

**What Can Near Synonyms Tell Us**

Lian-Cheng Chief, Chu-Ren Huang, Keh-Jiann Chen, Mei-Chih Tsai\*, Lili Chang

Academia Sinica, \*National Sun Yat Sen University

**Proceedings of the LFG98 Conference**

The University of Queensland, Brisbane

Miriam Butt and Tracy Holloway King (Editors)

1998

CSLI Publications

<http://www-csli.stanford.edu/publication>

## What Can Near Synonyms Tell Us

Lian-Cheng Chief, Chu-Ren Huang, Keh-Jiann Chen, Mei-Chih Tsai\*, Lili Chang

Academia Sinica, \*National Sun Yat Sen University

### Abstract

This study examines near synonyms and tries to extract the contrasts that dictate their semantic and associated syntactic behaviors. A near synonym pair of Chinese verbs, *fangbian* and *bianli*, which mean ‘to be convenient’, is under examination. Corpus data reveal some important but opaque distributional differences between this synonym pair that are hard to be recognized solely by intuition. This study demonstrates how the corpus data can be a useful tool that helps understanding the interaction between syntax and semantics.

### 1. Introduction

The aim of this paper is to find the semantic features that determine the relevant syntactic behaviors. Tsai et al (1996a; 1996b; 1997), in their recent studies of comparing near synonyms of Chinese verbs, claim that some basic semantic components or features can predict the different syntactic behaviors of near synonyms. One of the successful examples is the comparison of the near synonym pair *gaoxing* and *kuai* ‘happy vs. glad’. They propose two features, [ $\pm$  effect] and [ $\pm$  control], to account for the different syntactic behaviors of this pair of synonyms. In this study, we follow the same methodology to find other semantic features that can predict the syntactic patterns. The near synonym pair, *fangbian* and *bianli*, which mean ‘convenient’, is under examination to extract other relevant semantic features. We demonstrate that the lexical conceptual profile is one of these semantic features that determine the relevant syntactic behavior of near synonyms. It is hoped that each proposed semantic feature would contribute to the understanding of the interaction of syntax and semantics. This paper is organized as follows. First, we introduce our methodology in section 2. Second, we discuss the syntactic behaviors and the distributional differences of this pair of synonyms in section 3. The final section summarizes what near synonyms can tell us.

## 2. Methodology

Our approach is corpus-aided. In addition to the syntactic variations that can be easily recognized by our intuition, some implicit or opaque distributional differences in terms of syntactic functions that cannot be known simply by intuition are extracted from Sinica Corpus. Specifically, we believe that introspection is incomplete and distributional information is important in contrastive studies of near synonym. The aim is try to find out the differences between the near synonym pairs. We follow the approach adopted in Tsai et al (1997: 35). The first step is to determine distributional differences in syntactic patterns. The second step is to deduce the semantic features from the syntactic phenomena. Finally, we test the semantic features in new syntactic frames.

Under this approach, a few semantic features have been discovered. For example, [ $\pm$  effect] can properly account for the distinctions between *lei* and *pijuan* ‘tired’, and *gaoxing* and *kuaile* ‘happy or glad’. In the case of *lei* and *pijuan*, it accounts for why *lei* can be a resultative complement, while *pijuan* cannot. In the case of *gaoxing* and *kuaile*, it explains why *gaoxing* can be associated with the sentential-final particle *le*, whereas *kuaile* cannot. This is because *gaoxing*, with the feature [ $\pm$  effect], represents a change of state triggered off by some cause. In addition, [ $\pm$  telic] is used to explain the differences between *quan* and *shuifu* ‘persuade’. [ $\pm$  control] distinguishes the difference between *gaoxing* and *kuaile*.<sup>1</sup> Liu (1997) also employs the same methodology to account for the distinction between three Mandarin verbs of ‘build’, *jian*, *zao*, and *gai*. All the previous studies prove that the semantic components properly account for the syntactic differences of the near synonyms. In other words, these studies offer the evidence that syntactic behaviors can be predicted from lexical semantics. This is also the point that the present study tries to support.

