

International Journal of Computer Processing of Oriental Languages  
© Chinese Language Computer Society &  
World Scientific Publishing Company

## **Chinese Word Sketch and Mapping Principles: A Corpus-Based Study of Conceptual Metaphors Using the BUILDING Source Domain**

SHU-PING GONG

*Graduate Institute of Linguistics  
National Taiwan University  
No.1, Roosevelt Road, Section 4, Taipei 106, Taiwan  
[d91142001@ntu.edu.tw](mailto:d91142001@ntu.edu.tw)*

KATHLEEN AHRENS\*

*Graduate Institute of Linguistics & DFFL  
National Taiwan University  
No.1, Roosevelt Road, Section 4, Taipei 106, Taiwan  
[kathleenahrens@yahoo.com](mailto:kathleenahrens@yahoo.com)*

CHU-REN HUANG

*Institute of Linguistics  
Academia Sinica  
No. 128, Sec. 2, Academia Road, Nankang, Taipei 115, Taiwan  
[churen@sinica.edu.tw](mailto:churen@sinica.edu.tw)*

Received (Day Month Year)  
Revised (Day Month Year)  
Accepted (Day Month Year)

This study examines whether Mapping Principles, which govern how concrete concepts in a source domain are mapped to abstract concepts in a target domain, are realized in large-scale corpus data. In particular, we propose a frequency-based collocational approach to determine mapping principles based on the most productive mapping. In addition, in contrast to the target-domain-oriented method used in previous studies, we employ the source-domain-oriented method to search for lexical mappings from a particular source domain, i.e. BUILDING, to a number of target domains. The corpus data demonstrates that mapping principles exist in a source-target domain pairing. In addition, we found that different target domains select the source domain of BUILDING

---

\* Corresponding author: Graduate Institute of Linguistics, National Taiwan University, No.1, Roosevelt Road, Section 4, Taipei 106, Taiwan.

## 2 Gong et al.

for different underlying reasons. Our study follows the lexical analysis of conceptual metaphors proposed by the Conceptual Mapping Model and helps us better understand how semantic networks of lexical words are represented in the lexicon via conceptual mappings.

*Keywords:* Mapping Principles; Frequency; Collocation, Source/Target Domain, Properties

### 1. Introduction

Recently, there has been growing interest in examining how to identify and determine conceptual metaphors [1, 2,] source domains [3] and lexical mappings between a source domain and a target domain. In particular, Ahrens [4, 5] has proposed an intuition-based approach to determine lexical mappings in a source-target domain pairing by analyzing metaphorical expressions in terms of entities, qualities, and functions that can map between a source and a target domain. This approach allows us to uncover the underlying reason that the source domain (i.e. the concrete concept) is mapped conceptually to the target domain (i.e. the abstract concept) and to use the most salient meaning to determine the mapping principles. For example, in the two metaphorical expressions for AN IDEA IS A BUILDING, listed in (1) and (2), 架構 "frame", 鬆散 "loose" and 根基 "base" can be analyzed as referring to the structure of a building. Thus, the underlying reason, or Mapping Principle (MP) of AN IDEA IS A BUILDING, postulated is: *an idea is understood as a building because a building involves a (physical) structure and ideas involve (abstract) organization.*

- (1) 這個 理論 架構 很 鬆散。  
this theory frame very loose  
"The frame of this theory is very loose."
- (2) 你的 論點 根基 是 什麼？  
your argument base be what  
"What is the foundation of your argument?"

Discovering the mapping principles in a source-target domain pairing through linguistic data analysis allows predictions to be made concerning the processing of conventional and novel metaphors [4, 5]. Off-line psycholinguistic studies have demonstrated that novel metaphors that follow MPs were rated more acceptable and interpretable than conventional metaphors. Additionally, novel

metaphors that do not follow MPs were rated less acceptable and interpretable than novel metaphors that do follow MPs.

This intuition-based method, however, is difficult to apply to the cases when a conceptual metaphor has several lexical mappings but each mapping is equally salient. For example, the two collocates, i.e. 灌溉 "to irrigate" and 萌芽 "to sprout" of the source domain of PLANT are mapped to the abstract concept of love in Mandarin Chinese. In this case, it is not clear whether the mapping principle has to do with the notion of irrigation or with the concept of sprouting because we cannot tell which mapping is more salient.

