國立交通大學

外國文學與語言學研究所

碩士論文

歧義現象的多層次分析架構:
由中文動詞出發

A Multi-layered Resolution for Disambiguation:
Insight from Mandarin Verbs

研 究 生: 徐雅苓
指導教授: 劉美君 教授

中華民國九十五年六月

歧義現象的多層次分析架構:

由中文動詞出發

A Multi-layered Resolution for Disambiguation:

Insight from Mandarin Verbs

研 究 生: 徐雅苓　　Student: Ya-Ling Hsu

指導教授: 劉美君　　Advisor: Mei-Chun Liu

國立交通大學

外國文學與語言學研究所

碩士論文

A Thesis

Submitted to Graduate Institute of Foreign Literatures and Linguistics

College of Humanity and Social Science

National Chiao Tung University

In Partial Fulfillment of the Requirements

for the Degree of

Master of Arts

June 2006

Hsinchu, Taiwan, Republic of China

中華民國九十五年六月

多義現象的多層分析架構: 由中文動詞出發

學生: 徐雅苓　　指導教授: 劉美君

國立交通大學語言與文化研究所

摘　要

本篇論文的研究重點，是提出一個以語料庫語言學為基礎的多層次架構，來探究多義詞的多義現象，進而建立一套區辨語意的自動標注系統。透過不同語言學理論，例如框架語意學 (參見 Fillmore 和 Atkins 1992), 構式語法 (參見 Goldberg 1996) 以及話語分析 (參見 Hopper 和 Thompson 1980)，以期提供一個以語言學為出發點的語義檢索機制。多義詞作為詞彙的本質之一，不失為了解句法、語意、及語用三者互動關係的一個關鍵。雖然前人已提供許多不同的研究方向來探究多義詞的多義性，包括類別特徵分析方法、原型理論、框架理論、以及關係理論等，但是仍缺乏一個有系統且具可行性的方法。近來的研究如 Liu 和 Wu (2004)，他們提出以語意框架的觀點為基礎來檢視多義性，他們認為語詞的多義性就如 Fillmore 和 Atkins (1992) 所定義的一樣，是被定義在不同的框架之下。借重不同的框架成份以及其不同的語法表現，Liu 和 Wu (2004)依循著「一個語意，一個框架」的假設，使我們看到了，語意的不同可以歸結於動詞所屬的不同框架概念。然而，這樣的語意界定方法，似乎沒有辦法區辨一個多義動詞的不同語意，當他們屬於不同框架概念，卻有相同框架成分及語法表現的時候。以中文動作動詞「拿」為例，其中兩個語意就帶有相同的框架成分及語法表現，如例子 (1)，

(1) Agent>V>Theme:

a. …病人[Agent] **拿** 著健保卡[Theme]上門… (語意 1 '持')

b.…我[Agent]可不可以順道 **拿** 個研究學位[Theme]？... (語意 2 '得/取')

因此，由例子 (1) 我們可以預測，區辨多義詞只靠框架理論是不足夠的。當框架成分無法提供足夠的資訊來決定語意時，還有什麼是我們沒有考慮到的部分呢？本文中所要提出的架構，則將兩個重要可變因素考慮進來：配搭組合和語境依存。本文主要目的在於提出一個多層次的分析架構，來定義多義詞在不同語法表現中的適當語意。這個多層次的分析方法，依據以下三個步驟可以作為一個語義區辨的模組：(1) 以框架為依據的區辨方法 (2) 以配搭組合為依據的區辨方法 (3) 以語境依存為依據的區辨方法。

本文研究主要來自中研院漢語平衡語料庫的自然語料。在文中的個案研究皆為高頻詞，但每個個案只採 200 筆語料作細部標記。使用語料庫的語料，主要是

因為語料庫的語料，提供了重要的語法語意分布趨向，這是母語說話者的直覺沒有辦法察覺到的。

首先，依據 FrameNet 的理論，在我們區辨模組的第一步驟，是把一個語料庫中的多義詞依據其不同的語意框架概念，而定義為不同的語意；其主要的區分方式，則是依據不同的框架成分及其主要語法表現，來區分成不同的語意組。當第一個步驟無法成功區辨語意，也就是當碰到不同語意卻帶有相同框架成分及其相同的語法表現時，我們則需進入模組中的第二個區辨步驟—配搭組合。在這個步驟中，我們所須注意的是那些和非核心論元的搭配詞組；這些非核心論元的搭配詞組依據不同的詞類可再作分類，如副詞、形容詞、時態標記等。進而我們將會發現，多義詞的不同語義，和這些非核心論元會有不同的固定搭配關係。然而，當搭配組合的方法也無法提供更進一步的資訊時，我們則需要進到第三步驟—語境依存；在這個步驟中，我們將搜尋在跨語句的語境當中，是否有和多義詞不同語意相關的詞語。多義詞和不同語義的連結，主要是建立在它們之間語義或語用上的相關；在 SUMO 中，我們確實是可以搜尋到它們之間的連結。我們將以四個中文單詞動詞為例—走、拿、聽、看，以論證本文所提出的模組。

在本文中，藉由所提出的機制，除了重新定義多義性之外，也成功的提供電腦區辨系統，一個以語言學為基礎的有效的語義區辨模組。

A Hybrid Resolution for Polysemy:
Insight from Mandarin Verbs

Student: Ya-Ling Hsu Advisor: Dr. Mei-Chun Liu

Graduate Institute of Foreign Literatures and Linguistics
National Chiao Tung University

Abstract

This study explores how multiple senses of polysemous words could be distinguished. It proposes a hybrid and corpus-based linguistic model and specifies the procedures to build an automatic tagger for sense disambiguation based on Mandarin verbs. It seeks to provide a linguistically motivated solution for detecting meaning with the aid of linguistic theories such as Frame Semantics (Fillmore and Atkins 1992), Construction Grammar (Goldberg 1996) and discourse analysis (Hopper and Thompson 1980). Being an essential property of the lexicon, polysemy is the key to understanding the interplay between syntax, semantics and pragmatics. Although polysemy has been investigated in a number of approaches, including classical feature analysis, prototype theory, frame-based approach, relational approach, and so on, a systematic and applicable solution is still lacking. Recently, working on Mandarin lexical semantics, Liu and Wu (2004) proposed a frame-based perspective in viewing polysemy as belong to different 'frames', which is defined by Fillmore and Atkins (1992). Making use of the distinctions in frame elements and their grammatical realizations, Liu and Wu (2004) is able to show that semantic differences may be attributed to different semantic frames the verb belongs to, following 'the one sense, one frame' hypothesis. However, there are cases where two separate meanings

of the same verb may show exactly the same surface patterns with the same sets of frame elements. For example, in the case of the motion verb *NA* 拿, two separate senses may end up with the same number and pattern of frame elements, as shown in (1):

(1) Agent < V <Theme:

a. …病人[Agent] 拿 著 健保卡[Theme] 上門… (sense 1 'carrying')

bing ren *na* *zhe bao jian ka* *shang men*

patien take ZHE health insurance card up door

'The patient carried the health insurance card to the counter.'

b.…我[Agent]可不可以順道 拿 個 研究學位[Theme]？(sense 2 'getting')

wo *ke bu ke yi shun dao* *na* *ge* *yan jiu xue wei*

I can not can by the way take CL research academic degree

'By the way, can I get an academic research degree?'

Therefore, it is apparent that a purely frame-based approach may be insufficient in dealing with polysemes. When frame elements fail to provide determining clues, what else should be taken into consideration? The model proposed in this study calls for consideration of two other variables: colloconstructions and contextual dependencies. This study aims to propose a hybrid multi-module solution to identify the most appropriate lexical sense in various expressions of a polyseme. The hybrid approach can be viewed as a sense disambiguating model based on three steps: 1) frame-based distinction, 2) colloconstruction distinction, and 3) contextual dependence distinction.

The study is based on naturally occurring data extracted from the Sinica Balanced Corpus, which is established by the CKIP (Chinese Knowledge and Information Processing) group at Academia Sinica and open to the public at the Internet site: http://www.sinica.edu.tw/ftms-bin/kiwi.sh/. Given the high frequency of occurrences of the target words, only 200 entries are examined closely for the discussion. Corpus data provide explicit and implicit distributional tendencies which

may go beyond native speaker's intuition.

Using corpus data as the input, the first step of the proposed model is to identify the senses of a polysemous word corresponding to the distinctions in semantic frames, following FrameNet. The extracted data from Sinica Corpus can be roughly classified into several frames by their basic patterns of expressing the core frame elements (arguments). When distinctions of frame elements and their basic patterns fail, senses are further identified by the second module - Colloconstrucion. In this step, attention is paid to the collocational patterns of non-core arguments. These non-core arguments can be classified into various syntactic categories, such as adverbials, adjectives, aspectual markers, and so forth. And frequent collocates, be it grammatical or lexical, will be identified with each individual sense. However, when colloconstruction fails to indicate any decisive cues, the third module - contextual information is called upon. In this module, the relevant contextual elements are thoroughly searched to establish a relational link within or cross clausal boundaries. The relational link may be established by any semantic/pragmatic associations between the polyseme and the contextual element that a larger semantic taxonomy, such as SUMO synsets (translated in BOW). To demonstrate the model, four sets of verbs (*zou* 走, *na* 拿, *ting* 聽, *kan* 看) will be used as illustrations. By redefining polysemy with operational mechanisms, this study successfully provides a linguistic model with theoretical validity to develop a computational system for sense disambiguation.

致謝

此篇論文可以順利完成要感謝許多一路上給予我支持與鼓勵的人。首先，要感謝的是我的指導老師劉美君教授。從我一進語言所開始，便參與老師和電信所共同執行的「卓越計畫」。藉由和不同領域的人的接觸、溝通，使我對語言學的應用有不同一層的認識。老師總是引領我們從不同的方向去思考問題、發掘問題、解決問題，進而提出一個完善的方法;她從不事先給予我們答案，而是啟動我的思考機制，激發了我對語言學不同議題的熱誠。另外，也很感謝老師提供了我參與國科會「中文動詞語義研究」計劃的機會，使我不僅僅是接觸其他領域，更是加強了自身語言學專門領域的背景知識，以及研究方法。也由於參與了這兩個計畫，在研究中，啟發我對本篇論文的靈感。老師對於我論文的指導始終不厭其煩，時時提醒我論證及寫作邏輯上的重要觀點，同時也適時的鼓勵我投稿，以吸收其他人的批評指正，讓我不至於陷入自我的迷失之中。老師的諄諄教誨，讓我得以順利完成論文。另外，也要感謝口試委員連金發老師以及劉辰生老師。連老師點出了我在寫作上的盲點，更提供了寶貴的意見，引領我朝不同的方向來思考，進而使我的論文可以更面面俱到。辰生老師對於漢語內部細微的觀察，提醒了我在語料分析上必要的謹慎，以及在語料的觀察上必須注意的正反兩面;對於辰生老師我也要特別的感謝，因為老師是我在語言學上的啟蒙老師，老師對語言學的熱忱，讓我感受到語言學殿堂中的奧妙。對於老師們的建議與指教，讓我在論文寫作上可以更精進，學生由衷的感謝。同時，對於教導過我的交大與清大老師們(許慧娟、林若望、潘荷仙、曹逢甫)，我也深深的感謝，因為你們的教導，使我在語言學的路上，擷取了不同語言學領域的知識。

在語言學的路上，也要謝謝一路幫我加油打氣的朋友。Alice 百忙之中還抽空幫我看資格考的論文，不僅是我球場上的教練，更是我學業上的私人教師；馥綾、珮婷、明芹、英翰、小花學姊不時的鼓勵我，聽我吐吐悶氣，讓我可以適時的抒發心情。也很感謝佩儀學姐，她不僅僅是生活中的好朋友，更是學術上值得學習的榜樣，有你的分享，讓我在語言學的分析討論上有了不少的啟發。當然，交大的同學及學弟妹們，尤其是一起做國科會計劃的亭儀、姿君、明輝、佳音、子玲、瓊羽，和你們一起討論、彼此切磋語言學上的種種，每每都讓我獲益良多。

最後，當然要感謝我的家人，謝謝我最敬重的爺爺，謝謝你的慈愛，雖然您還來不及分享我的一切，如今我順利完成修業，你是我想第一個告知的人。謝謝爸爸媽媽，你們無怨無悔的付出，使我無後顧之憂的在求學路上不斷發展，你們的鼓勵，永遠是支持我最大的力量。我親愛的妹妹們，你們的加油打氣，永遠是最窩心，令我感動不已的，謝謝你們。

僅將此本碩士論文，獻給我最敬愛的父母，徐慶旺先生、何麗燕女士，表達女兒由衷的感謝，爸爸、媽媽，謝謝你們!

Table of Contents

1 Introduction

This paper presents a corpus-based hybrid-analysis model on the issue of sense disambiguation via case studies of Mandarin ambiguous verbs. It seeks to find a solution for a complete investigation of the behavior of ambiguous words in corpus. To explore ambiguity, an investigation of a relevant issue, polysemy discussed in linguistics, needs to be considered. Being one essential part of lexicon, polysemy provides the key to understanding the syntactic and semantic properties of lexicon. Previous research has proposed many different perspectives to discuss this issue. However, these studies still provide insufficient explanations. This issue might seem extremely complicated and thorny, but there is a need for a complete investigation of polysemy. An overall probe into ambiguouous words' behavior will help advance the research of Mandarin linguistics in general and provide a practical solution for application in computational system. Consequently, following this issue the goal of this study aims to provide a hybrid-analysis module to identify the various expressions of ambiguous words for an in-depth reconsideration of ambiguity. The fundamental of this research follows what is claimed in Fillmore (1992:76):

> "… a word's meaning can be understood only with reference to a structured background of experience, beliefs, or practices, constituting a kind of conceptual prerequisite for understanding the meaning."

According to Fillmore's conceptual schema, Liu and Wu (2004) distinguished ambiguous words base on different cognitive structures (or "frames") which contains various particular category (or "core frame elements") and specific "lexical-syntactic patterns" by Fillmore (1992). However, the essential problem remains: Is frame-based analysis sufficient to account for all the expressions of ambiguous words? This is the central question for which this paper seeks to provide an answer.

1.1 Polysemy in general

In previous studies, lexical ambiguity is a heterogeneous phenomenon. That is, lexical ambiguity is caused at least by the following three crucial factors (Pustejovsky and Boguraev 1996: 6):

- ✧ Contrastive ambiguity, which is normally resolved by contextual and discourse knowledge;
- ✧ Complementary ambiguity (or logical polysemy), as resolved by co-composition in the syntactic context of the sentence; and
- ✧ Sense extensions, as mediated by lexical rules and specific conditions relating to the speaker and context.

In general, the first factor contributes to the appearance of homonymy such as the two interpretations of *bank* as in 'river bank' and 'financial bank', and vagueness, for instance, *news* in 'I read the news this morning' (news as press communiqué) and in 'I haven't heard any news about him since he left' (news as the information about him). Traditionally, polysemy has been distinguished from homonymy and vagueness. Generally speaking, homographs are unrelated words which are represented by the same word forms while polyemous words have semantically related word meanings. Another distinction is made between polysemy and vagueness. Given a huge number of discussions, the differences between polysemy and vagueness are still controversial (Cruse, 1986; Lakoff, 1987; Wierzbicka, 1990; Geeraerts, 1993.). Geeraerts (1993) claimed that the failure of reaching a consensus could be attributed to the question: what is our conception of meaning and of lexical semantics? Recent development in semantic research has generated wide interest in investigating the polysemy caused by the third factor above—sense extension rather than homonym or vagueness. Sense extension is basically divided into two types: (1) extended by various syntactic behaviors, (2) extended via metaphor or metonymy

(Lien 2000). The polysemy considered in this study involves syntactic variation. In the study of polysemy attributing to the sense extension, researchers have tried to define and establish relations of different meanings of polysemous words. In the early semantic paradigm, the major concern on plysemy was focused on how to define a polysemous lexical item or how to determine the number of senses in a polysemous word, but there remains the issue of these discussions (Ravin, 1990; Jackendoff, 1985; Fellbaum, 1998).

In linguistics, there are four major approaches dealing with the issue of polysemy: classical approach of semantic anaysis, prototype approach, frame-based approach, and relational approach. Within each approach, there are still some controversies and variations in analyzing polysemy. One point worth emphasizing in the classical approach is that meaning is viewed as consisting of a set of decomposed semantic features by necessary and sufficient conditions (Katz 1972). However, there is a danger of an infinite increase of senses of a polysemous word which is identified by infinite semantic features.

The assumption of prototype approach is that meaning is defined in the concept that meaning exhibits family resemblance and is linked to mental representations, cognitive models and bodily experience. As a direct consequence, Rosch (1977) demonstrated that people categorize objects on the basis of resemblance of the objects to a prototypical member of the category. But the problem is that without constraints, meanings can be infinitely related to each other by resembling features, so that senses of irrelevant polysemous words may end in linking to each other.

Recently, Fillmore (1992) proposes a cognitive analysis based on frame semantics. In this theory, a word's meaning is understood within structured background knowledge. Thus, word senses are not directly related to each other but are defined by common background frames. Further, he investigated words'

meanings by their realization in different syntactic patterns with different highlighted categories (as core frame elements) from a large corpus and builds a frame-based online dictionary (FrameNet). Following Fillmore's Frame Semantics (1992), sense identification is not just defined by a speaker's intuition but comes from the real utilization of natural language. The concept of FRAME refines the notion of polysemy, as Fillmore reconsiders polysemy extending from a semantic frame into a new domain. However, sense identification of polysemy is still questioned when different senses of polysemous words occur in the same realization with the same core frame elements.

More recently, Fellbaum (1998) has proposed that words are constructed depending on their meanings by the remains of their semantic relation or the semantic network to which they belong. However, senses of a polysemous word which occur in the semantic network might be very distant from each other. For example, in WordNet, there are three senses of the noun *ash*: (1) the residue (2) timber trees (3) ash trees, the first one is in the structure as plant material while the other two as a woody plant (Yale and Claudia 2000). These senses are distant because their semantic relation cannot be linked by proximity in WordNet.

In sum, from the classical approach (Katz 1972), to the prototype approach (Jackendoff 1985, Lakoff 1987, Taylor 1989), to the frame-based approach (Fillmore and Atkins 1992) and finally to the relational approach, polysemy has gone through a long history and has been studied from different perspectives. In contrast, Mandarin polysemy still awaits detailed discussion. Research on Mandarin polysemy will be introduced below.

1.2 Polysemy in Mandarin

Due to different theoretical interests, the focus of previous studies on Mandarin polysemy can be characterized as follows:

- ✧ Grammaticalization: polysemy happens when "lexical items come in certain linguistic context to serve grammatical functions, and one grammaticalized, continue to develop new grammatical functions." (Su 2002) Then, various functions contribute to different meanings of the lexical items (i.e. Liu 1994, Su 2002, Lai 2004).
- ✧ Metaphorical extension: metaphorical extension is the supporting evidence for meaning change through the linking between abstract to concrete (Lin and Ahrens 2005, Cao, Cai and Liu2001).

In general, earlier discussions encounter two problems. First, the identification of different senses of polysemous words in natural occurrences significantly influences the application in natural language processing and Chinese teaching. However, what is insufficient is the investigation of polysemous words in naturally occurring data. Second, the crucial factor causing polysemy that is often taken into consideration pertains to the association between meaning and its syntactic behavior. These problems are the most important issue in sense disambiguation. What is lacking and needs to be explored is how to identify the senses of disambigous words through their realization in natural occurrences.

More recently, following Fillmore *et al* (1992), Liu and Wu (2004) provided one of earliest studies discussing Mandarin polysemy respect to Frame-based approach. Instead of explaining what is the way meaning extending, they have shifted the focus to investigating the distinction of different senses of an ambiguous word in corpus. Based on frame-semantics, first, they define the senses of a polysemous word via different syntactic behaviors corresponding to basic patterns (BP) in FrameNet. Besides, in their paper, Liu and Wu (2004) also provided other evidences to support this distinction, for example, the collocation association and the semantic attributes of core frame elements, such as the combination of manner, the aspectual markers and

the negatives. A more detailed introduction of the study is presented in Chapter 2 of this paper.

In sum, Mandarin polysemy studies are in the preliminary stage. More comprehensive and extensive research is supposed to unveil the sense distinction.

1.3 Questions and solutions

Following the study of Liu and Wu (2004), this paper aims to provide more clear and complete discussion for sense disambiguation. Traditionally, in Mandarin, monosyllabic characters are one crucial source of polysemy. Verbs have been viewed as the category which carries various meanings in different syntactic expression, such as mentioned by Pustejovsky and Boguraev (1996) that verbal polysemous words, sometimes, lead to some complicated problems for lexical semantics. Besides, in sense disambiguation, monosyllabic words are more complicated. Therefore, in this research, case studies will be focused on the monosyllabic verbs. First of all, extracted data from the corpus are also defined based on frame-semantics as in Liu and Wu's study (2004). However, some cases of corpus data still remain problematic (as in (1) and (2)) that the sense of these two *ZOU*s can not be identified solely via core frame elements and basic patterns.

(1) a. 我[Self-mover]走在 大安 森林 公園[Area] (Sense 1 'walking')

wo *zou zai* *da an* *sen lin* *gong yuan*

I walk in Da An forest park

'I walked in Da An forest park.'

b. 我[Self-mover]走一趟 大安森林 公園[Area] (Sense 3 'visiting')

wo *zou yitang* *da an sen lin* *gong yuan*

I go once Da An forest park

'I visited at Da An forest park'

(2) a. 我 腳 好痠， 我[Self-mover]沒辦法 走了(Sense 1 'walking')

wo jiao hao suan, wo mei ban fa zou le

my feet so limp, I cannot walk LE

'My feet are so limp that I can not walk anymore.'

b.火車 早 就 開走了，我們[Self-mover]沒辦法 走了(Sense 4 'leaving')

huo che zao jiu kai zou le, wo men mei ban fa zou le

train already drive away LE we cannot walk LE

'The train has already driven away, and we can't leave.'