## 3. The Data

The data used in this study is taken from the Sinica Corpus (version 2.0), which contains 3.5 million tagged Chinese words.<sup>2</sup> In this corpus, we found 445 entries of *fangbian* and 125 entries of *bianli*. We first present their syntactic behaviors in section 3.1 and then their distributional differences in section 3.2.

---

<sup>1</sup> For the details, please refer to Tsai et al (1997).

<sup>2</sup> Sinica Corpus 3.0, which contains 5 million words, has been released on June 1998. The trial web version can be found at <http://www.sinica.edu.tw/ftms-bin/kiwi.sh>.

### 3.1 The Near Synonym Pair: *Fangbian* and *Bianli*

The near synonym pair *fangbian* and *bianli* are used to define each other in many dictionaries. In addition to the similarities in meaning, these two verbs are seemingly syntactically parallel. For instance, both of them have the transitive and intransitive usage, can be nominal modifiers, and undergo nominalization. In this section, we introduce the syntactic behaviors.

#### 3.1.1 The Transitive/Intransitive Alternation

*Fangbian* and *bianli* have both the transitive and intransitive usage. Sentences (1) and (2) show the intransitive usage of these two verbs.

- (1) 停車 方便  
tingche **fangbian**  
parking convenient  
'To park (here) is convenient.'
- (2) 交通 便利  
jiaotong **bianli**  
traffic convenient  
'The transportation is convenient.'

In addition to the intransitive usage, they also have transitive usage as shown in sentence (3) and (4).

- (3) 設置 辦事處 方便 民眾 出國 觀光  
shezhi banshichu **fangbian** minzhong chuguo guanguang  
establish office convenient people go-abroad visit  
'Establishing an office makes it convenient for people to go abroad and travel.'
- (4) 修改 許多 法規 便利 山民 墾植  
xiugai shuduo fagui **bianli** shanmin kenzhi  
modify many rule convenient mountain-people cultivate  
'Modifying many rules makes it convenient for the aboriginal to cultivate.'

In the intransitive usage, both *fangbian* and *bianli* take a proposition as subject. In the transitive usage, they take a propositional object. Usually, the proposition subject or object propositions are represented by a clause, a verb phrase, or a complex nominal element. The proposition is what described as convenient. However, there is a striking difference between them. The proposition object of *fangbian* can undergo inversion as

in (5a) and (5b), while *bianli* does not allow such alternation.

(5a) 理想的場地是鄰近工作地點，方便員工參加  
lixiang de changdi shi linjin gongzuo didian, fangbian yuangong canjia  
ideal DE place be near work place convenient worker join  
'An ideal location is where it is near the working place and convenient for workers to join (the meeting).

(5b) 理想的場地是鄰近工作地點，員工參加方便  
lixiang de changdi shi linjin gongzuo didian, yuangong canjia fangbian  
ideal DE place be near work place workers join convenient  
'An ideal location is where it is near the working place and convenient for workers to join (the meeting).

(6a) 有各種產品，便利消費者選購  
you gezhong changpin bianli xiaofeizhe xung-gou  
have various product convenient consumer choose-buy  
The varieties of products makes it convenient for consumers to choose from.

(6b) \*有各種產品，消費者選購便利  
you gezhong changpin xiaofeizhe xung-gou bianli  
have various product convenient consumer choose-buy

We will try to account for this phenomenon in section 4.

### 3.1.2. Other Syntactic Functions of *fangbian* and *bianli*

In addition to verbal predicates, these two near synonyms can also appear as nominal modifiers and undergo nominalization. (7) and (8) illustrate *fangbian* and *bianli* as nominal modifiers.

(7) 方便的資訊  
fangbian de zixung  
convenient de information  
easily-accessible information

(8a) 便利的方式  
bianli de fangshi  
convenient de way  
convenient way

(8b) 便利商店  
bianli shangdian  
convenient store  
convenient store

(9) and (10) show that when this pair of near synonyms appears as a nominal elements.

