Computational Linguistics and Chinese Language Processing
Vol. 18, No. 4, December 2013, pp. 45-62

A Semantic-Based Approach to Noun-Noun Compound Interpretation

You-shan Chung* and Keh-Jiann Chen*

Abstract

In this project, we have studied Chinese noun-noun compounds (NNCs) and have found that N1 and N2 are linked either by semantic roles assigned by events (complex relations) or by static relations (simple relations), including meronymy, conjunction, and the host-attribute-value relation. Using data from the FrameNet and E-HowNet, we have found that, for NNCs of either type, the major semantic relations between the two components are limited enough to allow computational implementation. Regarding simple relations, most conjunction pairs have been listed in E-HowNet,and so are host-attribute-value sets. The E-HowNet Taxonomy also makes identification of meronymy possible. As for NNCs involving complex relations, each component's semantic role, along with the events that assign these roles, can be restored through mappings to corresponding frame elements (FEs) in entity and to event frames and lexical units (LUs) in FrameNet's frames, respectively, that represent the concept the NNC conveys.

Keywords: Noun-noun Compounds, Automatic Interpretation, Extended HowNet (E-HowNet), FrameNet

1. Introduction

Noun-noun compounds (henceforth NNC) are compounds composed of two nouns. For example:

麵包刀 *mianbao-dao* 'bread knife'
衛星城市 *weixin-chengshi* 'satellite city'
金融股 *jinrong-gu* 'stocks in the financial sector'
秋蟹 *qiu-xie* 'autumn crab'

* Institute of Information Science, Academia Sinica, Taipei, Taiwan
E-mail: {yschung, kchen}@iis.sinica.edu.tw

腳踏車輪胎 *jiao-ta-che luntai* ‘bicycle tire’
卵石地板 *luan-shi diban* ‘pebble floor’
鐘錶 *zhong-biao* ‘clock and watch’
鐵桌 *tie-zhuo* ‘iron table/desk’
車速 *che-su* ‘car speed’

While the part-of-speech (POS) of NNCs usually is nominal, their interpretations seem so diverse that some researchers even contend that they are completely determined by context (*e.g.* Dowty, 1979; reviewed in Copestake & Lascarides, 1997).

Nevertheless, the majority of researchers believe that there is at least some degree of regularity in NNC interpretation. This regularity is often reported to be at least partially universal as well (Levi, 1978; Søgaard, 2005). There are three popular theories along these lines, which are not mutually exclusive. First, there is a limited set of semantic relations between the two component nouns, N1 and N2 (Levi, 1978; as well as computational works that implemented her theory, *e.g.* Copestake & Lascarides, 1997; Søgaard, 2005; Copestake & Briscoe, 2005; Huang, 2008). Second, N1 and N2 are the arguments of an event that bridges them and by which they are assigned semantic roles (Levi, 1978; Leonard, 1984; Ryder, 1994). Third, the two component nouns sometimes are linked through similarity in some aspect, resulting in metaphorical readings.

Nevertheless, these accounts generally have the following four problems. First, the semantic relations they proposed or adopted tend to be not specific enough. Levi (1978), for instance, proposed nine semantic relations between N1 and N2, which she called Recoverably Deleted Predicates (RDP), including CAUSE, HAVE, MAKE, USE, BE, IN, FOR, FROM, and ABOUT. These RDPs, however, appear to be too general to be informative, especially with prepositional ones like IN and FOR, as NNCs linked by the same preposition belong to the same semantic categories only in a very broad sense.

Second, some of the studies resolve only limited or sporadic semantic categories, while others are questionable in terms of their correct prediction rate. For example, the fourteen semantic relations Li and Thompson (1981) proposed do not seem to make up a meaningful and discrete inventory of semantic relations, while Huang’s (2008) combinational patterns for three major categories of physical objects (*i.e.* animals, plants, and artifacts) are each based on the analysis of only six morphemes, raising concerns about generality.

The third problem is that the classifying criteria mostly are left unaccounted for; thus, they appear arbitrary. For example, Levi (1978) sees the two components of *lemon peel* and *apple seed* as linked by the predicates HAVE and FROM, respectively, but such a distinction

between the two NNCs may not be without controversy.