A corpora-based method has been proposed [6, 7, 8] to overcome the problem that the intuition-based method encounters. In particular, Ahrens et al. [6, 7] integrated the Conceptual Mapping Model [4, 5] with an ontology-based knowledge representation (i.e. SUMO) and WordNet to determine mapping principles between target-source domain pairings. They proposed that each source-target domain pairing has a prototypical instance of mapping as indicated by a lexical item that is frequently mapped compared to other mappings. In particular, Ahrens et al. [7] defined two numerical criteria for determining the MP. The first is that the minimum number of tokens for the most frequent mapping is ten metaphoric instances. The second is that for the most frequent mapping to reach consensus for a mapping principle requires 30% of the tokens. For example, 38% (i.e. 39 tokens) of 102 instances are mapped with the lexical item 建設 "construction" for the metaphor ECONOMY IS A BUILDING [7]. The most frequent mapping 建設 "construction" is higher than 10 tokens, i.e. 39 tokens, and the proportion of this mapping is also higher than 30%, i.e. 38%. So, the MP of ECONOMY IS A BUILDING postulated is: *Economy is a building because buildings involve a structure and economy involves an (abstract) structure.*

One frequently used way to search for lexical mappings in source-target domain pairings has been to use the target-domain-oriented method. Previous studies [4, 5] have narrowed a target domain to a particular lexical item, e.g. 經濟 economy, and then searched for lexical mappings from a target domain to a number of source domains [4, 5, 6, 7]. However, this target-domain-oriented method does not explain how and why a particular source domain maps to different target domains. Ahrens [4, 5] found that the different aspects of a source domain are selected by a number of target domains. For example, her studies [4, 5] show that both lexical items "ideas" and "love" are described in terms of the source domain of FOOD, but the two concepts select distinct aspects of FOOD for their Mapping Principles. The aspect of digestion is selected by IDEA while the property of taste is selected by LOVE. This comparison suggests that even

though a particular source domain may be repeatedly mapped to different target domains, different aspects of the source domain are selected by target domains for distinct underlying reasons. Few corpus-based studies, however, investigate the issue of what properties a target domain selects from a source domain. In this research, we employ the source-domain-oriented method, i.e. searching for metaphors from the source domain to extract lexical mappings, which allows us to expand the number of target domains that use the same source domain and to better understand why different target domains select the same source domain.

In this study, we are going to use a large-scale corpus, i.e. the Gigaword Corpus via the interface of Chinese Word Sketch, to examine whether a mapping principle exists in a source-target domain pairing. In particular, we propose a frequency-based collocational approach [6, 7] to determine mapping principles based on the most frequent lexical item shared by a source domain and a target domain and to search for its lexical mappings from the source domain of BUILDING 建築物 to many target domains. The corpus data will demonstrate that certain lexical mappings in the source-target domain pairings occur more frequently than others [8]. This will support the two proposals of the Conceptual Mapping Model: (1) a mapping principle underlies each source-target domain pairing; (2) different conceptual aspects of the source domain of BUILDING contribute to different target domains.

## 2. Two Corpora

In this section, we introduce the two corpora we used for this research: Chinese Word Sketch and the Academia Sinica Bilingual Ontological WordNet.

### 2.1. The Chinese Word Sketch

The first corpus we used in this study is Chinese Word Sketch ([http://corpora.fi.muni.cz/chinese\\_all/](http://corpora.fi.muni.cz/chinese_all/)), a corpus processing system that was developed in 2005 [9]. It was constructed by loading the Gigaword Corpus into the Sketch Engine [10]. The Gigaword Corpus contains about 1.12 billion Chinese characters, including 735 million characters from Taiwan's Central News Agency, and 380 million characters from China's Xinhua News Agency. Components of the Sketch Engine are found in the Chinese Word Sketch, including Concordances, Word Sketch, Thesaurus and Sketch Difference.

Chinese Word Sketch can provide collocational information based on a large-scale corpus. It offers researchers information about a keyword's functional distribution and the collocations in the corpus. The functional distribution

includes: subject, object, prepositional object, and modifier. The collocations provide information on how frequently a keyword occurs with a particular word.

Chinese Word Sketch also produces a thesaurus list for adjectives, nouns or verbs [10]. The synonym items are automatically extracted based on common patterns of syntactic structures. This function is able to demonstrate a cluster of lexical words that have similar meanings but are different in linguistic form. For example, Table 1 shows six synonym candidates produced by the thesaurus for the lexical item 建築物 "building".

Table 1. Synonymous words for 建築物 "building".

