Computational Linguistics and Chinese Language Processing
vol. 3, no. 1, February 1998, pp. 45-60

Meaning Representation and Meaning Instantiation for Chinese Nominals[1]

Kathleen Ahrens[*], Li-li Chang[+], Ke-jiann Chen[+], Chu-Ren Huang[+]

Abstract

The goal of this paper is to explicate the nature of Chinese nominal semantics, and to create a paradigm for nominal semantics in general that will be useful for natural language processing purposes. We first point out that a lexical item may have two meanings simultaneously, and that current models of lexical semantic representation cannot handle this phenomenon. We then propose a meaning representation that deals with this problem, and also discuss how the meanings involved are instantiated. In particular we posit that in addition to the traditional notion of sense differentiation, each sense may have different meaning facets. These meaning facets are linked to their sense or to other meaning facets through one of two ways: meronymic or metonymic extension.

1. Introduction

Lexical ambiguity resolution is a central concern of natural language processing [Small et al., 1988]. The traditional way of looking at the problem is to list the various meanings that a word has, and write a rule-based program to pick the appropriate meaning for the context. Both Categorical Grammar and Montague Semantics, for example, assume that meanings are discrete and that there is a one-to-one correspondence between a lexical item and its meaning translation. The discrete meaning hypothesis provides the conceptual basis for most of the previous literature on ambiguity resolution and semantic resolution. In short, ambiguity resolution is viewed as trying to choose from several discrete meanings that share the same linguistic form (i.e. lexical form). While this approach can provide an algorithm to identify an appropriate meaning in a given context, it cannot account for novel uses of lexical items.

---

* National Taiwan University, Taipei, Taiwan, R.O.C. E-mail: ahrens@ms.cc.ntu.edu.tw
+ Academia Sinica, Nankang, Taipei, Taiwan, R.O.C.
1. This paper is jointly authored. The names of the authors are listed in alphabetical order.

More recent work addresses this problem. Pustejovsky's [1995] Generative Lexicon provides a framework (i.e. qualia structure) for possible meanings, and discusses under what conditions which meaning will be chosen (i.e semantic coercion). His account is especially useful in dealing with the creative use of words in novel contexts, an area that had been previously ignored due to the assumption that either a) the novel usage could be listed if necessary, and b) often it was deemed not necessary to list these novel meanings because they occurred so rarely.

However, one issue that Pustejovsky and others have yet to account for is the fact that lexical meaning can be **actively complex**. All models of lexical ambiguity resolution assume that only one solution exists in a given context. In fact, what we will show is that more than one meaning can co-exist in the same context. A lexical item is actively complex if it allows simultaneous multiple interpretations. We will propose a meaning representation for lexical items that captures this complexity.

In addition, although Pustejovsky provides the framework to exclude the possible meanings, he cannot predict the relationship among the meanings, nor allow for cases where different meanings seem to exist simultaneously. Within the general theory of the Generative Lexicon, Copestake and Briscoe [1995] deal with meaning extension by either underspecification or lexical rules, which also implies that only one meaning can be expressed at any given time.

In our account, we will demonstrate that meaning can be predicted from its context by the interaction of a) the semantic class of the item, and b) its possible meaning extensions. Our account has the advantage of being able to account for a wider range of linguistic data, including puns and polysemous uses, in addition to novel extensions. Our account also has the advantage of being both computationally parsimonious, as well as conceptually intuitive.

Our paper is divided as follows: in section 2, we will first present background information and definitions concerning the different kinds of ways that meanings can vary. In section 3, we will present our arguments for the active complexity of lexical meaning, present a representation that can handle active complexity, and also give reasons for the conceptual intuitiveness of the model. In section 4, we will discuss the meaning extensions that have been found to date. Section 5 discusses the hierarchical information that is passed from a semantic class to an individual item of that class. Section 6 summarizes our findings and suggests future areas of research.