The last problem is that bridging does not seem to be eventive or by prepositions in the following three situations: first, the host-attribute-value relation (*e.g.* 鐵桌 *tie-zhuo* 'iron table/desk,' 車速 *che-su* 'car speed') with two special subclasses, where N1 denotes time (*e.g.* 秋蟹 *qiu-xie* 'autumn crab') or N1 denotes space (*e.g.* 倫敦地鐵 *Lundun-ditie* 'London Underground'); second, meronymy, or part-whole relation (part-whole: *e.g.* 雙底船 *shuang-di chuan* 'double-bottom,'; whole-part: *e.g.* 腳踏車輪胎 *jiao-ta-che luntai* 'bicycle tire'); and third, conjunction (*e.g.* 鐘錶 *zhong-biao* 'clock and watch,' 禮樂 *li-yue* 'manners and music') .

Before we go on, we need to explain the definition of Chinese NNCs adopted in this study. Unlike in English, formal similarity in Chinese does not entail a shared POS. For example, the first component in 希臘國歌 *xila guo-ge* 'the national anthem of Greece,' 希臘菜 *xila-cai* 'Greek dish,' and 月費 *yue-fei* 'monthly fee' corresponds to adjective forms in their English equivalents. Nevertheless, we include these various forms in our analysis since such formal differences do not reflect conceptual differences, as Levi (1978) has argued for this at length and also included adjectives in her analysis of what she called "complex nominal," or "NNCs" in our terms.

Addressing the aforementioned four problems, we used a knowledge base that we believe could help decide the precise semantic relations for both event-linked and non-event-linked NNCs, which is FrameNet (https://framenet.icsi.berkeley.edu/fndrupal/). In essence, the theory behind FrameNet is that lexical units (LU) evoke concepts represented by "frames," which are each composed of a set of frame elements (FE), *i.e.* the overtly-realized semantic roles assigned by the frame's LUs. Some LUs evoke entity concepts, while others evoke eventive ones. Since many entities in FrameNet have frames, we think it might be possible to map more NNC-productive N2s in our database, along with the NNCs they derive, to corresponding entity frames in FrameNet.

We have two research questions. First, with a corpus and FrameNet, we investigate whether there are only limited bridging verbs and semantic relations between the two component nouns of a NNC. Second, are there semantic relations between N1 and N2 that do not involve bridging events?

2. Complex Relations

As mentioned in the Introduction, many researchers hold that an NNC's component nouns are the arguments of an event that bridges them and by which they are assigned semantic roles. Levi (1978) regards all N1s and N2s as subjects and objects of nine linking predicates, with

one component entity doing something to the other. Below are her examples[1] and their Chinese equivalents:

CAUSE: *e.g.* malarial mosquitoes (瘧蚊 *nue-wen*)

HAVE: *e.g.* picture book (圖畫書 *tuhua-shu*), apple cake (蘋果蛋糕 *pingguo dangao*), gunboat (砲艇 *pao-ting*), industrial area (工業區 *gongye-qu*), imperial bearing (貴族氣質 *guizu-qizhi*)

MAKE: *e.g.* honeybee (蜜蜂 *mi-fong*), daisy chains (雛菊鍊 *chuju-lian*)

USE: *e.g.* steam iron (蒸氣電熨斗 *zhengqi dian-yundou*), solar generator (太陽能發電機 *taiyang-neng fadian-ji*)

BE: *e.g.* target structure (目標結構 *mubiao-jiegou*), ceiling price (天價 *tian-jia*), queen bee (女王蜂 *nu-wang fong*), satellite nation (衛星國家 *weixing-guojia*), phantom limb (幽靈肢 *youling-zhi*)

IN: *e.g.* field mouse (田鼠 *tian-shu*), autumnal rains (秋雨 *qiu-yu*)

FOR: *e.g.* horse doctor (馬醫 *ma-yi*), arms budget (武器預算 *wuqi-yusuan*), nasal mist (鼻腔噴霧 *bi-qiang pen-wu*)

FROM: *e.g.* olive oil (橄欖油 *ganlan-you*), test-tube baby (試管嬰兒 *shi-guan yinger*), apple seed (蘋果籽 *pingguo-zi*), rural visitors (鄉間訪客 *xiangjian fang-ke*)

ABOUT: *e.g.* tax law (稅法 *shui-fa*), criminal policy (刑事政策 *xingshi-zhengce*)

Levi says NNCs are all linked by one of the nine predicates, with the two components being their arguments; however, we believe that some NNCs simply involve more static relations and some relations are not covered by the above nine predicates. One instance that involves a missing static relation is, for example, the highly-productive shape relation, *e.g.* dragon boat (龍舟 *long-zhou*). In the following sections, we will use evidence of both language instinct and FrameNet data to support the distinction between simple and complex relations.