Target words	Candidate synonyms	
建築物 "building"	建築物 "a building"	公寓 "an apartment"
	房子 "a house"	住房 "a residence"
	住宅 "a residence"	國宅 "a house"

## 2.2. The Academia Sinica Bilingual Ontological WordNet

The second corpus we used is the Academia Sinica Bilingual Ontological WordNet, a Chinese-English bilingual translation equivalent database (<http://ckip.iis.sinica.edu.tw/CKIP/ontology/>) that has been in development by Academia Sinica since 2002. This corpus integrates WordNet with the Suggested Upper Merged Ontology (SUMO) and allows users to search and map English or Chinese lexical meanings as defined in WordNet with a concept node on the SUMO ontology. WordNet (<http://wordnet.princeton.edu/>) is an on-line lexical database of English. Different grammatical categories of the same lemma are grouped into sets of synsets. Thus, a keyword in this corpus is interlinked in terms of conceptual-semantic and lexical relations.

In addition, SUMO (<http://ontology.teknowledge.com/>) is a shared upper ontology developed by the IEEE Standard Upper Ontology Working Group. Similar to a dictionary or glossary, it includes higher level concepts and an associated structure that enable computers to process its content. It consists of a set of general concepts, relations, and axioms and formalizes the description of structures of knowledge domains. The upper ontology allows users to identify a concept either in the physical node or in the concrete node, which allows us to determine the source domain's concreteness based on this information [3].

In brief, in this study, we used several functions implemented in the two corpora to determine mapping principles, including the collocational patterns of

lexical items, synonyms of a particular word, semantic meanings in WordNet, and the SUMO nodes.

### 3. Method

In this section, we describe how we search for lexical mappings from a particular source domain, i.e. BUILDING, to many target domains. The most productive mapping in a source-target domain pairing serves as the mapping principle. The overview of our method for extracting lexical mappings is that we seek to extract verbs and nouns that occur with the BUILDING lexical items and use the verbs and nouns to look for any abstract concepts that occur with those lexical items from the source domain of BUILDING.

First, we translate BUILDING into Chinese 建築物 "building". Second, we select the synonym candidates automatically produced by the thesaurus list for the keyword 建築物 "building" in order to list a cluster of words that have similar meanings to "building".

Third, we use the Chinese Word Sketch to automatically select the top 25 verbs that most frequently use these seven BUILDING lexical items as a Subject (i.e. items listed in the Subject\_of), the top 25 verbs as an Object (items listed in the Object\_of categories) and the top 25 lexical items as a Possessor (i.e. items listed in the Possession category). Table 2 demonstrates that when 建築物 "building" functions as a Subject, it collocates with the verb 倒塌 "to collapse"; when 建築物 "building" functions as an Object, it collocates with the verb 老舊 "old". When 建築物 "building" functions as a Possessor, it collocates with 窗戶 "a window". Some lexical items repeatedly occur with the seven BUILDING items, such as 倒塌 "to collapse", 修繕 "to repair", etc. We have removed these repeated words.

Table 2: 建築物 "building" functions as three syntactic categories and their collocates

Syntactic categories	Lexical words
建築物 "building" as a Subject	倒塌 "to collapse"
建築物 "building" as an Object	老舊 "old"
建築物 "building" as a Possessor	窗戶 "a window"

We select words as a Subject, an Object and a Possessor in line with Ahrens' [4, 5] method which proposes that each source domain, as a reflection of our real world knowledge, can be delimited as being entities, qualities, and functions.

Based on the same analogy, we focus on lexical items that use BUILDING as a Subject, an Object and a Possessor. These three syntactic categories can reflect the functions, qualities, and entities of the source domains.

Fourth, we use the Chinese Word Sketch to check whether these verbs and nouns often occur with the BUILDING concepts. In Table 3, for example, even though the lexical item 申購 "to purchase" collocates with the lexical word 國宅 "a house", 申購 "to purchase" frequently collocates with a number of lexical words that are not associated with the BUILDING domain, such as 人 "people", 基金 "a fund", 身份證 "an identification card", etc. We have removed these ambiguous words and only select verbs and nouns that are associated with the BUILDING concept.