The senses of the verb *ZOU* in (1) are similar to *walking*, and *visiting* in English, and in (2) are similar to *walking and leaving* in English. In FrameNet, according to their basic patterns with core frame elements they should be classified into the same domains because both *ZOU*s in (1a) and (1b) share the basic pattern: Self-mover <*ZOU*< Area, and both *ZOU*s in (2a) and (2b) share the basic pattern: Self-mover < *ZOU*. However, by native speakers' intuition and according to other components in the context, these *ZOU*s should be identified as different meanings. What cannot be explained is why these different expressions of *ZOU* 走 belong to the same frame but denote different meanings, contradicting the frame-based approach. Therefore, there might be something insufficient in Frame-based analysis. Thus, as Liu and Wu (2004) proved in their study, collocational association could provide more information for further distinction of ambiguous words. This is similar to the second module-colloconstruction proposed in this paper. Colloconstruction refers to a specific lexical item categorized with syntactic characteristic bearing certain semantic properties that frequently co-occurs with the target sense. But, what remains unclear in Liu and Wu's research (2004) is that there is no explicit definition and criteria of their collocational association. Besides, it is also found that some cases denote different senses sharing the same collocational association. It would seem, therefore, that further investigations are needed in order to distinguish ambiguous words more

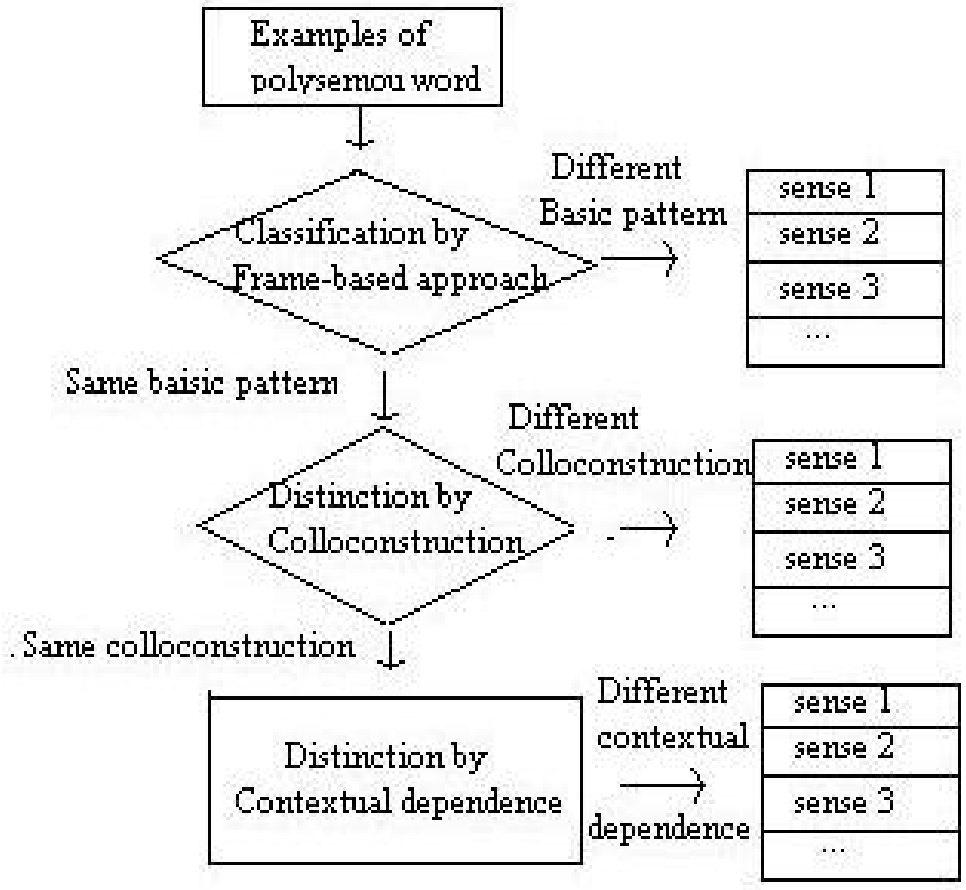
completely. In the previous studies, linguists have point out that the word meaning does not fully exist within lexical items, but some of the meaning components are generated by lexical relation in context (Saeed 1997:53; Cruse 2004). Other supports coming from psycholinguists, such as in Swenny's experiment (1979), point out that relevant lexical items in context can facilitate lexical decision of a ambiguous word. Therefore, the information from the context is thought to be significant for the further distinction and each sense is believed to display certain features linking to the relevant lexical in Context.

1.4 Model for sense disambiguating

A hybrid sense disambiguating model is proposed in this study to provide complete sense distinction of Mandarin ambiguity. The resolution formula is schematically represented in the model diagram as (3) below. In view of the need for investigating real language use, the analysis of this paper adopts corpus-based approach. Following the study of Liu and Wu (2004), the first module of sense distinction is also based on frame semantics. Following the "one sense, one frame" hypothesis, multiple senses of polysemous words are identified via different core frame elements and basic patterns. The senses of an ambiguous word either can be distinguished in this module or cannot be successfully defined and then, disambiguated in the second module-colloconstruction. Colloconstruction is different from collocation and construction. Colloconstruction refers to co-occurring patterns which might be collocates or possible constructions while collocation refers to co-occurring lexical items, and construction refers to a meaning unit. With the information of co-occurring categorical collocation with specific semantic meaning, distinctive colloconstruction may be found to further distinguish multiple senses. But, if colloconstruction fails, further distinctive features should be investigated. That is, in order to further disambiguate multiple senses which can not be identified in

module 1 and module 2, a third step—module 3: contextual dependency is necessary. In previous works, semantic contexts are divided into two types, word context, and sentential context (Johnson-Laird *et al* 1989). Word context refers to the background knowledge within the lexical system itself while sentential context means information from outside the lexicon. The third module, so called contextual dependency refers to the sentential contexts, that is, we abstract the relevant information from beyond the lexical background knowledge from the context. In this module, we look for the semantic or pragmatic relevant lexical items across clausal boundaries to help further distinction. What is presented in this paper can be viewed as a complete and detailed analysis on recent efforts and advances in the research of Mandarin ambiguity.

(3) The model of sense disambiguation



1.5 Organization of the study

To provide more detailed analysis of the issue—ambiguity—this research is composed of eight sections. The first section gives the brief but insightful introduction of this study. The second section is an overview of the background, the method and the theoretical perspectives from both analytic linguistics and computational linguistics involving the issue of ambiguity (or polysemy). The third Section gives an overall introduction of the three analytic steps (also three modules) and proposes a model of a sense disambiguator. The focus of section four and five is to postulate the three analytic steps with two case studies, motion verbs *ZOU* 走, and *NA* 拿; section six and seven provide another kind of verb—perception verbs—*TING* 聽 and *KAN* 看 as illustrations. The final part of this paper wraps up the study with a conclusion and invites more academic discussion and interest on the issue of ambiguity in Mandarin.

2. Literature review

In analytic (or theoretical) approaches to semantics, polysemy has been discussed for a long time but it is still puzzling nowadays. This issue is not only concerned with semantic linguistics, but also with an unsolved problem, sense disambiguation, in computational linguistics. In this section four major analytic approaches to polysemy, classical approach, relational approach, and prototype approach are presented. Since the corpus-based approach gradually becomes the major approach in analysis of linguistic research, it is also adopted in this paper. After the representation of the analytic approaches, a brief introduction of the corpus-based approach is given. Further, in order to apply this hybrid analytic disambiguating model in a computational system, an overview of previous studies involving polysemy in computational linguistics is also included. Finally, a brief generalization of previous research is presented with some critiques.

2.1. Four major analytic approaches that involve polysemy

A lexical item is presented in terms of a process of cognitive abstraction. In order to explain this process, the trend of semantic approaches is leaded by two principles, sometimes, with opposite viewpoints: first, generalization (or reducing) of polysemy, and second, distinction (or increasing) of polysemy. Depending on generalization, linguists try to generalize the discussion of polysemy in order to make an explanation of the theory more convincing. While, according to distinction, an accounting of the semantic details of polysemy, researchers try to find out as many distinctions as possible (Yale and Claudis, 2000). These diverse principles, classical approach, prototype approach, frame-based approach and relational approach provide four major different perspectives to polysemy and are introduced in the following sub-sections.

2.1.1 Classical approach

According to the classical approach, it is traditionally assumed that an individual entity is composed of a set of cognitive categories. For example, in the sentence, *John is a man*, if John possesses a number of necessary and sufficient properties (also features of this approach) that define the category *man*, he is a man. Following this concept, a new semantic explanation of the classical approach is developed by Katz (1972). He claims that when giving the sense of a word a conceptual schema should be provided rather than discussing the relationship of the meaning to the word. In this scheme, word knowledge is decomposed into numerous meaning features (which Katz called "conceptual categories") by necessary and sufficient conditions. In his principle, a lexicon consists of semantic components; and related senses might share some semantic features. For instance, *chair* might be decomposed into *object* and *physical*; besides, *chair*, *bottle*, and *window* may share the same semantic marker, *object*. Moreover, in this schema, even a distinctive semantic feature could be a significant hint to distinguish different senses of polysemous words.

However, the principle of Katz's claim brings about some problems of polysemy. First, infinite semantic features may generate infinite senses. Further investigating Katz's theory, Ravin (1990) proposes that "there are no clear criteria for which aspects of a real world situation are relevant to the semantics of a particular verb, but there is a methodology for determining which aspects ought to be semantically represented." Second, the classical approach does not emphasize how the semantic components can help us disambiguate polysemous words when different senses are realized in the same expression, that is, there is no mention about the syntactic behaviors of lexical items in the approach. Following Ravin's statement, a methodology is necessary and will be given in the following section to define the senses of polysemous words.

2.1.2. Prototype approach

In the classical approach, the view of word meanings as consisting of necessary and sufficient conditions has been questioned, especially in the philosophy of language. For example, Wittgenstein (1958) claims:

> The idea that in order to get clear about the meaning of a general term one had to find the common element in all its applications has shackled philosophical investigation for it has not only led to no result, but also made the philosopher dismiss as irrelevant the concrete cases, which alone could have helped him to understand the usage of the general term.

In Wittgenstein's (1953) famous discussion of the meaning of the word *game*, he concluded that categories of meanings are familiarly resembled. This approach is further introduced in psychology by Rosch (1977). She demonstrated that people categorize objects not depending on necessary and sufficient conditions, but by relying on the resemblance of these objects to the prototypical members. In her studies, Rosch did find that people categorize objects by the concept of prototypes. For example, in the Danni culture, they have only two color categories—one represents all light, warm colors and the other represent all dark, cool colors. Rosch found that in most conditions, they recognized prototypical red, yellow and white as being in the first category, and prototypical blue, and black in the second category. Rosch also claimed that there are two prototypical models: in the first one, a single prototype contains the largest number of characteristic features and in the second one several prototypes each contain a different set of characteristic features. Linguists have adopted the second one to deal with polysemy.

With the concept of prototypes, Fillmore (1982) proposed one of the earliest discussions about prototypical meanings. He defined a word's meaning by the components it resembles. When the meaning encompasses all of the components,

the use of the word is the most prototypical. When the meaning has none of the components, the use of the word is not prototypical. And when the meaning holds some of the components, the use of the word is peripherally prototypical. Taylor (1989) gives a more direct emphasis on connecting polysemy with the concept of prototypes: 'if different uses of a lexical item require, for their explication, reference to two different domains, or two different sets of domains, this is a strong indication that the lexical item in question is polysemous.' For example, *school* can be understood as the education of children as well as the administrative structure of a university which can be classified in different domains, thus, it can be viewed as a polysemous word. Further, Taylor adds another type of prototypical category—one without central meaning. For example, *over* can express a static relation of being vertical without contact with the reference, as in "the apple is over the table"; or a dynamic relation of being vertical without contact with the reference as in "the plain flew over the country." In *walk over the blocks* expresses a dynamic relation without involving contact, and so on (Ravin and Leacock 2000). With these prototypical categories, word meaning is defined by the resemblance in the prototype. However, in this approach there is no clear discussion of how to distinguish the meanings of polysemous words.

2.1.3. Frame-based approach

In addition to prototype concepts, Fillmore *et al* (1992) proposed frame semantics in which a word's meaning is defined by a cognitive frame—when one word's expression is compatible in this frame, it denotes the meaning of this frame. A frame is determined by our background knowledge and experience with the lexicon. That is, a lexical meaning is identified by a structured cognitive schema in our mind. Based on this notion, Fillmore built a frame-based online dictionary in which different senses of polysemous words are linked to various cognitive structures (or "frames"),

and the knowledge of the frame is encoded by the words. In his "frame-based" approach, the concept of "frames" makes it helpful in reconsidering polysemy. For example, it is known that there are two senses of the verb RISK which are RISK as "put at risk" and RISK as "face the risk of." These two senses occur in different contexts where they are found in very different syntactic structures, thus, the two senses vary from each other by their specific syntactic behaviors. Therefore, different usages of the verb RISK might be necessary to help identify the specific sense of it. The interrelations between Frame and Syntax, thus, become a very important issue in Fillmore's studies, and by this, a different concept of each sense helps distinguish polysemous words. Nevertheless, this approach still cannot account for the situation when the different senses of a polysemous word appear in the same expression.

2.1.4. Relational approach

Relation models are widely used to form a semantic network. In these models, words are organized depending on the semantic relations between their meanings. Similar to the prototype approach, the relational approach also deals with semantic fields. Word relations according to this approach include Synonymy, Atonymy, Hypernymy, and Hyponymy and so forth. Synonymy can be defined as when two words can be substituted for each other in a context without changing the meaning of the clause. Atonymy is defined as substitution of two words in a context that have opposite meaning in the phrase. Hypernymy (superordinate relation), also called IS A relation is the linkage between lexical items in a specific-general relationship. Hyponymy, the opposite relation to Hypernymy is the association between lexical items in a general-specific relationship. The relational approach is ideal for inferring, especially the transitive properties of word relation. For example, the hypernym of *book* is *publication*, and the hypernym of *publication* is *piece of work*. Because of

the transitive relation, an assumption could be concluded that the hypernym of *book* is *piece of work*.

Based on the relational approach, the online dictionary, WordNet (Fellbaum, 1998), was developed by George Miller and his colleagues at Princeton University. As a source of related words for target sense in queries, WordNet indeed provides an improved solution. For example, looking up the word *board* in the noun hierarchy of WordNet, the 'lumber' sense of *board* could be detected by the hint of its related word *nail*, *hammer*, and *carpenter*. When talking about the polysemous verb, however, in this network, no information about syntactic relations is given. As Ravin and Leacock (2000) stated, "…most relational approaches maintain the classical division of sense distinction for polysemous words but they do not decompose the meaning of concepts". Further, the relation of meanings of polysemous words might be far distant, thus their meaning relation cannot be defined by the semantic relation in the semantic network.

2.2 Corpus-based approach on polysemy

According to Douglas et al., there are four essential characteristics for defining what the corpus-based approach is (cf. Douglas, Susan, Randi, 1998:3-4):

- λ It is empirical, analyzing the actual patterns of use in natural texts;
- λ It utilizes a large and principled collection of natural texts, known as a "corpus," as the basis for analysis;
- λ It makes extensive use of computers for analysis, using both automatic and interactive techniques;
- λ It depends on both quantitative and qualitative analytical techniques.

With these features, corpus data can expose the general distribution information of lexical items that native speakers will not readily ascertain by intuition. Rather than discuss what is theoretically possible in language, the significance of the approach,

corpus-based approach is that it is more concerned about how exactly the language is used in daily life.

Naturally occurring data coming from corpus gives previous approaches a brand-new perspective to re-investigate research. The advantage of the corpus-based approach is that it provides a large database of naturally occurring data and from the data, observational generalization and significant statistic analysis can be more convincing. The range of variation in language is more honestly represented in the corpus. Further, naturally occurring data show distributional tendencies for linguistic analysis. As the target issue, polysemy can be more effectively solved by looking at corpus distribution of polysemous words. The studies of Fillmore *et al.* (1992) and Liu (2004) convincingly showed the merit of corpus-based approach in analyzing polysemous words.

From the corpus data, Fillmore *et al.* (1992) generalized numerous semantic concepts (frames) depending on their different sets of categories (core frame elements) realized in syntactic behaviors (basic patterns). For example, in the frame of *risk*, the core frame elements include Chance, Harm, Victim, Valued Object, (Risky) Situation, Deed, and Actor. These core frame elements are realized as different syntactic patterns, such as, the core frame elements *value object* (VO), and *situation* (Sit) are realized as: ^{VO}{NP} ^{Sit} {Prep NP} for the example(Fillmore *et al* 1992:87):

He was being asked to risk.

VO {he}

Sit {being asked to risk}

In reference to polysemy, Fillmore *et al* (1992) claimed that in addition to the sense extension by metaphor and so on, if the verb *risk* is realized as "put at risk" in one context but as "face the risk of" in another, these must be taken as evidence for different senses of the verb. Combining grammatical characteristics with semantic

properties Fillmore *et al* (1992) presented two kinds of polysemy, one kind "resulting from a transfer of a semantic frame to a new domain (through metonymy or metaphor, for example) and the kind that reflects merely the accommodation of a word to different syntactic patterns" (Fillmore *et al* 1992). Both are discussed in this study.

Following Fillmore *et al* (1992), Liu and Wu (2004) provide one of the earliest studies applying frame-based approach with respect to Mandarin polysemy. The goal of their paper was to investigate how meanings are different or related to each other by the case study of encoding verb *biao3 shi4* 表示. Through this case study, Liu and Wu (2004) showed that differences among the identical forms of a lexical (*biao3 shi4* 表示, for example) can be explained by a systematic matching via a "conceptual schema"

According to HowNet[1], Liu and Wu (2004) found that there are three definitions of *biao3 shi4* 表示, express, expression, and show emotion. Using WordNet, among the 7 senses of *express* listed, only 4 are linked to Mandarin *biao3 shi4* 表示. They are *express* as in *She showed her disappointment*, *verbalize* as in *She expressed her anger*, *state* as in *Could you express this distance in kilometers*", and *convey* as in *His voice carried a lot of anger*. Data from the corpus show that *biao3 shi4* 表示 can be similar to English *say, point out, state, add, describe, explain, note, affirm, chuckle, mutter, tell, express, refer to, show, indicate, mean, and represent*. The problem is, how many meanings does *biao3 shi4* 表示 have and what principles are used for distinction? Based on the syntactic behaviors, they classified the data into three groups (Liu and Wu, 2004):

[1] HowNet is an on-line common-sense knowledge base unveiling inter-conceptual relations and inter-attribute relations of concepts as connoting in lexicons of the Chinese and their English equivalents.

Group 1-*biao3 shi4* 表示 1

(4). 李先生**表示**：「這不過是做好分內的事」。

‘Mr. Li **says**, ‘‘I just did what I’m supposed to do.’’’

(related English equivalents: *say, point out, state, add, explain, note, affirm, chuckle, mutter, tell* and so on.)

Group 2-*biao3 shi4* 表示 2

(5). 李先生**表示**同情。

‘Mr. Li **expressed** his sympathy.’

(related English equivalents: ***express***, *show* and so on.)

Group3-*biao3 shi4* 表示 3

(6). a. 我這麼說並不表示我不重視可能的弊端

Saying so doesn’t **mean** that I am being taking lightly the possibility of creating abuses.

b. 鮮花**表示**愛情。

‘Fresh flowers **represent** love.’

c. 一支國旗**表示**一萬尾烏魚，

One flag **represents** 10,000 grey mullet.

(related English equivalents: *mean, show, represent, indicate, carry* and so on.)

Each group can correspond to different frames in FrameNet by linking them to their meaning in English. That is, *biao3 shi4* 表示 1 corresponding to *say* in English is in Statement frame which includes core frame elements Speaker, and Message; *biao3 shi4* 表示 2 corresponding to *express* is in Encoding frame which includes core frame elements Speaker, and Message; *biao3 shi4* 表示 3 corresponding to *represent* is in Evidence frame which includes core frame elements Sign, and Message. In

addition to frame information, Liu and Wu (2004) also provide more evidence to support this classification by collocation associations from four parts, the semantic attributes of core frame elements, the combination with manner, the aspectual markers and the negatives. Depending on their criteria, to some degree, polysemous words can be explained and defined well. However, as discussed in this paper, it is found that although these three senses of *biao3 shi4* 表示 belong to three different frames, they have similar sets of core frame elements. For example, in *biao3 shi4* 表示 1 and *biao3 shi4* 表示 2, they both contain Speaker and Message. The problem is when these two core frame elements are realized as the similar syntactic behavior (same patterns), how could the classification be completed? This is the central concern in my study.

2.3 Computational linguistics

Computer applications which involve handling the content of natural language need to be concerned with the issue of polysemy. In recent works, the main focus in Natural Language Processing (NLP) was on collocations, i.e. target lexical item co-occurs with preferring lexical items. However, only searching for collocation raises the problem that the co-occurring lexemes found in corpus data are usually unexpected. There are two considerable problems of collocation based on statistical methods, first, "low precision" and second, difficulty in dealing with "rare collocations" (Li et al., 2005). Moreover, collocational patterns provide a lack of adequacy of grammatical descriptions. To extend collocational analysis Collostructions are proposed (Anatol and Gries 2003). Collostructing is to attract lexemes which are associated with a particular construction; the combination of a collexeme and a collostruct is referred to as a collostruction. However, this device faces the problem that the extracted collexeme and a collostruct might be unexpected as well. Moreover, if no collexem and collostruct association is found, how do we

disambiguate polysemous words? As mentioned above, therefore, determining the correct sense of a query word by detecting collexem or collostruct is unlikely to be successful.

More recently, another important method which can help disambiguate is the finding of topic and local components. A study of 'Disambiguating Highly Ambiguous words'(Towell and Voorhees 1998) explores contextual representations by using neural networks to extract both topical and local contexts and combining the results of the two networks into a single word sense classifier. The topical component refers to the word co-occurring with the specific sense of a target word frequently, while the local component contains the syntactic information of the sense. This method has similar concerns to theoretical approaches that combine semantic information and syntactic realization.

Although utilizing topical and local components to help identify word senses seems more accurate, there is another perspective which Towell and Voorhees (1998) did not consider. It is found that in previous studies there is insufficient information in each study to provide a highly accurate disambiguator with convincing theoretical dependence. In this paper, a hybrid model based on the frame-semantic approach combined with syntactic and pragmatic (discourse) properties is provided. In searching for the most effective way of investigating polysemy, a hybrid analysis could provide a theoretical dependence conduit module to build a sense diambiguator.

2.4 Summary of previous works

The studies reviewed in this section all deal with polysemy from different points of view. Among these approaches, Katz's classical schema (1972), and Fellbaum's WordNet (1998) give a clear explanation of polysemy. However, they established the relationship between word meanings without investigating naturally occurring language. In addition, they were not concerned with the effect of syntactic

behavior on the lexical meaning. In contrast, by combing the semantic category and syntactic behavior, frame-based approach (Fillmore *et al*, 1992) investigation of corpus data provides a more convincing discussion of polysemy. But, still unsolved are the further distinctions which need respective collocational and pragmatic association. In Mandarin, besides the study of Liu and Wu (2004), few touch upon the issue of polysemy investigating corpus data. The problem remaining is how to disambiguate polysemous words completely; this is the major focus in this paper.

3. Methodology and data

3.1. Data

The data, in this study, are extracted from the Academia Sinica Balanced Corpus (Sinica corpus). The Sinica corpus, containing a total 5 million words with part-of-speech tagging (Huang *et al* 1996), is the largest balanced corpus containing both written and spoken contemporary Mandarin data. This corpus was established by the Chinese Knowledge and Information Processing (CKIP) group at Academia Sinica, Taiwan, and it is open to research through the Internet: http://www.sinica.edu.tw/ftms-bin/kiwi.sh/. In this corpus, over 200 entries are found for each case study, but due to limited time only 200 entries are tagged in detailed in this paper. Other websites utilized in this study include FrameNet (by Fillmore 1992): http://framenet.icsi.berkeley.edu/, and Sinica BOW (The Academia Sinica Bilingual Ontological Wordnet by sinica research group): http://bow.sinica.edu.tw/.

3.2 Methodology

In searching for a method of investigating ambiguity, corpus-based approach convincingly shows the advantage of looking at corpus distribution of ambiguous words. In addition to the obvious syntactic variations which can be easily dected by native speakers' intuition, in corpus data, there are some implicit distributional differences which are not directly recognized by speaker intuition. Therefore, depending on the corpus-based approach, this paper intends to further explore semantic and syntactic relations within the senses of an ambiguous word in the corpus.