3. Motivating Simple Relations

Besides event-bridging relations, we propose simple relations, where N1 and N2 are not interacting participants of an event. Despite their shared syntactic and semantic properties, instances of simple relations have not been recognized as a distinct category, as observed by Liu (2008) and by Chung and Chen (2010).

[1] Only NNCs within the scope of this paper are listed.

We identified three types of simple relations, as opposed to complex ones:

(1) N1 and N2 denote two of the three elements of a host-attribute-value set

(a) Temporal N1

N1 denotes time:

e.g. 晨霧 *chen-wu* 'morning mist' (value+host), 秋蟹 *qiu-xie* 'autumn crab' (value+host), 午夜列車 *wuyie-lieche* 'midnight train' (value+host)

N1 denotes frequency:

e.g. 月費 *yue-fei* 'monthly fee' (value+host)

(b) Locational N1

e.g. 希臘菜 *xila-cai* 'Greek dish' (value+host), 倫敦地鐵 *Lundun-ditie* 'London Underground' (value+host), 台北人 *Taibei-ren* 'Taipei people' (value+host)

(c) Others

e.g. 鐵桌 *tie-zhuo* 'iron table/desk' (value+host), 法式 *fa-shi* 'French-style' (value+attribute), 電價 *dian-jia* 'electricity price' (host+attribute), 金塊 *jin-kuai* 'gold bricks' (host+value), 衣服堆 *yifu-dui* 'heap of clothes' (host+value), 車速 *che-su* 'car speed' (host+attribute)

(2) Meronymy (*i.e.* part-whole relation)

N1 denotes part; N2 denotes whole:

e.g. 雙底船 *shuang-di chuan* 'double-bottom,'

N1 denotes whole; N2 denotes part:

e.g. 腳踏車輪胎 *jiao-ta-che luntai* 'bicycle tire,' 腸道 *chang-dao* 'intestine canal'

(3) Conjunction

e.g. 手腳 *shou-jiao* 'hands and feet,' 鐘錶 *zhong-biao* 'clock and watch,' 警民 *jing-min* 'the police and the people'

In (1a), the N1 usually denotes the value of the semantic role "time" of an event related to the N2. In 午夜列車 *wuyie-lieche* 'midnight train' and 秋蟹 *qiu-xie* 'autumn crab,' the temporal values are 午夜 *wuyie* 'midnight' and 秋 *qiu* 'autumn,' respectively. The two NNCs either can be elaborated to mean 'trains that travel at midnight' and 'crabs that reach maturity in autumn,' or can be simply put as 'trains at midnight' and 'crabs in autumn,' omitting the events. In (1b), locational N1s usually denote place names. (1a) and (1b) are similar in that understanding of the NNCs does not depend on figuring out the bridging events

that decide the semantic roles of the component nouns.

It should be noted, however, that the nature of the event that takes place in the time or space denoted by N1 can be less than straightforward. Sometimes, this indeterminacy is caused by the meaning shift of individual components. Take 秋葵 *qiu-kui* 'okra' for example. Even native speakers may have no idea what happens to the N2 '葵' *kui* in autumn (*i.e.* 秋 *qiu* 'autumn'). This is because 葵 *kui* may not be as familiar a vegetable to modern people as it was when the compound was coined. Sometimes, meaning extension allows multiple readings of a word. For example, in antiquity, when international travel was essentially impossible, 希臘人 *xila-ren* 'Greeks' usually lived and stayed in Greece, but nowadays 希臘人 *xila-ren* 'Greeks' and 希臘菜 *xila-cai* 'Greek dishes' can reach far beyond the national borders.

Nevertheless, while the bridging event can be obscure or diverse, NNCs with temporal or locational N1s share one common characteristic: Some bridging event(s) exists, but it does not have to be clearly identified to enable sufficient understanding.

