation process” (Steen et al. 2010b, p. 17). However, the two parts of the compound can also be separated by the aspect marker *le* 了, (as in *shēng-le-qì* 生了气 produce-ASP-gas), a measurement phrase composed of a numeral and a classifier (e.g. *shēng-yī-dùn-qì* 生一顿气 produce-one-CL-gas), or a modifier of the object constituent (e.g. *shēng-tā-de-qì* 生他的气 produce-3SG-GEN-gas ‘be angry at him/her’).⁷ The compound even allows inversion of the verb and the object, behaving like a regular verb-object phrase. In these cases, *shēng* 生 and *qì* 气 are all taken as separate lexical units with their own basic, literal meaning (i.e. *produce* and *gas*), which contrasts with the emotion-related contextual meaning; thus, both are coded as metaphor-related words.

Compared to VOCs, RVCs are subject to a higher degree of regularity in terms of the separability of their parts. RVCs are essentially V–V complex verbs in which V₂, or the postverb expresses a certain result of the action or process designated by V₁ (Packard 2000; Ross 1990). Four major types of result can be expressed by an RVC: cause (e.g. *dǎ-pò* 打破 hit-broken ‘to break’), direction (e.g. *zǒu-chū* 走出 walk-out ‘to walk out’), achievement (e.g. *mǎi-dào* 买到 buy-arrive ‘to manage to buy’), and phase (e.g. *yòng-wán* 用完 use-finish ‘to use up’) (Li and Thompson 1981). In general, RVCs do not allow intervening constituents such as aspect markers or measure words between V₁ and V₂, a property suggesting that RVCs grammatically function more like a single verb than VOCs (Thompson 1973). Nonetheless, the parts of some directional RVCs can be split (Li and Thompson 1981; Ross 1990; Shi 2002). To illustrate, the two morphemes in *ná-lái* 拿来 bring-come ‘to bring’ can be separated by an aspect marker and the direct object of the RVC, as in *ná-le-shuǐ-lái* 拿了水来 bring-ASP-water-come ‘brought water’; in this case, the verbs *ná* 拿 and *lái* 来 would be segmented as separate words with their own POS tags in LCMC.

It is worth noting that some RVCs are not segmented as one word but two separate words in LCMC. These compounds are not recognized as a gestalt verb by the frequency dictionary incorporated in the word segmentation system used by LCMC (cf. McEnery and Xiao 2004; Xiao 2017). Given the productive formations of RVCs in Chinese, some members in this natural class are expected to be more idiomatic and lexicalized than others. For example, while both *zhǎn-kāi* 展开 spread-open ‘to spread out’ and *ná-kāi* 拿开 take-away ‘to take away’ share the same postverb, the former is processed in LCMC as one word but the latter, two. The RVC *zhǎn-kāi* 展开 indeed exhibits a higher degree of lexicalization in that it has developed out of the original sense a non-compositional meaning, namely, ‘to start/launch.’ The more lexicalized status of *zhǎn-kāi* 展开 is also reflected in its entrance into CWN and the two Chinese dictionaries consulted. By comparison, the

⁷ It is also possible, albeit extremely rare, for other types of Chinese compounds to be used discontinuously, or “ionized” (Chao 1968) (e.g. *yōumò* 幽默 ‘humor’ → *yōu le tā yī mò* 幽了他一默 ‘to make a joke with him’ and *kāngkǎi* 慷慨 ‘generous’ → *kāng tā rén zhī kǎi* 慷他人之慨 ‘generous with other people’s goods’). Such pseudo-VOCs differ from regular VOCs in two key aspects. First, the two parts of pseudo-VOCs in ionization are not independent words and thus cannot be examined separately for their contextual and basic meanings. Second, the two parts of pseudo-VOCs do not have a verb-object relation, so these forms would not have the dual status of a morphologically complex word and a syntactic phrase as true VOCs do (cf. Packard 2000). Given such internal properties of pseudo-VOCs, in cases of ionization, their two split parts are treated as a single lexical unit rather than separate ones.

meaning of *ná-kāi* 拿开 is always compositional, and the RVC has not entered any of the lexical tools adopted. The RVCs in which V_1 and V_2 are segmented as separate words are reasonably assumed to be less idiomatic combinations. For these cases, again we followed the word segmentation as provided by LCMC and analyzed them as composed of separate lexical units, on the grounds that Chinese speakers may need to parse the compounds into their component parts to establish their conceptual relation.

Even for the RVCs segmented in LCMC as a single word, not all of them are listed as such in CWN or the two dictionaries adopted. For the unfound RVCs, we analyzed their concordance citations sampled from multiple balanced Chinese corpora to determine their more basic meaning(s) (cf. Deignan 2015). One such example is *chuán-chū* 传出 pass.on-out, which, based on our concordance analysis, has two major senses: ‘(sounds) to emit’ and ‘(news) to get out.’ Contextual uses of the RVC in the latter sense were considered metaphorical because this sense can be contrasted and understood by comparison with the former, more concrete sense. The same guidelines were also applied to other types of lexical gaps (see the example of *xīn-zhōng* 心中 ‘at heart’ in Sect. 4).

A key identifying structural feature of RVCs is that most of them can occur in the ‘potential form’ (Li and Thompson 1981), which involves the insertion of a morpheme, either the affirmative *de* 得 or the negative *bù* 不, between the two verb constituents to indicate that the action or process conveyed by V_1 can or cannot achieve the result conveyed by V_2 . RVCs used in the potential form are usually segmented into three words in LCMC, a practice that has consequences in terms of the identification of basic and contextual meanings of the verb constituents. For example, while the RVC *gǎn-shàng* 赶上 catch.up-up ‘to catch up with’ is treated as a single lexical unit, in which the postverb can only have an attainment reading, its negative potential form, *gǎn-bù-shàng* 赶不上 catch.up-NEG-up ‘unable to catch up with’ is taken as three separate lexical units. In the latter case, *-shàng* 上 is segmented as an independent lexical unit, so the attainment meaning can be contrasted with its more basic sense of physical ascending, and therefore the postverb here should be considered an MRW.

3.4 Annotation categories

In terms of distinct types of MRWs, following MIPVU, we annotated indirect metaphor, direct metaphor, implicit metaphor, and metaphor flags (MFlag). We also maintained the subcategory of “When in Doubt, Leave it in” (WIDLII) in MIPVU to accommodate lexical units with ambiguous readings (literal/basic vs. metaphorical/non-basic) in the Chinese texts as well as problematic cases that cannot be resolved after discussion among annotators. The first and predominant type of linguistic metaphor annotated is indirect metaphor. As an initial example, consider the use of the word *zhàn* 占 ‘to occupy’ in (4), taken from a news article about college-graduate salespeople in an electronics store. The contextual meaning of the verb, ‘to take up (a proportion),’ can be understood in contrast with its basic meaning, ‘to use a space,’ and thus the word was marked as an indirect MRW.