(9) 聯繫 上 的 方便  
 lianxi shang de fangbian  
 communicate in de convenience  
 convenience of communicating

(10) 生活 的 便利  
 shenghuo de bianli  
 living de convenience  
 convenience in living

As shown in this section, it seems that *fangbian* and *bianli* can be used interchangeably. However, the statistics shown by the corpus demonstrate that they distribute very differently.

### 3.2 Distributional Differences

In this section, we examine the distributional differences of this pair of near synonyms extracted from the Sinica Corpus. Our approach is to search all the instances of *fangbian* and *bianli* in the corpus and then classify each occurrence according to their syntactic function such as verbal predicates, nominal modifiers, verbal modifiers, and nominals. Second, we calculate the appearances of the transitive and intransitive alternation of their verbal predicate usage. Third, we classify them in terms of the object type they take. The results demonstrate that the contrasts between them are clearly displayed in the distributional differences.

#### 3.2.1 Distributional Differences in Terms of Syntactic Functions

Table 1 illustrates their distribution in terms of syntactic functions.

Table 1 Distributional Differences in terms of Syntactic Function

	Verbal Predicates	Nominal Modifiers	Verbal Modifiers	Nominalization
Fangbian 445	77%	7%	5%	10%
Bianli 125	44%	34%	0%	22%

From Table 1, some contrasts between *fangbian* and *bianli* can be clearly found. First, *bianli* cannot be used as verbal modifier, whereas *fangbian* can. Second, when used as a nominal modifier, *bianli* is much more preferred than *fangbian*. These two pieces of

evidence give to two questions. First, why can't *bianli* be used as verbal modifier? Second, why is *bianli* often selected when people try to express the concept that something is convenient?

### 3.2.2 Distributional Differences in terms of Transitive/Intransitive Alternation

The distributional differences in table 2 show that *fangbian* more often appears in intransitive form; while *bianli* shows no such preference. In addition, when used as transitive verbs, *fangbian* predominantly takes a sentential object.

Table 2 Distributional Differences in terms of Transitive/Intransitive Alternation

	Transitive	Intransitive
Fangbian 342	31%	69%
Bianli 55	53%	47%

Table 3 Distributional Difference in terms of the Type of Object

	Sentential or Verbal Object	Complex Nominal Object
Fangbian 107	90%	10%
Bianli 29	62%	37%

### 3.2.3 Negation

In the corpus, we also found that *bianli* cannot be modified by the negative marker *bu* 'not'. In other words, it tends not to be negated as shown in table 4.

Table 4. Co-occurrence with negative marker *bu* 'not'

	Negation (preceded by <i>bu</i> 'not')	all instances
<i>fangbian</i>	44	445
<i>bianli</i>	0	125

This also give us the other question that why *bianli* cannot be negated syntactically.

### 3.3 Summary

These distributional differences extracted from corpus not only give us a clear picture of their contrast in usage but also show the inadequacy of the present definition in dictionaries. Though they are used to define each other in many dictionaries, the description of their variations is ignored by those lexicographers. That is, there is the preference of the main function of *fangbian* and *bianli* in different contexts and usage.

## 4. Explanation

To account for the observed contrasts displayed by the distributional differences, we propose that two semantic factors, (i) beneficial role and (ii) lexical conceptual profile, determine the different syntactic patterns of this pair of near synonyms.

### 4.1 Beneficial Role

From the evidence presented in section 3, we summarize that there are at least four major differences between *fangbian* and *bianli*. First, *bianli* never appears as a verbal modifier. Second, *bianli* occurs as transitive verbs in most cases. Third, when they are used as transitive verbs, 90% of the *fangbian* instances takes sentential and verbal object. Fourth, *bianli* cannot be negated. To account for these variations, we propose that the profile of the event structure of *fangbian* is on the description of the whole proposition event, while that of *bianli* is on the description of the beneficial role of the event. In other words, *fangbian* profiles the whole proposition event, whereas *bianli* profiles the beneficial role of the event. The following pair of sentences illustrates this.