Finally, (1c) consists of host-attribute-value relations other than time and space. As argued by Chung and Chen (2010) in line with Liu (2008), objects and events are characterized by the attributes they have, and attributes are characterized in turn by values. For the examples in (1c), the morphemes 式 *shi* 'style,' 價 *jia* 'price,' and 速 *su* 'speed' are attributes and 鐵 *tie* 'iron' and 法 *fa* 'French' are attribute-values of material and style, respectively. In other words, both objects and events (collectively called "hosts") generally are associated with some attributes and attributes are associated with values. For example, artifacts, which are a subclass of objects, have the attribute "material," and "iron" is a kind (value) of material.

Given that N1 usually specifies N2, it is natural for value and host, value and attribute, and host and attribute to form NNCs in order to modify the host and attribute or to name the relevant host of an attribute.

As for (2), in 雙底船 *shuang-di chuan* 'double-bottom' and 腳踏車輪胎 *jiao-ta-che luntai* 'bicycle tire,' N1 and N2 are not interacting participants of an event. Likewise, in (3), N1 and N2 assume parallel roles in situations like 手腳看起來很乾淨 *Shou-jiao kan-qilai hen ganjing* 'Hands and feet look tidy,' 修理鐘錶 *xiuli zhong-biao* 'repair a clock (watch),' 警民合作打擊犯罪 *Jing-min hezuo daji fanzui* 'The police and the people join hands to fight crime.'

4. Mapping NNCs to FrameNet's Frames

We chose NNC-productive N2s (*i.e.* those that form NNCs with various types of N1s) from our Prefix-Suffix Database (http://140.109.19.103/affix/), sorted them according to their semantic categories and the situations their derived NNCs described, and matched these situations with FrameNet's frames.

To the extent that frames represent concepts, to map NNCs to frames is to identify the concepts NNCs convey. The corporal data to date have indicated that N2s of nine semantic categories are NNC-productive. They are: "people," "people of different vocations," "food," "clothing," "container," "vehicle," "wealth," "text," and "road." We have listed the most common relations between N1 and N2 for each category at the appendix. These categories each can be mapped to one or more entity frames, where the N2 is represented by an FE that usually has the same name as the frame itself and the N1 by another FE of the frame. Below are some examples of such mappings. (Frame names have all capital letters, while FEs have only the initial letters as capital letters.)

Simple relation (subclass: host-attribute-value)

FOOD

N1-N2=Material-Food

e.g. 玉米餅 *yumi-bing* 'corn cake,' 綠豆糕 *lu-dou gao* 'green beans cake,' 牛肉湯 *niu-rou tang* 'beef soup,' 奶茶 *nai-cha* 'milk tea,' 蘋果汁 *pingguo-zhi* 'apple juice,' 花生醬 *huasheng-jiang* 'peanut butter'

CLOTHING

N1-N2=Material-Clothing

e.g. 草鞋 *cao-xie* 'straw shoes,' 木鞋 *mu-xie* 'wooden shoes,' 皮鞋 *pi-xie* 'leather shoes,' 膠鞋 *jiao-xie* 'plastic shoes,' 豹皮帽 *bao-pi mao* 'leopard-skin hat,' 毛衣 *mao-yi* 'sweater,' 布衫 *bu-shan* 'cotton shirt'

Simple relation (subclass: meronymy)

VEHICLE_SUBPARTS

N1-N2=Part-Whole

e.g. 雙底船 *shuang-di chuan* 'double-bottom,' 鐵殼船 *tie-ke chuan* 'iron ship'

BUILDING_SUBPARTS

N1-N2=Whole-Building_part

e.g. 院牆 *yuan-qiang* 'yard wall,' 屋簷 *wu-yian* 'roof'

Complex relation

PEOPLE_BY_VOCATION

N1-N2=Person-Type

e.g. 弓箭手 *gong-jian-shou* 'archer,' 樂師 *yue-shi* 'musician,' 水電工 *shui-dian-gong* 'utilities technician'

MONEY

N1-N2=Buyer-Money

e.g. 家長費 *jiazhang-fei* 'parental fee'

N1-N2=Goods-Money

書款 *shu-kuan* 'money for buying books,' 田租 *tian-zu* 'land rent'

The above mappings show that NNCs that involve simple (as well as complex) relations correspond to FE pairs in FrameNet's entity frames. Take 玉米餅 *yumi-bing* 'corn cake' for example. The NNC can be mapped to FOOD, with the N2 餅 *bing* 'cake' denoting the FE "Food" and the N1 玉米 *yumi* 'corn' denoting "Material," which is another FE of the frame.