- (4) 像 小李 这样 的 大学生 营业员 ,
xiàng xiǎolǐ zhèyàng de dàxuéshēng yíngyèyuán,
 like Xiao.Li such NOM college.student salespeople
 我们 商场 里 共 有 22 名 ,
wǒmen shāngchǎng lǐ gòng yǒu 22 míng,
 1PL store inside in.total have 22 CL
 占 营业员 的 近 1/3 [A14]
zhàn yíngyèyuán de jìn 1/3
 occupy salespeople NOM close 1/3
 ‘In our store, we have a total of 22 college-graduate salespeople like Xiao Li, representing close to one-third of the salespeople.’

Compared to English, Chinese has a much richer inventory of classifiers, which can also be systematically annotated via MIPVU. Previous studies (e.g. Huang and Ahrens 2003; Tai 1994; Zhang 2007) have observed that the same classifier may have more than one use/meaning, depending on the co-occurring nominals. Example (5), taken from a fiction text, shows an indirect use of the classifier, *piàn* 片, which in this context describes the bustling noises in the hotel yard. Such contextual use can be understood as a metaphorical extension of the classifier’s basic meaning, which, according to CWN, refers to concrete, flat and thin entities (e.g. *bǐnggān* 饼干 ‘cookie’ and *cháyè* 茶叶 ‘tea leaf’) (cf. Tai and Chao 1994). Another MRW in this example in need of explanation is the word *shuì-xǐng* 睡醒 ‘to wake up.’ While its basic meaning is associated with a human activity, here the verb takes an inanimate subject, Huangyuan City. Given such violation of selection restriction (cf. Dorst 2011), the verb receives a code for indirect metaphor, added with a note of personification.

- (5) 在 黄原城 还 没有 睡醒
zài huángyuán-chéng hái méiyǒu shuì-xǐng
 in Huangyuan-city yet NEG wake.up
 之前, 东关 这个 旅社 的
zhīqián, Dōngguān zhègè lǚshè de
 before Dongguan this hotel ASSOC
 院子 里 就 一 片 熙熙攘攘 了 [K22]
yuànzi lǐ jiù yī piàn xīxīrǎngrǎng le
 yard inside then one CL bustling ASP
 ‘Before the Huangyuan City woke up, the yard of the Dongguan hotel was already bustling.’

To illustrate how the second category of MRWs, direct metaphor, can be manifested and annotated in Chinese texts, consider (6), taken from an academic text on the environmental benefits of forests. In this example, the author overtly compares forests to vacuum cleaners, via a direct use of the word *xīchénqì* 吸尘器 ‘vacuum cleaner,’ whose referent is clearly “incongruous” (Cameron 2003; Charteris-Black 2004) with its co-text, along with the metaphor flag *hǎobǐ* 好比 ‘can be compared to.’ Also, the topic shift here spans more than one lexical unit. In MIPVU, all content words within the stretch of directly used language expressing a cross-domain mapping are marked for direct metaphor. In this case, both the adjective *tiānrán* 天然 and the noun *xīchénqì* 吸尘器 were marked for direct metaphor. Example (7), drawn from a fiction text, showcases another flagged direct metaphor that involves an explicit comparison between a woman’s waist and hemp stalks. In this case, the word *mágǎn* 麻杆 ‘hemp stalk,’ followed by the metaphor signal *side* 似的 ‘-like,’ is used in its direct meaning to introduce an incongruous local referent into the discourse. In fact, the two metaphor flags in examples (6) and (7) belong to distinct lexical categories (i.e. verbs and particles, respectively) and thus differ in their relative position to the directly expressed linguistic metaphors.

- (6) 森林 好比 “天然 的 吸尘器”，
sēnlín *hǎobǐ* “*tiānrán* *de* *xīchénqì*”，
 forest can.be. natural NOM vacuum.cleaner
 compared.to
- 树叶， 特别 是 多 毛 的 树叶，
shùyè, *tèbié* *shì* *duō* *máo* *de* *shùyè*,
 leaf particularly COP plenty hair NOM leaf
 具有 很 好 的 吸 尘 能力 [J05]
jùyǒu *hěn* *hǎo* *de* *xī* *chén* *nénglì*
 have very good NOM absorb dust ability
- ‘Forests are just like “natural vacuum cleaners.” Leaves, particularly hairy leaves, have very good dust-absorbing ability.’

- (7) 沙 新 此时 忘 了 ,
Shā Xīn cǐshí wàng le,
 Sha Xin at.this.moment forget ASP
 当年 谈 恋爱 时 就 爱 她
dāngnián tán liànài shí jiù ài tā
 in.those. years talk love during just love 3sg
 那 麻秆 似的 细 腰 [K14]
nà mágǎn sìde xì yāo
 that hemp.stalk like slender waist
 ‘At this moment Xin Sha forgot that when they were going out, he simply liked her hemp stalk-like slender waist.’

According to the original MIPVU protocol for English, implicit metaphor can take two forms: by substitution or by ellipsis. To exemplify implicit meaning by substitution in Chinese discourse, consider example (8), which is excerpted from an academic text on the adaption of higher education institutions in China to the needs of the socialist market economy. In this case, the pronoun *tā* 它 in the last clause serves as a cohesive device and points to recoverable metaphorical material. That is, based on the co-occurring verb *fǎnyìng* 反映 ‘to reflect,’ the antecedent of the pronoun can be traced back to the nouns, *qíngyǔbiǎo* 晴雨表 ‘barometer’ and *jìngzi* 镜子 ‘mirror,’ both of which are metaphor-related words involved in the cross-domain comparison of higher education institutions to the two tools. Therefore, *tā* 它 receives a code for implicit metaphor. While theoretically Chinese also allows for implicit metaphor by ellipsis, we did not find any occurrence of this type in our data. It remains to be explored whether and, if so, how this form of implicit metaphor works in Chinese discourse.

- (8) 高校 是 社会 的 “晴雨表”，
gāoxiào shì shèhuì de “qíngyǔbiǎo”，
 higher. COP society ASSOC barometer
 education.
 institution
- 是 社会 的 一 面 镜子，
shì shèhuì de yī miàn jìngzi，
 COP society ASSOC one CL mirror
 它 灵敏地 反映 着 社会 [J44]
tā língmǐnde fǎnyìng zhe shèhuì
 3SG keenly reflect ASP society
- ‘Higher education institutions are a barometer and a mirror of the society; it keenly reflects the society.’

