

What Can Near Synonyms Tell Us?¹

Lian-Cheng Chief*, Chu-Ren Huang*, Keh-Jiann Chen*, Mei-Chih Tsa[†] Li-li Chang*

Abstract

This study examines a near synonym pair *fangbian* and *bianli*, 'to be convenient/' and extracts the contrasts that dictate their semantic and associated syntactic behaviors. Corpus data reveal important but opaque distributional differences between these synonyms that are not readily apparent based on native speaker intuition. In particular, we argue that this synonym pair can be accounted for with a lexical conceptual profile. This study demonstrates how corpus data can serve as a useful tool for probing 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 of the near synonym pair *fangbian* and *bianli*. Tsai *et al.* [1998 & 1999], in their recent comparative studies of near synonymous Chinese verbs, claim that basic semantic components or features can predict the different syntactic behaviors of near synonyms. One example is their comparison of the near synonym pair *gaoxing* and *kuaile* 'happy vs. glad.' Tsai *et al.* [1998] proposed two features, [\pm effect] and [\pm control], to account for the different syntactic behaviors of these synonyms. In this study, we use the same methodology to find other semantic features that can predict syntactic patterns. The syntactic patterns of the near synonym pair *fangbian* and *bianli*, which mean 'to be convenient,' are examined to extract relevant semantic features. We demonstrate that the lexical conceptual profile is one semantic feature that determines the relevant syntactic behavior of the near synonym pair. It is hoped that each proposed semantic feature will contribute to our understanding of the interaction between syntax and semantics. This paper is organized as follows. First, we introduce our methodology in section 2.

Then, we discuss the syntactic behaviors of and the distributional differences between these synonyms in section 3. The final section summarizes the information that near synonyms can give us.

* Academia Sinica

[†] National Sun Yat Sen University

¹ An earlier version of this paper appeared in the electronically published Proceedings of the LFG98 Conference (Miriam Butt and Tracy Holloway King Eds, <http://www-csli.stanford.edu/LFG/3/lfg98.html>). We would like to thank Kathleen Ahrens for her detailed comments on several versions of this paper. We are also grateful to participants of the conference as well as colleagues at CKIP, Academia Sinica for their helpful comments.

2. Methodology

Our approach is corpus-aided. In addition to the syntactic variations that can be easily recognized by means of our intuition, implicit or opaque distributional differences in terms of syntactic functions that cannot be discerned simply by means of intuition were extracted from the Sinica Corpus. Specifically, we believe that introspection is incomplete, and that distributional information is important in contrastive studies on near synonyms. Our aim is to try to determine the syntactic and semantic differences between members of near synonym pairs. We follow the approach adopted by Tsai et al. [1999]. 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.

Through this approach, several semantic features have been discovered. For example, $[\pm\text{effect}]$ can account for the distinctions between *lei* and *pijuan* 'tired,' and *gaoxing* and *kuaile* 'happy or glad.' In the case of *lei* and *pijuan*, $[\pm\text{effect}]$ accounts for why *lei* can be a resultative complement while *pijuan* cannot. In the case of *gaoxing* and *kuaile*, $[\pm\text{effect}]$ explains why *gaoxing* can be associated with the sentential-final particle *le*, whereas *kuaile* cannot. This is because *gaoxing*, with the feature $[\pm\text{effect}]$, represents a change of state triggered by some cause. In addition, $[\pm\text{telic}]$ is used to explain the differences between *quan* and *shuifu* 'persuade.' $[\pm\text{control}]$ distinguishes between *gaoxing* and *kuaile*². Liu [1997] also employs the same methodology to account for the distinctions among three Mandarin verbs of 'build,' *jian*, *zao*, and *gai*. These previous studies demonstrate that semantic components account for the syntactic differences between the members of near synonym pairs. In other words, these studies offer evidence that syntactic behaviors can be predicted based on lexical semantics. This is also the point that the present study aims to support.

3. The Data

The data used in this study were taken from the Sinica Corpus (version 2.0), which contains 3.5 million tagged Chinese words³. In this corpus, we found 445 entries of *fangbian* and 125 entries of *bianli*. We will first present their syntactic behaviors in section 3.1 and then their distributional differences in section 3.2.

² For details, please refer to Tsai et al. [1999]

³ Sinica Corpus 3.0, which contains 5 million words, was released in June of 1998. It can be found at <http://www.sinica.edu.tw/ftms-bm/kiwi.sh>.

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

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

3.1.1 The Transitive/Intransitive Alternation

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

(1) 停 車 方便
tingche **fangbian**
parking convenient
'Parking (here) is convenient.'