For NNCs of complex relations, besides an entity frame, the N1 usually can be mapped to another event frame, a point we will return to in Section 6.

5. Results

We have two findings attested to by the behavioral patterns of the nine semantic categories of N2s and their derived NNCs. First, NNCs generated by N2s of the same semantic category mostly correspond to one or a few conceptually-related frames. Second, some of the relations mapped are simple and some are complex, with N2 categories varying in their tendencies to denote simple and complex semantic relations.

5.1 Mapped to Entity Frames, Bridged by a Few Events, and Involving Limited Semantic Relations

We noticed that, when N1 and N2 are bridged by events, they usually can be mapped to both an entity frame and one or more event frames. We also found that common bridging events that link N1s to a N2 for each semantic category of N2 are limited.

For example, some of the NNCs the N2 category "money" derives include 中資 *zhong-zi* 'China capital,' 車款 *che-kuan* 'money for buying a car,' and 所費 *suo-fei* 'institute fund,' which we identified to belong to the entity frame "MONEY," where the N1s in the above three examples can be mapped to the FE "Use" and the N2 to "Money." Meanwhile, we found these N1s labeled as FEs in at least two event frames, which are "COMMERCE_BUY" and

"COMMERCE_SELL, where the three N1s 中 *zhong* 'China,' 車 *che* 'car,' and 所 *suo* 'institute,' correspond to the FEs Buyer, Goods, and Seller, respectively, are all core FEs of the event frames. Since the range of LUs and FEs for each frame usually is limited, the range of possible interpretations is more or less restricted for each NNC.

Below are all of the LUs and some of the FEs of these two event frames. (Not all non-core FEs are listed.)

COMMERCE_BUY

LUs: *buy.v, purchase_(act).n, purchase.v*

Core FEs: Buyer, Goods, Seller

Non-core FEs (not exhaustively listed): Manner, Means, Money, Purpose, Purpose_of_Goods, etc.

COMMERCE_SELL

LUs: *auction.n, auction.v, retail.v, retailer.n, sale.n, sell.v, vend.v, vendor.n*

Core FEs: Buyer, Goods, Seller

Non-core FEs (not exhaustively listed): Manner, Means, Money, Rate, Unit, etc.

5.2 N2 Categories Vary in Tendencies to Involve Simple and Complex Relations

Most of the N2 categories we have analyzed so far have produced both simple and complex-type NNCs. Below are two entity frames, CLOTHING and VEHICLE, which correspond to the N2 categories "clothing" and "vehicle". Each frame has at least one simple and one complex relation, which differ in frequency. The simple ones are labeled with their subclasses (and FEs[2]); the complex ones are labeled with the relevant FEs, which refer to the FEs that occur most or second-most often. (The frame names have all capital letters, while FEs only have initial capital letters.)

[2] FrameNet sometimes has FEs that we consider the simple type as well.

CLOTHING (衣 *yi* 'clothes,' 服 *fu* 'clothes,' 裝 *zhuang* 'clothes,' 帽 *mao* 'hat,' 鞋 *xie* 'shoes,' etc.)

Most frequent semantic relation: simple_host-attribute-value

The relevant FE(s) of N1

- As realized in CLOTHING (entity frame): Material
- (Not realized in event frames)

e.g. 草鞋 *cao-xie* 'straw shoes,' 木鞋 *mu-xie* 'wooden shoes,' 皮鞋 *pi-xie* 'leather shoes,' 膠鞋 *jiao-xie* 'plastic shoes,' 豹皮帽 *bao-pi-mao* 'leopard-skin hat,' 毛衣 *mao-yi* 'sweater'

Second-most frequent semantic relation: complex (*i.e.* eventive)

The relevant FE(s) of N1

- As realized in CLOTHING (entity frame): Wearer
- As realized in WEARING (event frame): Wearer

e.g. 女鞋 *nu-xie* 'women's shoes,' 僧鞋 *seng-xie* 'monk's shoes,' 法衣 *fa-yi* 'judge's robe,' 官服 *guan-fu* 'official robe,' 童裝 *tong-zhuang* 'children's clothes,' 學士服 *xue-shi-fu* 'Bachelor's gown'

VEHICLE (車 *che* 'vehicle,' 船 *chuan* 'ship,' etc.)