2. Background

In this paper we devise a meaning representation for nominals (and Chinese nominals in particular) such that all meaning aspects of a noun are dealt with parsimoniously. Nouns, at first glance, do not seem to warrant representational complexity. When one is asked to think of a noun, one commonly thinks of a concrete object, such as 'paper'. When asked to define it, one could reply that it is a thin, white, rectangular object (appearance) made from the pulp of trees (origin) that people nowadays use to write and print on (function). But 'paper', even if we do not talk about its additional meanings in compound items such as 'wrapping paper', 'tissue paper' etc., has a variety of meanings including: a piece of paper, a newspaper, the office where a newspaper is written, and an academic paper. This phenomenon is not language specific. For example, in Mandarin Chinese, the word 雜誌 'magazine' can refer to the physical object (1a), or the information contained within (1b), or the publishing house (1c).

(1a) 他 手 上 拿 了 本 雜誌。

ta shou shang na le ben zazhi
he hand on hold asp. CL magazine
'He is holding one magazine in his hand.'

(1b) 我們 從 雜誌 中 得到 許多 寶貴的 資料。

women cong zazhi zhong dedao xuduo baoguide ziliao
we from magazine within obtain many precious data
'We have obtained a lot of precious data from magazines.'

(1c) 美國 各 大 雜誌 無不 挖空 心思 爭取 採訪 機會。

meiguo ge da zazhi wubu wakong xinsi zhengqu caifang jihui
America every big magazine do dig-empty mind fight fro interview chance
'Major American magazines fight for interview opportunities.'

Nor is this phenomenon limited to words relating to items that may contain information such as papers and magazines. The word 天 'tian' in Chinese can refer to the sky (2a), God (2b), weather (2c), time (2d), day(s) (2e), or nature (2f). The word 刀 'dao' can refer either to the whole knife (3a), or only to the blade of a knife (3b). The word 梅花 'meihua' can refer either to the plum-flower blossom (4a), or the plum-flower plant (4b). The word 白菜 'baicai' can refer to either the round raw vegetable (5a), or the soft

cooked mass (5b).

(2a) 抬頭 望 著 湛藍的 天。

taitou wang zhe zhanlande tian
raise head watch asp. blue sky
'Raise one's head and look at the blue sky.' ('Tian' refers to sky.)

(2b) 中國人 說 福 自 天 來。

zhongguoren shuo fu zi tian lai
Chinese say happiness from sky come
'Chinese say, happiness comes from heaven.' ('Tian' refers to God/heaven.)

(2c) 天 冷 時 別 忘 了 加 件 衣服。

tian leng shi bie wang le jia jian yifu
sky cold time not forget asp. add CL clothes
'Don't forget to put on more clothes when the weather is cold.'
('Tian' refers to weather.)

(2d) 天 不 早 了。

tian bu zau le
sky not early particle
'It is not early.' ('Tian' refers to time.)

(2e) 他 在 這裡 待 了 一 整 天。

ta zai zheli dai le yi zheng tian
he in here stay asp. one whole sky/day
'He has stayed here for one whole day.' ('Tian' refers to day(s).)

(2f) 人類 是 大部分 動物 的 天敵。

renlei shi dabufen dongwu de tiandi
human being is most animal 's natural enemy
'Human beings are the natural enemy of almost all animals.' ('Tian' refers to nature.)

(3a) 我 向 他 借 了 一 把 刀。

wou xiang ta jie le yi ba dao
I from him borrow asp. one CL knife
'I borrowed a knife from him.' ('Dao' refers to the whole cutting instrument.)

(3b) 這 把 刀 很 利。

zhe ba dao hen li
this CL knife very sharp
'The knife is very sharp.' ('Dao' refers to the cutting edge.)

(4a) 一 朵 梅花

yi duo meihua
one CL plum-flower
'a plum-flower blossom' ('Meihua' refers to the blossom.)

(4b) 一 棵 梅花

yi ke meihua
one CL plum-flower
'a plum-flower plant' ('Meihua' refers to the whole plant.)

(5a) 一 棵 白菜

yi ke baicai
one CL Chinese cabbage
'a Chinese cabbage' ('Baicai' refers to the vegetable plant.)

(5b) 一 盤 白菜

yi pan baicai
one CL Chinese cabbage
'a dish of Chinese cabbage' ('Baicai' refers to the cooked dish.)