Following the approach adopted by Liu *et al* (2004), the first step of the proposed model is to identify the senses of an ambiguous word corresponding to FrameNet. In this step, the extracted data from Sinica Corpus is roughly classified into several

groups by their various collocations of core arguments. Using the Chinese-English translations in the BOW online dictionary, these different groups are related to various senses in English corresponding to Chinese. It is found that each sense of the ambiguous word extracted from the corpus did relate to different frames in FrameNet. The various core argument collocations of each sense are also similar to various basic patterns with core frame elements in the different frames in FrameNet. A small tag corpus with core frame elements of four case studies is established for this study[2].

Sense is further identified by the second step—Colloconstrucion. In this module, first, a search for categorical collocation from Sinica Corpus is executed. The range of the collocations is set up between 5 lexical units preceding and following the target verb. Then, only the co-occurring categorical types of non-core arguments are addressed. These non-core arguments are concluded to be various categorical collocates. Within each categorical collocate, some lexical items with specific meanings are found to frequently appear with different senses of the target verb.

In advance of defining the sense of problematic examples in the second step, the third module—contextual information—is necessary. In this step, the relevant lexical items are scrutinized in the context of where the target word exists. The way to look for relevant lexical items is to investigate them in previous or following clauses, that is, the relevant items would be found across clausal boundaries (usually within the range of one clause in the front or back of the target clause). The relation between the target word and the relevant items is associated by their semantic similarities. The semantic similarities are established by relating to the related wordnet synsets, a set of near-synonyms, in BOW.

---

[2] See appendix I.

4. Case study of the motion verb *ZOU* (走)

The motion verb *ZOU* 走 is an ambiguous word with high frequency. Liu and Lien (2004) mentioned that in Taiwanese (which is considered to be a dialect of Mandarin) *ZOU* 走 has multiple meanings varying from its "conceptual structure" and "semantic structure". Therefore, *ZOU* 走 can be utilized as a representative case for research on polysemy. In Sinica Corpus, there are more than 1000 entries for *ZOU* 走. However, for the purpose of economy only 200 entities are tagged in detail. But all the data are discussed and investigated in this case study.

4.1 Frame-based Sense identification

According to the model proposed in this paper, by the first step, most examples of *ZOU* 走 from Sinica Corpus can be tentatively classified into four major groups based on their different collocations with core arguments. Adopting the Chinese-English translations in the BOW online dictionary, these four groups can be related to various senses in English corresponding to Chinese: sense 1 as 'walking' *zoulu* (走路), sense 2 as 'moving' *yidong* (移動), sense 3 as 'visiting' *canfang* (參訪), and sense 4 as 'leaving' *likai* (離開). The distribution percentage for each sense is presented in the table below.

(8) Percentage of 200 Entries of *ZOU*

	Percentage (%)
Sense 1: walking	49.5
Sense 2: moving	13.5
Sense 3: visiting	5
Sense 4: leaving	32

As can be seen, sense 1 'walking' occurs most frequently (as shown in Table (8))

and denotes a specific physical action so that it is assumed to be cognitively salient and prototypical. The other senses, according to Fillmore's Frame Semantics (1992), are transferred from the SELF-MOTION domain (Frame) to other domains (Frames). The process of how to transfer from one source domain to the target domain is not discussed in this paper. It is more important to investigate how the different frames with their varying basic patterns consisting of core frame elements can help distinguish senses.

In this investigation, each sense of *ZOU* 走 did relate to different frames with core argument collocations corresponding to various basic patterns with core frame elements in FrameNet[3]. For example, sense 1 'walking' is contained in SELF-MOTION frame, sense 2 'moving' is in MOTION frame, sense 3 'visiting' is included in ARRIVING frame, and sense 4 'leaving' is in the DEPARTING frame. The classification of *ZOU* 走 depending on basic patterns with core frame elements is shown in Table (9) –(12).

(9) Basic Patterns with Core Frame Elements of Sense 1 'Walking'

Sense	Frame and Frame Elements	Frame Elements	No.	Basic Pattern with Core Frame Elements and Examples	(%)
Sense1: *ZOU LU* (走路) 'walking /going'	SELF-MOTION	Area, Goal, Path, Source, Self-mover,	BP1	**Self-mover < *[4] < Path** 羊男和姊妹倆[Self-mover]一齊**走**在林間小路上[Path]。	33.02
			BP2	**Self-mover < *** 我們[Self-mover]紛亂，疲憊地**走**著	25.47

[3] In Mandarin, VerbNet is the only Frame-based searching engine, but it is still under construction. Therefore, in this paper sense identification is via FrameNet through Chinese-English translations

[4] The asterisk '*' represents the target verb of each case study.

Sense	Frame and Frame Elements	Frame Elements	No.	Basic Pattern with Core Frame Elements and Examples	(%)
Sense1: *ZOU LU* (走路) 'walking /going'	SELF-MOTION	Area, Goal, Path, Source, Self-mover, Duration, Direction	BP3	**Self-mover < Direction < *** …一個酒鬼，他[Self-mover]半醒半睡地往上[Direction]**走**。…	12.26
			BP4	**Self-mover < * < Area** ..你[Self-mover]一個人**走**在威尼斯聖馬可廣場上[Area]，…	7.6
			BP5	**Self-mover < path < *** [Self-mover]帶球滿街[path]**走**。	6.6
			BP6	**Self-mover< * < Goal** …[Self-mover]再**走**著就到了遊塞納河的遊船碼頭 PontdeL [Gaol]…	3.77
			BP7	**Self-mover < * < Direction** 七人[Self-mover]這時所**走**的方向[Direction]，早已不是李文秀平日去師父居所的途徑。	2.83
			BP8	**Self-mover < Area < *** 我[Self-mover]在滿街水兵和軍官們中間[Area]**走**著	2.83
			BP9	**Self-mover < * < Duration** [CNI/Self-mover]又**走**了十來分鐘[Duration]，終於到了小敏的家 (CNI: co-referential Null identity)	2.83
			BP10	**Self-mover < Path < *** 我[Self-mover]..挑了僻靜的街道[Path]慢慢地**走**。	1.89
			BP11	**Self-mover < Path < Goal** 如果從倫敦清晨出發，[Self-mover]**走**M１A１公路[Path]下午３點左右便可抵達愛丁堡[Goal]。	0.9

(10) Basic Patterns with Core Frame Elements of Sense 2 'Moving'

Sense	Frame and Frame Elements	Frame Elements	No.	Basic Pattern with Core Frame Elements and Examples	(%)
Sense2: *YI TONG* (移動) 'moving'	MOTION	Goal, Source, Theme, Direction, Path	BP12	**Theme < *** 他與電腦對奕，…電腦[Theme]每**走**一步，聲彥就得全盤摸一遍…	40
			BP13	**Theme < Path < *** 這個車隊 [Mover]要沿著峭壁中間的公路[Path]往前**走**三公里…	25
			BP14	**Theme < Direction < *** …整個時代[Theme]要往哪裡[Direction]**走**才有希望…	25
			BP16	**Theme < Source < Goal <*** 車隊[Theme]從臺北[Source]往宜蘭[Goal]**走**…	5
			BP17	**Theme < Area < Path < *** 西方式的民主政治[self-mover]，在中國大陸[Area]還有極其長遠的路[Path]要**走**	5

(11) Basic Patterns with Core Frame Elements of Sense 3 'Visiting'

Sense	Frame and Frame Elements	Frame Elements	No.	Basic Pattern with Core Frame Elements and Examples	(%)
Sense3: *GUAN CAN* (參觀) 'visiting'	ARRIVING	Area, Goal, Self-mover, Source	BP5	**Self-mover < * < Area** 我[Self-mover]今天其實打算**走**一趟〝金洋村〞[Area],看看有沒有機會一訪〝神秘湖〞。	**50**
			BP8	**Self-mover < Area < *** 再來我們[Self-mover]到南橫[Area]**走**一趟	**30**
			BP18	**Area < Self-mover < *** 塔克金溪縱谷與司馬庫斯部落[Area]遙遙相望。我[Self-mover]希望下次有機會去**走**一趟。	20

(12) Basic Patterns with Core Frame Elements of Sense 4 ‘Leaving’

Sense	Frame and Frame Elements	Frame Elements	No.	Basic Pattern with Core Frame Elements and Examples	(%)
Sense4: *LI KAI* (離開) ‘leaving’	DEPARTING	Self-mover, Source, Area	BP2	**Self-mover < *** 我突然驚醒，她在與大家辭別，[CNI/Self-mover: 辭別<S>]果真要走了！	**95.3**
			BP19	**Source < * < Self-mover** 確實，幾年前香港[Source]移民走了一批高級職員[Self-mover]	3.1
			BP20	**Self-mover < * < Area** 回頭一望家三遠，[CNI/Self-mover]不知何事走他鄉[Area]	1.6

As shown in Table (9)-(12), sense 1 ‘walking’ is defined by the basic patterns in the SELF-MOTION frame: Self-mover < * < Path, Self-mover < *, Self-mover < Direction < *, Self-mover < path < *, Self-mover < * < Area, Self-mover < * < Goal, Self-mover < * < Direction, Self-mover < Area < *, Self-mover < * < Duration, Self-mover < Path < *, and Self-mover < Path < Goal; sense 2 ‘moving’ is identified by the basic patterns in the MOTION frame: Theme < *, Theme < path <*, and Theme < Direction < *, Theme < Source < Goal < *, and Theme < Area < Path < *; the meaning of ‘visiting’(sense 3) is determined by the basic patterns in the ARRIVING frame: Area <Self-mover < *, Self-mover < Area < *, and Self-mover < * < Area; sense 4 linked to the meaning of ‘leaving’, is according to the basic patterns in the DEPARTING frame: Self-mover＜*, Source < *< Self-mover and Self-mover < * < Area .

However, Table (9)-(12) also show the problem that some cases cannot be disambiguated by frame-based distinction. That is, it is found that different frames may have similar basic patterns with core frame elements. For illustration, the following instances are presented:

(13) a. 我[Self-mover]在　紅樹林　　裡[Area]走 (Sense 1 ‘walking’)

wo *zai* *hong shu lin* *li* *zou*

I in the mangrove inside walk

‘I walked in the mangrove’

b. 他[Self-mover]到　　紅樹林[Area]走一趟 (Sense 3 ‘visiting’)

ta *dao* *hong shu lin* *zou* *yi tang*

he goes to the mangrove walk once

‘He visited the mangrove.’

(14) a. 我[Self-mover]走在　大安森林　　公園[Area] (Sense 1 ‘walking’)

wo *zou zai* *da an* *sen lin* *gong yuan*

I walk in Da An forest park

‘I walked in Da An forest park.’

b. 我[Self-mover]走 一趟　大安森林　公園[Area] (Sense 3 ‘visiting’)

wo *zou yitang* *da an sen lin* *gong yuan*

I go once Da An forest park

‘I visited at Da An forest park’

(15) a. 我腳　好痠，我[Self-mover]沒辦法　　走了(Sense 1 ‘walking’)

wo jiao hao suan, wo *mei ban fa* *zou le*

my feet so limp, I cannot walk LE

‘My feet are so limp that I can not walk anymore.’

b. 火車早就開走了，我們[Self-mover]沒辦法走了(Sense 4 ‘leaving’)

huo che *zao* *jiu* *kai zou le,* *wo men* *mei ban fa* *zou le*

train already drive away LE we cannot walk LE

‘The train has already driven away, and we can't leave.’

As can be seen in examples (13a) and (13b), sense 1 ‘walking’ and sense 3 ‘visiting’ share the same pattern: Self-mover < Area < * and in example (14) they

share the same pattern: Self-mover < * < Area[5]. In examples (15a) and (15b) sense1 'walking' and sense 4 'leaving' show the similar problem that they share the pattern: Self-mover < * .

Why would different senses belonging to different Frames share the same basic pattern? The reason is that in frame semantics, ambiguity is caused by the transfer of one domain to another domain, but these two domains may not have totally different sets of core frame elements. That is, there might be some basic elements which are shared in various domains. As a consequence, in the realization of core frame elements, relative frames may have similar expressions. In such a situation, determining how to distinguish the different senses carrying similar basic patterns needs to be further explained. The next section provides a solution to solve this problem.

4.2 Colloconstructional distinction

When frame semantic information is insufficient, word senses can only be defined by a careful examination of colloconstruction. As Liu and Wu (2004) have mentioned, collocational association is also an important anchor for sense disambiguation. Adopting their findings, the second module—colloconstruction—is proposed in this paper for the further sense distinction. In this module, first, a search for categorical collocation from Sinica Corpus is executed (see 16 A)[6]. Then, various categorical collocates of *ZOU* 走 are found. Several categorical collocates with high frequency are found from the Table (16A) (the shaded statistic data present

[5] The same basic patterns also represented in the shaded areas in Table (9), (11) and (12).

[6] The statistics and the categories in table (16A) are adopted from Sinica Corpus. The first acronym of the categorical label represents the traditional syntactic categories (such as V(erb), N(oun), P(reposition), and so on (see appendix IV for all the abbreviations of category) except for the D and Di which means Adverb and aspectual markers, respectively. The second alphabet of the categorical label specifies the subpart characteristics of the categories.

more frequently co-occurring with target verb), such as D + *ZOU* 走, *ZOU* 走 + Di, D + *ZOU* 走 +Di and so forth (see (16B)). Looking into all the data of *ZOU* 走 in detail, the highly frequent patterns are generalized as shown in table (16C): (a) Adv(erb)[7] + V (also D + V in (16B)), (b) V + Comp(lement) (also V + VH/VC/NeuNf in (16B)), (c) V + Asp(ectual marker) (also V+ Di in (16B).), and (d) Adv(erb) + V + Asp(ectual marker) (also D + V + Di in (16B)). The distribution of these four major patterns varies from sense to sense (see 16 C).

(16)

A. Major Categorical Collocates of *ZOU* 走

category	Left 5	Left 4	Left 3	Left 2	Left 1	Target verb	Right 1	Right 2	Right 3	Right 4	Right 5	Total	%
VA	63	85	68	**111**	23	1951	18	**109**	57	67	55	2607	12.57
D	203	202	**205**	227	**687**	0	99	**146**	**182**	228	249	2428	11.70
Na	260	209	271	262	242	0	79	224	194	261	243	2245	10.82
Nh	96	116	132	155	169	0	25	91	108	93	106	1091	5.26
VH	94	122	**87**	83	**55**	0	65	**127**	**105**	88	81	907	4.37
P	82	77	**90**	**189**	15	0	**155**	33	45	57	56	799	3.85
VC	91	101	**91**	**88**	10	21	2	36	57	67	79	643	3.10
Di	39	52	46	67	**33**	0	**237**	1	29	37	33	574	2.77
Nf	58	64	51	52	17	0	2	**111**	94	42	53	544	2.62

[7] In order to search for the significant syntactic patterns, we generalize the highly frequent patterns according to (16B) in terms of traditional syntactical categories (as (16C)). However, the generalization of highly frequent syntactic collocates must come from the statistic information in table (16A) and (16B). The same procedures are also adopted in the following three case studies.

T	29	37	32	23	4	0	**270**	22	31	24	39	511	2.46
Nc	56	58	50	38	40	0	24	54	58	61	67	506	2.44
Neu	54	49	43	23	0	0	**106**	86	34	54	35	484	2.33
Others													

B. High frequency non-core arguments of *ZOU* 走

No.	Left 3	Left 2	Left 1	Target verb	Right 1	Right 2	Right 3
1	**D**	**D**	**D**	**走**	**Di**	**D**	**D**
2	**VC**	**P**	**VH**	**走**	**T**	**VH**	**VH**
3	**P**	**VA**	**Cbb**	**走**	**P**	**Nf**	**VC/VA**
4	**VH**	**VC**	**Di**	**走**	**Neu**	**VA**	
Others							

C. Distribution of non-core arguments of Four Senses of *ZOU* 走

	Adv + *ZOU*	*ZOU* + Comp.	*ZOU* + Asp	Adv + *ZOU* + Asp
Sense 1 'walking'	25.25 %	1.01 %	12.12 %	2.02 %
Sense 2 'moving'	33.33 %	11.11%	0	0
Sense 3 'visiting'	40 %	100 %	0	0
Sense 4 'leaving'	29.69 %	0	37.5 %	25 %

a. Adv + *ZOU*

I. 我索性抱了鏡子，挑了僻靜的街道慢慢地走。(sense 1 walking)

II. 其二則是近年來製造業不斷自動化或漸走技術密集的道路，致使多餘

的製造業勞工自該業中釋放出來。(sense 2 moving)

III. 而今又是一個多霧的天氣，但又有誰願意伴我[Self-mover]再走一趟石門[Area]？(sense 3 visiting)

IV. 妹妹在一旁說：鬧水災了，再不快走就完了！(sense 4 leaving)

b. *ZOU* + Comp.

I. 阿曼道：「這般大風雪中，諒他也走不遠，勉強掙扎，非死在雪地中不可。(sense 1 walking)

II. 嘉裕。燁甲特，以及昨日低檔有守的彥武是否能同時走強，則有待觀察。(sense 2 moving)

III. 塔克金溪縱谷與司馬庫斯部落遙遙相望。我希望下次有機會去走一趟。(sense 3 visiting)

c. *ZOU* + Asp

I. 走著走著，珮珍知道自己到家了，卻不敢進屋(sense 1 walking)

II. 現在，我要走了，希望下次再來的時候，住的是一片安全美麗的大地(sense 4 leaving)

d. Adv + *ZOU* + Asp

I. 迷宮般的巷道迂迴曲折，我們[Self-mover]紛亂。疲憊地走著 (sense 1 walking)

II. 老師聽了馬上吩咐班長管秩序，交代完畢後，就匆匆忙忙的走了。(sense 4 leaving)

Within the dominant collocation type in each sense, we may find the most crucial lexical groups to help identify the sense, and that would be discussed individually in the following sub-sections.

4.2.1 Sense disambiguation: sense 1 vs. sense 3

As mentioned in section 4.2, sense 1 ‘walking’ and sense 3 ‘visiting’ show the problem that they occur in the same basic patterns: Self-mover < Area < * and

Self-mover < * < Area in some cases (such as in examples (13) and (14)). In corpus, several examples are also found, such as in examples (17) and (18)

(17) a. 我[Self-mover]在 滿街 水兵 和 軍 官 們 中 間[Area]

wo *zai man jie* *shui bing han jun guan men zhong jian*

I in the stree soldiers and militaries between

走著(sense 1 'walking')

zou zhe

walk ZHE

'I am walking in the street full of soldiers and the military.

b. 再來我們[Self-mover]到南橫[Area]走一趟 (sense 3'visiting')

zai lai wo men *dao nan hen* *zou yi tang*

then we go to Nanhen walk once

'Then, we visited at Nanhen.'

(18) a. 東尼[Self-mover]走在他身後 (sense 1 'walking')

dong ni *zou zai t ashen hou*

Tony walk in his after

'Tony walked after him.'

b. 我[Self-mover]今天其實打算走一趟〝金洋村〞[Area] (sense 3 'visiting')

wo *jin tian qi shi* *da suan zou yi tang jin yang cun*

I today in fact plan walk once 'Jin Yan village'

'Today, in fact, I planed to take a visit at 'Jin Yan village.'

In order to distinguish sense 1 and sense 3 in these cases, we need to go into the second step—colloconstruction. In corpus, it is found that sense 3 'visiting' usually co-occurs with the syntactic collocate: V + Comp. while sense 1 'walking' frequently co-occurs with the syntactic collocate: Adv + V(as shown in (16C)). When look into each collocate, sense 3 'visiting' constantly appears with the verbal measure words,

yitang/yizao 一趟/一遭 'once', while sense 1 'walking' does not have any specific lexical collocation either in the pattern Adv + V or V + Adv. Consider the following usages of sense 3 'visiting' from the Sinica Corpus:

(19) a. 到 小人國 走一遭，…

dao xiao ren guo zou yi zao

in 'Xiao Ren Guo' zou yi zao

'to take a look at Xiao Ren Guo'

b. 我 今天 其實 打算 走一趟 "金洋村",…

wo jin tian qi shi da suan zou yi tang jin yang cun

I today in fact plan walk once 'Jin Yang village'

'Today, in fact I planned to visit 'Jin Yan village.'

c. 農委會 建議 民眾， 何妨 走一趟 休閒農場，…

nong wei hui jian yi min zhong he fang zou yi tang xiu xian nong chang

The COA suggest people why not walk once recreation farm

'The Council of Agricultural suggests people to visit the recreation farm.'

These examples tell us that the measure words, such as *yitang/yizao* 一趟/一遭 'once', are the crucial indicators to trigger the sense 'visiting' of *ZOU* 走 while these collocates are not significant to sense 1 'walking' (consider the Table in (20)).

(20) Frequency of Co-occurring with y*itang and yizao*

collocates	*ZOU* + *yitang*	*ZOU* + *yizao*	Total
Sense 1: walking	71.42%	28.58%	100%
Sense 3: visiting	0	0	0

And the colloconstructions of sense 3 found in the corpus are:

ZOU 走 + Frequency adjunct

{ *yitang* 一趟
  yizao 一遭
  yihue 一回
  …
}

In this case, although sense 1 'walking' and sense 3 'visiting' share the same patterns: Self-mover < Area < * and Self-mover < * < Area, they are re-exemplified in (21) and (22). The Colloconstructions within the basic patterns of sense 3: Self-mover + Area + *ZOU* + *frequency adjunct* and Self-mover < *ZOU*+ *frequency adjunct* < Area help distinguish the two senses in these xamples.

(21) a. 我[Self-mover]在 紅 樹林 裡[Area]走 (Sense 1 'walking')

wo *zai* *hong shu lin* *li* *zou*

I in the mangrove inside walk

'I walked in the mangrove'

b. 他[Self-mover]到 紅樹林[Area]走一趟 (Sense 3 'visiting')

ta *dao* *hong shu lin* *zou* *yi tang*

he goes to the mangrove walk once

'He visited the mangrove.'

(22) a. 我[Self-mover]走在 大安森林 公園[Area] (Sense 1 'walking')

wo *zou zai* *da an sen lin* *gong yuan*

I walk in Da An forest park

'I walked in Da An forest park.'

b. 我[Self-mover]走一趟 大安森林 公園[Area] (Sense 3 'visiting')

wo *zou yitang* *da an sen lin* *gong yuan*

I go once Da An forest park

'I visited at Da An forest park'

4.2.2 Sense disambiguation: sense 1 vs. sense 4

More data in Sinica Corpus could be found to indicate that sense 1 'walking' and sense 4 'leaving' share the pattern: Self-mover < * , as illustrated in example (23).

(23) a. 我[Self-mover]必須 再踏步 走 (sense 1 'walking')

wo *bi xu* *zai ta bu* *zou*

I have to again march

'I have to march again.'

b. 胡適[Self-mover]先 叫 車 走了 (sense 4 'leaving')

hu shi xian jiao che zou le

Shi, Hu first call car walk LE

'Shi, Hu hired a car to leave first.'