(11a) 設置 辦事處 方便 民眾 出國 觀光  
shezhi banshichu **fangbian** minzhong chuguo guanguang  
establish office convenient people go-abroad visit  
'Establishing an office makes it convenient for people to go abroad and travel.'

(12a) 修改 許多 法規 便利 山民 墾植  
xiugai shuduo fagui **bianli** shanmin kenzhi  
modify many rule convenient mountain-people cultivate  
'Modifying many rules makes it convenient for the aboriginal to cultivate.'

In sentence (11a) whose main verb is *fangbian*, the profile is on the whole embedded event "people go abroad and visit". The syntactic evidence as shown in sentence (11b) and (12b) support this argument.

(11b) 設置 辦事處 民眾 出國 觀光 方便  
shezhi banshichu minzhong chuguo guanguang **fangbian**  
establish office people go-abroad visit convenient  
'Establishing an office makes it convenient for people to go abroad and travel.'

(12b)\* 修改 許多 法規 山民 墾植 便利  
xiugai shuduo fagui shanmin kenzhi **bianli**  
modify many rule mountain-people cultivate convenient  
'Modifying many rules makes it convenient for the aboriginal to cultivate.'

As shown in sentence (11b) the post-verbal element, the proposition event, can be inverted to pre-verbal position; whereas in sentence (12b) such inversion is not allowed. In contrast to sentence (11a), in sentence (12a) whose main verb is *bianli* the profile is on the beneficial role (the aboriginal) of the embedded event (cultivate). In other words, the profile of sentence (12a) is on *the aboriginal* who cultivate rather than the event “cultivate”. Following this explanation, we deduce a semantic feature to show the contrast between this pair of near synonym [ $\pm$  beneficial role]. Specifically, the beneficial role in the event structure of *bianli* is most prominent and important. In contrast, there is no beneficial role in the event structure of *fangbian*, or its status is trivial. In short, the meaning of this pair of near synonyms is ‘to be convenient’, but the concept of convenient is on different level. For *fangbian*, it means that the way to perform the action is convenient; whereas for *bianli*, it means that it is convenient for someone to perform the action.

#### 4.2 Profile on Event vs. Profile on Beneficial Role

The notion that lexical conceptual profile is on different sub-part of an event properly accounts for the contrasts between *fangbian* and *bianli*. First, we mentioned that *bianli* cannot function as a verbal modifier. In other words, when people want to describe that certain event is easily conducted, they will choose *fangbian* to express this concept. Why is this so? Since the lexical conceptual profile of *fangbian* is on the proposition event, *fangbian* can easily modify a verb, thus functions as a verbal adjunct. In other words, profile of the whole proposition event is the inherent meaning of *fangbian*. In contrast, the lexical conceptual profile of *bianli* is on the beneficial role of the proposition event; therefore, *bianli* cannot be used to modify a verb. Because *bianli* does not profile the event.

The data in corpus show that *bianli* cannot be negated whereas *fangbian* can be negated by the negative marker *bu* ‘not’. Our proposed semantic features also properly explain this. First, the profile of *fangbian* is on the whole sub-event, the scope of negation can cover the whole sub-event. Therefore, *fangbian* can be easily negated. That is, the embedded predicate is negated. In contrast, the profile of *bianli* is on the beneficial role rather than the whole sub-event, so *bianli* cannot be negated. In order to profile on the beneficial/causee role, the whole proposition must be presupposed. Hence, the presupposition cannot be negated/cancelled. The controlled subject cannot be negated. The second explanation is also about the beneficial role. Since there is a

beneficial role, *bianli* has the meaning “causing the event to be very convenient for the beneficial role”. Specifically, the semantic of *bianli* denote positive meaning. It makes the beneficial role in good condition. It would be semantic anomalous, if the predicate is negated.

### 4.3 Syntactic Patterns

Based on the two semantic features, the scope of focus and beneficial role, we propose that *fangbian* and *bianli* have different syntactic frames.