The examples we have given above are all examples of polysemy, which is when a word has several, related meanings. But meanings can also be unrelated, as in the case of the two meanings for 'bank' (i.e. 'financial institution' and 'land on the side of a river'). A

noun that has two unrelated meanings is referred to as homonymous. Meanings for a word can also be vague or underspecified. An example of this in English is 'aunt' which can refer to someone's parent's sister, where the gender as to the parent is unspecified. (The parent's gender in other languages, such as Mandarin, is important and specified.) The difference as to whether a word is ambiguous or polysemous depends on the perceived relationship (or lack thereof) between the meanings. The distinction between vagueness and polysemy 'involves the question whether a particular piece of semantic information is part of the underlying semantic structure of the item, or is the result of a contextual (and hence pragmatic) specification' [Geerarts 1993:228].

This definition, however, cannot be applied as straightforwardly as it appears. Consider example (1) above. It could be the case that there is no underlying semantic structure for the three meanings (that is, they are vague), and that context alone 'brings out' these meanings. But 1) intuitively these meanings seem to have an underlying structure, and 2) nouns of a similar semantic class (i.e. magazines and newspapers) have similar meanings, which indicates that an underlying structure exists. If it is the case that the pieces of semantic information are part of the underlying structure of the item, then we must deal with the paradoxical situation (given the definition above) that these different meanings are brought out in different contexts.

Tuggy [1993] points out that ambiguity, polysemy and vagueness are better dealt with on a continuum, rather than as sets with discrete boundaries. The prototypical case of ambiguity is where well-entrenched and salient semantic structures are associated with the same phonological representation, and there is no clear subsuming semantic schema. The prototypical case of vagueness is where the meanings are not well-entrenched, and there is a clear subsuming semantic schema (as in the case of parent's sister for 'aunt'). Polysemy is viewed as being in between these two extremes, with there are well-entrenched and salient semantic structures associated with the same phonological representation, but there is also a subsuming schema.

3. Meaning Representation

3.1 Active Complexity of Lexical Items

The above discussion has assumed that one meaning is chosen in a given context. But that is not necessarily the case. There are two types of active complexity in natural language. The first is 'triggered complexity' and involves puns. For example, in (6) either liquor and shipyard is possible as the meaning of port, but it is also possible for both meanings to exist at the same time.

(6) After the accident, the captain went straight for the port.

Example (6) can mean that a) the captain went straight for shore (but humorously implies that the captain was so shook up as to need a drink), or b) that he went straight for his bottle of liquor and also towards the shore (although this is much less likely since this interpretation is not seen as humorous).

The phenomenon in example (6) is a pun. Puns are a humorous play on ambiguous words. Because puns are used for special linguistic purposes (such as humor), and because it is the effect of co-existing meanings that creates the humor, this phenomenon has not previously been considered to be relevant to lexical semantic analysis and lexical representation. The complexity is triggered since it must be initiated by the speaker.

Second, in Chinese, nouns can be actively complex, even when there is no pun or vagueness intended. This is 'latent complexity.' In (7), for example 'book' must be understood as both a physical object, and as information.

(7) 張三 在 翻閱 那 一 本 書。

Zhangsan zai fanyue na yi ben shu.
Zhangsan PROG turn page/read that one CL book
'Zhangsan is turning the pages of the book and reading it.'

In fact, such latent complexity also exists in English nominal semantics. It is well-known that words referring to building apertures, such as door or window are often lexically ambiguous with the structure built to block that aperture. Thus, door in (8) could only refer to the structure, while door in (9) can only refer to the aperture. However, both the aperture and structure's meanings exist simultaneously for both the English and Chinese sentence in (10).

(8) The door is heavy.

(9) John walked in the door.

(10) 門 很 寬

men hen kuan
door very wide
'The door is very wide',

We think this kind of data presents the strongest argument against representing

nominal semantics as discrete meaning translations, and for representing nominal semantics as structured meanings connected by conceptual links, such as the qualia structure in Pustejovsky's Generative Semantics. However, since we have shown that different but related meanings can coexist in the same context, Pustejovsky's formulation where related meanings are represented as different attribute value pairs in a feature matrix is inadequate since only one attribute value pair can be picked in each context. We posit that these related meanings are like the facets of a three-dimensional object, such as a diamond, where the meaning instantiation could be a straightforward single facet or multiple connected facets, depending on the context.