Relying solely on basic patterns with core frame elements, it is difficult to tell the differences between instances of sense 1 'walking' (as in example (23a)) and sense 4 'leaving' (as in example (23b)) because they occur with the same core frame elements and basic pattern. In order to disambiguate sense 1 and sense 4 in these kinds of sentence, we go into the next step—colloconstruction. Table (16C) tells us that sense 4 'leaving' dominantly occurs with the syntactic pattern: V + Asp, while sense 1 'walking' usually co-occurs with the syntactic pattern: Adv + V. When look into the syntactic pattern, it is found that the verb-final aspectual marker *le* 了 might be a crucial anchor to identify sense 4 because it has quite distinct distributions within clauses containing sense 4 'leaving' as can be seen in Table (24). However, sense 1 'walking' does not have any significant lexical collocation. The verb-final aspectual marker *le* 了 is not a crucial indicator for sense 1 (as in (24)). For illustration, please consider the following Examples (25a), (25b) and (25c) from Sinica Corpus:

(24) Frequencies of *ZOU LE* 走了 within The Clauses of Sense 1 and Sense 4

Sense	Co-occur with verbal *le* 了
'leaving'	40/64 (62.5%)
'walking'	9/99(9.1%)

(25) a. 美美 說 了一大串，然後 頭 也不回 的走了。(sense 4 'leaving')

mei mei shuo le yi da chuan ran hou tou ye bu hui de walk le

Mei Mei say LE a string then head also not back DE walk LE

' Mei-Mei said a major string of words, and then she left without turning back.

b. 沙爾索 也走了， 他特意 留了幾支

sha er suo ye zou le ta te yi liu le ji zhi

Saelso also walk LE he especially leave LE few

大麻煙 給我 (sense 4 'leaving')

da ma yan gei wo

hemp smokes give me

'Saelso also left, and he especially left a few cigarrets for me.'

c. 他…匆匆忙忙 的告別走了 (sense 4 'leaving')

ta cu cu mang mang De gao bie zou le

he in a hurry DE farewell leave

' He said goodbye in a hurry and left.'

The colloconstruction of sense 4 is:

ZOU 走 + verbal *le* 了

As shown in (24), the possible anchor *le* 了 indeed has a notable frequency co-occurring with the sense 4 'leaving'. This observation seems to point out the fact that most examples of sense 4 'leaving' co-occurring with *le* '了' forms a colloconstruction - *ZOU* 走+ *le* 了, which helps us identify the different senses as exemplified in example (26a) and (26b).

(26) a. 我[Self-mover]必須　再踏步　走 (sense 1 ‘walking’)

wo *bi xu* *zai ta bu* *zou*

I have to again march

‘I have to march again.’

b. 胡適[Self-mover]先　叫　車 走了 (sense 4 ‘leaving’)

hu shi xian jiao che zou le

Shi, Hu first call car walk LE

‘Shi, Hu hired a car to leave first.’

However, in (24), it also presents that sense 1 ‘walking’ also appears with verbal *le* 了 in a few cases (9.1%) and it also appears with the same Colloconstruction: *ZOU* 走 + verbal *le* 了 within the same basic pattern-Self-mover < *ZOU* 走 + *le* 了. For instance, note the following examples:

(27) a. **…**，我[Self-mover]沒辦法　走了，… (sense 1 ‘walking’)

wo *mei ban fa* *zou LE*

I can not walk LE

‘I can not leave.’’

b. **..**，你[Self-mover]就要走了，…　(sense4 ‘leaving’)

ni *jiu yao zou le*

you then want walk LE

‘You are going to leave.’

Comparing (27a) and (27b) above, sense 1 ‘walking’ and sense 4 ‘leaving’ are almost identical in surface structure as they share the follows:

Shared Core Frame Elements: self_mover

Shared Basic Pattern: Self-mover <*

Shared Colloconstruction: *ZOU* 走+ *le* 了

As illustrated in (27a) and (27b), it is difficult to distinguish sense 1 and sense 4

when they are identical in surface structure, therefore, we have to go into the next step—contextual dependence. The further distinction of these senses is discussed in the next section.

In what way, exactly, can Colloconstruction help disambiguate? The answer is: when frame-based semantic roles and patterns fail to disambiguate the two senses of a polysemous word. This resolution conforms to the perspective of Emergent Grammar, as Firth (1957) contended that "…usage patterns of lexical forms can best be examined by looking at 'the company' they keep". However, given the dynamic nature of word usage, colloconstruction alone may still not be flexible enough to distinguish subtle differences of the senses of a polysemous word. Then, the problem is assumed to be solved in the next step.

4.3 Contextual dependence distinction

Since *ZOU* 走 in (27a) and (27b) denote different meanings, there might be some distinctive properties between these two clauses. In order to distinguish these two senses, additional information from the larger context is needed. In this paper, the linkage between the target sense and the relevant lexical items within the context is built by relating to the related wordnet synsets in BOW via their semantic similarity. For instance, in order to further identify the senses of (27a) and (27b), a search for relevant lexical items should be made. And, in these two examples, it is found that the relevant items, *hao3 lei4* 好累 'tired' and *hui2 lai2* 回來 'come back' might be a crucial relevant items in the re-exemplified example (28a) and example (28b), respectively[8].

[8] Thank Prof. Liu for pointing that in sentence (28a), *ZOU* also can present the meaning, 'leaving'in the sentence such as *媽好累，要走了* 'mom is so tired that mom need to leave.' In this case, we cannot say 累 'tired' is still relavent to sense 1 'walking.' There is no perfect solution to account for this problem. However, what we found is that there is a tendency between a sense and its synsets.

(28)a. 我說: 媽好累， 我[Self-mover]沒辦法 走了，要回去了。(Sense 1: walking)

wo shuo ma hao lei wo mei ban fa zou le yao hui qu le

I say mom so tired I can not w alk LE want go back LE

‘ I said: Mom is so tired that I cannot walk anymore and I want to go back.’

b. 我死之後，你[Self-mover]就要走了，永遠不會 回來 了(Sense 4: leaving)

wo si zhi hou ni jiu yao zou le yong yuan bu hui hui lai le

I die after you want walk le forever never come back LE

‘After I die, you will leave and never come back.’

The relational linkage between *hao3 lei4* 好累 ‘tired’ and the sense of *ZOU* in (28a) or between *4 hui2 lai2* 回來 ‘come back’ and the sense of *ZOU* in (28b) is established by their similar semantic properties. In BOW, a group of English synsets of sense 1 ‘walking’ are found and by translation we can find the corresponding Chinese related wordnet synsets, such as the following list (see the English related wordnet synsets in Appendix II):

(29)

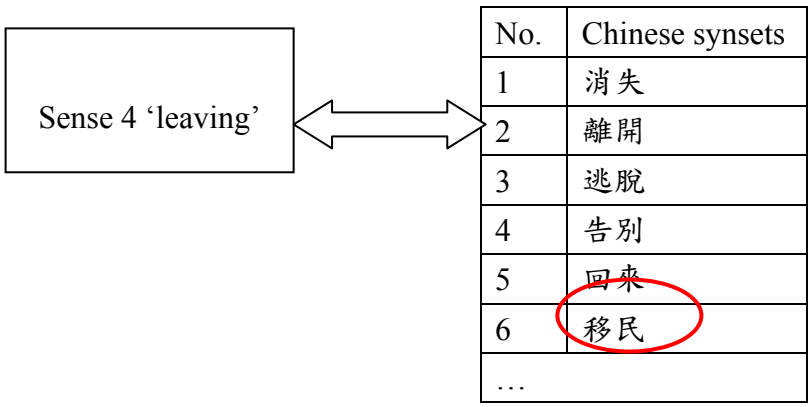
Sense 1 ‘walking’ ⟺

No	Chinese synsets
1	徒步旅行
2	拖著腳走
3	疲弱的
4	閒逛
5	散步
6	沉重的走
…	

Also, there are a number of related wordnet synsets of sense 4 ‘leaving’ in BOW:

That is sense 1 ‘walking’ tends to relate to the synsets meaning *tired*、*step*、*drag and* so on and sense 4 ‘leaving’ tends to relate to the synsets meaing *back*、*disappear*、*goodbye* and so forth.

(30)



Why can *hao3 lei4* 好累 ‘tired’ be the significant relevant item in (28a) to help identify the sense 1 ‘walking’? The reason is that *hao3 lei4* 好累 ‘tired’ is corresponding to *pi2 ruo4 de* 疲弱的 ‘exhausted’ in the related wordnet synsets of sense 1 (see (29)) because they both denote ‘tired for physical limitation’. Similarly, *hui2 lai2* 回來 ‘come back’ is the significant relevant item because it is identical to *hui2 lai2* 回來 ‘come back’ in the wordnet synsets of sense 4 (see (30)) denoting ‘returning’. As such, the sense of *ZOU* 走 in example (28a) can be identified as ‘walking’, and the sense of *ZOU* 走 in example (28b) can be defined as ‘leaving’.

5 The case of the motion verb *NA*(拿)

Another verb which is plysemous and high frequency is Motion verb *NA* 拿. In the Sinica Corpus, there are more than 1000 occurrences of *NA* 拿. Again for economy, only 200 entries are tagged and all the data are examined in the study as well

5.1 Frame-based sense distinction

Initially, most examples of *NA* 拿 from Sinica Corpus can be roughly classified into six major different groups by collocations of core arguments. Again, through the Chinese-English translations in the BOW online dictionary, these six groups can be related to various senses in English corresponding to Chinese: sense 1 as 'carrying' *chi* (持), sense 2 as 'getting' *de* (得) , sense 3 as 'utilizing' *yong* (用), sense 4 as 'giving' *nacho* (拿出), sense 5 as 'deciding' *jueding* (決定), and sense 6 as 'treating' *dui* (對). The following Table (31) shows the distribution percentage for each sense.

(31)Percentage of 200 Occurrences of *NA*

	Percentage (%)
Sense 1: carrying	51.5%
Sense 2: getting	26.5%
Sense 3: utilizing	18 %
Sense 4: taking out	2 %
Sense 5: deciding	1 %
Sense 6: treating	1 %

Also, in FrameNet, each sense of *NA* 拿 does relate to different frames according to the core arguments collocation which is corresponding to the various basic patterns with core frame elements. For example, sense 1 'carrying' is in BRINGING Frame, sense 2 'getting' is contained in TAKE Frame, sense 3 'utilizing' is included in USE Frame, sense 4 'taking out' is in GIVING Frame, sense 5 'deciding' is included in DECIDING Frame, and sense 6 'treating' is contained in INTENTIONALLY_ACT Frame. The classification of *NA* 拿 depending on Frame analysis is shown in Tables

(32)-(37).

(32) Basic Patterns with Core Frame Elements of Sense 1 ‘Carrying’

Sense	Frame	Frame Elements	No.	Basic Patter with Core Frame Elements and Examples	(%)
Sense 1: *CI* (持) ‘carrying’	Bringing	Agent, Theme, Carrier, Source	BP21	**Agent < * < theme** 好多媽媽[Agent]紅著眼，拿著手帕[Theme]，感動的氣氛帶動著每個人。	80.6
			BP22	**Agent < Carrier < * < Theme** 店內小朋友[Agent]一個個手上[Carrier]拿著一本書[Theme]，就著昏暗的燈光認真讀著。	15.5
			BP23	**Agent < * < Theme < Goal** 男方[Agent]占卜得到吉兆以後，拿著禮物[Theme]到女家[Goal]報喜	1.9
			BP24	**Carrier < * < Theme** 右手[carrier]拿了一把可折疊式的粉紅色小傘[Theme]	1
			BP25	**Agent < * < Carrier** 王質[Agent]拿到手裡[Carrier]看看	1

(33) Basic Patterns with Core Frame Elements of Sense 2 ‘Getting’

Sense	Frame	Frame Elements	No.	Basic Patter with Core Frame Elements and Examples	(%)
Sense2: *DE* (得) ‘getting’	Taking	Agent, Them, Source	BP1	**Agent < * < Theme** 弟弟想盡辦法好好讀書，[CNI/Agent:讀書<S>]拿好成績[Theme]，才能在父母眼中與哥哥有所區別。	75.4
			BP2	**Agent < Source < * < Theme** 他[Agent]一個月跟人家[Source]拿一萬[Theme]還不知羞恥！	18.9
			BP3	**Agent < theme < *** 雖然指導老師[Agent]都沒有鐘點費[Theme]可拿，卻依然辦得有聲有色。	3.8
			BP4	**Source < Theme < *** 所有的地下鐵車站和火車站[Source]均有免費的地圖及時刻表[Theme]可拿	1.9

(34) Basic Patterns with Core Frame Elements of Sense 3 ‘Using’

Sense	Frame	Frame Elements	No.	Basic Patter with Core Frame Elements and Examples	(%)
Sense 3: *YONG* (用) ‘utilizing’	Use	Agent, Instrument, Purpose, Role, Theme	BP26	**Agent < * < Instrument < Purpose** 第二天看到那隻蛇，老師[Agent]拿那個電魚的[Instrument]去電他[Purpose]	94.4
			BP27	**Agent < *< Instrument < Role** 不順的時候，我[Agent]不應該拿你們[Theme]當出氣筒[Role]	5.6

(35) Basic Patterns with Core Frame Elements of Sense 4 ‘Giving’

Sense	Frame	Frame Elements	No.	Basic Patter with Core Frame Elements and Examples	(%)
Sense 4: *NA CHU* (拿出) ‘taking out’	Giving	Donor, Recipient, Theme	BP28	**Donor < * < Theme < Recipient** 父親卡里夫[Donor]拿了兩千元[Theme]給他[Recipient]	50
			BP29	**Donor < *<Theme** Ｐｏｎｔ—Ｒｏｙａｌ是路易十四[Donor]拿錢[Theme]出來蓋的	50

(36) Basic Patterns with Core Frame Elements of Sense 5 ‘Deciding’

Sense	Frame	Frame Elements	No.	Basic Patter with Core Frame Elements and Examples	(%)
Sense 5: *JUE DING* (決定) ‘deciding’	Deciding	Cognizer, Decision	BP30	**Cognizer < * < Decision** 我[Cognizer]當時也拿不定主意要不要嫁他[Decision]	100

(37) Basic Patterns with Core Frame Elements of Sense 6 ‘Treating’

Sense	Frame	Frame Elements	No.	Basic Patter with Core Frame Elements and Examples	(%)
Sense 6: *DUEI* (對) ‘treating’	Intentional_act	Agent Act	BP31	**Agent < * < Act** 全家人[Agent]拿你沒辦法[Act]，你真是頑皮	100

As shown in (31), sense 1 ‘carrying’ is assumed to be cognitively salient and

prototypical. First, because sense 1 occurs most frequently and second, because it denotes a specific physical action. Other senses are viewed as extensions of sense 1 'carrying'. Following the procedures of my model, these senses are also specifically defined by basic patterns with core frame elements. As shown in (32)-(37), sense 1 'carrying' is defined by the basic patterns in the BRINGING Frame: Agent < * < Theme, Carrier < * < Theme < Goal and Agent < Carrier < * < Theme; sense 2 'getting' is specifically defined by the basic patterns in the TAKING Frame: Agent < * < Theme, Source < Theme < Agent < * , Agent < Theme < * , and Agent < Source < * < Theme; sense 3 'utilizing' is identified by the basic patterns in the USE Frame: Agent < * < Instrument < purpose, and Agent < * < Instrument < role, sense 4 'taking out' is specified by the basic patterns in GIVING Frame: Donor < * < Theme < Recipient, and Donor < *<Theme; sense 5 'deciding' is defined by the basic patterns in Deciding Frame: Cognizer < * < Decision; sense 6 'treating' is identified by the basic patterns in INTENTIONALLY_ACT Frame: Agent < * < Act. All the basic patterns and examples can be seen in Tables (32)-(37). In this case study, the frame-based distinction exactly separates sense 3 'utilizing', sense 4 'taking out', sense 5 'deciding', and sense 6 'treating' from the remaining two senses.

However, a similar problem is shown in (32) and (33) by frame-based distinction. That is, different frames have similar basic patterns with core frame elements (as the shaded areas). Such as, the pattern: Agent < * < Theme occurs in both of sense 1 'carrying' and sense 2 'getting'. For illustration, please consider the following examples:

(34)a.…病人[Agent]拿 著 健保卡 [Theme] 上門… (sense 1 'carrying')

bing ren *na* *zhe* *bao jian ka* *shang men*

patien take ZHE health insurance card up door

'The patient carried the health insurance card to the counter.'

b.…我[Agent]可不可以順道　拿個 研究 學位[Theme]？... (sense 2 ‘getting’)

wo *ke bu ke yi shun dao* *na ge* *yan jiu xue wei*

I can not can by the way take CL research academic degree

‘By the way, can I get an academic research degree?’

In order to distinguish these two senses which share the same basic pattern with core frame elements, colloconstruction is adopted.

5.2 Colloconstructional distinction: sense 1 vs. sense 2

In order to disambiguate the senses in (34a) and (34b), a search for non-core argument collocations is necessary. In this step, a search for categorical collocation from Sinica Corpus is executed by the same token. Then, variant categorical collocates of *NA* 拿 are obtained (35A) and are generalized as a few highly frequent categorical collocates: D + V, V + Di, or D + V + Di and so forth. Further examining all the data of *NA* 拿, the highly frequent patterns are concluded the following three major types as shown in Table (35C): (a) Adv (erb) + V (also D +V in (35B)), (b) V+ Asp (ectual marker) (also V + Di in (35B)), (c) Adv (erb) + V + Asp (ectual marker) (also D + V +Di in (35B)). Besides, these syntactic patterns have different distribution in each sense. (see (35 C))

(35)

A. Categorical Collocates of Five Senses of *NA* 拿

category	Left 5	Left 4	Left 3	Left 2	Left 1	Target verb	Right 1	Right 2	Right 3	Right 4	Right 5	Total	%
Na	147	147	165	149	147	0	290	252	231	193	154	1875	17.07
VC	55	57	**41**	24	5	969	9	**98**	**74**	110	99	1541	14.03
D	84	90	**90**	**117**	**365**	0	**44**	**65**	**88**	95	78	1116	10.16

category	Left 5	Left 4	Left 3	Left 2	Left 1	Target verb	Right 1	Right 2	Right 3	Right 4	Right 5	Total	%
Di	21	19	16	6	5	0	**287**	1	3	6	16	380	3.46
VH	71	47	**50**	**31**	**12**	0	**23**	38	**35**	41	29	377	3.43
Nf	30	33	27	**36**	**10**	0	17	**111**	**75**	18	20	377	3.43
P	29	26	26	**40**	6	64	10	25	25	29	37	317	2.89
Neu	23	17	27	15	0	0	**96**	**66**	15	16	16	291	2.65
VA	25	33	**29**	30	**14**	0	4	31	18	38	34	256	2.33

B. High Frequency Non-core Argument of *NA* 拿

No.	Left 3	Left 2	Left 1	Target verb	Right 1	Right 2	Right 3
1	**D**	**D**	**D**	**拿**	**Di**	**Nf**	**D**
2	**VH**	**P**	**VA**	**拿**	**Neu**	**VC**	**Nf**
3	**VC**	**Nf**	**VH**	**拿**	**D**	**Neu**	**VC**
4	**VA**	**VH**	**Nf**	**拿**	**VH**	**D**	**VH**
Others							

C. Distribution of Non-core Arguments of Five Senses of *NA* 拿

	Adv + V	V + Asp	Adv + V + Asp
Sense 1 'carrying'	7.8 %	58.25 %	2.9 %
Sense 2 'getting'	41.5 %	7.5 %	0
Others			

a. Adv + V

I. 我聽見皮皮的叫聲，趕緊拿一片葉子讓皮皮爬上來(sense 1 carrying)

II. 貝克在一九八九年連拿溫布頓及美國公開賽冠軍時，其英姿勃發意氣昂揚，令人側目。(sense 2 getting)

b. V + Asp

I. 就吃完一餐飯以後，大家拿著電子計算機出來，打一打，互相掏錢(sense 1 carrying)

II. 年輕的王文輝，已累積許多經歷拿了新加坡政府獎學金出國獲取經濟學學士(sense 2 getting)

c. Adv. + V + Asp.

I. 捷運局長賴世聲在巡視工地時，也拿了一塊大的貝殼化石帶回辦公室留念。(sense 1 carrying)

Within the major collocation types of each sense, sense 1 usually co-occurs with aspectual markers while sense 2 constantly appears with adverbs. However, within the collocating adverbs, no crucial lexical collocates are found to identify sense 2 'getting'. While within aspectual markers, it is found that the aspectual marker *zhe* 著 has significant distribution in examples of sense 1 'carrying'. Therefore, the aspectual marker *zhe* 著 might be a crucial anchor triggering the 'carrying' sense of *NA* 拿 (consider (36)). In other words, sense 1 'carrying' co-occurs with the aspectual marker *zhe* 著 with considerable frequency. This combination of sense 1 'carrying' and aspectual marker *zhe* means that sense 1 'carrying' puts emphasis on the progressive state of the physical action while sense 2 'getting' does not.

(36) Frequency of Co-occurring with *zhe* 著

collocates	*NA* + *zhe*
Sense 1: carrying	47.1 %
Sense 2: getting	0

In consequence, although sense 1 'carrying' and sense 2 'getting' share the same core frame element pattern as: Agent < *NA* + *zhe* < Theme exemplified in (34a) and

(34b), the Colloconstruction of sense 1:

NA 拿 + *zhe* 著

helps distinguish the two senses which are realized in the same basic pattern: Agent < * < Theme. For example, as the following examples are shown:

(37) a. 莉絜[Agent]拿著 江如華的 名片[Theme](sense 1 'carrying')

li jie *na zhe* *jiang ru hua* *de* *ming pian*

Li-jie take ZHE Ru-hua Jiang DE name card

'Li-jie is taking Ru-hua Jiang' name card.'

b. 他[Agent]拿著 照片 與 簽名球[Theme]…(sense 1 'carrying')

ta *na zhe* *zhao pian* *yu* *qian ming qiu*

he take ZHE picture and autographed ball

'He is carrying the picture and the autographed ball.'

c. 她[Agent]拿著 一小段 的竹竿[Theme]…(sense 1 'carrying')

ta *na zhe* *yi xiao duan* *de zhu gan*

she take ZHE one small piece DE bamboo

'She is carrying one a segment of bamboo.'

However, in (36), another problem that emerges is how to distinguish sense 1 'carrying' from sense 2 'getting' when sense 1 does not co-occur with aspectual marker—*zhe* 著 (about 52.9%), such as the following examples:

(38) a. …賴世聲[Agent]…也拿 一塊大 的 貝殼化石[Theme]……(sense 1 'carrying')

lai shi sheng *ye na* *yi kuai da* *de* *bei ke hua shi*

Lai Shi-sheng too take one piece big DE shell fossil

'Shi-sheng Lai also carried a piece of big shell fossil'

b. … 艾德華[Agent] 愈 老 愈 俏，

ai de hua *yu lao yu qiao*

Edward more old more fascinating

全場　　獨拿　　廿 六　分[Theme] (sense 2 ‘getting’)

quan chang　*du na*　　*er shi liou fen*

whold field　lonely take　26 scores

‘Edward is the older and the more fascinating for he himself alone got 26 scores.’

Comparing examples (38a) and (38b) above, sense 1 ‘carrying’ and sense 2 ‘getting’ are almost identical in surface structure in that while they share the following, they occur without any specific collocational association:

Shared Core Frame Elements: Agent, Theme

Shared Basic Pattern: Agent < * < Theme

In the same way, we have to go into the next step to find contextual relevant items which can help disambiguate these two senses when they both occur without any distinctive collocates. The further distinction would be discussed in the next section.