Three annotators were recruited in the manual annotation process. About one-fifth of the overall data were used for the inter-annotator reliability test, the focus of the next section. The rest of the data were annotated by the most experienced annotator and checked by at least one more annotator participating in the reliability experiment. Before we performed quantitative data analysis on the distribution of MRWs, all the files had undergone at least five passes of checking, which helped remedy most, if not all, inconsistencies in the annotation. Errors in word segmentation and POS tagging were also corrected. A post hoc process of correction was conducted to make sure that problematic or easily-overlooked words were treated consistently across the files.

4 Inter-annotator reliability test

The reliability of MIPVU as modified to apply to Chinese texts was evaluated by comparing annotations performed by three annotators independently. The annotators are native speakers of Mandarin Chinese and were doctoral students in the field of applied linguistics during the annotation period. One of the annotators was responsible for devising the annotation scheme and the other two novice annotators

Table 4 Results of the inter-annotator reliability test

	Academic (2004 words)	Fiction (2007 words)	News (2004 words)	Overall (6015 words)
Fleiss' Kappa	.846	.814	.838	.840

received multiple training sessions from the more experienced annotator to be familiarized with the purpose of the annotation as well as the annotation scheme before the test.

Approximately 2000 lexical units from each register (i.e. 20% of the overall data) were randomly selected from the 30,000-word corpus as samples for the reliability test. Since non-indirect metaphor (including direct metaphor, implicit metaphor and WIDLII) were rare in the samples, this study, following Steen et al. (2010a, b), only focused on the issue of whether a lexical unit is metaphor-related for the reliability test. Given the binary categorical distinction, Fleiss's kappa was adopted as the statistical measure of the extent of agreement among the three annotators. Table 4 presents the results of our inter-annotator reliability test.

The results show that metaphorically used words can also be reliably identified in Chinese texts with the aid of the adjusted version of MIPVU. The three annotators agreed on whether a word is metaphor-related for 96.5% of the cases. The mean Kappa value in our reliability test is comparable to the values reported by Steen et al. (2010a, b) for English and Dutch and by Badryzlova et al. (2013) for Russian. The majority of disagreement cases can be attributed to coder error and can be quickly resolved through discussion. In particular, lexical units whose contextual, non-literal meanings are much more frequent than their basic meanings are easily overlooked. Another type of MRWs easily overlooked are expressions that can be analyzed as involving either personified (metaphorical) or metonymic uses. As a case in point, consider the use of the word *zhǐchū* 指出 point-out 'to point out' in (9). The only and basic sense of the verb, according to CWN, refers to a human activity: “指明特定事件, 使他人瞭解” ‘to point out a specific matter in order to make others understand.’ The selection restriction of the verb that requires a human agent is violated by having a non-human entity, a book, in subject position. Therefore, the context here allows for double possibility of metaphor (the personification of the book) and metonymy (the book standing for its author). Following the original protocol of MIPVU, we also marked such ambiguous words as MRWs with a note of possible personification.

- (9) 《怀孕 指南》一 书 指出： 孕妇 在 怀孕 期间
huáiyùn zhǐnán yī shū zhǐchū yùnfù zài huáiyùn qījiān
 pregnancy guide one book point.out expectant. in pregnancy period
 mother

最 重要 的 是 必须 保持 精神 愉快 [K01]
zuì zhòngyào de shì bìxū bǎochí jīngshén yúkuài
 most important NOM COP have.to maintain spirit cheerful

‘The book *A Guide to Pregnancy* points out that the most important thing for expectant mothers during pregnancy is to have to maintain a happy spirit.’

We also examined the extent to which lexical units processed as individual words in LCMC but not found in CWN or the dictionaries could cause disagreement. Among the 6015 lexical units in the reliability test files, 228 (excluding numerals and proper nouns) were not listed in any of the dictionaries, representing 3.89% of the data. Out of the 228 tokens, only 7 cases (3.1%) exhibited disagreement among the three annotators. The main reason why these unfound words are mostly unproblematic for annotation is that their basic meanings are easily identifiable. Some are reduplicated words, like *tián-tián* 甜甜 sweet-sweet as used in *tián-tián-de-xiào* 甜甜地笑 sweet-sweet-ADV-smile ‘smile sweetly.’ The basic and contextual meanings of lexical units involving reduplication were determined according to the meaning definitions of its base form, in this case, *tián* 甜. Here, the reduplicated word is coded as an indirect MRW because its contextual ‘kind, pleasant’ meaning can be understood in comparison with its basic meaning of sugary taste (analogous to the usage of its English counterpart *sweet*).

The rest of the unfound words are mostly monosemous compounds, whose basic meanings can be easily determined as the sum of the basic meanings of individual components. Examples range from nominal compounds (e.g. *zhú-yǐ* 竹椅 bamboo-chair ‘bamboo chair’), to verb-object compounds (e.g. *jìn-chéng* 进城 enter-city ‘to enter the city’), to negated modals (*bù-néng* 不能 NEG-can ‘cannot’). The real problematic cases are those complex lexical units whose gestalt basic meanings are not compositional of their parts. To determine the basic sense of these lexical units, we analyzed concordance citations sampled from multiple balanced Chinese corpora (cf. Deignan 2015). One example is *xīn-zhōng* 心中, a space word composed of the noun *xīn* 心 referring to the body organ ‘heart’ and the localizer *zhōng* 中 meaning ‘inside.’ When the compound is used in a context like (10), one may consider it an MRW since its contextual meaning does not refer to the physical location of being inside one’s heart (organ). However, further concordance analysis indicated that the word here should not be annotated as metaphorical because its contextual meaning ‘at heart’ is the only possible interpretation of the compound and, therefore, is also its basic meaning.

- (10) 南浦 大桥 成为 上海 人民 心中 的 丰碑 [A33]
nánpǔ dà-qiáo chéngwéi Shànghǎi rénmin xīn-zhōng de fēngbēi
 Nanpu big-bridge become Shanghai people heart-inside NOM monument
 ‘Nanpu Bridge has become a monument in the heart of Shanghainese people.’

In short, the reliability test revealed largely smooth application of MIPVU to Chinese texts. For an insignificant proportion of words not found in the dictionaries and causing disagreement in annotation, these difficult cases were resolved through discussion based on analyses of concordance citations sampled from additional corpus data.