3.2 Meaning Representation

The meaning representation that we select is quite straightforward, but differs from other representations in several crucial respects. First, words are listed (following Chinese lexicographic tradition) in terms of their orthographic representation (i.e. the stroke order of the Chinese characters.) Then the senses for each word are listed. The phonological representations are associated with each sense listing, and may or may not be the same. Second, the sense differentiation includes senses that are related (polysemous senses) as well as unrelated (homonymous senses). There is no attempt in this representation to distinguish clearly between those meanings that are polysemous or homonymous. This is because speakers tend to draw their own conclusions about the relationships between senses (e.g. many speakers see a relationship between 'ear of corn' and 'ear that you hear with', although there is no historical or semantic relationship whatsoever [Lyons 1977]).[2] Third, and most importantly, our lexical representation has **meaning facets** located within each sense. Meaning facets reflect an aspect of a sense. For example, in (11)-(14) we show examples of words with one sense, of which there are two to three different meaning facets.

(11) 雜誌 **--Sense$_1$**: **ZAZHI *magazine*** -- meaning facet$_1$: *physical object*
-- meaning facet$_2$: *information contained within*
-- meaning facet$_3$: *institution that publishes magazine*

(12) 刀 **--Sense$_{1:}$ DAO *knife*** -- meaning facet$_1$: *physical object*
-- meaning facet$_2$: *the blade of it*

2. However, if a study was run on native speakers to find out their understanding of the relative closeness of relationship among meanings, this information could be incorporated into our representation by simply indicating which senses should be grouped together.

(13) 梅花 --**Sense$_1$: MEIHUA *plum flower***

-- meaning facet$_1$: *physical object: the blossom*

-- meaning facet$_2$: *the whole plant contains the blossom*

(14) 白菜 --**Sense$_1$: BAICAI *Chinese cabbage***

-- meaning facet$_1$: *physical object: the vegetable*

-- meaning facet$_2$: *the cooked form of it*

In (15) we give an example of a word with four different senses, of which one has three different meaning facets.

(15) 天 --**Sense$_{1:}$ TIAN *sky*** -- meaning facet$_1$: *sky as a physical object (that can be viewed)*

-- meaning facet$_2$: *God/heaven*

-- meaning facet$_3$: *weather*

--Sense$_2$: TIAN *time*

--Sense$_3$: TIAN *day*

--Sense$_4$:TIAN *nature*

How do we decide whether a certain meaning is a sense or a meaning facet? A meaning facet is an extension from a particular sense. It has the following three properties: 1) it can appear in the same context as other meaning facets, 2) it is an extension from a core sense or from another meaning facet (unless it is the core sense), 3) nouns of the same semantic classes will have similar sense extensions to related meaning facets. Individual senses, on the other hand, 1) cannot appear in the same context (unless the complexity is triggered), 2) have no core sense from which it is extended, or it is very hard to concisely define what the core sense would be, and 3) no logical/conceptual links can be established between two senses, non can the link between two senses be inherited by class of nouns.

For example, in (16) below, we can see that the meaning of sky (as a physical object) and God can appear in the same context, as can sky (as a physical object) and weather (17), sky (as a physical object), God, and weather (18). Thus, they are all different meaning facets of the first sense in (15).

(16) 有 人 開始 不 敬 天 也 不 拜 天 了。

you ren kaishi bu jing tian ye bu bai tian le
there are person begin not respect sky and not worship sky particle
'There are people who ceased to respect heaven or to worship heaven.'
('Tian' refers to both sky and God/heaven.)

(17) 天 放晴 了。

tian fangqing le
sky become sunny particle
'It became sunny.' ('Tian' refers to both sky and weather.)

(18) 農民 長久 靠 天 依 地 的 生活。

nongmin changjiou kau tian yi di de shenghuo
farmer long depend sky depend ground DE live
'Farmers have long lived a life that depends on heaven and earth.'
('Tian' refers to sky, God, and weather.)

The above examples also demonstrate that only one sense can occur in any given context. The sense of 'time' or 'day' or 'nature' is not available in any one of the above contexts.[3] Only meaning facets of a particular sense can be available in the same context. Context, in effect, selects which sense is made available. Context may also select a particular meaning facet, as in (2a)- (2c), but it does not necessarily have to, because context may activate several meaning facets at once, as in (16) - (18).

What aspects of context help to pick a sense or a meaning facet? Verbs and prepositions are usually instrumental in determining which meaning can occur in which context. For example, in the above instance, the meaning of 'God' can only occur with volitional verbs and cannot occur with verbs having to do with pure locative. The type of contextual information that picks out one sense or one meaning facet is an important area of future research.