Table 3: Lexical items that are/not associated with the concept of BUILDING

Not associated exclusively with BUILDING	Associated with BUILDING
申購 "to purchase"	新建 "newly built"
消費者 "a consumer"	陽台 "a balcony"

Fifth, we select the concepts that occur with the verbs and nouns in the BUILDING domain from the Chinese Word Sketch and use SUMO to define the concepts as abstract concepts. For example, the verb 重建 "to rebuild" occurs with two lexical items 觀念 "idea" and 對話 "conversation". Figure 1 illustrates that SUMO classifies the lexis 觀念 "idea" in the domain of PROPOSITION 命題 and its super-class contains a node 抽象的 "abstract", which suggests that it is an abstract concept. On the other hand, the lexical item 對話 "conversation" is classified in the domain of COMMUNICATION 溝通 and its super-class involves a node 物質的 "physical", which suggests that it is a concrete concept. We remove the concrete concepts in order to focus on abstract concepts that are described metaphorically in terms of concrete concepts of the BUILDING source domain.

Sixth, we define these abstract concepts and their target domain knowledge on the basis of the senses and explanations that are defined in WordNet. For example, the lexical items 案件 "a legal case", 問題 "a question", 危機 "a crisis", 謎 "a riddle", and 難題 "a difficult problem" are identified as the target domain of PROBLEM because the senses and explanations for those lexical items in WordNet include the word "problem", as shown in Table 4. Even though this step has to be made manually, it allows us to define the target domain knowledge in a less subjective manner because we are constrained by the WordNet senses.

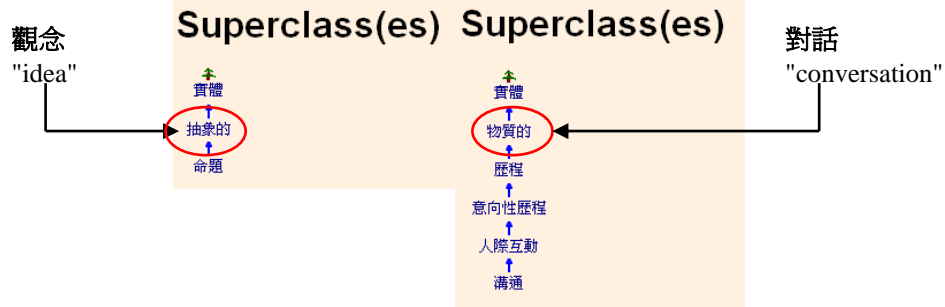


Figure. 1. Hierarchical structures for the lexical items 觀念 "idea" and 對話 "conversation" in SUMO

Seventh, we add up the frequencies of the same collocates within the same conceptual metaphor. For example, within PROBLEM IS A BUILDING, the lexical item 材料 "materials" collocates 6 times with 案件 "a case" and 10 times with 問題 "a problem" in the Gigaword Corpus. Thus, the total frequency for the lexical word 材料 "materials" within PROBLEM IS A BUILDING is 16 times. Finally, we postulate mapping principles based on the most productive lexical item.

Table 4: Abstract concepts that contain "a problem" or "a difficulty" are determined as the target domain of PROBLEM

Words	WordNet Senses	WordNet Explanations
案件	case	A <u>problem</u> requiring investigation
問題	<u>problem</u>	A state of <u>difficulty</u> that needs to be resolved
危機	crisis	An unstable situation of extreme danger or <u>difficulty</u>
謎	riddle	A <u>difficult problem</u>
難題	puzzle	A particularly baffling <u>problem</u> that is said to have a correct solution

#### 4. Data Analysis, Results and Discussion

In this section, we present the corpus data and discuss whether mapping principles underlie the source target domain pairing. We also discuss what

properties of the source domain of BUILDING are selected by different target domains.

The corpus data show that there are seven target domains that use the BUILDING source domain, including IDEA, RELATIONSHIP, DIGNITY, REPUTATION, SPIRIT, LIFE, and PROBLEM. In Tables 5-11 below, the total number of lexical items is given at the end of each table. Frequency indicates the number of tokens for collocates in the Gigaword Corpus and the percentage refers to the percentage of the number of collocates compared to the total number of lexical mappings. The Mapping Principle for each conceptual metaphor is postulated according to the most productive collocate.