Most frequent semantic relation: complex (*i.e.* eventive)

The relevant FE(s) of N1

- As realized in VEHICLE (entity frame): Use
- As realized in BRINGING (event frame): Theme

e.g. 娃娃車 *wa-wa-che* 'kindergarten school bus,' 砂石車 *sha-shi che* 'gravel truck,' 客船 *ke-chuan* 'passenger ship,' 貨船 *huo-chuan* 'cargo ship'

Second-most frequent semantic relation: (complex, simple[3])

Complex

The relevant FE(s) of N1

- As realized in VEHICLE (entity frame): Means-of-propulsion
- (Not realized in event frames)

e.g. 電車 *dian-che* 'trolley bus,' 人力車 *ren-li-che* 'rickshaw'

Simple_meronymy

[3] For VEHICLE, the complex and simple relations have about the same second-highest frequencies.

The relevant FE(s) of N1

- As realized by VEHICLE (entity frame): Part
- (Not realized in event frames)

e.g. 雙底船 *shuang-di-chuan* 'double-bottom,' 鐵殼船 *tie-ke-chuan* 'iron ship'

5.3 The Coverage of the Identified Semantic Relations

As shown in Table 1, the average coverage of the semantic relations that FrameNet and E-HowNet have is 94.2% for the 1,153 compositional NNCs in the Prefix-Suffix Database. Below is the individual coverage of each N2 category.

Table 1. The average coverage of the semantic relations for the nine semantic categories of N2

Category	**Coverage**
Road	40/40 (100%)
Text	121/121 (100%)
People	241/243 (99.2%)
People of Different Vocations	46/48 (95.8%)
Wealth	72/72 (100%)
Container	411/427 (96.3%)
Food	60/86 (69.8%)
Clothing	42/47(89.3%)
Vehicle	53/69 (76.8%)
Mean	**1086/1153 (94.2%)**

Table 2 shows the average coverage of the three and five most frequent semantic relations. For the mapped percentage of each fine-grained relation for the nine categories, please refer to Appendix B.

We found that the top three most frequent semantic relations account for about eighty percent of the NNC instances. Meanwhile, the five most frequent relations on average have about 8% better coverage than the top three.

Table 2. The average coverage of the three and five most frequent semantic relations

Category \ Coverage	Top3	Top5
Road	67.5%	92.5%
Text	100%	100%
People	86.8%	94.2%
People of Different Vocations	83.3%	89.5%
Wealth	100%	100%
Container	94.7%	96.3%
Food	69.8%	69.8%
Clothing	72.2%	89.2%
Vehicle	49.2%	65.1%
Mean	**80.4%**	**88.5%**

Nevertheless, we noticed individual differences among N2's categories, with "food" and "vehicle" having a much lower coverage than others. Also, although we considered compositional NNCs only, there are still some relations that we lack labels for in FrameNet and E-HowNet. Some of these instances include metaphors, *e.g.* 野雞車 *yie-ji che* 'unlicensed car,' 霸王車 *bawang-che* 'unpaid ride'; apposition, *e.g.* 酒吧車 *jiuba-che* 'bar van,' 袍服 *pao-fu*, 'robe,' 靶船 *ba-chuan* 'target ship'; and those whose N1 indicates a general "use" relation unlike the other fine-grained mappings, *e.g.* 商輪 *shang-lun* 'merchant vessel,' 交通船 *jiaotung-chuan* 'commuter ship.'

6. Discussion

In this section, we will relate the two findings to our two research questions.

First, are there only limited bridging verbs and semantic relations between the two component nouns?

In the nine categories we investigated, the NNCs' bridging verbs, as well as the possible semantic roles that N1s and N2s take, are very limited, with an average coverage of over ninety percent. Even the least covered category "food" has 69.8% of its instances accounted for.

These findings support previous studies proposing that N1 and N2 often are bridged by events (Levi, 1978; Leonard, 1984; Ryder, 1994), that bridging events are limited (Levi, 1978;

Copestake & Lascarides, 1997; Copestake & Briscoe, 2005), and that the semantic relations are limited as well (Søgaard, 2005; Huang, 2008).

Second, are there semantic relations between N1 and N2 that do not involve bridging events?