---

3. 'Time' might be viewed as a meaning facet of the sense 'sky', as shown by the identical strings in (i) and (ii).

(i) [$_{s}$ 天 [$_{VP}$ 黑 了]] 。
tian hei le
sky dark particle
'The sky turned dark.'

(ii) [$_{S}$[$_{VP}$[$_{V}$ 天 黑] 了]]] 。
tianhei le
sunset particle
'The sun has set (i.e. it is late).'

However, the interpretation in (i) is a subject-predicate sentence, while the interpretation in (ii) involves a disyllabic lexical item. Thus, these two sentences are structurally different and no latent complexity is involved.

3.3 Conceptual Adventages

Viewed from this perspective, context always plays a role in determining which meaning is chosen, whether the word is ambiguous, polysemous, or vague. Tuggy's meaning models were two dimensional. But we suggest that a 3-dimension model allows for a greater understanding of the relationship between meaning and context. Imagine a multi-faceted object, such as a cube. Imagine that there is a core in the center of the cube, and that there are lines that radiate out to each of the six surfaces (i.e. this would be the case for a word that had six senses). The core represents the orthographic representation of the word, and each surface represents a different sense of the word and its associated phonological representation (i.e. the information that is bolded in our lexical representation above). Furthermore, from each surface of the cube, there may also be (dotted) lines that radiate out to additional surfaces, which are the facets of that particular sense (i.e. the non-bolded information in our lexical representation above). Thus, when context turns the cube so that one particular sense surface is shown to a light source (i.e. the hearer) then light is reflected from only that surface, and only that sense is computed. In the case, however, where context turns the cube so that a sense surface that has meaning facets extending from it is shown to a light source, the light can reflect off of any one, or any combination of the meaning facets, just as light can reflect from the different facets of a diamond. Our representation, then, is not only computationally adequate, it is also conceptually intuitive.

In what follows we present the types of links that can occur in noun meaning representations, and we also present the underlying schema for the information contained in each meaning facet.

4. Meaning Links

In our model the meaning representation is structured, and the structure is built upon meaning links. One implication of this model is that semantic classes in a semantic hierarchy will inherit both traditional semantic features as well meaning link structures. Lexical semantic issues will therefore be defined in terms of 1) lexical senses, 2) the possible meaning links of their sense classes and 3) constraints on meaning extensions through these links.

The relationship between a sense and its meaning facets is an area that deserves in depth research and analysis. What follows is a preliminary report of our findings to date. We have found that there are two main ways that meaning facets can extend either from a sense or from another meaning facet: meronymic and metonymic extensions.

4.1 Meronymic extensions

Meronymic extensions involve both the whole standing for part, and part standing for whole. We observe that meronymic extensions are driven by cognitive and conceptual saliency. For example, in (3b) knife actually refers to the blade of the knife. This meronymic extension is motivated by the fact that 'blade' is the locus of cutting, and the most salient function of knife. We also observe that such cognitively driven extensions are not sensitive to blocking effects. For instance, the instance of the specific term 刀刃 'blade' does not block us from saying 'the knife is sharp' as in (3b). Our speculation here is that only conventionalized usages are subject to blocking effects since blocking is the result of (competing) conventions.

In the case of part standing for whole, cognitive saliency is again the prime motivator of the extension. For example, in the case of (19), plum-flower stands for the whole plum tree. The plum flower with its color and scent and endurance in cold weather is the most cognitively salient aspect of the plum tree (for Chinese).

(19) 院子 裡 有 許多 梅花

yuanzi li you xuduo meihua
garden inside exist many plum-flower
'There are many plum-flowers in the garden.'

4.2 Metonymic Extensions

Metonymic extensions are different from meronymic extensions in that the extended meaning is related to the origin of the basic sense, but is not inherent to the basic sense (cf. the part-whole relation above). Metonymic extensions are typically driven by certain eventive relationships such as the ones encoded in Pustejovsky's qualia structure. Unlike meronymic extensions, metonymic extensions are often sensitive to blocking effects. For instance, the grinding extension allows the individual terms to refer to a mass produced from that individual. For example, in (5b) the basic meaning ' 白菜 baicai' refers to the cabbage plant, but after the grinding extension it refers to a mass noun. But in the case of rice ' 米 mi', the grinding extension does not work, because there is a term ' 飯 fan' (cooked rice) already.