The lexical items in the IDEA target domain include 觀念 "an idea", 內容 "contents", 意見 "an opinion", 論證 "a proof", and 資料 "data". In the case of AN IDEA IS A BUILDING (Table 5), the underlying reason has to do with foundations because the lexical item 基礎 "a foundation" is the most frequent collocate (92.7%). The Mapping Principle postulated is: *An idea is understood as a building in that buildings involve a (physical) foundation and ideas involve an (abstract) base.*

Table 5: AN IDEA IS A BUILDING

	Metaphorical Collocates	Frequency	%
Functions	毀損 "to damage"	14	2.5
Qualities	老舊 "aged"	19	3.4
Entities	基礎 "a foundation"	524	92.7
	材料 "a material"	8	1.4
Total		565	100

The lexical item in the RELATIONSHIP target domain only includes 關係 "relationship". In the case of RELATIONSHIP IS A BUILDING (Table 6), the underlying reason has to do with foundation (99%) because 基礎 "a foundation" is the most productive collocate. The Mapping Principle postulated is: *A relationship is understood as buildings in that buildings involve a (physical) foundation and a relationship involves an (abstract) base.*

When we compare A RELATIONSHIP IS A BUILDING to AN IDEA IS A BUILDING, the concept of 基礎 "a foundation" is mapped to the IDEA and RELATIONSHIP domains. However, some metaphoric instances are mapped to certain target domains. The lexical word 構築 "construct" is uniquely mapped to RELATIONSHIP but not to IDEA while the lexical words 毀損 "to damage", 老

舊 "aged" and 材料 "a material" are mapped to IDEA and not to RELATIONSHIP.

Table 6: A RELATIONSHIP IS A BUILDING

	Metaphorical Collocates	Frequency	%
Entities	基礎 "a foundation"	1237	99.0
Functions	構築 "construct"	12	1.0
Total		1249	100

The lexical items in the DIGNITY target domain include 自尊 "self-respect", 威信 "prestige", 尊嚴 "dignity" and 信心 "confidence". In the case of DIGNITY IS A BUILDING (Table 7), the underlying reason is postulated to be related to demolishing because 受損 "to demolish" has the highest proportion (62.8%). The Mapping Principle postulated is: *Dignity is understood as a building in that buildings are demolished when a physical attack occurs and dignity is demolished when a verbal attack occurs.*

Table 7: DIGNITY IS A BUILDING

	Metaphorical Collocates	Frequency	%
Functions	受損 "to demolish"	60	68.2
	重建 "to reconstruct"	28	31.8
Total		88	100

The lexical items in the REPUTATION target domain include 名聲 "reputation", 名譽 "fame", 信譽 "prestige", 商譽 "goodwill", 聲望 "prestige", and 聲譽 "reputation". In the case of REPUTATION IS A BUILDING (Table 8), the underlying reason has to do with destruction because 受損 "to demolish" is the most frequent lexical word (90%). The Mapping Principle postulated is: *Reputation is understood as a building because buildings are demolished when a physical attack occurs and reputation is demolished when a verbal attack occurs.* The corpus data show that the collocates of REPUTATION always relate to the notion of demolishing or damage.

When the abstract concepts of dignity and reputation are discussed metaphorically in terms of BUILDING, the two Mapping Principles involve the concept of demolishing. This may suggest that DIGNITY and REPUTATION are often expressed with a more negative connotation.

Table 8: REPUTATION IS A BUILDING

	Metaphorical Collocates	Frequency	%
Functions	受損 "to demolish"	319	90.0
	毀損 "to demolish"	14	4.0
	損毀 "to damage"	7	2.0
	損壞 "to damage"	7	2.0
	毀壞 "to damage"	7	2.0
Total		354	100

The lexical item in the SPIRIT target domain only includes 心靈 "mind." In the case of SPIRIT IS A BUILDING (Table 9), the underlying relationship concerns reconstruction because the lexical item 重建 "to reconstruct" is the most prototypical lexical item (97.5%). The Mapping Principle postulated is: *Spirit is understood as a building in that buildings are able to be reconstructed when building materials are ready and spirit is able to be reconstructed when the human emotion is ready to recover.* Interestingly, the lexical items 窗戶 "a window" and 門窗 "doors and windows" are mapped to SPIRIT but are not mapped to other abstract domains discussed previously.