5 Corpus data analysis

In this section, we examine the distribution of metaphor-related words (MRWs) across the three written registers represented in the annotated Chinese corpus, namely, academic discourse, fiction, and news. We first present the overall distribution of all MRWs in the three registers, and then look more closely at the distribution of MRWs by lexical category.⁸ We also compare the distribution of MRWs in the Chinese corpus with that in the English corpus reported by Steen et al. (2010b).

5.1 Metaphor across registers

Of all 30,012 lexical units in the Chinese sample, only 3370 (11.2%) are related to metaphor. In other words, on average, approximately one in every nine words is metaphor-related. Table 5 summarizes the distribution of MRWs in the individual texts across the three registers represented in the Chinese corpus. On average, the academic texts contain the highest proportion of MRWs per sample (mean = 16.2%), while the news texts contain the lowest (mean = 8%). Levene's test revealed no significant difference in the variances of the proportion of MRWs among texts in the three registers (Levene statistic = 2.813, $p = .067$). A one-way analysis of variance revealed significant differences in the mean proportion of MRWs among texts in the three registers ($F(2,72) = 32.943$, $p < .001$). Post hoc pairwise comparisons using Tukey's honestly significant difference post hoc test showed significant differences in the mean proportions of MRWs between academic and fiction texts ($p < .001$) as well as between academic texts and news texts ($p < .001$). However, no significant difference was found between fiction and news texts ($p = .408$).

The finding that academic and fiction texts exhibit larger variance in metaphor density than news texts can be accounted for by their register-specific properties. The topics or themes of academic texts vary considerably in their level of abstractness and the authors' subjectivity in the writing. Texts on the relations between law and science, for example, would contain more MRWs than one on the history of paper. As for fiction texts, while they are often composed of a mixture of narration and dialogues, some use a lot more direct speech than others. Our sample shows that the fiction texts with more dialogues resemble conversation more (cf. Shutova et al. 2013; Steen et al. 2010b) and use fewer metaphor-related words. Finally, the original texts in the academic and fiction registers are much longer than the new texts in LCMC. Their greater length may also privilege the use of extended metaphor, in which the same groups of metaphor-related words would repeatedly occur in the sample, increasing the metaphor density of particular texts.

Table 6 summarizes the distribution of MRWs in the Chinese corpus and in the VU Amsterdam Metaphor Corpus (VUAMC) by register, with all samples in each register analyzed as one large sample. This is done to facilitate comparison with results reported by Steen et al. (2010b). A Chi square analysis showed a significant

⁸ In all analyses, tokens tagged as MFlags are not considered MRWs, but those tagged as WIDLLI are so that unwarranted exclusion of linguistic metaphors can be avoided.

Table 5 Mean proportion of MRWs per sample by register in the Chinese corpus

	Academic	Fiction	News
Total number of samples	30	20	25
Mean number of tokens per sample	333.4	500.2	400.2
Mean number of MRWs per sample	54.43	47.45	31.52
Mean proportion of MRWs per sample	.162	.096	.080
Standard deviation of mean proportion	.045	.045	.028

Table 6 Comparison of distribution of MRWs in the Chinese corpus and VUAMC by register

	Academic	Fiction	News
Chinese corpus			
Total number of MRWs	1633	949	788
Total number of tokens	10,002	10,004	10,006
Proportion of MRWs	.163	.095	.079
VUAMC			
Total number of MRWs	9120	5293	7342
Total number of tokens	49,314	44,648	44,792
Proportion of MRWs	.185	.119	.164
Log likelihood ratio	22.14*	42.09*	465.46*

* Denotes $p < .0001$

difference in the proportion of MRWs across the three different registers in the Chinese corpus ($\chi^2(2) = 404.143$, $p < .0001$). Pairwise Chi square analyses with Bonferroni correction (which corrects the alpha value to $.05/3 = .0167$) further revealed significant differences between academic texts and fiction ($\chi^2 = 208.207$, $p < .0001$), between academic texts and news ($\chi^2 = 335.914$, $p < .0001$), as well as between fiction and news ($\chi^2 = 16.377$, $p < .0001$). These results provide additional evidence for cross-register variation in the distribution of MRWs. Steen et al. (2010b) also reported significant cross-register variation in the distribution of MRWs. Similar to the results from our Chinese corpus, the results from VUAMC also show a significant difference between academic texts and fiction ($\chi^2 = 795.265$, $p < .0001$) and between academic texts and news texts ($\chi^2 = 73.884$, $p < .0001$). However, while in the Chinese corpus news texts have the lowest proportion of MRWs, in VUAMC they have a significantly higher proportion of MRWs than fiction ($\chi^2 = 379.302$, $p < .0001$). In addition to this cross-linguistic difference in cross-register variation in the distribution of MRWs, the log-likelihood ratios reported in the last row of Table 6 also indicate that the proportion of MRWs is significantly higher in VUAMC than in the Chinese corpus for all three registers, with the largest difference found in news texts.

Table 7 summarizes the distribution of different types of MRWs in the Chinese corpus by register. A Chi square analysis showed a significant difference in the

Table 7 Lexical units in relation to metaphor, divided by register

	Academic	Fiction	News	Total
Indirect metaphor	1597	835	747	3179
Direct metaphor	9	67	19	95
Implicit metaphor	3	6	0	9
WIDLII	24	41	22	87
Total MRWs	1633	949	788	3370
MFlags	4	25	12	41
Non-MRWs	8365	9030	9206	26,601
Total tokens	10,002	10,004	10,006	30,012

proportion of different types of MRWs across the three different registers ($\chi^2(6) = 123.291$, $p < .0001$). Academic texts have a significantly higher proportion of indirect metaphors than both fiction ($\chi^2 = 105.518$, $p < .0001$) and news ($\chi^2 = 15.519$, $p < .0001$), and news has a significantly higher proportion of indirect metaphors than fiction ($\chi^2 = 24.564$, $p < .0001$). Fiction has a significantly higher proportion of direct metaphor than both academic texts ($\chi^2 = 89.009$, $p < .0001$) and news ($\chi^2 = 19.772$, $p < .0001$), and news has a significantly higher proportion of direct metaphor than academic texts ($\chi^2 = 16.086$, $p < .0001$) as well. Implicit metaphors are rare in all three registers. While relatively few lexical units are coded as WIDLII in all three registers, academic texts have a lower portion of such cases than fiction ($\chi^2 = 19.875$, $p < .0001$) and news ($\chi^2 = 4.985$, $p = .026$).