4.3 Partial list of Meaning Links

We give here a partial list of the meaning links found to date. We also provide the list of semantic classes that we have found to inherit these links.

I. Meronymic Extensions

1. Whole for part
   a. whole → functional part {semantic class: artifacts, buildings}
   b. whole → sentiently salient part {semantic class: body parts}
2. Part for whole
   a. conceptually salient part → whole {semantic class: fruit, flower}

II. Metonymic Extensions

1. agentivization
   a. information media → information creator {semantic class: publications}
2. product instantiation
   a. institution → product {semantic class: manufacturer, trademarks}
3. grinding
   a. individual → mass {semantic class: vegetables, fruits}
4. portioning
   a. information media → information {semantic class: publications}
   b. container → containee
   c. body part → function
5. space mark-up
   a. landmark → space in vicinity {semantic class: locations, landmarks}
   b. structure → aperture {semantic class: doors, windows}
   c. institution → locus {semantic class: institutions}
6. time mark-up
   a. event → temporal period
   b. object → process
   c. locus → duration

A summary of the links used in the lexical representation of the words we define in this paper is given below (cf. ex. 11): First, the meaning links between the different facets of 'zhazhi' (magazine) are as follows: the first meaning link refers to the concept of magazine as a physical object, the second meaning link is a metonymic extension that relates information media to information, and the third meaning link is a metonymic extension that relates information media to information creator. The link between the two facets of 'dao' (knife) (cf. ex. 12) are that the first link refers to the concept of knife as a physical object (in its entirety), and the second link is a meronymic extension (whole for part) to the meaning facet of 'blade.' The link between the two facets of 'meihua' (plum-flower) (cf. ex. 13) are that the first link refers to the conceptually salient notion of plum-flower, and the second link is a meronymic extension (part for whole) to the

meaning facet of plum tree. The link between the two facets of 'baicai' (cabbage) (cf. ex. 14) is that the first link refers to the individual head of cabbage, and the second link is a metonymic extension (grinding) to the meaning facet of a 'dish of cabbage'. The links among the facets of 'tian' (sky) (cf. ex. 15) are that the first link refers to sky as a physical object, the second link is a metonymic extension of space mark-up, and the third link is a meronymic extension of whole extending to the sententially salient part.

We have found that these two types of links (i.e. meronymic and metonymic extensions) are the most productive among meaning extensions. This might be because these types of extensions refer only to the knowledge concerning the lexical item itself. Metaphorical extensions, on the other hand, map a domain of knowledge that does not have anything to do with the lexical item onto the domain of knowledge surrounding the lexical item. Thus, metaphorical extensions are clearly conceptually more complex than metonymic and meronymic extensions, and will be the focus of future research.

5. Meaning Inheritance

Another important issue in lexical semantics is the semantic class. Traditionally, the taxonomic hierarchies are discussed in terms of ISA relationship and inherited features, such as humanness and animacy [Chen and Cha 1988, Sowa 1993]. However, this simplistic traditional model (such as Schank's well-known semantic network) have difficulties when certain nodes do not necessarily inherit all the features from the higher nodes. For example, an ostrich is a bird, but it cannot inherit the feature of [+flight] because it does not fly. Default override of inheritance is computationally plausible though costly.

The other problem with traditional semantic hierarchies has to do with multiple inheritance. For instance, it is intuitive to classifiy '籃球 lan-qiu' (basketball) as a physical object. However, it is also clearly an abstract event (i.e. the basketball game). Hence there is cross-taxonomic paradox, which is usually accounted for with the computationally costly mechanism of multiple inheritance [Briscoe et al., 1993].

In our model, both kinds of inheritance problems disappear since what a semantic class shares is a partial structure of semantic links. That is, we will annotate meaning links to a semantic class, and these links will be inherited by all the members of the class.[4] In the case of '球 qiu' (ball), it inherits the metonymic link of a round physical object and extends to the game played with the object. This explanation is more parsimonious since it reduces the costly computation of multiple inheritance and makes most

4. Of course, the lexicon would have to specify any blocking effects where the linking does not apply.

cases of the local overriding of inheritance unnecessary. It is also conceptually powerful in allowing richer semantic representation. For instance, the semantic class of flowers will inherit the meronymic extension of part for whole.