5.3 Contextual dependence distinction

Since local information (within clause) is insufficient for sense disambiguation in (38), additional information from across clausal boundaries is necessary. In this step, each sense is assumed to be associated with other relevant lexical items through semantic linkage in Contextual Dependence. In these two examples, crossing clauses, it is found that the relevant items, *dai4 hui2* 帶回 ‘bring back’ and *de2 fen1* 得分 ‘getting score’ might be the crucial relevant items as in the re-exemplified examples (39a) and (39b), respectively.

(39) a. 賴世聲[Agent]… 也拿　一塊　大的　貝殼　化石[Theme]，

lai shi sheng　*ye na*　*yi kuai*　*da de*　*bei ke*　*hua shi*

Lai Shi-sheng　too take one piece big DE　shell fossil

說要　帶回　辦公室　留念…(sense 1 ‘carrying’)

shou yao *dai huei* *ban guong shi* *liou nian*

say want bring back office keep memorize

'Shi-sheng Lai also carried a piece of big shell fossil, and he said that he wanted to bring it home to be as a souvenir.'

b. 艾德華[Agent] 愈老 愈俏， 全場 獨拿 廿六分[Theme]，

ai de hua *yu lao yu qiao* *quan chang* *du na* *er shi liou fen*

Edward more old more fascinating whold field lonely take 26

寫下 本季 個人 最高 得分 …(sense 2 'getting')

xie xia ben ji *ge ren* *zuei gao de fen*

scores write down this season personal highest score

'Edward is the older and the more fascinating for he himself alone got 26 scores and this season recorded his personal highest score'.

The relational linking between *dai4 hui2* 帶回'bring back' and the sense of *NA* 拿 in (39a) or between *de2 fen1* 得分 'getting score' and the sense of *NA* 拿 in (39b) is established by their similar semantic properties. In BOW, a number of English related wordnet synsets belonging to sense 1 'carrying' are found and by translation we can find the corresponding Chinese related wordnet synsets in Table (40) (see the English related wordnet synsets in Appendix I):

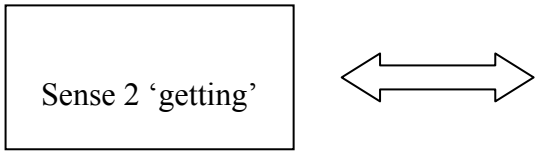
(40)

Sense 1 'carrying' ⟺

No	Chinese synsets
1	保留以備將來之用
2	以腹帶固定
3	保護的 安全的
4	攜帶 帶著
5	轉移
…	

Also, there are various synsets of sense 2 ‘getting’ in BOW:

(41)



No.	Chinese synsets
1	獲得
2	得到
3	接受
4	侵吞
5	取用
…	

How can we say that *dai4 hui2* 帶回‘bring back’ is a crucial relevant item in (39a) to help identify the sense 1 ‘carrying’? The reason is that *dai4 hui2* 帶回 ‘bring back’ is similar to *dai4 zhe* 帶著 ‘bring’ in the related wordnet synsets of sense 1 (consider (40)) because they both denote ‘the act of bringing’. Similarly, *de2 fen1* 得分 ‘getting score’ is the significant relevant item because it is linked to *de2 dao4* 得到 ‘obtaining’ in the related wordnet synsets of sense 2 (see (41)) denoting ‘getting’. In sum, in this step, through contextual information, the sense of *NA* 拿 in example (39a) can be identified as ‘carrying’, and the sense of *NA* 拿 in example (39b) can be defined as ‘leaving’.

6 The case of the perception verb *TING* (聽)

Perception verb *TING* 聽 is also a polysemous word with high frequency. In this case study, various syntactic behaviors and semantic properties which differ from motion verbs are investigated. In the Sinica Corpus, there are more than 1000 occurrences of *TING* 聽. Again, we tag only 200 entries in detailed but examine the overall data.

6.1 Frame-based sense distinction

Preliminarily, most examples of *TING* 聽 from Sinica Corpus could be roughly divided into four major groups by their collocations of core frame elements. In addition, via the Chinese-English translations in BOW, these four groups can be related to various senses in English corresponding to Chinese: sense 1 as 'listening' *TING* (聽), sense 2 as 'perceiving' *TING (dao)*(聽到)) , sense 3 as 'hearing about' *TING (guo)*(聽過), and sense 4 as 'obeying' *TING hua* (聽話). The percentages of distribution of each sense are shown in Table (42).

(42)Percentage of 200 occurrences of *TING*

	Percentage (%)
Sense1: listen	16
Sense2: perceive	64
Sense3: hear (about)	13.5
Sense4: obey	6.5

Also, each sense of *TING* 聽, in FrameNet, related to different frames according to core argument collocation corresponding to the various basic patterns with core frame elements in different frames. For example, sense 1 'listening' is in PERCEPTION_ACTIVE frame, sense2 'perceiving' is included in PERCEPTION_EXPERIENCE frame, sense 3 'hearing about' is contained in HEAR frame, and sense 4 'obeying' is in COMPLIANCE frame. The classification of *TING* 聽 depending on Frame analysis is shown in (43)-(48)

(43) Basic Patterns with Core Frame Elements of Sense 1 ‘Listening’

Sense	Frame	Frame Elements	No.	Basic Patter with Core Frame elements and Example	(%)
Sense 1: **聽** ‘listening’	PERCEPTION_ ACTIVE	Perceiver Body_part Phenomenon Manner	BP32	**Perceiver < *** 聽不懂的地方可以問你。‖好的，星期四我正好有空，我一定去**聽**。‖除了你跟我以外，我還約定平跟其他的同學，一塊兒去聽	45.71
			BP33	**Perceiver < * < Phenomenon** 大家認真**聽**每一個人在關心什麼、對公司有什麼建議	28.57
			BP34	**Phenomenon< Perceiver < *** 消極？音樂我現在連聽都不願意**聽**了。	14.29
			BP35	**Body_part < *** 曉該是早晨的意思，處處也學過，是各處的意思，聞大概是用耳**聽**不是用鼻子聞味。	5.71
			BP36	**Perceiver < Phenomenon < *** 現在不要把它當做一種混亂的聲音來**聽**，把它當做本來就應該有的聲音，那麼心裡就不會受影響。	2.86
			BP37	**Medium < Perceiver < *** 一卷錄音帶他竟然重覆**聽**了將近八十遍，然後理智地思考，選擇投入慈濟	2.86

(44) Basic Patterns with Core Frame Elements of Sense 2 ‘perceiving’

Sense	Frame	Frame Elements	No.	Basic Patter with Core Frame elements and Example	(%)
Sense 2: **聽（了）** ‘perceiving’	PERCEPTION_ EXPERIENCE	Perceiver Depictive Agent Phenomenon Reaction	BP38	**Phenomenon < Perceiver < * < Depictive** 憶山東兄弟』這首詩，也是因為我想念家鄉，才寫下來的。」我**聽**了，真有點同情他呢！	27.13

Sense	Frame	Frame Elements	No.	Basic Patter with Core Frame elements and Example	(%)
Sense 2: **聽（了)** 'perceiving'	PERCEPTION_ EXPERIENCE	Perceiver Depictive Agent Phenomenon Reaction	BP39	**Perceiver < * < Phenomenon** 事實上我們在剛才節目當中，**聽**了很多的同學講出心裡的話	23.26
			BP40	**Perceiver < *** 不論小孩子問什麼問題，我們都**聽**清楚，然後適當的給他回答	19.39
			BP41	**Phenomenon < Perceiver < * <reaction** 有人說要捐錢給政府，買飛機大砲來保衛國家。大家**聽**了，都喊著要捐錢。	8.52
			BP42	**Perceiver < * < Phenomenon < Depictive** 我**聽**了一下午的風聲蟲鳴，我內心仍有揮之不去的迷惑	8.52
			BP43	**Phenomenon < Perceiver < *** 每當王鎮華老師與爸爸在討論時，他就在旁邊**聽**。在一本筆記上，他零星寫下他對未來房子的一些想法	4.65
			BP44	**Perceiver < * < Phenomenon < reaction** 山民**聽**了這個消息，歡天喜地的走了。	3.87
			BP45	**Agent < Phenomenon< Perceiver < *** 丘子章就放了新疆各族的情歌集子給大家**聽**，其中也有秦腔，和寧夏的牧歌	3.1
			BP46	**Agent < Medium < Perceiver < *** 塞住耳朵，已聽不到了，但是同學們一直鼓勵他，放錄音帶給他**聽**，拿書給他看。	0.77
			BP47	***< Perceiver** 極深的內心翻騰、痛楚。有過這樣遭遇的人需要傾吐，需要「被**聽**」、被支持、被知道、被了解、被接受	0.77

(45) Basic Patterns with Core Frame Elements of Sense 3 ‘hearing about’

Sense	Frame	Frame Elements	No.	Basic Patter with Core Frame Elements and Examples	(%)
Sense3: 聽到 ‘Hearing about’	HEAR	Hearer Message Topic Speaker	BP48	**Hearer < * < Message** 摩俄說：大塔山。大塔山？我怎麼從來沒**聽**過這座山的名字。印象中是玉山。陽明山。太魯閣。	25.92
			BP49	**Hearer < * < Speaker < Message** 我覺得太不像樣，正打算把東尼拉起來，卻**聽**東尼說：「…去他的臭錢…我好想妳…」	22.22
			BP50	**Hearer < * < Topic** 嗯！我覺得這也是有影響。‖崇洋的關係，變得大家都，寧願去**聽**西方歌劇，‖習慣了把西洋文化抬的比較高。‖	22.22
			BP51	**Hearer < * < Speaker** 聽說你最近很臭屁哦。‖沒有啊，你**聽**誰講啊，我那有那個啊?	11.11
			BP52	**Speaker < Message < Hearer < *** 我小時候常和她一同去牧羊，她唱了很多歌給我**聽**，還說了很多故事。好幾年不見，想不到她…她竟死了。	7.4
			BP53	**Message < Hearer < *** 裡面有很多叫做公主、少爺的，有沒有聽過，不知道大家有沒有**聽**過？‖有。‖沒有一個人搖頭的。	7.4
			BP54	**Speaker < Hearer < *** 龍說：「不行。」雨神問：「為什麼？」龍說：「讓我來解釋給你**聽**吧！」	3.7

(46) Basic Patterns with Core Frame Elements of Sense 4 ‘obeying’

Sense	Frame	Frame Elements	No.	Basic Patter with Core Frame Elements and Examples	(%)
Sense 4: **聽話** ‘obeying’	COMPLIANCE	Norm Protagonist State of affairs	BP55	**Protagonist < * < State of affairs** 「你幾時肯好好**聽**我一番話？講沒兩句拍了桌子就走，不拐彎抹角成嗎？」	38.46

Sense	Frame	Frame Elements	No.	Basic Patter with Core Frame Elements	(%)
Sense 4: **聽話** 'obeying'	COMPLIANCE	Norm Protagonist State of affairs	BP56	**Protagonist < * < Norm** 遇到有民眾釣魚時，陳主任表示，校警會出面勸導，如果民眾不**聽**勸導，校方也很難做事。	38.46
			BP57	**State of affairs < Protagonist < *** 蘇魯克在一旁叫道：「蘇普，傻小子，別進去！」蘇普卻那裡肯**聽**？ 李文秀見到他這般癡情的模樣，心中又是一酸	23.07

As shown in (42), sense 1 'listening' is assumed to be perceivably salient and prototypical. Because, first, it occurs most frequently and second, it denotes a specific physical perception. Other senses are viewed as extensions of sense 1 'listening'. Following the procedures of the model provided in this paper, these senses are also specifically defined by basic patterns with core frame elements. As shown in (43)-(46), sense 1 'listening' is defined by the basic patterns in the PERCEPTION_ACTIVE Frame: perceiver < *, Perceiver < * < Phenomenon, Phenomenon < Perceiver < *, body part < *, Perceiver < Phenomenon < *, and Medium < Perceiver < *; sense 2 'perceiving' is specifically defined by the basic patterns in the PERCEPTION_PASSIVE Frame: Phenomenon < Perceiver < * < Depictive, Perceiver < * < Phenomenon, Perceiver < *, Phenomenon < Perceiver< * < reaction, Perceiver < * < Phenomenon < Depictive, Phenomenon < Perceiver < *, Perceiver < * < Phenomenon < Reaction, Agent < Phenomenon < Perceiver < *, Agent < Means < Perceiver < *, and * < Perceiver; sense 3 'hearing about' is identified by the basic patterns in the HEARING Frame: Hearer < * < Message, Hearer < * < Speaker < Message, Hearer < * < Topic, Hearer < * < Speaker, Speaker < Message < Hearer < *, Message < Hearer < *, and Speaker < Hearer < *; and sense 4 'obeying' is specified by the basic patterns in COMPLIANCE Frame: Protagonist <

* < State of affairs, Protagonist < * < Norm, State of affairs < Protagonist < *. In this case study, the frame-based distinction exactly separates sense 3 'hearing about', and sense 4 'obeying', from the other two senses. However, a similar problem is shown in (47) and (49) by frame-based distinction. That is, different frames have similar basic patterns with core frame elements. For example, the patterns: Perceiver < * < Phenomenon, Phenomenon < Perceiver < * < Depictive, and Phenomenon< Perceiver < * occur in both of sense 1 'listening' and sense 2 'perceiving'. For illustration, please consider the following examples:

(47) Perceiver < * < Phenomenon

a. 丘子章[Perceiver]…很 願意 **聽** 她發脾氣[Phenomenon]。 (sense 1 'listening')

qiu zi zhang hen yuan yi TING ta fa pi qi

qiu zi zhang very willing listen she angry

'Qiu, zi-zhang would like to listen to her being angry very much.'

b. 我們[Perceiver]**聽** 了 很多 的 同學

wo men TING le hen duo de tung xue

we hear LE very much DE classmate

講 出 心裡 的話[Phenomenon]？(sense 2 'perceiving')

jiang chu xin li de hua

speak out heart DE words

'We heard a lot of classmates speak the words of the heart'

(48) Perceiver < *

a…好的， 星期四 我 正好 有空， 我[Perceiver]一定去**聽**(sense 1 'listening')

hao de xing qi si wo zheng hao you kong wo yi ding qu TING

OK Thursday I just have free I must go to listen

'Ok, I just have free time on Thursday, and I will go to listen.'

b. 沒有 考 上 大學 開始， 父親 就 不 跟他 講話[Phenomenon]。

mei you kao shang da xue kai shi fu qin jiu bu gen ta jiang hua

No pass college begin father nerver with him talk

我[Perceiver]**聽** 了 很難過[Depictive] (sense 2 'perceiving')

wo TING LE hen nan guo.

I hear LE very sad.

'Since he did not pass the college entrance exam, father doesn't talk to him any more.I am very sad to hear that'

(49) Phenomenon < Perceiver < *

a.「妳 這樣 說， 未免 太 消極 了吧？」「消極？音樂[Phenomenon]

ni zhe yiang shou wei mian tai xiao ji le ba xiao ji yien yue

you such say rather too negative Le ba negative music

我[Perceiver]現在 連 聽 都 不願意 **聽** 了。」(sense 1 'listening')

wo xian zai lian ting do bu yuan yi TING LE

I right now even listen all not willing listen Le

"It isn't too negative?""Negative? " "Music ? I do not ever want to listen."'

b. 每當 王鎮華 老師 與 爸爸 在 討論[Phenomenon]時，

mei dang wang zhen hua lao shi yu ba ba zai tao lun shi

every when Wang, zhen-hua teacher and father in discuss when

他[Perceiver]就 在 旁邊 **聽**(sense 2 'perceiving')

ta jiu zai pong bian TING

he just be side hear

'Whenever teacher Wang, Zhen-Hua talks with father, he just hears nearby.'

As can be seen in (47)-(49), sense 1 'listening' and sense 2 'perceiving' are realized in the same basic pattern with core frame elements. In order to distinguish these two senses, we need to go into the second step—colloconstruction distinction.

6.2 Colloconstructional distinction: sense 1 vs. sense 2

Following the procedures of the proposed model, exploring for non-core argument categorical collocations is executed to further disambiguate senses (see (50A)). Several major categorical collocations are scrutinized as in (50B). In this case study, various senses are especially associated with different syntactic patterns which can be concluded as three major types as: (a) D(Adv.) + V + D(Adv.), (b) V(VK, VH) + V(*Ting*), (c) V + Di (Asp), and (d) V + Comp.(VE, VH, VC). The distributions of the collocations of sense 1 and sense 2 are shown in (50C) for further distinction.

(50)

A. Categorical Collocates of Four Senses of *TING* 聽

category	Left 5	Left 4	Left 3	Left 2	Left 1	Target verb	Right1	Right 2	Right 3	Right 4	Right 5	Total	%
VE	50	87	65	**32**	22	1829	13	**275**	**110**	94	87	2664	13.49
Na	249	277	269	222	200	0	269	191	318	276	254	2525	12.78
D	210	170	**126**	**226**	**634**	0	**158**	**191**	**206**	198	217	2336	11.83
Nh	115	96	138	188	271	0	228	84	83	94	94	1391	7.04
VH	133	106	**108**	**81**	**28**	0	**39**	**81**	**98**	92	94	860	4.35
Di	20	24	16	9	5	0	**501**	11	28	24	28	666	3.37
VC	97	86	**72**	27	11	9	21	**66**	**56**	67	78	590	2.99
T	52	69	**73**	17	0	0	**60**	34	51	27	31	414	2.10
Nb	24	35	28	50	79	0	57	35	23	25	38	394	1.99
VA	42	56	44	31	7	0	3	55	40	50	48	376	1.90
P	40	42	43	**47**	**32**	0	6	26	34	42	50	362	1.83

VK	27	22	23	20	**36**	55	5	60	37	29	44	358	1.81
Others													

B. High Frequency Non-core Argument of *TING* 聽

No.	**左 3**	**左 2**	**左 1**	**關鍵詞**	**右 1**	**右 2**	**右 3**
1	**D**	**D**	**D**	**聽**	**Di**	**VE**	**D**
2	**VH**	**VH**	**VK**	**聽**	**D**	**D**	**VE**
3	**VC**	**P**	**P**	**聽**	**T**	**VH**	**VH**
4	**T**	**VE**	**VH**	**聽**	**VH**	**VC**	**VC**
Others							

C. Distribution of Non-core Arguments of Five Senses of *TING* 聽

	Adv +*TING*	V +*TING*	*TING* +Asp	*TING* + Comp.
Sense 1 'listening'	15.63	53.13	5 %	6.25 %
Sense 2 'perceiving'	23.1 %	0	39.1 %	16.36 %

(51) Adv +*TING*

a. 什麼是天籟？就是你聽的時候要用心地**聽**，這是第一步。(sense 1 'listening')

She me shi tian lai jiu shi ni ting de shi ho yiao yong xin de TING zhe shi di yi bu

What is sounds of nature is you listen when want carefully TING this is first step.

'What is sounds of nature? That is at the moment when you heartly listen to. This is the first step'

b..「好極了，快告訴我，我乖乖地聽！」東尼嬉皮笑臉的說(sense 2 'perceiving')

hao ji le kuai gao su wo wo guai guai de TING dong ni xi pi xiao lian de shou

'exellent quick tell I I be well-behaved hear' Tony laughing face DE say

'"Exellent, quick to tell me, I will be well-behave to hear !" Tony said with laughing face.

(52) V + *TING*

a. 如果有旅行團，他一定參加。有音樂會就去聽，有畫展也去看。 (sense 1 'listening')

ru guo you lv xing tuan ta yi ding can jia you yin yue hue jiu qu TING you hua zhan ye qu kan

if have tourist group, he must attend have concert go to listen have art exhibition also go to see

'If there is a tourist group, he must attend. If there is a concert, he will go to listen, and it there is art exhibition, he will also went to see.'

(53) *TING* +Asp

a. 濟公點上香..唸道：大家聽著，王婆請我吃粉湯.. (sense 1 'listening')

ji guong dian shang xiang nian dao da jia TIGN zhe wang po qien wo chi fen tang

Taoist light the joss-stick and say: everybody listen ZHE Ms.Wang treat me the powder soup

'The Taoist light the joss-stick and say: Everybody listen, Ms. Wang treat me the powder soup'

b. 不知道那一陣風雨，又把園中的花打落了多少。我聽了，覺得這首詩確實很美(sense 2 'perceiving')

bu zhi dao na yi zhen fong yu you ba pu zhong de hua da luo le duo shao wo TING le jue de zhe shou shi que shi hen mei

don't know which gust rains and winds, again BA garden de flower blow off LE how much I hear feel this poem is really very beautiful

'*Don't know which gust of rains and winds blow off how much of the flowers in the garden again*.When I heard this poem, I thought it is really very beautiful'

(54) *TING* + Comp.

a. 我不便再問下去，沙爾索只要一開口，總是滔滔不絕。我**聽**不下去，又插不進口。(sense 1 'listening')

Sa er suo zhi yao yi kai kou zuong shi tao tao bu jue wo TING bu xia qu you cha bu jin kou

Saelso if only open mouth always blather, I listen any more again insert not interpose

'When Saelso opens his mouth, he always blathers. I cannot listen anymore, but I cannot interpose,neither.'

b. 我**聽**得懂！鳥當然不是真的會掉下來，你比喻得很好！(sense 2 'perceiving')

wo TING de dong niao dan ran bu shi zhen de hue diao xia lai ni bi yu de hen hao

I hear understand bird of course not really can drop down you analogy very good

'I heard and I can understand! The bird, of course, would not drop down and you have a good metaphor for it.'

Within these collocation types, sense 1 'listening' usually co-occurs with preceding predicates while sense 2 'perceiving' often appears with aspectual markers and adjunct complements. Since co-occurring preceding predicates are significant to sense 1, some crucial lexical collocates might be found to help identify sense 1 'listening' of *TING*. Aspectual markers and adjunct complements might equally help specify sense 2 'perceiving' of *TING*. It is found that among the collocating preceding predicates of sense 1 'listening', the lexical items with volitional meaning occupy most parts, such as 用心 'attentively', 去 'go to', and 願意 'volitionally'. In contrast, there is no specific lexical item found in the collocating adverbs of sense 2 'perceiving'. Nevertheless, within the two significant collocations of sense 2 'perceiving' (aspectual markers and with adjunct complements), some special lexical items are detected, such as the aspectual marker 了 *LE* denoting the completeness of actions, and adjunct complements 清楚 *qing chu* 'clear' denoting the resulting state.