Table 9: SPIRIT IS A BUILDING

	Metaphorical Collocates	Frequency	%
Functions	重建 "to reconstruct"	577	97.5
Entities	窗戶 "a window"	9	1.5
	門窗 "doors and windows"	6	1.0
Total		592	100

The lexical item in the LIFE target domain only includes 生活 "life". For LIFE IS A BUILDING (Table 10), the relationship concerns reconstruction because 重建 "to reconstruct" is the only and highest lexical item (100%). The Mapping Principle postulated is: *Life is understood as a building in that buildings are able to be reconstructed when building materials are ready and life of a disadvantaged minority is able to be reconstructed when the social policies are established to help them.*

Table 10: LIFE IS A BUILDING

	Metaphor	Frequency	%
Functions	重建 "to reconstruct"	274	100
Total		274	100

When we compare SPIRIT IS A BUILDING to LIFE IS A BUILDING, the source domain of BUILDING repeatedly collocates with the concepts of spirit and life. The two Mapping Principles involve the concept of reconstruction. This may suggest that SPIRIT and LIFE are often expressed linguistically with a more positive connotation.

The lexical words in the PROBLEM target domain include 案件 "a case", 問題 "a question", 危機 "a crisis", 奧秘 "a mystery", 僵局 "a deadlock", 謎 "a riddle", and 難題 "a problem". For A PROBLEM IS A BUILDING (Table 11), the relationship concerns a housekey because the highest proportion of 鑰匙 "a key" (66%). The Mapping Principle postulated is: *A problem is understood as a building in that buildings need keys for entering a house and problems need keys for solving the difficulty.* In addition, when we compare A PROBLEM IS A BUILDING to other conceptual metaphors, the lexis "keys" is uniquely mapped to PROBLEM but never mapped to IDEA, DIGNITY, SPIRIT, etc.

Table 11: A PROBLEM IS A BUILDING

	Metaphorical Collocates	Frequency	%
Entities	鑰匙 "a key"	31	66.0
	材料 "a material"	16	34.0
Total		47	100

The corpus data presented above suggest that different underlying reasons do occur to characterize how a concrete concept is used to describe an abstract concept. In particular, the prototypical mapping for each conceptual metaphor can reflect why a source domain is repeatedly mapped to multiple target domains. This supports the proposal of Conceptual Mapping Model [4, 5] that a mapping principle exists in each source-target domain pairing.

Table 12 shows how and why different target domains select the same source domain of BUILDING. Indeed, different aspects of the source domain of BUILDING contribute to different target domains.

Table 12: Four aspects of a building are used for different source domains

Aspects of a building	Source domains
Foundations	AN IDEA, A RELATIONSHIP
Destruction	DIGNITY, REPUTATION
Reconstruction	SPIRIT, LIFE
Keys	A PROBLEM

The corpus data show that when IDEA is discussed metaphorically in terms of BUILDING, the aspect of foundation is emphasized. On the other hand, when the PROBLEM domain is discussed metaphorically in terms of BUILDING, the notion of a key is used. The emphasis for IDEA is on foundations in order to indicate an initial stage of an action or event. On the other hand, the property of keys of a house is used for the target domain of PROBLEMS in order to explain how to solve a problem.

In addition, the corpus analysis demonstrates how people use metaphors in daily discourse. For example, the metaphorical usages for the concepts of reputation and dignity are more frequently expressed with a negative connotation. The aspect of demolishing is borrowed from the BUILDING source domain to express the two concepts, i.e. reputation and dignity in association with negative outcomes. On the other hand, the metaphorical usages for the concepts of spirit and life are more frequently expressed with a positive connotation. Another aspect, i.e. construction, is borrowed from of the BUILDING source domain to positively frame two concepts, i.e. spirit and life.

This study uses a frequency-based collocational method to identify the mapping principles. We search for lexical mappings from the source domain of BUILDING to a number of target domains. The analysis of lexical mappings between a source domain and a target domain show that mapping principles do underlie a source-target domain pairing. We also find that the different aspects of a particular source domain may sometimes map to different target domains, e.g. with respect to BUILDING, the notion of foundations is selected by IDEA and that of keys of a building is selected by PROBLEM.

## 5. Conclusion

In this study, we use a large-scale corpus, i.e. the Gigaword Corpus in combination with the Chinese Word Sketch, to examine relationships between source and target domain pairings. We propose to use a source-domain-oriented

method to search for lexical mappings from a particular source domain to a number of target domains. This approach seeks to avoid the limitations of previous studies which employed the target-domain-oriented method but did not allow researchers to compare what properties of a particular source domain might be selected by distinct target domains.