These results are largely consistent with the results Steen et al. (2010b) reported for VAUMC in several ways. First, in both the Chinese corpus and VAUMC, indirect metaphor is the dominant type of metaphor in all registers and is more frequent in academic texts than in news and fiction. Second, in both corpora, direct metaphors are used more frequently in fiction and news than in academic texts. Third, in both corpora, a higher proportion of direct metaphor is found to be accompanied by a higher proportion of MFlags. However, implicit metaphor appears to be more frequent in VUAMC than in the Chinese corpus. The relative low frequency of implicit metaphor, expressed by substitution only, in Chinese texts can be partly attributed to the salient feature of zero anaphora in Chinese grammar (cf. Li and Thompson 1981; Tao 1996). In other words, pronouns substituting for metaphorically used nominal lexical units tend to be grammatically omitted in Chinese discourse if the referents can be understood from the context.

5.2 Metaphor across lexical categories

Table 8 summarizes the overall distribution of MRWs by lexical category in the Chinese corpus, without differentiation of register. A Chi square analysis indicated significant differences in the proportions of tokens that are MRWs across different lexical categories ($\chi^2(8) = 2,506.195$, $p < .0001$). As the results show, verbs account for the greatest proportion (39.9%) of all MRWs in the corpus, followed by

Table 8 Overall distribution of MRWs by lexical category in the Chinese corpus

	Tokens	MRWs	Proportion
Verbs	7234	1343	.186
	.241	.399	
Nouns	8046	542	.067
	.268	.161	
Prepositions	1398	406	.290
	.047	.120	
Localizers	585	286	.489
	.019	.085	
Pronouns ^a	1942	232	.119
	.065	.069	
Adjectives	1367	227	.166
	.046	.067	
Classifiers	823	97	.118
	.027	.029	
Adverbs	2228	15	.007
	.074	.004	
Others	6389	222	.035
	.213	.066	
Total	30,012	3370	.112

For each lexical category, proportion = MRWs/tokens, the percentage under each token count = tokens/total tokens, and the percentage under each MRW count = MRWs/total MRWs

^a In LCMC, lexical units used as determiners are also tagged as pronouns

nouns (16.1%), prepositions (12.0%), and localizers (8.5%). However, the lexical categories with the highest proportions of MRWs are localizers (48.9%) and prepositions (29.0%).

The finding that prepositions, localizers, and verbs are more frequently related to metaphor in Chinese texts than lexical units in other parts of speech is consistent with what Cameron (2003) has found in her English data. Specifically, in terms of their grammatical forms, 47% and 34% of linguistic metaphors in spoken classroom discourse are verb- and preposition-metaphors, respectively (Cameron 2003, p. 89). Linguistic metaphors in the form of prepositions in English may involve two word classes in their Chinese counterparts, namely, prepositions and localizers. Specifically, locative phrases in Chinese generally use both the preposition *zài* 在 and localizers (e.g. *shàng* 上 ‘above, on top of,’ *xià* 下 ‘under, below,’ *qián* 前 ‘in front of,’ and *hòu* 后 ‘in back of, behind’). A locative phrase with a figurative use of preposition in English (e.g. *in this process*) would thus involve not only a metaphorically used preposition but also a non-literal localizer in its Chinese counterpart (i.e. *zài-zhègè-guòchéng-zhōng* 在这个过程中 at-this-process-amid). It is no surprise that both localizers and prepositions are highly metaphorical in Chinese texts.

Table 9 Distribution of MRWs by lexical category and register in the Chinese corpus

	Academic			Fiction			News		
	Tokens	MRWs	Proportion	Tokens	MRWs	Proportion	Tokens	MRWs	Proportion
Verbs	2408	615	.255	2591	420	.162	2235	308	.138
	.241	.377		.259	.443		.223	.391	
Nouns	3059	301	.098	2104	129	.061	2883	112	.039
	.306	.184		.210	.136		.288	.142	
Prepositions	568	247	.435	402	62	.154	428	97	.227
	.057	.151		.040	.065		.043	.123	
Localizers	192	145	.755	193	66	.342	200	75	.375
	.019	.089		.019	.070		.002	.095	
Pronouns ^a	383	107	.279	1101	76	.069	458	49	.107
	.038	.066		.110	.080		.046	.062	
Adjectives	567	109	.192	431	64	.148	369	54	.146
	.057	.067		.043	.067		.037	.069	
Classifiers	111	25	.225	284	34	.120	428	38	.089
	.011	.015		.028	.036		.043	.048	
Adverbs	603	5	.008	961	4	.004	664	6	.009
	.060	.003		.096	.004		.066	.008	
Others	2111	79	.037	1937	94	.049	2341	49	.021
	.211	.048		.194	.099		.234	.062	
Total	10,002	1633	.163	10,004	949	.095	10,006	788	.079

For each lexical category, proportion = MRWs/tokens, the percentage under each token count = tokens/total tokens, and the percentage under each MRW count = MRWs/total MRWs

^a In LCMC, lexical units used as determiners are also tagged as pronouns

Table 9 summarizes the distribution of MRWs by lexical category in the three different registers in the Chinese corpus. Compared to academic texts, both news and fiction have a higher proportion of MRWs that are verbs or classifiers but a lower proportion of MRWs that are nouns and prepositions. In addition, compared to fiction, news texts have a lower proportion of MRWs that are verbs or pronouns but a higher proportion of MRWs that are nouns, prepositions, localizers, adjectives, or classifiers. What is especially noteworthy is that in academic texts, 75.5% of localizers, 43.5% of prepositions, and 27.9% of pronouns are MRWs. These proportions are all substantially higher than the corresponding proportions in fiction and news texts. In particular, the frequent metaphor-related use of demonstrative pronouns and determiners (i.e. *zhè* 这 ‘this’ and *nà* 那 ‘that’) can be linked to their crucial role as cohesive devices for abstract referents in academic prose (Gray 2010).

We will now take a closer look at the distribution of MRWs by lexical category in each of the three registers and compare our results against those from VUAMC reported by Steen et al. (2010b). For each register, a Chi square analysis shows significant differences in the proportions of tokens that are MRWs across the different lexical categories (for academic texts, $\chi^2(8) = 1438.122$, $p < .0001$; for

fiction, $\chi^2(8) = 483.297$, $p < .0001$; for news texts, $\chi^2(8) = 722.883$, $p < .0001$). These results are in line with the significant association between lexical category and relation to metaphor in the same three registers in VUAMC reported by Steen et al. (2010b). However, there are several differences between the Chinese corpus and VUAMC in terms of the distribution of MRWs by lexical category in each of the three registers. In academic texts, the largest group of lexical units related to metaphor in VUAMC is the group of prepositions, followed by nouns and verbs, whereas in the Chinese corpus, this is the group of verbs, followed by nouns and prepositions. Prepositions are also the most often metaphorical in VUAMC, with 42.5% of them related to metaphor. In the Chinese corpus, localizers are the most often metaphorical instead, with 75.5% of them related to metaphor. For the other two registers, in VUAMC, the top three categories with the highest number of MRWs in both fiction and news are verbs, prepositions, and nouns. In the Chinese corpus, the top three categories are verbs, nouns, and pronouns in fiction and verbs, nouns, and prepositions in news texts. Also in both registers, it is prepositions in the VUAMC and localizers in the Chinese corpus that have the highest percentage of MRWs.