(55) Frequency of Co-occurring with Volitional Markers, Aspectual Marker *le* and Perception Boundary Marker

	Volitional Marker + V	V + 了 *le*	V + Perception Boundary Marker (Result)
Sense 1 'listening'	68.42%	2.63%	5%
Sense 2 'perceiving'	0	32.26%	13.7%

These collocates are indicative to sense 1 and sense 2 respectively because they are semantically relevant to the target senses. In the case of sense 1 'listening', adverbial collocates are specific to sense 1 'listening' that being in Perception_active frame, sense 1 denotes volitional meanings related to those collocates. On the other hand, in the examples of sense 2 'perceiving', the aspectual marker *LE* and the adjunct complements are specific to sense 2 only. The reason for the linking of sense 2 to its significant collocates is that being in Perception_passive frame, sense 2 'perceiving' calls for a boundary to complete the perceiving acts. For this reason, it makes sense to believe that the collocates-aspectual marker *Le* and resulting complements are kind of connection with the meaning of sense 2. Consequently, by colloconstruction module, it is found that markers of volition (願意, 用心) help disambiguate sense 1 from sense 2, while markers of perception boundary (了, 懂, 清楚) could tell the differences between sense 1 and sense 2. According to the colloconstructions, it is observed that the major difference between sense 1 and sense 2 is the degree of control—sense 1 has a higher degree of control while sense 2 has a lower degree of control. Furthermore, according to the various structural behaviors, sense 1 and sense 2 must necessarily be distinguished. In other words, via the separation, the different realization of these two meanings can be seen. With this semantic connection, though sense 1 'listening' and sense 2 'perceiving' share the same core frame element patterns (Perceiver < * < Theme, Perceiver < *, and

Phenomenon < Perceiver < * as exemplified in (56a-58a) and (56b-58b)), they can be distinguished by their specific Colloconstructions:

✧ Volitional marker(s) + *TING* 聽

✧ *TING* 聽 + Perception boundary marker(s)

- TING 聽 + *le* 了
- *TING 聽 + Resulting*

For illustration, Perceiver < 願意 'willing' + * < Phenomenon (for sense 1)and Perceiver < * + 了 < Phenomenon (for sense 2) in (56), Perceiver < 一定去 'must to'+ * (for sense 1) and Perceiver < * + 了 (for sense 2) in (57), Phenomenon < Perceiver < 願意 'willing' + * (for sense 1) and Phenomenon < Perceiver < * + 了(for sense 2) in (58), please see the following instances:

(56) Perceiver < 願意 + * < Phenomenon

a. 丘子章[Perceiver]…很願意**聽**她發脾氣[Phenomenon]。… (sense 1 'listening')

Perceiver < * + 了 < Phenomenon

b. 我們[Perceiver]**聽**了很多的同學講出心裡的話[Phenomenon]？... (sense 2 'perceiving')

(57) Perceiver < 一定去+ *

a…. 好的，星期四我正好有空，我[Perceiver]一定去**聽** (sense 1 'listening')

Perceiver < * + 了

b. 他從十年前沒有考上大學開始，父親就不跟他講話[Phenomenon]。我[Perceiver]**聽**了很難過[Depictive] (sense 2 'perceiving')

(58) Phenomenon < Perceiver < 願意 + *

a.「妳這樣說，未免太消極了吧？」「消極？音樂[Phenomenon]我[Perceiver]現在連聽都不願意**聽**了。」(sense 1 'listening')

Phenomenon < Perceiver < * + 了

b. 不論小孩子問什麼問題[Phenomenon]，我們[perceiver]都**聽**清楚，然後適當的給他回答(sense 2 'perceiving')

However, in (59), another problem might emerges: how to distinguish sense 1 'listening' from sense 2 'perceiving' when they do not co-occur with indicative colloconstructions, such as in the following examples:

(59) a. 我 自己 有了 孩子 以後，我 也是這樣 跟 他們講。

wo zhi ji you le hai zi yi hou wo ye shi zhe yiang gen ta men jiang

I myself have LE child hereafter I also is like this with they talk

結果 他們 好像 **聽** 不進去， 就說，媽媽，你又來了(sense 1 'listening')

jie guo ta men hao xiang TING bu jien qu jiu shou ma ma ni you lai le

result they seem to listen not go in then said mother you come again

'After I by myself have the child; I also talk to them like thisis like this. But as a result, they seem not to listen to me and say, mother, you came again'

b. …關鍵 不是 你說 了什麼，而是 孩子 **聽** 懂 了什麼，

guan jian bu shi ni shou le she mo er shi hai zi TING dong le she mo

key is not you say what but is child hear understand what

感受 到了什麼(sense 2 'perceiving')

gan shou dao le she mo

feel arrive what

'The key is not what you said, but what the child heard, understood, and felt.'

Investigating examples (59a) and (59b) above, sense 1 'listening' and sense 2 'perceiving' are almost identical in surface structure in that they share the following and occur with the same collocational association:

Shared core frame elements: Perceiver

Shared basic pattern: Perceiver < *

Shared colloconstruction: Perceiver < * < Result

Obviously, another problem arises in example (59) when (59a) and (59b) both carry both maker of perception boundary: how do we identify which sense of *TING* is

implied in (59a) and (59b)? In order to answer this question, we have to go into the third step to find contextually relevant items which can help disambiguate these two senses. This further distinction is discussed in the next section.

6.3 Contextual dependence distinction

When examining example (59a) and (59b), local information (within clause) is insufficient for sense disambiguation. Therefore, supplementary information across clausal boundaries is considered, and then we go into the third module—Contextual dependence. Equally, in this step, each sense might be associated with relevant lexical items via certain semantic linking. In (59), by crossing clauses, the relevant items, *jiang* 講 'speaking' and *gan shou* 感受 'reception' are the crucial relevant items as in the re-exemplified example (60a) and (60b), respectively.

(60) a. 我 自己 有了 孩子 以後，我 也是這樣 跟 他們講。

wo zhi ji you le hai zi yi hou wo ye shi zhe yiang gen ta men jiang

I myself have LE child hereafter I also is like this with they talk

結果 他們 好像 聽 不進去， 就說，媽媽，你又來了(sense 1 'listening')

jie guo ta men hao xiang TING bu jien qu jiu shou ma ma ni you lai le

result they seem to listen not go in then said mother you come again

'After I by myself have the child; I also talk to them like thisis like this. But as a result, they seem not to listen to me and say, mother, you came again'

b. 關鍵 不是 你說 了什麼，而是 孩子 聽 懂 了什麼，

guan jian bu shi ni shou le she mo er shi hai zi TING dong le she mo

key is not you say what but is child hear understand what

感受 到了什麼(sense 2 'perceiving')

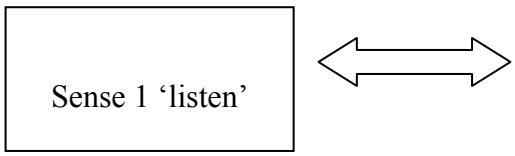
gan shou dao le she mo

feel arrive what

'The key is not what you said, but what the child heard, understood, and felt.'

The relational linking between *jiang* 講 'speaking' and the sense of *TING* 聽 in (60a) or between *gan shou* 感受 'reception' and the sense of *TING* 聽 in (60b) is established by their similar semantic properties. In BOW, a number of English related wordnet synsets belonging to sense 1 'listening' are found and by translation we can find the corresponding Chinese related wordnet synsets, shown in (61) (see the English related wordnet synsets in appendix I):

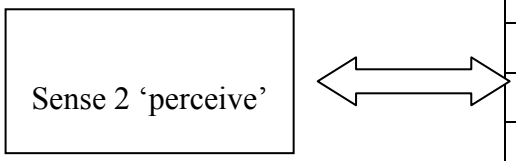
(61)



No	Chinese synsets
1	(注意)聽
2	(說話)聲音
3	(音樂)聲音
4	注意、留心
…	

Also, there are various related wordnet synsets of sense 2 'perceiving' in BOW (shown in 62):

(62)



No	Chinese synsets
1	感知 感覺
2	接收
3	察覺
4	明白
5	注意
…	

An explanation of the connection follows. How does *jiang* 講 'speaking' which is a impressive item, in (60a) help identify sense 1 'listening'? *Jiang* 講 'speaking' is similar to *(shuo hua) sheng yin de jie shou* (說話)聲音的接收 'speech perception' of the related wordnet synsets of sense 1 (see (61) because they both denote 'sound'. Equally, *gan shou* 感受 'reception' is an indicative lexical item

because it is linked to *gan zhi* 感知 'sense perception' in the synsets of sense 2 (see (62)) and they both denote 'perceiving'. In sum, in this module, via contextual information, *TING* 聽 in example (60a) is identified as 'listening', and in example (60b) it is specified as 'perceiving'.

7 The case of the motion verb *KAN* (看)

KAN 看 is another perception verb and polysemous word with high frequency. In the Sinica Corpus, there are also more than 1000 occurrences of *KAN* 看. As with other polysemous words, only 200 entries are tagged with core frame elements. In order to completely investigate the syntactic and semantic properties, we explore the overall data of *KAN* 看.

7.1 Frame-based sense distinction

Preliminarily, most examples of *KAN* 看 from Sinica Corpus can be roughly divided into four major groups by their collocations of core frame elements. In addition, via the Chinese-English translations in BOW, these four groups can be related to various senses in English corresponding to Chinese: sense 1 as 'seeing' *KAN* DAO(看), sense 2 as 'watching' *KAN* (細看), sense 3 as 'scritiny' *SI KAN* (觀察), sense 4 as 'visiting' *TAN WAN* (探望), and sense 5 as 'depending' YI ZHOU(依照) The distribution percentage of each sense are shown in (62).

(62)Percentage of 200 Occurrences of *KAN*

	Percentage (%)
Sense1: seeing	59
Sense2: watching	6.5
Sense3: scrutinizing	29.5
Sense4: visiting	3
Sense5: depending	2

Each sense of *KAN* 看, in FrameNet, related to different frames according to core argument collocation corresponding to the various basic patterns with core frame elements in different frames. For example, sense 1 'listening' is in PERCEPTION_ACTIVE frame, sense 2'perceiving' is included in PERCEPTION_EXPERIENCE frame, sense 3 'hearing about' is contained in HEAR

frame, sense 4 ‘obeying’ is in COMPLIANCE frame. The classification of *KAN* 看 depending on Frame analysis is shown in (63)-(67)

(63) Basic Patterns with Core Frame Elements of Sense 1 ‘Seeing’

Sense	Frame	Frame Elements	No.	Basic Patter with Core Frame Elements and Examples	(%)
Sense 1: 看 SEEING	PERCEPTION_ ACTIVE	Perceiver Body_part Phenomenon Theme Agent	BP55	**Perceiver < * < Phenomenon** 這次旅行，我雖然**看**了不少令人驚奇的景色，但是印象最深刻的，還是燕子口。	41.52
			BP56	**Perceiver < * < Theme** 選取他看待藝術品的態度（角度）？或者更理論化地說，我們在**看**一個物體時，到底如何選取我們看它的角度？	24.58
			BP57	**Phenomenon < Perceiver < *** 還是要談，但事實上美方官員也瞭解有些要求只是做給國會議員**看**，不會真對我造成壓力。	16.1
			BP58	**Theme < Perceiver < *** 難怪你說聽不懂，他說的話我都記下來了，我的筆記可以借給你**看**。‖那真是太好了，我這就拿去影印。	9.3
			BP59	**Agent < Perceiver < * < Theme** 平時他看到的是你給他看的，而你給他**看**的只是一個臉譜而已，因應這個機，你就站在你的立場來扮演你	2.5
			BP60	**Perceiver < *** 寶貝女兒瞪了他一眼，斬釘截鐵的回敬她老爹一句：「我不可以**看**，因報上明明標著兒童不宜。」	2.5
			BP61	**Perceiver <Theme < *** 這份閑心。老人當真生氣了：這麼好的一些姑娘，你為什麼誰也**看**不上？你不願安家？我不能。普新民非常慎重地宣布：因為我	0.88

Sense	Frame	Frame Elements	No.	Basic Patter with Core Frame Elements and Examples	(%)
Sense 1: 看 SEEING	PERCEPTION_ ACTIVE	Perceiver Body_part Phenomenon Theme Agent	BP62	**Agent < Theme < Perceiver < *** 聽不到了，但是同學們一直鼓勵他，放錄音帶給他聽，拿書給他**看**。這位同學在生命的最後旅程中，就靠我的書來支撐他的精神。	0.88
			BP63	**Phenomenon < Body_part < *** 熱到超過五十度，「到五十度時，空氣中的熱流蒸騰，用眼睛都**看**得見，真的是熱浪呀！」	0.87
			BP64	**Body_part < *** 小鷹很害怕，站在那兒，眼睛向山下**看**著，一動也不敢動，也不敢把翅膀展開。	0.87

(64) Basic Patterns with Core Frame Elements of Sense 2 ‘Watching’

Sense	Frame	Frame Elements	No.	Basic Patter with core frame elements	(%)
Sense 2: (注意)看 WATCHING	PERCEPTION_ EXPERIENCE	Perceiver(passive) Phenomenon Theme	BP65	**Perceiver < * < Theme** 他從小就喜歡看鵝。有一次，他在全神貫注地**看**鵝，特別注意到鵝的長頸。	46.14
			BP66	**Perceiver_ < * < Phenomenon** 他跑過去仔細一**看**，被挖過的地方露出一個好大好大的卵大約有十來斤重。	38.46
			BP67	**Perceiver < *** 丘子章很想一個個去擁抱他們，告訴他，他會好好去**看**。去聽。去記下來	7.7
			BP68	**Theme < Perceiver < *** 請問陳先生還有什麼要討論的問題嗎? 陳：這份草約請您仔細地**看**一次。	7.7

(65) Basic Patterns with Core Frame Elements of Sense 3 ‘Scrutinizing’

Sense	Frame	Frame Elements		Basic Patter with core frame elements	(%)
Sense3: 觀察 SCRUTINIZING	SCRUTINY	Cognizer Ground Phenomenon	BP69	**Cognizer < Ground <*** 從過去廿五年的發展來**看**，結合繪畫中的寫實與文學中的隱喻，才是張振宇創作的主要	49.15
			BP70	**Cognizer < * < Phenomenon** 樓「我知道，你也早點回去休息，我**看**你也很累了」如深情的注視著她的男朋友，這兩天最累的就屬他	37.29
			BP71	**Phenomenon < Cognizer < *** 由於他不相信藝術有任何意義或不變的價值，因此在他**看**來，博物館與景仰大師的群眾都是荒謬的	6.76
			BP72	**Cognizer < Ground < * < Phenomenon** 第一講「走過人生的關卡」，是從人的自然生命角度來**看**人生的關卡。人在成長過程中有所謂的血氣：包括少年的血氣	3.9
			BP73	**Cognizer < *** 讚詞的，不止是田家英和那位高級法院院長，連彭德懷都是這麼**看**的。	3.9

(66) Basic Patterns with Core Frame Elements of Sense 4 'Visiting'

Sense	Frame	Frame Elements	No.	Basic Patter with core frame elements	(%)
Sense 4: 拜訪 VISIT	ARRIVING	Goal Theme Cotheme Ground	BP74	**Theme < Ground < * < Goal** 我曾去巴黎**看**羅丹，專門儲藏羅丹的作品就有一個博物館。	50 %
			BP75	**Theme < * < Goal** 好，你需要什麼，就打電話給我，我明天再來。‖謝謝你來**看**我，再見。	33.33%
			BP76	**Theme < Cotheme < * < Goal** 聽小玲的同學林又華說你們住在這兒，所以我今天帶小玲來**看**你們。	26.67%

(67) Basic Patterns with Core Frame Elements of Sense 5 ‘Depending’

Sense	Frame	Frame Elements	No.	Basic Patter with core frame elements	(%)
Sense 4: 依據 DEPENDING	Contingency	Determinant Outcome		**Outcome < * < Determinant** 一個事業，公營不一定效率低，民營也不一定效率高，全看經營的方法怎麼樣。	100%

As can be seen in (62), sense 1 ‘seeing’ is observed to be perceivably salient and prototypical. First, it occurs most frequently and second, it denotes a specific physical perception, while other senses are taken as sense extensions from sense 1. These major senses are further distinguished by the model provided in this paper. In first module—frame-based distinction, these senses are specifically defined by the basic patterns with core frame elements. As shown in (63)-(67), sense 1 ‘seeing’ is defined by the basic patterns in the PERCEPTION_EXPERIENCE Frame: Perceiver < * < Phenomenon, Perceiver < * < Theme, Phenomenon < Perceiver < *, Theme < Perceiver < *, Agent < Perceiver < * < Theme, Perceiver < *, Perceiver < Theme < *, Agent < Theme < Perceiver < *, Phenomenon < Body_party < *, and Body_part < *; sense 2 ‘watching’ is specifically defined by the basic patterns in the PERCEPTION_ACTIVE Frame: Phenomenon < * < Theme, Perceiver < * < Phenomenon, Perceiver < *, and Theme < Perceiver < *; sense 3 ‘scrutinizing’ is identified by the basic patterns in the SCRUTINY Frame: Cognizer < Ground <*, Cognizer < * < Phenomenon, Phenomenon < Cognizer < *, Cognizer < Ground < * < Phenomenon, and Cognizer < *; sense 4 ‘visiting’ is specified by the basic patterns in ARRIVING Frame: Theme < Ground < * < Goal, Theme < * < Goal and, Theme < Cotheme < * < Goal; and sense 5 ‘depending’ is defined by the basic pattern in CONTINGENCY frame: Outcome < * < Determinant. In this case study, the frame-based distinction separates sense 3 ‘scrutinizing’, sense 4 ‘depending’, and sense 5 ‘visiting’ from the other two senses. Nevertheless, the same problem found

in (63) and (64) by frame-based distinction is that different frames have similar basic patterns with core frame elements. For example, the patterns: Perceiver < * < Phenomenon, Phenomenon < * <Theme, Perceiver < * and Theme < Perceiver < * occur in both sense 1 'seeing' and sense 2 'watching'. For illustration, please consider the following examples:

(68) **Perceiver < * < Phenomenon**

a. 這次旅行，我[Perceiver]雖然 看了 不 少 令人 驚奇 的景色[Phenomenon]，

zhe ci lv xin wo *suei ran KAN le bu shao lien ren jieng qi de jieng se*

this trip I though see LE not few astonishing view

但是 印象 最 深 的，還 是 燕子口。… (sense 1 'seeing')…

dan shi yin xiang zuei shen de hai shi yian zi kou

but impression most deep still barn swallow

'On this trip, although I saw a few astonishing views, what I was deeply impressed with was the barn swallow'

b. 他[Perceiver]跑 過去 仔細 看， 被 挖 過 的 地方

ta *pao guo qu zi xi* *KAN* *bei ua guo de* *di fang*

He run to carefully watch BE dug DE place

露出 一個 好大 的卵[Phenomenon]大約 有 十 來斤 重。(sense 2 'watching')

lou chu yi ge hao da de ruan *da yue you shi lai jien zhong*

basset one-CL very giant DE egg about have ten kilograms heavy

'He ran and carefully watched the place where they dug a giant egg which weight about ten kilograms.'

(69) **Perceiver < * < Theme**

a…老師 撫掌 說： 請 從 便！然後 在 黑板 上 寫一些

lao shi fu zhang shou qien cong bian *ran hou zai hei ban* *shang xie yi xie*

Teacher stroke hand say please take easy then on blackboard up wrote some

我們[Perceiver]看 不 懂 的文字[Theme]。(sense 1 ‘seeing’)

wo me *KAN* *bu dong* *de wen zi*

we see not understand DE characters.

‘The teacher stroked her hand to say: Please take it easy, and then wrote down some characters which we saw but could not understand.’

b.他 從小 就喜歡 看 鵝。 有一次，他[Perceiver]在 全 神 貫 注

ta cong xiao *jiu xi huan kan e* *you yi ci* *ta* *zai quan shen guan zhu*

He since childhood like watch goose once he in concentrate

地 看 鵝[Theme]特別 注意 到 鵝的 長頸。(sense 2 ‘watching’)

de KAN e *te* *bie zhu yi dao e* *de chang jing*

watching goose particulaly attentiongoose’s long neck

‘He has liked watch geese since his childhood. Once he concentrated watching geese, he particularly paid attention to the long neck.’

(70) **Theme < Perceiver < ***

a. 李 白 回 頭 看看 白 帝 城[Theme]，已 經 掩沒

Li bai huei tou *kan kan* *bai di cheng* *yi jing* *yian mo*

Li, Bai turn head look Bai Di City already cover

在 彩 雲 中 間，[CNI/Perceiver]看 不 見了。(sense 1 ‘seeing’)

zai *cai yuen* *zhong jian* *KAN* *bu jian le*

in clouds iner see disapper LE

‘Li, Bai turned back looking at Bai Di City which was already coverd by clouds and could not be seen.’

b. 這 份 草 約[Theme]請 您[Perceiver]仔細地 看 一次(sense 2 ‘watching’)

zhe fen *cao yue* *qin* *nin* *zi xi de KAN yi ci*

this CL draft please you carefully watch once

Please carefully read this draft once.

(71)**Perceiver < ***

a. 他一定　永　遠　也 不 會　幸　福，因 為他[Perceiver]

ta yi ding yong yuan　ye bu huei xing fu　yin wei ta

he must　forever　either can happiness　because he

永　遠　也 不 可　能　**看**　得 清 晰(sense 1 'seeing')

yong yuan ye　bu ke neng KAN de qing xi

forever　also not　can　see　DE clear

'He cannot have happiness forever because he can not clearly see forever.'

b. 丘 子 章　很 想　一 個個 去 擁　抱　他 們，告 訴 他，

qiou　zi zhang hen xiang　yi ge ge　qu　yong bao ta men　gao su ta

Qiou, zi-zhang very want one by one　to　hug　them,　tell him,

他[Perceiver]會　好　好 去 **看**。 去 聽。去　記　　下 來(sense 2 'watching')

ta　huei hao hao qu KAN　qu ting　qu　ji　xia lai

he　will　well　to see　to listen　to remember down.

'Qiou, Zi-Zhang really wanted to hug them one by one, tell him and he will give a good look, listen to and remember.'

As can be seen in (68)-(71), sense 1 'seeing' and sense 2 'watching' appear with the same basic patterns with core frame elements. Therefore, to further distinguish these two senses, it is necessary to go into the second module—colloconstruction distinction.

7.2 Colloconstructional distinction: sense 1 vs. sense 2

In order to further distinguish sense 1 and sense 2, as in (68)-(71), first, a search for categorical collocations is necessary. Exploration can be automatically executed from Sinica Corpus as in (72 A), and the non-core argument categorical collocations can be concluded, as in (72B).