To understand NNCs of simple relations does not require the identification of what one component entity does to the other. FrameNet data also suggest that bridging events sometimes are absent. We say this because we found that, among the NNCs that a N2 derives, N1s that involve complex relations usually can be mapped to FEs in eventive frames that the bridging event represents, while those that involve simple ones do not. For example, although "Material" is a productive static FE in the entity frame CLOTHING, it is not among the FEs of the eventive frame DRESSING, which describes the process and state of putting and having clothes on. In contrast, "Wearer" and "Body_location," which are also FEs of CLOTHING but involve complex relations, also assume FEs in DRESSING as arguments of LUs like "dress-up" and "put-on." Such distributional differences of FEs mean that the N1s represented by them are also distributed differently, resulting in NNCs contrasting in simple and complex terms.

While the simple-complex distinction also is attested to in a corpus-based framework like FrameNet, it seems that it is not recognized as a distinct class in Levi's widely-adopted system. While it appears that Levi (1978) considers some simple NNCs under the predicate HAVE, the status of other simple NNCs is unclear. For example, *imperial bearing* is classified as an instance of HAVE and paraphrased as 'have the bearing of an emperor.' Nevertheless, it seems that HAVE does not cover all the simple relations, as she defines the predicate as roughly corresponding to the semantic roles of "productive," "constitutive," and "compositional," which do not exhaust all simple relations. Moreover, some simple instances fall under her other predicates. For example, *apple seed* is considered an instance of FROM. We think FrameNet as a mapping means helps sort simple NNCs under semantic relations like Levi's predicates.

With regards to implementation, the findings indicate that simple and complex NNCs should be processed differently. For simple NNCs, host-attribute-value sets, place names, temporal expressions, and conjunction pairs to some degree can be exhaustively listed, as we have done in our knowledge base, Extended-HowNet, reducing identification of simple relations to table-checking. The E-HowNet taxonomy can also detect meronymy relations. For complex NNCs, the inventory of LUs and their argument FEs in FrameNet's frames narrows down the possible interpretations of NNCs.

We believe such mappings can complement the inadequacies of frameworks like Levi's (1979). First, the designation of FrameNet makes NNCs' readings more specific, as frames use fine-grained FEs and LUs are real words. Similarly, classification can be FE-based. For example, *lemon peels* and *apple seed* both belonging to the FE pair Whole-Part can be a reason for them to be grouped under the same predicate; for example, HAVE. Another classifying criterion is the simple-complex distinction. For example, to analyze the example in a different way, NNCs of the HAVE type can be defined as being made up of FE pairs like Whole-Part or Part-Whole and belonging to the simple type. Along the same vein, her IN category may involve NNCs with N1s of the FEs Time and Location, which in turn define the simple subclass of time and space. Finally, since frames are motivated by semantic and syntactic differences between words, they are expected to grow in coverage with more words' behaviors analyzed and new frames annotated.

7. Conclusion

The current study shares the insights with previous researchers that NNCs usually describe a limited range of situations and that the meaning of an NNC is compositional, while putting forth the idea that the range of semantic relations for event-bridging NNCs usually is clustered around the head, *i.e.* N2. We attained such findings by mapping the situations sorted by N2's semantic categories to frames from FrameNet, which is based on corpus-attested thematic patterns. We also noted that N1 and N2 sometimes are bridged in non-eventive ways. Both eventive and non-eventive cases can be interpreted through mapping to resources like FrameNet and E-HowNet.

Acknowledgement

This research was supported in part by the National Science Council under a Center Excellence Grant NSC 99-2221-E-001-014-MY3.

References

Chung, Y. S., & Chen, K. J. (2010). Analysis of Chinese morphemes and its application to sense and part-of-speech prediction for Chinese compounds. *ICCPOL 2010*, California, USA, 2010.

Copestake, A., & Lascardies, A. (1997). Integrating symbolic and statistical representations: The lexicon pragmatics interface. In *Proceedings of the 35th Annual Meeting of the ACL and 8th Conference of the EACL (ACL-EACL'97)*, Madrid, 1997, 136-43.

Huang, H. J. (2008). *Strategies in comprehending Mandarin Chinese noun-noun compounds with animals, plants, and artifacts as constituents*. MA thesis. National Cheng-Kung University, 2008.