6 Conclusions

As noted by the developers of MIP(VU) (e.g. Pragglejaz Group 2007; Steen et al. 2010a, b) and researchers who have applied MIP(VU) (e.g. Nacey 2013), a major challenge of adopting the procedure in metaphor research is the labor-intensive, time-consuming nature of the annotation process. Such weakness is nonetheless outweighed by the strengths of the procedure. The present study confirms that MIPVU, originally devised for identifying linguistic metaphor in English discourse, can be flexibly adjusted and reliably applied to Chinese texts. How the unit of analysis, ‘lexical unit’, is operationalized needs to be adapted to Chinese-specific lexical types (e.g. verb-object compounds and resultative verb compounds as well as reduplicated words) based on the word segmentation tool and part-of-speech tagging provided by the corpus available. Special treatments for the meaning definitions provided by the Chinese lexical tools are called for when sense conflation is suspected. Analyses of concordance citations sampled from multiple Chinese corpora may be required to determine the basic meanings of problematic cases whose sense descriptions in the Chinese WordNet and/or the reference dictionaries are inadequate or unavailable.

More importantly, building a corpus comprised of texts with manual annotation of metaphor enables the researchers to investigate a set of research questions pertinent to the three-way relationship between metaphor density, lexical category, and register. This study corroborates previous research pointing out that among the three written registers under examination, academic texts have the highest density of metaphor-related words. The finding that Chinese verbs, prepositions and localizers feature prominently as metaphor-related words in the corpus provides further support for earlier findings regarding English discourse. A closer scrutiny of the data also reveals the connection between different types of metaphor and register. The

pattern that indirect metaphors are more common in academic texts whereas direct metaphors are more frequent in fiction has been found in both English and Chinese texts. Finally, some lexical categories may be especially susceptible to metaphor use in a particular register. For example, the pronouns and determiners in academic texts are more likely to be related to metaphor, given that these constituents are more commonly used as cohesive devices to refer to abstract entities in academic discourse. The findings of this study can thus complement the fruitful results of previous research on the cross-domain mappings of specific metaphorical expressions in Chinese.

It follows that annotating linguistic metaphors in discourse based on the variants of the same procedure, namely, MIPVU, allows researchers to examine metaphor in language use from a cross-linguistic perspective. One of the findings (albeit tentative) of this study is that metaphor density appears to be much lower in Chinese (11.2%) than in English written texts (15.68%). The size of the corpus needs to be expanded in the future to see if our preliminary results still hold. In addition, natural conversation texts are preferably included to explore whether there are quantitative and/or qualitative differences in metaphor use between written and spoken Chinese. Given the success of our adjusted application of MIPVU to Chinese texts, English-Chinese parallel corpus data can also be annotated in the future to systematically examine the relationship between specific discourse features of Chinese and metaphor uses (e.g. the pervasiveness of zero anaphora in Chinese and the frequency of implicit metaphor). The metaphor-annotated Chinese corpus also makes it possible to compare metaphor density in English and Chinese texts using alternative measures (e.g. by the proportion of sentences containing MRWs) to paint a more comprehensive picture of the similarities and differences in the frequency and distribution of linguistic metaphor in the two languages.

Due to opaque word boundaries in Chinese, the same strings of characters may be tokenized differently by distinct word segmentation tools, and such discrepancies can systematically lead to different annotation results. The methodological issues discussed and the solutions provided in this paper for the adaptation of MIPVU to Chinese, given the lexical resources adopted in our project, are by no means exhaustive or conclusive. It is also important to keep in mind that MIPVU is designed to capture metaphorical uses at the word level; the present metaphor-annotated Chinese corpus thus does not offer information about metaphorical meanings below or beyond the word level, such as those of characters. Still, like the VU Amsterdam Metaphor Corpus, the corpus presented here can no doubt serve as an important language resource for researchers to investigate patterns of metaphor in usage that would otherwise be difficult to unveil without such full-text annotations. Ultimately, as pointed out by Veale et al. (2016), the texts manually annotated for metaphor will likely be used as training data for developing machine learning systems for automated metaphor detection, contributing to the advancement of computational approaches to figurative language.⁹

⁹ Work on automated metaphor detection using the VUAMC is already under way (e.g. Do Dinh and Gurevych 2016; Dunn 2013; Haagsma and Bjerva 2016).

Acknowledgements We are grateful to the anonymous reviewers for their insightful comments and suggestions on earlier drafts of this manuscript. This project was funded by a Gil Watz Early Career Professorship in Language and Linguistics to the first author. Special thanks go to Chan-Chia Hsu and Eric Po-Chung Lin for their participation in the inter-annotator reliability test. We would also like to thank Haiyang Ai for his assistance at the early stage of this project. Any remaining inadequacies are our sole responsibility.