(72)

A. Categorical Collocates of Five Senses of *KAN* 看

category	Left 5	Left 4	Left 3	Left 2	Left 1	Target verb	Right1	Right 2	Right 3	Right 4	Right 5	Total	%
Na	745	813	753	1028	569	0	815	626	872	807	791	7819	15.01
D	485	384	**375**	**503**	**1801**	0	**432**	**496**	**523**	515	535	6049	11.61
VC	267	269	**226**	**187**	**90**	3513	19	**258**	**211**	210	259	5509	10.57
Nh	215	230	257	273	623	0	526	251	191	167	194	2927	5.62
VE	107	95	119	74	**135**	1307	7	**256**	121	129	128	2478	4.76
VH	289	286	**246**	**203**	68	4	**117**	**231**	**299**	286	274	2303	4.42
P	284	214	**225**	**235**	49	0	64	122	123	125	146	1587	3.05
Di	62	55	56	30	56	149	**622**	22	32	56	52	1192	2.29
Nf	128	138	141	104	42	0	19	135	**178**	140	132	1157	2.22
Nc	104	102	106	110	88	0	88	90	117	95	114	1014	1.95
VA	137	142	116	110	**89**	1	26	67	90	97	111	986	1.89
Neu	97	100	94	36	11	0	91	121	98	104	98	850	1.63
T	94	104	96	49	1	0	**186**	63	76	63	68	800	1.54
Others													

B. High Frequency Non-core Argument of *KAN* 看

No.	左 3	左 2	左 1	關鍵詞	右 1	右 2	右 3
1	D	D	D	看	Di	D	D
2	VH	P	VE	看	D	VC	VH
3	VC	VH	VC	看	T	VE	VC
4	P	VC	VA	看	VH	VH	Nf
Others							

C. Distribution of Non-core Argument of Five Senses of *KAN* 看

	Adv + *KAN*	(Adv.)V + *KAN*	*KAN* + Asp	*KAN* + Comp.
Sense 1 'seeing'	8.55	5.13	12.82	16.24
Sense 2 'watching'	61.54	30.76	0	0

a. Adv + *KAN*

i. 這是一條金帶子！我知道你愛看，就讓你盡情的看。(sense 1 'seeing')

ii. 他從小就喜歡看鵝。有一次，他在全神貫注地看鵝，特別注意到鵝的長頸(sense 2 'watching')

b. (adv.)V + *KAN*

i. 李文秀驚魂未定，轉頭看那姓全的強人時，只見他直挺挺的躺在地上。(sense 1 'seeing')

ii. 丘子章很想一個個去擁抱他們，告訴他，他會好好去看。(sense 2 'watching')

c. *KAN* + Asp

i. 籠子裡有好幾隻小狗。我看了很想要，你的奶奶就買了一隻給我。 (sense 1 'seeing')

d. *KAN* + Comp.

i. 有一陣子聽不到蟋蟀的叫聲，屋裡的光線很暗，也看不清楚蟋蟀在哪兒。 (sense 1 'seeing')

From Tables (72A) and (72B) above, the non-core argument categorical collocates can be further generalized as the following four distinctive types: (a) D (Adv) + *KAN*, (b) (Asp) V + *KAN*, (c) *KAN* + Di (Asp), and (d) *KAN* + Comp

(VC,VE,VH) as in (72C). Here, to be more significant, predicates (VH, VE) are put together as complements as different senses have various distributions of categorical collocations. Subsequently, in order to disambiguate sense 1 and sense 2, an investigation of their syntactic collocates must be practiced. The distributions, following some illustrations are shown in (72C).

Among these collocation types, sense 1 usually co-occurs with aspectual markers and complements, while sense 2 often appears with adverbs or preceding predicates. Further, within those co-occurring aspectual markers and adjunct complements, some crucial lexical collocates of *KAN* are found to help identify sense 1 'seeing'. For example, some indicative lexical items are detected, such as the aspectual marker 了 *LE* denoting the completeness of actions, and adjunct complements 清楚 *qing chu* 'clear' denoting the resulting state. Adverbs or preceding predicates which are significant to sense 2, might equally help specify sense 2 'watching' of *KAN*. For instance, within the collocating adverbs or preceding predicates, the lexical items with volitional meaning or with eye actions are the majority, such as 仔細 *zi xi* 'carefully', 去 *qu* 'go to', 凝神 *ning shen* 'concentrate', 虎視眈眈(的) *hu shi dan dan* 'glowering', 眼巴巴(的) *yan ba ba* 'anxiously'. The distribution of the significant collocates of sense 1 and sense 2 can be seen below:

(73) The distribution of the specific collocates of sense 1 and sense 2

	Eye actions + V	Volitional marker + V	V + 了 *le*	V + Result
Sense 1 'seeing'	0	0	10.26 %	16.23 %
Sense 2 'watching'	7.3 %	84.62	0	0.2 %

These collocates are indicative and semantically relevant to the target senses,

sense 1 and sense 2 for several reasons. On one side, the verbal *LE* and the results are significant to sense 1 'seeing' because being in Perception_passive frame, sense 1 'seeing' needs a boundary for the action of perceiving. For this reason, the collocates-aspectual marker *Le* and resulting complements, which both denote a boundary of an action, have a connection with sense 1. On the other side, in the case of sense 2 'watching', those volitional markers (adverbial or preceding verbal collocates) are related to sense 2 'watching' for sense 2 being in Perception_active frame, denotes volitional meaning which is equally relevant to those markers. As a result, by colloconstruction module, the markers of perception boundary (了, 清楚) tell the differences between sense 1 and sense 2 while markers of volition (仔細, 凝神) help disambiguate sense 2 from sense 1. With this semantic connection, though sense 1 'seeing' and sense 2 'watching' share the same core frame element pattern as Perceiver < * < Phenomenon, Perceiver < * < Theme, Theme < Perceiver < * and, Perceiver < *, as exemplified in (74a-77a) and (74b-77b), they can be distinguished by their specific Colloconstructions:

- ✧ *KAN* 看+ Perception boundary marker(s)
  - a. *KAN* 看+ *le* 了
  - b. *KAN* 看+ *Result*
- ✧ Volitional marker(s) + *KAN* 看
- ✧ Eye actions + *KAN* 看

For illustration, considering the following instances, Perceiver < * + 了 < Phenomenon (for sense 1)and Perceiver < 虎視眈眈 + * < Phenomenon (for sense 2) in (74); Perceiver < * + 不懂 < Theme (for sense 1) and Perceiver < 全神貫注地 + * < Theme (for sense 2) in (75); Theme < Perceiver < * + 不見 (for sense 1) and Theme < Perceiver <仔細地+ * (for sense 2) in (76); Perceiver < * +清晰 (for sense

1) and Perceiver < (好好)去+ * (for sense 2) in (77):

(74) Perceiver < * + 了 < Phenomenon

a. 這次旅行，我[Perceiver]雖然**看**了不少令人驚奇的景色[Phenomenon]，但是印象最深刻的，還是燕子口。… (sense 1 'seeing')

Perceiver < 虎視眈眈 + * < Phenomenon

b. 老師[Perceiver]也可以不必勞心傷神，虎視眈眈地**看**著學生考試[Phenomenon]了。... (sense 2 'watching')

(75) Perceiver < * + 不懂 < Theme

a. 老師撫掌說：請從便！然後在黑板上寫一些我們[Perceiver]**看**不懂的文字[Theme]。(sense 1 'seeing')

Perceiver < 全神貫注地 + * < Theme

b. 有一次，他[Perceiver]在全神貫注地**看**鵝[Theme]，特別注意到鵝的長頸(sense 2 'watching')

(76) Theme < Perceiver < * + 不見

a. 李白回頭看看白帝城[Theme]，已經掩沒在彩雲中間，[CNI/Perceiver]**看**不見了。(sense 1 'seeing')

Theme < Perceiver <仔細地+ *

b.這份草約[Theme]請您[Perceiver]仔細地**看**一次。(sense 2 'watching')

(77) Perceiver < * +清晰

a. 他一定永遠也不會幸福，因為他[Perceiver]永遠也不可能**看**得清晰(sense 1 'seeing')

Perceiver < (好好)去+ *

b. 丘子章很想一個個去擁抱他們，告訴他，他[Perceiver]會好好去**看**。去聽。去記下來(sense 2 'watching')

Nevertheless, as can be seen in (78), another problem emerges. The distinction

between sense 1 ‘seeing’ from sense 2 ‘watching’ not only depend on the two previous modules because they co-occur with same basic patterns with core frame elements, but also with the same indicative colloconstructions, such as in the following examples:

(78) a. 宋 明 理學 家 對 世 道 人 心 的影響， 反 而遠

Song ming li xue jia duei xhi dao ren xin de ien xiang fan er yuan

Song Ming scientist to public morals DE influence, on the contrary far

不 如 小 說 戲劇來 得 深 入 普 遍。 因為 理學 家 的 著 作[Theme]

bu ru xiao shou xi ju lai de shen ru pu pian yin wei li xue jia de zhu zuo

not as novel dramacome DE deep into widespread. Because scientist DE work

只有 知 識分子[Perceiver] 才 能 **看** 得 懂， 一 般 民 眾

zhi you zhi shi fen zi cai nan KAN de dong yi ban min zhong

only intellectuals just can see DE understand, common people

是 沒 有 辦法 了解 的。 (sense 1 ‘seeing’)

shi mei you ban fa liou jie de

is not have way understand DE.

‘In the Song and Ming dynasties, Scientists’ effects on the public morals was far behind the influenceof dramas and novels. Because only the intellectuals could undertand the Scientist’s works, and general people could not undertand them at all.’

b. 許多 家 長 帶 孩子 來 諮 詢：教 科書[Theme]，

xiu duo jia zhang dai hai zi lai zi xiuen jiao ke shu

many parents bring children come consult: school book

孩子[Perceiver]總 是 **看** 不下 去，注意力 無 法集中

hai zi zong shi KAN bu xia qu zhu yi li wu fa ji zhong

children always watch not down go attention cannot concertrate

煩躁，愛發脾氣。… (sense 2 ‘watching’)

fan zao ai fa pi qi

irritable, easy to be angry

‘Many parents bring their children for a consultation: Children always have trouble reading their school books, cannot concentrate, and are irritable and easily angered’

Examining examples (78a) and (78b) above, sense 1 ‘seeing’ and sense 2 ‘watching’ are almost identical in surface structure as they share the following features:

Shared core frame elements: Perceiver, Theme

Shared basic pattern: Theme < Perceiver <*

Shared colloconstruction: Theme < Perceiver < * < Result((得)懂/下去)

In example (78a) and (78b) both sense 1 and sense 2 of *KAN* 看 are not only realized in the same basic pattern but also carry the same marker of perception boundary (result). That is , they appear with the same colloconstruction. Then, to distinguish sense 1 from sense 2 in (78a) and (78b), respectively, the next step is adopted to search for contextual relevant items. The further distinction is discussed in the next section.

7.3 Contextual Dependence Distinction

In example (78), local information (within the clause where the target verb exists) is not distinctive enough for sense disambiguation. Then, the next step—Contextual dependence, exploring of extra information across clausal boundaries is necessary. Likewise, in this module, each sense is assumed to be associated with some specific relevant lexical items through semantic linking. Following this criteria, in (79), the relevant items, *liao jie* 了解 'understand' and *zhu yi li* 注意力 'attention', are detected as crucial relevant items for sense 1 'seeing' and sense 2 'watching' shown in the re-exemplified example (79a) and (79b), respectively.

(79) a. 宋 明 理學 家 對 世 道 人 心 的影響， 反 而遠

Song ming li xue jia duei xhi dao ren xin de ien xiang fan er yuan

Song Ming scientist to public morals DE influence, on the contrary far

不 如 小 說 戲劇來 得 深 入 普 遍。 因為 理學 家 的 著 作[Theme]

bu ru xiao shou xi ju lai de shen ru pu pian yin wei li xue jia de zhu zuo

not as novel dramacome DE deep into widespread. Because scientist DE work

只有 知 識分子[Perceiver] 才 能 看 得 懂， 一 般 民 眾

zhi you zhi shi fen zi cai nan KAN de dong yi ban min zhong

only intellectuals just can see DE understand, common people

是 沒 有 辦法 了解 的。 (sense 1 'seeing')

shi mei you ban fa liou jie de

is not have way understand DE.

'In the Song and Ming dynasties, Scientists' effects on the public morals was far behind the influenceof dramas and novels. Because only the intellectuals could undertand the Scientist's works, and general people could not undertand them at all.'

b. 許多 家 長 帶 孩子 來 諮 詢：教 科書[Theme]，

xiu duo jia zhang dai hai zi lai zi xiuen jiao ke shu

many parents bring children come consult: school book

孩子[Perceiver]總 是 看 不下 去，注意力 無 法集中

hai zi zong shi KAN bu xia qu zhu yi li wu fa ji zhong

children always watch not down go attention cannot concertrate

煩躁，愛發脾氣。… (sense 2 'watching')

fan zao ai fa pi qi

irritable, easy to be angry

'Many parents bring their children for a consultation: Children always have trouble reading their school books, cannot concentrate, and are irritable and easily angered'

Via similar semantic features, the linking between *liao jie* 了解 'understand' and the sense of *KAN* 看 in (79a) or *zhu yi li* 注意力 'attention' and the sense of *KAN* 看 in (79b) is established. In BOW, a number of English synsets belonging to sense 1 'seeing' are searched and through translation they are equated to some Chinese lexical items which are relevant to sense 1 'seeing', as in (80) (see the English synsets in appendix I):

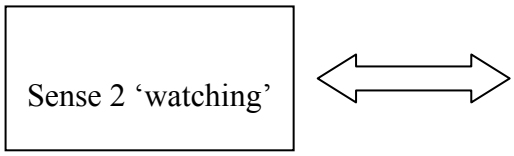
(80)

Sense 1 'seeing' ⟺

No	Chinese synsets
1	碰見
2	接收
3	來自經驗的
4	明白,了解
5	辨出
…	

Also, there are also some relevant Chinese lexical items of sense 2 ‘watching’ by exploring English synsets in BOW:

(81)



No.	Chinese synsets
1	察覺
2	凝視
3	(注意)看
4	留心、留意
…	

Next, we have to account for the linking between sense1 and its synsets and the connection between sense 2 and its synsets. The lexical item *liao jie* 了解 ‘understand’ is an indicative anchor in (79a) to help identify sense 1 ‘seeing’. The reason could be that *liao jie* 了解 ‘understand’ is similar to *ming bai/liao jie* 明白/了解 ‘catch’ within the synsets of sense 1 (consider (80)) for they both denote ‘become aware of something’. Equally, *zhu yi li* 注意力 ‘attention’ is an indicative lexical item because it is linked to *zhu yi/liou xin* 注意/留心 ‘attend’ within the synsets of sense 2 (see (81)) and they both denote ‘pay attention (to)’. Consequently, in this module, via contextual information, the sense of *KAN* 看 in example (79a) is identified as ‘seeing’, and in example (79b) as ‘watching’.

8. Conclusion

In this paper, a preliminary model of disambiguating polysemous words has been presented. Given the principle of economy, it is assumed that not all the senses of a polysemous word have equal weights and require exactly the same procedure for sense identification. Therefore, three steps are called upon in a sequence when needed. The first step focuses on frame-based information regarding participating frame elements and their expressions, and in most cases, senses can be distinguished in this step. However, in a few cases one basic pattern with core frame elements is shared by different senses. In this case, second step—colloconstruction is proposed to search for further informative syntactic adjuncts to help sense disambiguate. In the second step, word senses are distinguished beyond the expression of core arguments and a detailed lexical as well as grammatical association patterns are sought. However, those word senses sharing the same basic pattern(s) and same colloconstruction(s) or sharing the same basic pattern(s) without finding of distinctive colloconstruction(s) are still unsolved. Thus, the third step—contextual dependence is proposed. The final step to help disambiguate polysemous words is by contextual dependency cues. Through semantic properties, the relevant lexical items are investigated to trigger the target sense. By establishing a linkage to BOW, complicated senses of polysemous words can be identified through connecting the relevant lexical items to the synsets in BOW by their shared meanings. In this step, discourse-level factors are utilized with a clear measure of their semantic relations, just as Biq (1988) stated "Any effort to systematically identify and explain the different types of usage has to consider not only sentential entities but also entities which are outside of the sentences/ proposition yet relevant to the discourse."

This model is to be applied in computation systems. Therefore, automatic disambiguation is a crucial part. In the preliminary stage, this frame-based model

has the frame elements manually tagged. However, automatic tagging of frame elements is not impossible. For instance, first, through generalizing the categorical composition of each frame elements (Part of Speech (or POS) from Sinica Corpus); a set of categorical groups corresponding to the frame element is explored. For example, the frame element self-mover consists of various categories [N], [N][Conj][N] or other categorical sets. The frame elements Area is also composed of a number of categorical compositions, such as [Nc], [Nec] or other categorical sets (see (82)):

(82) Self-mover < * < Area

[N]	[Nc]
[N][Conj][N]	[Nec]
. ..	...

These sets of categorical groups also tell the various structures of each frame element. The automatic tagging program can be designed following the procedures above. Further, in the second and third modules, there are also two available searching tools, Sinica corpus and Ontology in BOW. These two searching engines help automatic rummage for intra-clausal colloconstructions and inter-clausal relevant synsets in module 2 and 3 respectively. Combining these three parts, automatic sense disambiguation might be realized. This is an issue for further studies.

By redefining polysemy with operational mechanisms, this study has provided a linguistic model with theoretical validity to develop a computational system for disambiguation. Although, this study is by no means exhaustive, it nevertheless bears some significant implications on both theoretical linguistics as well as computational linguistics:

a) The solution integrates syntax, semantics and pragmatics in a step-by-step manner and make linguistic theories more accessible for computational

applications.

b) With a corpus-based multi-module approach, this model can be universally applied in other languages with the three clearly defined steps.

c) Word senses can be systematically detected via the three steps, incorporating existing linguistic theories that are interactive in nature.

d) As comprehensive investigations of Mandarin lexical semantics are under way (Liu 2002, 2004) and a bilingual ontological Wordnet (Sinica BOW) is also available (Huang et al 2004), the proposed model may offer a workable resolution to develop a computer system dealing with polysemy resolution.

References

Adriaens, Geert, Steven L. Small, Garrison W. Cottrell, and Michael K. Tanenhaus. 1988. Lexical ambiguity resolution: perspectives from psycholinguistics, neuropsychology, and artificial intelligence. San Mateo, Calif.: Morgan Kaufmann Publishers.

Anatol, S., and Stefan Th. Gries. 2003. Collostructions: investigating the interaction of words and constructions. International Journal of Corpus Linguistics 2. 209-43.

Biber, Douglas, Susan Conrad and Randi Reppen 1998. Corpus linguistics: investigating language structure and use. Cambridge: University Press.

Bybee, Joan L. 1985. Morphology: a study of the relation between meaning and form. Amsterdam, Philadelphia: J. Benjamins.

Cau, Feng-Fu, Li-Zhong Cai, and Xiu-Ying Liu. 2001. Shen Ti Yu Pi Yu: Yu Yan Yu Ren Zhi De Shou Yao Jie Mian. Taipei: Wen He.

Chief, Lian-Cheng, Chu-Ren Huang, Keh-Jiann Chen, Mei-Chih Tsai, and Li-Li Chang. 2000. What can near synonyms tell us? Computational Linguistics and Chinese Language Processing 5.1. 47-60

Cruse, Alan. 1986. Lexical semantics. Cambridge: Cambridge University Press.

—— 2004. Meaning in language: an introduction to semantics and pragmatics. New York: Oxford University Press.

Edmonds, Philip and Graeme Hirst. 2002. Near-Synonymy and Lexical Choice. Computational Linguistics, 28. 2. 105-144

Fauconnier, Gilles. 1994. Mental spaces: aspects of meaning construction in natural language. Cambridge, NY, USA: Cambridge University Press.

Fauconnier, Gilles. 1997. Mappings in thought and language. Cambridge, NY, USA: Cambridge University Press.

Fauconnier, Gilles, and Mark Turner. 2002. The way we think: conceptual blending and the mind's hidden complexities. New York: Basic Books.

Fellbaum, Christiane. 1998. WordNet: an electronic lexical database. Combridge, Mass.: MIT Press.

Feng, Heng. 2000. Ci Lei Wei Tan. Taipei: Da Qian Press.

Fillmore, Charles J. 1976. Frame semantics and the nature of language. In Annals of the New York Academy of Sciences: Conference on the Origin and Development of Language and Speech, Volume 280. 20-32.

——1982. Towards a descriptiove framework for spatial deixis. In R.J. Jarvella and W. Klein (eds.), Speech, place and action. New York:Wiley.

——and B. T. S. Atkins. 1992. Towards a frame-based organization of the lexicon: The semantics of RISK and its neighbors. In Lehrer, A and E. Kittay (eds.), Frames, fields, and contrast: new essays in semantics and lexical organization. Hillsdale: Lawrence Erlbaum Associates. 75-102.

Firth, John Rupert. 1957. Papers in linguistics 1934-1951. London: Oxford University Press.

Gale, William A., Kenneth W. Church and David Yarowsky. 1992. One sense per discourse. Proceedings of the 4th DARPA Speech and Natural Language Workshop.

Geeraerts, Dirk. 1993. Vagueness's puzzles, polysemy's vagaries. Cognitive Linguistics, 4.3. 223-272.

Goldberg, Adele E. 1996. Conceptual structure, discourse, and language. Stanford, Calif.: CSLI Publications.

Gorfein, David S. 1989. Resolving semantic ambiguity. New York: Springer-Verlag.

Gries, Stefan Th. 2005. Syntactic priming: A corpus-based approach. Journal of Psycholinguistic Research 34. 365-99.

Hopper, Paul J., and Sandra A. Thompson. 1980. Transitivity in grammar and discourse. Language 2.251-99.

Hu, Yu-Shu and Xiao Fan 1995. Dong Ci Yan Jiu. He-nan: He Nan University Press.

Huang, Chu-Ren, Ru-Ying Chang, and Shiang-Bin. Lee. 2004. Sinica BOW (Bilingual Ontological Wordnet): Integration of Bilingual WordNet and SUMO. Paper presented at 4th International Conference on Language Resources and Evaluation (LREC 2004). Lisbon. Portugal. 26-28 May, 2004.

—— Keh-Jiann Chen and Li-Li chang. 1996. Segmentation standard for Chinese natural language processing. Proceedings of the 1996 International Conference on Computational Linguistics (COLING 96). August. Corenhagan, Denmark.

Ikegaya, Yuki, Yasuhiro Noguchi, and Satoru Kogure. 2005. Integration of dependency analysis with semantic analysis referring to the context. Proceedings of PACLIC 19, the 19th Asia-Pacific Conference on Language, Information and Computation. Academia Sinica, Taipei.

Jackendoff, Ray S. 1985. Multiple subcategorization and the theta-criterion: the case of *climb*. Natural Language and Linguistic Theory, 3. 271-295.

Johnson-Laird, P. N., R. M. J. Byrne, and P. Tabossi. 1989. Resoning by model: the case of multiple quantification. Psychological Review 96, 658-673.

Katz, Jerrold J. 1972. Semantic theory. New York: Harper and Row.

Lakoff, George. 1987. Women, fire, and dangerous things: what categories reveal about the mind. Chicago: University of Chicago Press.

Lai, Huei-ling. 2004a. Same constructions but different grammaticalization paths of Hakka *lau* and *tung*. Proceedings of Fifth International symposium on Taiwanese Languages and Teaching, 195-210. Prudence University, Taichung.

Lehrer, Adrienne, and Eva Feder Kittay. 1992. Frames, fields, and contrasts :new essays in semantic and lexical organization. Hillsdale, N.J.: L. Erlbaum Associates.

Li, W., Q. Lu, and R. Xu. 2005. Similarity based chinese synonym collocation extraction. Computational Linguistics and Chinese Language Processing 1. 123-44

Liu, Mei-chun. 1994. Discourse explanations for the choice of *jiu* and *cai* in Mandarin conversation. Chinese Language and Linguistics. 2, 671-709

___ 2002. Mandarin verbal semantics: a corpus-based approach. Taipei, Taiwan: Crane Publishing.

—— and Yiching Wu. 2004. A frame-based analysis of polysemy-verbs of Encoding in Mandarin. Paper presented in The 12th Annual Conference of the IACL. Nankai University. Tianjin, China.

Lien, Chin-fa. 2000. A frame-based account of lexical polysemy in Taiwanese. Language and Linguistics 1.1. 119-138. Beijing: Shang Wu Yin Shu Guan.