Leonard, R. (1984). *The Interpretation of English Noun Sequences on the Computer*, Amsterdam: North-Holland.

Levi, J. N. (1978). *The Syntax and Semantics of Complex Nominals*. New York: Academic Press.

Liu, C. H. (2008). *Xiandai Hanyu Shuxing Fanchou Yianjiu* (*現代漢語屬性範疇研究*). Chengdu: Bashu Books.

Ryder, M. E. (1994). *Ordered Chaos: The Interpretation of English Noun-Noun Compounds*, University of California Press, Berkeley, CA.

Søgaard, A. (2005). Where does the meaning of compounds and possessives come from? A contrastive view. *The 3rd International Conference in Contrastive Semantics and Pragmatics*, Shanghai, China.

Appendix A: Examples[4] of mappings of N2-based NNC categories to FrameNet's entity and event frames (To avoid visual cluster, subclasses of simple relations are indicated as numbered in Section 3)

Simple_(1c/1a[5])

Telic/Use[6] + Clothing

[4] In part because of limited space and in part for demonstrative purpose only, we did not list examples of two of the nine productive semantic categories, "vehicle" and "container," neither did we exhaust all the instances of the other seven categories.

[5] The N1s here can be seen as either spatial (1a) or an important attribute of PEOPLE (1c).

[6] Sometimes called "Type" in FrameNet.

Appendix B: The mapped percentage of each N1 semantic role for the nine categories:

("Others" refers to instances we could not map with existing semantic role labels from FrameNet and E-HowNet.)

N2 category	**Road**		**Text**	
N1's role /N1-N2 elation, type number, and percentage	Theme	10 (25%)	Text	114 (94.2%)
	Conjunction	10 (25%)	Medium	7 (5.8%)
	Meronymy	7 (17.5%)		
	Material	6 (15%)		
	Path	4 (10%)		
	Name	3 (7.5%)		
Total mapped instances	**40/40 (100%)**		**121/121 (100%)**	

N2 category	**People**		**People of Different Vocations**	
N1's role /N1-N2 relation, type number, and percentage	Origin	167 (68.7%)	Telic/Use	35 (72.9%)
	Ethnicity	32 (13.2%)	Place_of_ Employment	3 (6.2%)
	Affiliation	12 (4.9%)	Contract_Basis	2 (4.2%)
	Hobby	10 (4.1%)	Compensation	2 (4.2%)
	Vocation	8 (3.3%)	Ethnicity	1 (2.1%)
	Material	5 (2.1%)	Rank	1 (2.1%)
	Appearance	4 (1.6%)	Compensation	2 (4.2%)
	Time	3 (1.2%)		
	Others	2 (0.8%)	Others	2 (4.2%)
Total mapped instances	**241/243 (99.2%)**		**46/48 (95.8%)**	

N2 category	Wealth		Container	
N1's role /N1-N2 relation, type number, and percentage	Goods	59 (81.9%)	Material	224 (52.5%)
	Seller	7 (9.7%)	Content	166 (38.9%)
	Buyer	6 (8.3%)	Meronymy	14 (3.3%)
			Shape	4 (0.9%)
			Location	3 (0.7%)
			Others	16 (3.7%)
Total mapped instances	**72/72 (100%)**		**411/ 427 (96.3%)**	

N2 category	Clothing		Vehicle	
N1's role /N1-N2 relation, type number, and percentage	Material	17 (36.2%)	Theme	17 (24.6%)
	Wearer	10 (21.3%)	Means-of-propulsion	10 (14.5%)
	Sub_region	7 (14.9%)	Location	7 (10.1%)
	Conjunction	5 (10.6%)	Possessor	6 (8.7%)
	Location	3 (6.4%)	Meronymy	5 (7.2%)
	Others	5 (10.6%)	Shape	3 (4.3%)
			Itinerary	2 (2.9%)
			Conjunction	2 (2.9%)
			Material	1 (1.4%)
Total mapped instances	**42/47 (89.3%)**		**53/69 (76.8%)**	

N2 category	Food	
N1's role /N1-N2 relation, type number, and percentage	Constituent_Part	53 (61.6%)
	Conjunction	4 (4.7%)
	Shape	3 (3.5%)
	Others	26 (30.2%)
Total mapped instances	**60/ 86 (69.8%)**	