References

- Ahrens, K., Chung, S.-F., & Huang, C.-R. (2003). Conceptual metaphors: Ontology-based representation and corpora driven mapping principles. In *Proceedings of the ACL 2003 workshop on the lexicon and figurative language* (pp. 36–42). Stroudsburg, PA: Association for Computational Linguistics.
- Ahrens, K., Chung, S.-F., & Huang, C.-R. (2004). From lexical semantics to conceptual metaphors: Mapping principle verification with WordNet and SUMO. In *Proceedings of the fifth Chinese lexical semantics workshop* (pp. 99–106). Singapore: COLIPS.
- Badryzlova, Y., Shekhtman, N., Isaeva, Y., & Kerimov, R. (2013). Annotating a Russian corpus of conceptual metaphor: A bottom-up approach. In *Proceedings of the first workshop on metaphor in NLP* (pp. 77–86). Stroudsburg, PA: Association for Computational Linguistics.
- Cameron, L. (2003). *Metaphor in educational discourse*. New York/London: Continuum.
- Chao, Y. R. (1968). *A grammar of spoken Chinese*. Berkeley: University of California Press.
- Charteris-Black, J. (2004). *Corpus approaches to critical metaphor analysis*. New York: Palgrave Macmillan.
- Chen, K.-J., & Bai, M.-H. (1998). Unknown word detection for Chinese by a corpus-based learning method. *Computational Linguistics and Chinese Language Processing*, 3(1), 27–44.
- Chen, K.-J., & Liu, S.-H. (1992). Word identification for Mandarin Chinese sentences. In *Proceedings of the fourteenth conference on computational linguistics* (pp. 101–107). Stroudsburg, PA: Association for Computational Linguistics.
- Chiang, W.-Y., & Duann, R.-F. (2007). Conceptual metaphors for SARS: ‘War’ between whom? *Discourse and Society*, 18(5), 579–602.
- Chiu, S.-H., & Chiang, W.-Y. (2011). FIGHT metaphors in legal discourse: What is unsaid in the story? *Language and Linguistics*, 12(4), 877–915.
- Chung, S.-F., Ahrens, K., & Huang, C.-R. (2005). Source domains as concept domains in metaphorical expressions. *Computational Linguistics and Chinese Language Processing*, 10(4), 553–570.
- Chung, S.-F., & Huang, C.-R. (2010). Using collocations to establish the source domains of conceptual metaphors. *Journal of Chinese Linguistics*, 38(2), 183–223.
- Deignan, A. (2005). *Metaphor and corpus linguistics*. Amsterdam/Philadelphia: John Benjamins.
- Deignan, A. (2015). MIP, the corpus, and dictionaries: What makes the best for metaphor analysis? *Metaphor and the Social World*, 5(1), 145–154.
- Do Dinh, E.-L., & Gurevych, I. (2016). Token-level metaphor detection using neural networks. In *Proceedings of the fourth workshop on metaphor in NLP* (pp. 28–33). Stroudsburg, PA: Association for Computational Linguistics.
- Dorst, A. G. (2011). Personification in discourse: Linguistic forms, conceptual structures and communicative functions. *Language and Literature*, 20(2), 113–135.
- Duann, R.-F., & Huang, C.-R. (2015). When embodiment meets Generative Lexicon: The human body part metaphors in Sinica Corpus. In *Proceedings of the twenty-ninth Pacific Asia conference on language, information and computation* (pp. 396–403).
- Dunn, J. (2013). What metaphor identification systems can tell us about metaphor-in-language. In *Proceedings of the first workshop on metaphor in NLP* (pp. 1–10). Stroudsburg, PA: Association for Computational Linguistics.
- Gibbs, R. W. (Ed.). (2008). *The Cambridge handbook of metaphor and thought*. Cambridge/New York: Cambridge University Press.
- Goatly, A. (1997). *The language of metaphors*. New York: Routledge.
- Gong, S.-P., Ahrens, K., & Huang, C.-R. (2008). Chinese Word Sketch and mapping principles: A corpus-based study of conceptual metaphors using the BUILDING source domain. *International Journal of Computer Processing of Languages*, 21(1), 3–17.

- Gray, B. (2010). On the use of demonstrative pronouns and determiners as cohesive devices: A focus on sentence-initial this/these in academic prose. *Journal of English for Academic Purposes*, 9(3), 167–183.
- Haagsma, H., & Bjerva, J. (2016). Detecting novel metaphor using selectional preference information. In *Proceedings of the fourth workshop on metaphor in NLP* (pp. 10–17). Stroudsburg, PA: Association for Computational Linguistics.
- Han, C. (2014). *Metaphor and entertainment: A corpus-based approach to language in Chinese online news*. New York: Palgrave Macmillan.
- Hsieh, S. C.-Y. (2006). A corpus-based study on animal expressions in Mandarin Chinese and German. *Journal of Pragmatics*, 38, 2206–2222.
- Huang, C.-R., & Ahrens, K. (2003). Individuals, kinds and events: Classifier coercion of nouns. *Language Sciences*, 25(4), 353–373.
- Huang, C.-R., Chen, K.-J., Chen, F.-Y., & Chang, L.-L. (1997). Segmentation standard for Chinese natural language processing. *Computational Linguistics and Chinese Language Processing*, 2(2), 47–62.
- Huang, C.-R., Hsieh, S.-K., Hong, J.-F., Chen, Y.-Z., Su, I.-L., Chen, Y.-X., et al. (2010). 中文词汇网络：跨语言知识处理基础架构的设计理念与实践 [Chinese WordNet: Design, implementation, and application of an infrastructure for cross-lingual knowledge processing]. 中文信息学报 [Journal of Chinese Information Processing], 24(3), 14–23.
- Huang, C.-R., & Xue, N. (2015). Modeling word concepts without convention: linguistic and computational issues in Chinese word identification. In W. S.-Y. Wang & C.-F. Sun (Eds.), *The Oxford handbook of Chinese linguistics* (pp. 348–361). New York: Oxford University Press.
- Jing-Schmidt, Z., & Peng, X. (2017). Winds and tigers: Metaphor choice in China's anti-corruption discourse. *Lingua Sinica*. doi:10.1186/s40655-016-0017-9.
- Kövecses, Z. (2005). *Metaphor in culture: Universality and variation*. Cambridge/New York: Cambridge University Press.
- Kövecses, Z. (2010). *Metaphor: A practical introduction* (2nd ed.). New York: Oxford University Press.
- Lakoff, G., & Johnson, M. (1980/2003). *Metaphors we live by*. Chicago: University of Chicago Press.
- Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. New York: Basic Books.
- Lakoff, G., & Turner, M. (1989). *More than cool reason: A field guide to poetic metaphor*. Chicago/London: University of Chicago Press.
- Lee, L.-H., Hsieh, S.-K., & Huang, C.-R. (2009). CWN-LMF: Chinese WordNet in the lexical markup framework. In *Proceedings of the seventh workshop on Asian language resources* (pp. 123–130). Stroudsburg, PA: Association for Computational Linguistics.
- Li, C. N., & Thompson, S. A. (1981). *Mandarin Chinese: A functional reference grammar*. Berkeley: University of California Press.
- Lu, X. (2007). A hybrid model for Chinese word segmentation. *Journal for Language Technology and Computational Linguistics*, 22(1), 71–88.
- Lu, X. (2008). Improving part-of-speech guessing of Chinese unknown words using hybrid models. *International Journal of Corpus Linguistics*, 13(2), 169–193.
- Lu, L. W.-L., & Ahrens, K. (2008). Ideological influence on BUILDING metaphors in Taiwanese presidential speeches. *Discourse and Society*, 19(3), 383–408.
- McEnery, T., & Xiao, R. (2004). The Lancaster Corpus of Mandarin Chinese: A corpus for monolingual and contrastive language study. In M. Lino, M. Xavier, F. Ferreira, R. Costa & R. Silva (Eds.), *Proceedings of the fourth international conference on language resources and evaluation* (pp. 1175–1178). Lisbon: ELRA.
- Nacey, S. (2013). *Metaphors in learner English*. Amsterdam/Philadelphia: John Benjamins.
- Packard, J. L. (2000). *The morphology of Chinese: A linguistic and cognitive approach*. Cambridge/New York: Cambridge University Press.
- Partington, A. (2006). Metaphors, motifs and similes across discourse types: Corpus-Assisted Discourse Studies (CADS) at work. In A. Stefanowitsch & S. T. Gries (Eds.), *Corpus-based approaches to metaphor and metonymy* (pp. 267–304). Berlin/New York: Mouton de Gruyter.
- Pasma, T. (2012). Metaphor identification in Dutch discourse. In F. MacArthur, J. L. Oncins-Martínez, M. Sánchez-García, & A. M. Piquer-Piriz (Eds.), *Metaphor in use: Context, culture, and communication* (pp. 69–83). Amsterdam/Philadelphia: John Benjamins.
- Pragglejaz Group. (2007). MIP: A method for identifying metaphorically used words in discourse. *Metaphor and Symbol*, 22(1), 1–39.