——Xiou-ying, Liu. 2004b Taiwan South-Min yi dong dong ci '*Zou*'de duo yi xing ji gai nian jie gou: yu yi ien shen de tu jing. Yu Yan Xue Lun Cong. 29. 334-371.

Lin, Chienjer Charles and Ahrens, Kathleen. 2005. How many meanings does a word have? Meaning estimation in Chinese and English. In James Menitt and William S-Y Wang (eds.) Language Acquisition, Change and Emergence: Essays in Evolutionary Linguistics. City University of Hong Kong Press.

Lv, Shu-Xiang. 1992. Zhong Guo Wen Fa Yao Lue. Taiwan: Wei Shi Zhe Press.

Perter, Harder. 1996. Functional Semantics: a theory of meaning, structure and tense in English. New York: Mouton de Gruyter.

Pustejovsky, James. 1995. The generative lexicon. Cambridge, Mass.: MIT Press.

Pustejovsky, James and Branimir Boguraev. 1996. Lexical semantics: the problem of polysemy. New York: Oxford University Press.

Ravin, Yael. 1990. Lexical semantics without thematic roles. Oxford: Oxford University Press.

Robins, Robert Henry. 1979. A short history of linguistics. Bloomington: Indiana University Press.

Rosch, Eleanor. 1977. Human categorization, in N. Warren ed., Advances in cross-cultural psychology, vol. 7. London: Academic Press.

Saeed, John I. 1997. Semantics. Cambridge, Mass.: Blackwell Publishers.

Schutze, Hinrich. 1997. Ambiguity resolution in language learning: computational and cognitive models. Stanford, Calif.: CSLI Publications.

Su, Lily, I-Wen. 2002. What can metaphors tell us about culture? Language And Linguistics 3.3, 589-613.

Swinney, David A. 1979. Lexical access during sentence comprehension: (re) consideration of context effects. Journal of Verbal Learning and Verbal Behavior 18, 645-659

Taylor, John. 1989. Linguistic categorization: prototypes in linguistic theory. Oxford: Oxford University Press.

Towell, Geoffrey and Ellen M. Voorhees. 1998. Disambiguating highly ambiguous words. Computational Linguistics 24.1, 125-145

Tsai, Jia-Lin., Wen-Lian. Hsu, and Jebg-Woei. Su. 2002. Word sense disambiguation and sense-based NV event frame identifier. Computational Linguistics and Chinese Language Processing 1. 29-46.

Tsai, Mei-Chih. 2005. Word sense disambiguation based on syntactic construction. Computational Linguistics and Chinese Language Processing 10.4:483-494.

Biq, Yung-O 1988. From focus in proposition to focus in speech situation: *cai* and *jiu* in Mandarin Chinese. Journal of Chinese Linguistics 16.1. 72-108.

Wang, Huei. 2002. A study on noun sense disambiguation based on syntagmatic features. Computional Linguistics and Chinese Language Processing 2. 77-88.

Wierzbicka, Anna. 1990. Prototypes save: on the uses and abuses of the notion of prototypes in linguistics and related fields', in S. L. Tsohatzidis ed., Meanings and prototype: studies in linguistic categorization. London: Routledge and Kegan Paul.

Wittgenstenin, L. 1953. Philosophical investigations. Oxford: Basil Blackwell and Mott.

——1958. The blue and brown books. Oxford: Basil Blackwell and Mott.

Appendix I

Examples of tagged corpus (for the purpose of economy, not all tagged data are listed)

Case study of *ZOU* 走

	Sense 1 'walking'
	Self-mover < Area < *
1	丁同笑道：「是，是！」[CNI/Self-mover]轉身繞到門前[Area]，**走**了進去。小屋中陳設簡陋，但桌椅整潔，打掃得乾乾淨淨。丁同
2	少將楊在祥看見唸外交研究所的女兒楊瑞珊[Self-mover]穿著旗袍在台上[Area]**走**貓步，不禁嚇了一跳
3	滿滿的，武裝衛兵把守著碼頭入口。我[Self-mover]在滿街水兵和軍官們中間[Area]**走**著，聽他們用熟悉的粗話互相笑鬧著、喧囂著，一直來到碼頭邊
	Self-mover < *
1	怎麼辦？我今天小考考八十分，回去一定會被爸爸罵死。[CNI/Self-mover]**走**著走著，珮珍知道自己到家了，卻不敢進屋
2	六年級的何正輝說：「不過[CNI/Self-mover]邊**走**邊玩，很快就到了。」
3	怎麼沒有人跟我做朋友呢？沒關係，我再接再厲。[CNI/Self-mover]走走！走**走**走！小野豬看到一隻美麗的小白兔，小野豬走過去問小白兔
4	我吐吐舌頭，趕忙壓低了聲音，老王你發了！別生悶氣了[CNI/Self-mover]！**走**！我帶你去做皇帝，好好的爽一下！你看這是什麼？老王掏出一
5	趨勢是從「物」的收藏，變成以「人」為重點，[CNI/Self-mover]和群眾越**走**越近，參觀的觀眾也不斷往社會金字塔的下層擴散。
6	我知道牠一定是在謝謝我，便開心的哈哈大笑。[CNI/Self-mover]**走**著走著，又看到最美麗的孔雀，正悠閒的踱著方步。
	Others
	Self-mover < * < Path
1	讓國軍佔得銅鼓鎮，上峰至少要提拔他一級。獎酬也不會少的。[Self-mover]**走**升官發財之路[Path]比捏著小命鑽山林強多了。這還得感謝姓莫的給他
2	讓小女孩破繭而出，尋回她原先快樂的自我，共處時，我以自己[Self-mover]**走**過的心路歷程[Path]，來對她做好心理建設。我告訴她，從小，我是
3	覺得舞劇是培養情感表達及團隊精神最好的方法，且現在大家[Self-mover]都**走**創作路線[Path]，對於能跳如此精緻的古典芭蕾感到很興奮。藝專
4	覺得保護自我還是第一步，最簡單的事情，你[Self-mover]不要自己一個人去**走**暗巷子[Path]，當然如果你真的打算隱居起來那也就罷了。這些問題就
5	蘇聯美食節來促銷。不過，仍有不少飯店以平常心來對待。[Self-mover]一向**走**平實路線[Path]，不大搞活動噱頭的兄弟飯店，則一如往年只推出聖誕

	others
	Self-mover < * < Area
	覺得真的是時間是非常有限的，你以前都在做這件事情，你[Self-mover]忽然**走**一個，完全陌生的世界[Area]是滿難的。
	機場內有二處海關，從法國來的有特別一處海關，而其它國家[Self-mover]則**走**另一處[Area]。觀看日內瓦美景最佳去處就是白朗碼頭
	歐洲男人帥呆了的浪漫女子，請注意下面這個故事：當你[Self-mover]一個人**走**在威尼斯聖馬可廣場上[Area]，一位又高又帥的男子跟你搭訕，問你..
	others
	Self-mover < * < Duration
1	從「浮在空中」般的不確定，到踩穩自己的步伐，[CNI/Self-mover]**走**了五年多[Duration]，這期間，家人的態度也逐漸由反對懷疑轉為肯定支持
2	鏡子裡面才是我要走的方向，而眼前的現實卻只是一片虛無。[CNI/Self-mover]又**走**了十來分鐘[Duration]，終於到了小敏的家。
3	這波由底擴大到頭肩底的強勢走法，反彈時間[Self-mover]預估可**走**二個月左右[Duration]，做頭反轉後將以波段式下跌，而且下跌波段中可
	Self-mover < path < *
1	還是可以開車，安全帶不太好綁，‖帶球走是犯規的，[Self-mover]帶球滿街[path]**走**。‖好的，還有，好…好，前面這位。請拿著麥克風。‖我選擇
2	學的佛法有多深奧，文句有多美，而是[Self-mover]要知道人生的道路[Path]該怎麼**走**，舉手投足才不會有錯誤。佛陀在世時化導眾生，也是一種社會
3	樣東西去發展，自然本身就有它的規律，[self-mover]就有它的路[Path]可以慢慢的**走**。所以老子的思想從「知」著手，認為人有認知的天性，這一點
4	就到台灣了。‖這條路有多少英里？‖有七千英里。[Self-mover]這條路[Path]最常**走**。‖到台灣沒有比這條更近的路嗎？‖沒有。經過歐洲有一條路
5	酸了，腿也酸了，我已經走累了。抬頭看看，我還在山腰，[Self-mover]路[Path]才**走**了一半。向前，向前，奮勉向前。要爬山，就不怕艱難。不知道
	Others
	Self-mover< * < Goal
1	前置一張黃花梨圓後背交椅，這是主人迎接客人的起點。[CNI/Self-mover]再**走**幾步，便來到了由十二面屏風環繞的羅漢床坐榻[Goal]
2	我們先前還想就著大路，[CNI/Self-mover]大不了多**走**幾步，一定找得到山下的大街[Goal]。哪曉得山路彎彎曲曲的轉來轉去
3	還在附近三公尺深的河底發現了當年入侵巴黎維京人的棄船。[Self-mover]再**走**著就到了遊塞納河的遊船碼頭Ｐｏｎｔ ｄｅ Ｌ[Gaol]，它在１９７４年
4	臨時起意的，那天我正好放假，揹著心愛的相機在蘇黎士閒逛，[Self-mover]**走**著走著便進了車站[Goal]，本想獵取鏡頭，眼睛就不停地在相機的
	Self-mover < * < Path < Goal

	隨心所欲，一般機場內皆設有租車服務。如果從倫敦清晨出發，[Self-mover]走M１A１公路[Path]下午３點左右便可抵達愛丁堡[Goal]。M表示高速公路，
	Self-mover < * < Direction
	政大東語系洪士培的男友是日本人，她說雙方[Self-mover]已決定未來要走的方向[Direction]，男友可能先回日本再至國外深造，而她則留在台灣工作
	見左首和右首兩條路上都有淡淡的足跡。蘇魯克道：「四個[Self-mover]走左邊[Direction]的，三個走右邊的，待會兒再在這裡會合。」李文秀道：「
	這是人，不是鬼。然而那是誰？七人[CNI/Self-mover]這時所走的方向[Direction]，早已不是李文秀平日去師父居所的途徑。她突然想起：
	Self-mover < Direction < *
1	覺得這些教法真是至高無上的真理，他覺得身心輕安，[Self-mover]又再向前走。沒想到，離開這燒窯的房子沒有多遠，忽然間一頭牛衝出來，
2	繼續進步，到某種階段就停下了。當他發現自己[Self-mover]沒有希望再往上走時，就會產生自暴自棄的心理，而且表現在挑剔別人上。這種人
3	還算好，可也不像過去那樣無所不談、無話不講。有次她[Self-mover]在前面走，我和幾個人在後面說話，說的完全是跟她不相干的人和事，
4	據導遊表示，如果是旺季前來，[Self-mover]根本是被後面的人潮催促著往前走，但因為冬天是歐洲的旅遊淡季，大可放慢腳步，沈浸在一座座
5	領著我們進操場，每個小朋友都精神抖擻，[Self-mover]踏著整齊的步伐向前走。接著點燃聖火，一會兒，天空出現五彩繽紛的氣球和飛得很快
	Others
	Self-mover < Path < *
1	所以，[CNI/Self-mover]未來仍有很長的路[Path]要走，僅以簡單的研究心得供作進一步研究的基礎。註釋：１．根據
2	我索性抱了鏡子，[CNI/Self-mover]挑了僻靜的街道[Path]慢慢地走。還有一個多星期就要到聖誕節了，許多人家已經在房子周圍
	Sense 2 ‘moving’
	Theme < Source < Goal <*
	優良表現，特別計畫了兩天一夜的短程旅行。車隊[Theme]從臺北[Source]往宜蘭[]走，經南方澳到花蓮，右轉橫貫公路的太魯閣、天祥，經大禹嶺，
	Theme < Path < *
	與學生共組讀書會，在國外風行，在台灣[Theme]卻有條坎坷的路[Path]要走。師範大學英語系及東吳大學社會系的讀書會，都面臨人數不多
	，但含油脂量又不會這麼高。在油品的改善[Theme]上尚有一段距離[Path]可走，例如植物油及動物油的含量如何能攝取到最低的程度
	也就是說當我們的身心達到平衡的時候，[Theme]可能在性的路途[Path]上走得會比較平穩一點。那麼所以說在性的這個取捨方面，一方面

	還有一段很長的路要走，[Theme]如何在未來的資訊高速公路[Path]上**走**得平穩安全，端賴吸收既有的經驗，提早做準備。
	輛的車隊。這個車隊在仙跡岩停車，[Theme]要沿著峭壁中間的公路[Path]往前**走**三公里，才能看到滑坡的場面。仙跡岩過去三四百公尺，就遇到
	theme < Direction < *
	「博物館的生命是永久，中華文化[Theme]也一直在往前[Direction]**走**，」秦院長表示，為了顧及民國以來的藝術發展，故宮在民國
	孔子了解傳統之後，發現整個時代[Theme]要往哪裡[Direction]**走**才有希望，他成為當時最有資格來承擔一種使命
	自由化時，國內整車市場或是車輛相關行業[Theme]必然需朝此方向[Direction]**走**
	何勞我們西方人苦苦生產？現代的全球市場[Theme]也在朝這個方向[Direction]**走**。」這是最典型的態度，理所當然地使用開發中國家的廉價勞工
	另外，如名聲與地位[Theme]，到了一個頂點之後，就會往下[Direction]**走**。這些都是外在化、數量化、物質化的東西。假如以這些東西
	Theme < Area < Path < *
	觀點來看，西方式的民主政治[Theme]，在中國大陸[Area]還有極其長遠的路[Path]要**走**。中國政治將在人治的政權鬥爭中，緩緩變化；在接班的權力
	Theme < *
	至４０５０點二分之一拔檔，有低價可以撿回最好，[Theme]若持續**走**高，則於４４００點左右全部出清。空手投資人，不必在意高價
	好忙解釋一句道：這船[Theme]**走**著真像個搖籃，人給它擺得迷迷糊糊只想睡。
	他與電腦對奕，棋子以磁石吸附在棋盤上，電腦[Theme]每**走**一步，聲彥就得全盤摸一遍，有時摸得自己的棋子也東倒西歪
	夜晚，一切這麼寧靜安詳。我沿著溪岸走，月亮[Theme]也跟著我**走**。我走到那熟悉的楊柳樹下，在石凳上坐下來，看那靜靜的溪水
	聯合政府，若從這點來看，則新黨[Theme]事實上是跟著民進黨的政策在**走**。
	確定性的熟悉，和頑固性的信任感，歷史[Theme]一直以相同的腳步向前**走**著面目相似的步伐，當科技不能反省人文，人文未能追上科技時
	Others
	Sense 3 ‘visiting’
	Self-mover < * < Area
	夏天又到了，你想到何種避暑方法了嗎？農委會建議民眾[Self-mover]，何妨**走**一趟休閒農場[Area]，在綠山青水間戲水、避暑，清涼一夏！
	選擇呢？操作日本旅遊市場２０餘年的陳總解析道，選對時機，[Self-mover]**走**一趟日本[Area]，您將不會有入寶山空手而回之憾，他特地向消費大眾
	走遍整個石門。而今又是一個多霧的天氣，但又有誰願意伴我[Self-mover]再**走**一趟石門[Area]？上個星期妳給了我一個電話，問：會恨我頓時訝然了
	我[Self-mover]今天其實打算**走**一趟〝金洋村〞[Area],看看有沒有機會一訪〝神秘湖〞。

	他的鬼魂走出雕像，拖著瘦弱的身體，[CNI/Self-mover]再**走**一趟宿命的旅程[Area]。幽魅的鬼魂，口中哼著送葬的哀歌，一語道斷
	Self-mover < Area < *
	東海岸的東管處,順便到泰源幽谷看看,再來我們[Self-mover]到南橫[Area]**走**一趟,走了這一趟回來,我發現我們都已經不怕冷了
	有種難以言喻的美，繁複的生命景致讓人目不暇給，如果你[Self-mover]到紅樹林[Area]**走**一遭，就能夠同時進行賞樹、賞鳥、賞蟹等貼近自然的活動，到處
	帆船，乃荷蘭村內最具象徵性的展示物。小人國；[Self-mover]在小人國內[Area]**走**一遭，彷彿自己便是童話中的巨人格列弗。威簾塔（
	Area < Self-mover < *
	著塔克金溪縱谷與司馬庫斯部落[Area]遙遙相望。我[Self-mover]希望下次有機會去**走**一趟。鎮西堡靠近新光，標高超過２０００公尺，此地早晚溫差
	中國人常做的運動，[Self-mover]最好就是哪一天一大清早到公園[A]去**走**一趟。早晨在公園裡運動的，男女老少都有
	Sense 4 ‘leaving’
	self-mover < *
	師父教過我的，怎地忘了？」瓦耳拉齊喝道：「你[Self-mover]再不**走**，我要殺你了！」
	我ＫＴＶ和同伴們一定會感激不盡的。現在，我[Self-mover]要**走**了，希望下次再來的時候，住的是一片安全美麗的大地
	尼奧卻說：「你們玩吧，我和秀子[Self-mover]吃完了就**走**。」「何必太嚴肅呢？看看何妨？」
	急著想出來，聽聽屋子裡很久沒有聲音了，料想皇帝[Self-mover]大概已經**走**了，就一邊兒把頭伸出來一邊兒問：老頭子走了吧?
	*＜self-mover
	確實，幾年前香港移民**走**了一批高級職員[Self-mover]，形成真空，底下的人藉此機會升了上來。
	所以今天能決定中國將來之運命者，第一件事便是請**走**宋子文[Self-mover]，否則政府必然垮台
	Self-mover < * < Area
	滾滾古路長，滿目空雲待夕陽，回頭一望家三遠，不知何事[SCNI/elf-mover]**走**他鄉[Area]。」請問這是出家人的境界嗎？

Appendix II

Semantic Associates in Context (search from BOW)

A . Case Study of *ZOU* 走

Sense 1 of *ZOU*: 'walking'

Chinese	English	*gloss*
徒步旅行	**afoot(p), walking(a)**	*traveling by foot*
拖著腳走	**shuffle, scuffle, shamble**	*walk by dragging one's feet*
疲弱的	**limp, hobble, hitch**	*walk impeded by some physical limitation or injury*
閒逛	**putter, potter, potter around, putter around**	*move around aimlessly*
散步	**amble, mosey**	*walk leisurely*
沉重的走	**lumber, pound**	*move heavily or clumsily*

Sense 4 of *ZOU*: 'leaving'

Chinese	English	*gloss*
消失	**disappearance, disappearing**	*the act of leaving secretly or without explanation*
離開	**departure, going, going away, leaving**	*act of departing*
逃脫	**breaking away**	*departing hastily*
告別	**farewell, leave, leave-taking, parting**	*the act of departing politely*
回來	**return, homecoming**	*a coming to or returning home*
移民	**immigration, in-migration**	*migration out of a place*

B. Case Study of *NA* 拿

Sense 1 of *NA*: 'getting'

Chinese	English	*gloss*
獲得	**acquisition**	*the act of contracting or assuming or acquiring possession of something*
得到	**acquiring, getting**	*the act of acquiring something*
接受	**reception**	*the act of catching a pass in football*
侵吞	**annexation, appropriation**	*the act of making money*
取用	**annexation, appropriation**	*incorporation by annexation*

Sense 2 of *NA*: 'carrying'

Chinese	English	*gloss*
保留以備將來之用	**reserve, backlog, stockpile**	*something kept back or saved for future use or a special purpose*
以腹帶固定腰帶	**cinch, girth**	*tie a cinch around*
保護的　安全的	**protected, secure**	*kept safe or defended from danger or injury or loss*
攜帶　帶著	**carry**	*the act of carrying something*
轉移	**transplant, transplanting**	*the act of uprooting and moving a plant to a new location*

C. Case Study of *TING* 聽

Sense 1 of *TING*: 'listening'

Chinese	English	*gloss*
(注意)聽	**listening**	*the act of hearing attentively*
(說話)聲音	**speech perception**	*the auditory perception (and comprehension) of speech*
(音樂)聲音	**musical perception**	*the auditory perception of musical sounds*
注意、留心	**attend, pay attention**	*give heed (to);*

Sense 2 of *TING*: 'perceiving'

Chinese	English	*gloss*
感知 感覺	**sensory, sensorial**	*involving or derived from the senses*
接收	**perceive, comprehend**	*to become aware of through the senses*
察覺	**perceiving, perception**	*becoming aware of something via the senses*
明白	**catch**	*become aware of*
注意	**detected**	*perceived or discerned*

D. Case Study of *KAN* 看

Sense 1 of *KAN*: 'seeing'

Chinese	English	*gloss*
碰見	**come across**	*be perceived in a certain way*
接收	**perceive, comprehend**	*to become aware of through the senses*
來自經驗的	**experiential**	*derived from experience or the experience of existence*
明白,了解	**catch**	*become aware of*
辨出	**recognize**	*detect with the senses*

Sense 2 of *KAN*: 'watching'

Chinese	English	*gloss*
察覺	**detection**	*the act of detecting something*
凝視	**contemplation**	*a long and thoughtful observation*
(注意)看	**watch**	*observe with attention*
注意、留心	**attend**	*pay attention (to);*

Appendix III

Gloss of Categories corresponding to traditional categories (adopted from Sinica corpus)

普及化標記	現代漢語詞類標記	說明
A	A	(非謂)形容詞
ADV	D ,Da ,Dfa ,Dfb ,Dk	副詞
ASP	Di	時態標記
C	Caa ,Cbb	連接詞
DET	Nep ,Neqa ,Nes ,Neu	定詞
FW	FW	外文標記
M	Nf	量詞
N	Na ,Nb ,Nc ,Ncd ,Nd ,Nh	名詞
P	P	介詞
POST	Cab ,Cba ,Neqb ,Ng	後置詞
T	DE ,I ,T	語助詞
Vi	VA ,VB ,VH ,VI	不及物動詞
Vt	SHI ,VAC ,VC ,VCL ,VD ,VE ,VF ,VG ,VHC ,VJ ,VK ,VL ,V_2	及物動詞
NAV		名謂詞

現代漢語詞類標記	說明
A	非謂形容詞
D	副詞
Da	數量副詞
Dfa	動詞前程度副詞
Dfb	動詞後程度副詞
Dk	句副詞
Di	時態標記
Caa	對等連接詞，如：和、跟

Cbb	關聯連接詞
Nep	指代定詞
Neqa	數量定詞
Nes	特指定詞
Neu	數詞定詞
FW	外文標記
Nf	量詞
Na	普通名詞
Nb	專有名稱
Nc	地方詞
Ncd	位置詞
Nd	時間詞
Nh	代名詞
P	介詞
Cab	連接詞，如：等等
Cba	連接詞，如：的話
Neqb	後置數量定詞
Ng	後置詞
DE	的, 之, 得, 地
I	感嘆詞
T	語助詞
VA	動作不及物動詞
VB	動作類及物動詞
VH	狀態不及物動詞
VI	狀態類及物動詞
SHI	是
VAC	動作使動動詞
VC	動作及物動詞
VCL	動作接地方賓語動詞
VD	雙賓動詞
VE	動作句賓動詞
VF	動作謂賓動詞

VG	分類動詞
VHC	狀態使動動詞
VJ	狀態及物動詞
VK	狀態句賓動詞
VL	狀態謂賓動詞
V_2	有