(13) *fangbian* <SUBJ **XCOMP**>

(14) *bianli* <SUBJ **OBJ** XCOMP>

(13) and (14) show that *fangbian* has two roles, whereas *bianli* has three roles (with an additional beneficial role). The shadowed bold indicates the scope of profile. That is, the profile of the event of *fangbian* is the whole embedded event, whereas that of *bianli* is the object function, the beneficial role. As mentioned previously, this account has two advantages. First, *bianli* cannot be an adjunct of a verb because it does not profile an event. On the contrary, *fangbian* can easily modify a verbal predicate because its semantic inherently profiles an event. Second, *fangbian* rather than *bianli* can be negated because the scope of the negation can cover the whole sub-categorized XCOMP of *fangbian* but cannot cover the XCOMP of *bianli*.

Finally, this also accounts for the syntactic alternation of *fangbian* and the lack of such alternation of *bianli* as shown in sentences (5a) and (5b), repeated below.

(5a) 理想的場地是鄰近工作地點，方便員工參加  
 lixiang de changdi shi linjin gongzuo didian, fangbian yuangong canjia  
 ideal DE place be near work place convenient worker join  
 ‘An ideal location is where it is near the working place and convenient for workers to join (the meeting).

(5b) 理想的場地是鄰近工作地點，員工參加方便  
 lixiang de changdi shi linjin gongzuo didian, yuangong canjia fangbian  
 ideal DE place be near work place workers join convenient  
 ‘An ideal location is where it is near the working place and convenient for workers to join (the meeting).

(6a) 有各種產品，便利消費者選購  
 you gezhong changpin bianli xiaofeizhe xung-gou  
 have various product convenient consumer choose-buy  
 The varieties of products makes it convenient for consumers to choose from.

(6b) *有	各種	產品，	消費者	選購	便利
you	gezhong	changpin	xiaofeizhe	xung-gou	bianli
have	various	product	convenient	consumer	choose-buy

Sentences (5)-(6) demonstrate that post-verbal element of *fangbian* can undergo inversion whereas that of *bianli* cannot. Since the post-verbal elements of *bianli* have two roles, one of the roles cannot be inverted alone. On the contrary, *fangbian* has only one post-verbal element.<sup>3</sup> In brief, syntactic profile cannot contradict lexical conceptual profile.

#### 4.4 An Alternative Way to Look at the Distinction

The distinction between this pair of synonyms might have to do with the distinction between type and token of certain event. Since *fangbian* profiles the whole proposition event and *bianli* profiles the beneficial role of the event, *fangbian* might tend to be used to describe the generic event while *bianli* might tend to be used to describe the specific event. The profile of the event of *bianli* is on how the event affects the individual who performs the action. In the event of *fangbian*, the status of the individual is trivial. It is important that the manner/way to perform the action/event is convenient. Therefore, *fangbian* is about the comment of generic event. On the contrary, *bianli* focuses on the individual. It profiles how the individual performs the action in each event, so *bianli* tends to be used to describe specific event. In conclusion, we suggest the type and token is also the potential distinction between *fangbian* and *bianli*. *Fangbian* is about a group of events; that is, the type of the event. *Bianli* is about each single event; that is, the token of the event.

#### 4.5 Summary

From the distributional difference, we find out the contrasts between *fangbian* and *bianli* that are hard to be discovered solely by our intuition. We assert that two semantic factors determine the relevant syntactic behaviors of this pair of near synonyms. The lexical conceptual profile accounts for why *bianli* cannot function as an adjunct of verb and why *bianli* cannot be negated. The additional beneficial role of *bianli* explains the lack of syntactic alternation that *fangbian* allows. That is, the valency might predict the

---

<sup>3</sup> For the scope of this paper, we do not discuss which pattern (transitive/intransitive) of *fangbian* is the basic pattern nor do we discuss whether *fangbian* has two lexical entries or on lexical entry.

syntactic alternation. Finally, the concept of type and token might have to do with the preference of choosing one of the synonym pair in certain usage.