- Ross, C. (1990). Resultative verb compounds. *Journal of the Chinese Language Teachers Association*, 25 (3), 61–83.
- Semino, E., Heywood, J., & Short, M. (2004). Methodological problems in the analysis of metaphors in a corpus of conversations about cancer. *Journal of Pragmatics*, 36(7), 1271–1294.
- Shi, Y. (2002). *The establishment of modern Chinese grammar: The formation of the resultative construction and its effects*. Amsterdam/Philadelphia: John Benjamins.
- Shutova, E., Devereux, B. J., & Korhonen, A. (2013). Conceptual metaphor theory meets the data: A corpus-based human annotation study. *Language Resources and Evaluation*, 47(4), 1261–1284.
- Steen, G. J. (2007). *Finding metaphor in grammar and usage*. Amsterdam/Philadelphia: John Benjamins.
- Steen, G. J., Dorst, A. G., Herrmann, J. B., Kaal, A. A., & Krennmayr, T. (2010a). Metaphor in usage. *Cognitive Linguistics*, 21(4), 765–796.
- Steen, G. J., Dorst, A. G., Herrmann, J. B., Kaal, A. A., Krennmayr, T., & Pasma, T. (2010b). *A method for linguistic metaphor identification: From MIP to MIPVU*. Amsterdam/Philadelphia: John Benjamins.
- Stefanowitsch, A. (2006). Corpus-based approaches to metaphor and metonymy. In A. Stefanowitsch & S. T. Gries (Eds.), *Corpus-based approaches to metaphor and metonymy* (pp. 1–16). Berlin/New York: Walter de Gruyter.
- Stefanowitsch, A., & Gries, S. T. (Eds.). (2006). *Corpus-based approaches to metaphor and metonymy*. Berlin/New York: Mouton de Gruyter.
- Sun, C. (2006). *Chinese: A linguistic introduction*. New York: Cambridge University Press.
- Sun, M., & Zou, J. (2001). 汉语自动分词研究评述 [A critical appraisal of the research on Chinese word segmentation]. *当代语言学* [Contemporary Linguistics], 3(1), 22–32.
- Tai, J. H.-Y. (1994). Chinese classifier systems and human categorization. In M. Chen & O. Tzeng (Eds.), *In honor of William S.-Y. Wang: Interdisciplinary studies on language and language change* (pp. 479–494). Taipei: Pyramid Press.
- Tai, J. H.-Y., & Chao, F.-Y. (1994). A semantic study of the classifier zhang. *Journal of the Chinese Language Teachers Association*, 29(3), 67–78.
- Tao, L. (1996). Topic discontinuity and zero anaphora in Chinese discourse. In B. A. Fox (Ed.), *Studies in anaphora* (pp. 487–513). Amsterdam/Philadelphia: John Benjamins.
- Tay, D. (2015). Metaphor in case study articles on Chinese university counseling service websites. *Chinese Language and Discourse*, 6(1), 28–56.
- Thompson, S. A. (1973). Resultative verb compounds in Mandarin Chinese: A case for lexical rules. *Language*, 49(2), 361–379.
- Veale, T., Shutova, E., & Beigman Klebanov, B. (2016). *Metaphor: A computational perspective*. San Rafael, CA: Morgan & Claypool Publishers.
- Xiao, R. (2017). Lancaster Corpus of Mandarin Chinese. In R. Sybesma (Ed.), *Encyclopedia of Chinese language and linguistics*. Amsterdam: Brill.
- Yu, N. (1998). *The contemporary theory of metaphor: A perspective from Chinese*. Amsterdam/Philadelphia: John Benjamins.
- Yu, N. (2008). Metaphor from body and culture. In R. W. Gibbs (Ed.), *The Cambridge handbook of metaphor and thought* (pp. 247–261). Cambridge/New York: Cambridge University Press.
- Yu, N., & Jia, D. (2016). Metaphor in culture: LIFE IS A SHOW in Chinese. *Cognitive Linguistics*, 27(2), 147–180.
- Zhang, H. (2007). Numeral classifiers in Mandarin Chinese. *Journal of East Asian Linguistics*, 16(1), 43–59.
- Zhang, H.-P., & Liu, Q. (2002). 基于 N-最短路径方法的中文词语粗分模型 [Model of Chinese words rough segmentation based on N-shortest-paths method]. *中文信息学报* [Journal of Chinese Information Processing], 16(5), 1–7.

Chinese Dictionaries

- Commercial Press Dictionary Research Centre. (2010). *Xiandai Haiyu Xuexi Cidian* [A Learner's Dictionary of Modern Chinese]. Beijing: Commercial Press.
- Li, X. (Ed.). (2014). *Xiandai Hanyu Guifan Cidian* (3rd ed.) [A Standard Dictionary of Modern Chinese]. Beijing: Foreign Language Teaching and Research Press.

Chinese Lexical Resources

Academia Sinica Balanced Corpus of Modern Chinese (Sinica Corpus). <http://asbc.iis.sinica.edu.tw/>.

Chinese WordNet. <http://cwn.ling.sinica.edu.tw/>.

Chinese National Corpus. <http://www.encorpus.org/index.aspx>.

The Lancaster Corpus of Mandarin Chinese. <http://www.lancaster.ac.uk/fass/projects/corpus/LCMC/>.