We also propose a frequency-based collocational approach to determine mapping principles based on the most productive lexical mapping. For each conceptual metaphor, we found that a particular lexical mapping occurs more frequently than others. This supports the Conceptual Mapping Model that there is an underlying reason occurring in a source-target domain pairing. Furthermore, our data demonstrate that different target domains make use of different aspects of the same source domain. For example, four different aspects of the source domain of BUILDING were found to map to four distinct target domains. The Mapping Principle of AN IDEA IS A BUILDING has to do with the notion of foundation of a building. However, the Mapping Principle of PROBLEMS ARE BUILDINGS has to do with the notion of the key of a building. The Mapping Principle of REPUTATION IS A BUILDING has to do with the notion of the destruction of a building. Finally, the Mapping Principle of SPIRIT IS A BUILDING has to do with the reconstruction of a building. Four aspects of a building are selected by the seven different target domains.

To conclude, this corpus-based study follows the Conceptual Mapping Model's proposal that lexical mappings can be acquired through linguistic analysis between a source domain and a target domain. Further research can employ the same method discussed herein to explore more metaphors in Mandarin Chinese that use other source domains, such as FOOD, GAME, FIRE, etc. In addition, the manual steps discussed in this study may help provide ways to identify metaphorical mappings automatically/semi-automatically. It is hoped that in this way conceptual mappings will no longer be considered *ad hoc* results of source-target domain pairings but rather involve principled explanations based on prototypical mappings.

### **Acknowledgments**

This research was supported by grants from the National Science Council of Taiwan to the second author (NSC 95-2411-H-002 -037). We would like to thank Prof. I-wen Su, Prof. Hintat Cheung, Dr. Siaw-Fong Chung, and the participants of the 8<sup>th</sup> Chinese Lexical Semantics Workshop 2007 (CLSW 2007) at the Hong

Kong Polytechnic University for their comments on earlier versions of this work. Remaining errors are our sole responsibility.

### References

- [1] G. Steen, From linguistic to conceptual metaphor in five steps, *Metaphor in Cognitive Linguistics*, eds. R. W. Gibbs and G. Steen, Amsterdam: John, 1997, pp. 57-77.
- [2] A., Stefanowitsch, The function of metaphor: developing a corpus-based perspective, *International Journal of Corpus Linguistics*, 10(2), 2005, pp. 161-198.
- [3] S.-F. Chung, *A Corpus-Driven Approach to Source Domain Determination*, Unpublished Doctoral Dissertation, National Taiwan University, 2007.
- [4] K. Ahrens, When Love is not Digested: Underlying Reasons for Source to Target Domain Pairing in the Contemporary Theory of Metaphor, in Proceedings First Cognitive Linguistics Conference, Cheng-Chi University, Taipei, Taiwan, 2002, pp. 273-302.
- [5] K. Ahrens, Mapping Principles for Conceptual Metaphors, in *Researching and Applying Metaphor in the Real World*, eds. L. Cameron, A. Deignan, G. Low, and Z. Todd, John Benjamins. Forthcoming.
- [6] K. Ahrens, S.-F. Chung, and C.-R. Huang. Conceptual Metaphors: Ontology-based Representation and Corpora Driven Mapping Principles, in *Proc. ACL Workshop on the Lexicon and Figurative Language*, 2003, pp. 35-41.
- [7] K. Ahrens, S.-F. Chung, and C.-R. Huang. From Lexical Semantics to Conceptual Metaphors: Mapping Principle Verification with WordNet and SUMO, in *Recent Advancement in Chinese Lexical Semantics: Proceedings of 5th Chinese Lexical Semantics Workshop (CLSW-5)*, eds. J. Donghong, L. K. Teng, and H. Wang, Singapore, 2004, pp. 99-106.
- [8] C.-R. Huang, S.-F. Chung and K. Ahrens, An Ontology-based Exploration of Knowledge Systems for Metaphor, in *Ontologies: A Handbook of Principles, Concepts and Applications in Information Systems*, eds. K. Rajiv, R. Ramesh, and R. Sharman, Volume 14, Springer, 2006.
- [9] C.-R. Huang, A. Kilgarriff, Y. Wu, C.-M. Chiu, S. Smith, P. Rychly, M.-H. Bai, and K.-J. Chen, Chinese Sketch Engine and the Extraction of Collocations, in *Pro. Fourth SIGHAN Workshop on Chinese Language Processing*, October 14-15, Jeju, Korea, 2005, pp. 48-55.

16 Gong et al.

- [10] A. Kilgarriff, C.-R. Huang, P. Rychly, S. Smith, and D. Tugwell, Chinese Word Sketches, in *ASIALEX 2005: Words in Asian Cultural Context*, Singapore, 2005.