6. Conclusion: Implementation and Implications

Traditional methods of dealing with ambiguity and vagueness in natural language processing have been complicated by the on-line compilations that are usually necessary to deal with the 'additional' meanings created by the context. But our account postulates multiple senses and structured ways of linking additional meaning facets to the senses so that the information is all listed in the representation, and therefore easier to access. Our proposal is to have not only the different senses of a word listed, but also its different meaning facets. We claim that there are conceptual or logical relationships between the facets and their senses, as discussed in section 4.

The organization that we have proposed here is a shallow structure, with only two levels: the sense level and the meaning facet level. Both levels can be annotated with meaning links. Conceptually it is as explanatory as a theory where all the meaning links are structurally represented. This is because all represented meaning links can be traced, and a (semantic-class-based) meaning derivation tree can be established off-line. Moreover, not having an overt tree of meaning extensions allows us to avoid multiple-inheritance and blocking problems. A shallow structure also allows efficient access, reflecting the psychological reality that the depth of meaning derviation is not relevant in lexical access.

In this paper we propose a meaning representation for Chinese nominal semantics, as well as a paradigm for nominal semantics in general that will be useful for natural language processing purposes. We point out that a lexical item may have two meanings simultaneously, which current models of lexical semantic representation cannot handle. We call this phenomena 'active complexity.' There are two types of active complexity: 'triggered complexity' where the noun is purposely selected to be simultaneously ambiguous, and 'latent complexity' where the noun selected has two or more meanings coexist, but the effect is not humorous and was not selected for such an effect. We propose a meaning representation to account for this phenomena, and also discuss how the meanings involved are instantiated. We postulate that in addition to the traditional notion of sense differentiation, each sense may have different meaning facets. These meaning facets link to their sense or to other meaning facets through one of two ways: meronymic or metonymic extension. We also point out that instead of a traditional taxonomic relationship, what is being inherited in addition to semantic features is

meaning extensions/relations, such that words of the same semantic class have the same meaning extensions. Our representation, therefore, allows for predictions of meaning extensions from a semantic class.

The representation proposed here is the result of extensive corpus-based studies of the 200 most productive nominal endings in Mandarin [CKIP 1995]. These productive nominal endings in turn each derive scores of highly frequent nouns. Hence we have accounted for a substantial portion of Chinese noun usages. We have also provided detailed semantic representation of the nominal heads based on our proposed representation. This is a significant first step towards the comprehensive formal representation of Mandarin nominal semantics and is also the first step towards fully automated Mandarin Language Understanding.

References

Briscoe, E., J. Copestake, and V. de Paiva, *Inheritance, Defaults and the Lexicon*. Cambridge University Press, 1993.

Chen, K.-J., and C.-S. Cha, "The Design of a Conceptual Structure and Its Relation to the Parsing of Chinese Sentences," *Proceedings of 1988 International Conference on Computer Processing of Chinese and Oriental Languages* (*ICCPCOL*), 1988, pp. 428-431.

CKIP, "Contents and Explanations of Sinica Corpus," *CKIP Technical Report. 95-02*. Nankang: Academia Sinica, 1995.

Copestake, A., and T. Briscoe, "Semi-productive Polysemy and Sense Extension," *Journal of Semantics*, 12, 1995, pp.15-67.

Geerarts, D., "Vagueness's puzzles, polysemy's vagaries," *Cognitive Linguistics*, 4.3, 1993, pp. 223-272.

Lyons, J., *Semantics*, Cambridge University Press, 1977.

Pustejovsky, J., *The Generative Lexicon*, MIT Press, 1995.

Small, S., G. Cottrell, and M. Tanenhaus, *Lexical Ambiguity Resolution: Perspectives from Psycholinguistics, Neuropsychology and Artificial Intelligence*, Morgan Kaufmann Publishers, 1988.

Sowa, J., "Lexical Structure and Conceptual Structure," in *Semantics and the Lexicon*, Pustejovsky (Ed.), Kluwer Academic Publishers, 1993, pp. 223-262.

Tuggy, D., "Ambiguity, Polysemy and Vagueness," *Cognitive Linguistics*. 4.3, 1993, pp. 273-290.