## 5. What Can Near Synonyms Tell Us

The hypothesis that the syntactic behaviors of verbs are semantically determined has been supported by a series of studies of comparison of near synonym pairs. The present study can be viewed as one of the bricks to the architecture of lexical semantics in Mandarin Chinese, especially the framework proposed by Huang and Tsai (1997). The semantic features proposed to distinguish the relevant syntactic behaviors of near synonyms are lexical conceptual profile and beneficial roles. Lexical conceptual profile determines the syntactic function that a word can have and also the scope of negation in the embedded predicates. The presence and absence of beneficial role, that is, the number of roles, predicts the relevant syntactic alternation. So far, this series of studies (Tsai et al. 1996a; 1996b; 1997) have proposed many semantic features that properly explain the syntactic differences and predict the syntactic behaviors. For example, the distributional contrasts show that *lei* ‘tired’ can function as a resultative complement but absolute cannot undergo nominalization; whereas *pijuan* ‘tired’ never function as a resultative complement but can undergo nominalization (Tsai et al 1997). The semantic feature [ $\pm$ effect] explains their complimentary distribution in terms of these two functions. *Lei* has effect on the event, so it can occur as a resultative complement; while *pijuan* has no effect on the event, so it cannot function as a resultative complement. In addition, the semantic feature [ $\pm$ effect] also accounts for why *gaoxing* ‘happy’ can take a sentential object while *kuai* ‘happy’ cannot. The semantic feature [ $\pm$ telic] distinguishes the meaning of the synonym pair *shuifu* and *quan* ‘persuade’. *Quan* denotes an extensible or atelic event; whereas *shuifu* denotes a bounded or telic event. The semantic feature [ $\pm$ control] accounts for other contrast between *gaoxing* and *kuai* that *gaoxing* never occurs in wish sentences but in evaluational sentences, while *kuai* occurs in wish sentences but never appears in evaluational sentences. All these semantic features are the supporting evidence that sometimes syntax can be predicted from semantics.

If semantics can properly and nicely predict syntactic behaviors, then the pair of words that have the exactly same meaning should have exactly the same syntactic behaviors. Therefore, the syntactic differences of near synonyms indicate the existence of subtle semantic difference. However, these syntactic differences are hard to be discovered solely by our intuition. The present paper employs the corpus data to find the differences and then looks for the semantic explanation for the relevant syntactic

behaviors. These semantic differences are proved to determine the syntactic differences. In conclusion, the approach based on the comparison of synonyms and aided by corpus provides a new direction to understand the interaction between syntax and semantics in Mandarin Chinese.

## References

- Huang, Chu-Ren and M.-C. Tsai. 1997. From Near Synonyms to Event Structure: Corpus-based Studies of Mandarin Verbal Semantics. Paper Presented at Mini-Conference on Lexical Semantics. National Chung Cheng Univ., Nov. 25, 1997.
- Liu, Meichun. 1997. Lexical Meaning and Discourse Patterning – the three Mandarin cases of ‘build’. Paper Presented at the Third Conference on Conceptual Structure, Discourse, and Language. Boulder, Colorado.
- Liu, Meichun, C.-R. Huang, and C. Lee. 1998. When Endpoint Meets Endpoints: A Corpus-based Lexical Semantic Study of Mandarin Verbs of Throwing. Paper Presented at the 7th International Conference on Chinese Linguistics/The 10th North American Conference on Chinese Linguistics. June 16-18, 1998. Stanford Univ.
- Tsai, Mei-Chih, C.-R. Huang, and K.-J. Chen. 1996a. Yuliaoku weibian de yuyi xunxi choqu yu bianxi – yi jinyici yianyio weili. Proceeding of Rocling IX: 281-293.
- . 1996b. You jinyici bianyi biao zhun kan yuyi jufa zhi hudong (From near-synonyms to the interaction between syntax and semantics). Proceedings of 5th International Symposium on Chinese Languages and Chinese Linguistics (IsCLL 5): 167-180.
- Tsai, Mei-Chih, C.-R. Huang, K.-J. Chen, and K. Ahrens. 1997. Towards a Representation of Verbal Semantics – An Approach Based on Near Synonyms. Proceedings of Rocling X: 34-48. Taipei: Association for Computational Linguistics and Chinese Language Processing.