(2) 交通 便利
jiaotong **bianli**
traffic convenient
'Transportation is convenient.'

In addition to their intransitive usages, they also have transitive usages 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 travel abroad.'

(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 aborigines to cultivate [land].'

In their intransitive usages, both *fangbian* and *bianli* take a proposition as a subject. In their transitive usages, they take a propositional object. Usually, the propositional subject or propositional object is represented by a clause, a verb phrase, or a complex nominal element. The proposition describes what is convenient. However, the propositional 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 near the work 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 near the work 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 xuan-gou
have various product convenient consumer choose-buy

'The variety of product makes selection convenient for consumers.'

(6b)*有各種產品，消費者選購便利

you gezhong changpin xiaofeizhe xuan-gou bianli
have various product convenient consumer choose-buy

We will 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 the use of *fangbian* and *bianli* as nominal modifiers⁴.

(7) 方便的資訊

fangbian de zixun
convenient de information
easily-accessible information

⁴ However, we only found examples of *bianli* (but not *fangbian*) used in nominal compounds in the Sinica Corpus as shown below. We do not account for this difference in this paper.

便利商店
bianli shangdian
convenient store
convenience store

(8) 便利 的 方式
bianli de fangshi
convenient de way
convenient way

Examples (9) and (10) show that this pair of near synonyms can function as nominal heads.

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

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

As shown above in this section, it appears that *fangbian* and *bianli* can be used interchangeably. However, the statistics obtained from the corpus demonstrate that these syntactic patterns have different statistical distributions.

3.2 Distributional Differences

In this section, we will examine the distributional differences based on all the examples extracted from the Sinica Corpus. After searching for all the instances of *fangbian* and *bianli* in the corpus, we first classified each occurrence according to its syntactic function, such as nominal verbal predicate, nominal modifier, and verbal modifier. Second, we calculated the number of occurrences of transitive and intransitive alternations of these synonyms as verbal predicates. Third, we classified them in terms of the object types they take. The results demonstrate clear distributional differences.

3.2.1 Distributional Differences in Terms of Syntactic Functions

In this section, we will present the distributional differences in terms of syntactic functions. The range of syntactic functions of this near synonym pair can be illustrated by the previously given examples (1)-(10) as well as (11) below.

(11) 使用者 可以 更 方便 的 處理 事情

shiyungzhe keyi geng fangbian de chuli shiqing
 user can more convenient de manage thing
 'Users can manage things more conveniently.'

Four different functions are identified. First, verbal predicates are exemplified by (1)-(6). Second, nominal modifiers are given in (7) and (8). Third, (9) and (10) are instances of nominalization. Lastly, (11) is an example of a verbal modifier in which *de* is preceded by *fangbian* and followed by a head verb. We cannot find any example in which *bianli* is used in this way in our corpus, which also confirms our intuition.

Table 1 illustrates their distributions 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	44%	34%	0%	22%

In Table 1, some differences between *fangbian* and *bianli* can be found. First, *bianli* cannot be used as a verbal modifier, whereas *fangbian* can. Second, when used as a nominal modifier, *bianli* is preferred more than *fangbian*. These two pieces of evidence give rise to two questions. First, why can't *bianli* be used as a verbal modifier? Second, why is *bianli* often selected when people try to express the idea that some event is convenient?

3.2.2 Distributional Differences in terms of the Transitive / Intransitive Alternation

The distributional differences indicated in Table 2 show that *fangbian* more often appears in intransitive form (e.g. (1)) while *bianli* shows no such preference. In addition, when used as a transitive verb, *fangbian* usually takes a sentential object (e.g. (3)).

Table 2. *Distributional Differences in terms of the 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

From the corpus, we also find that *bianli* cannot be modified by the negative marker *bu* 'not,' as shown in Table 4.

Table 4. *Co-occurrence with Negative Marker bu 'not'*

	Negation (preceded by <i>bu</i> 'not')	Total instances
Fangbian	44	445
Bianli	0	125

This also gives rise to the second question as to why *bianli* cannot be negated syntactically.

3.3 Summary

The distributional differences extracted from the corpus not only give us a clear picture of their differences in usage, but also show the inadequacy of their present definitions in dictionaries. Though they are used to define each other in many dictionaries, their differences in terms of function and distribution are neither described nor explained.

4. Explanation

To account for the observed differences in syntactic distribution, we propose two semantic factors, (i) beneficial role and (ii) lexical conceptual profile. In other words, we propose that there is a beneficial role in the argument structure of *bianli*. Further, we point out that profiling different perspectives of an event nicely captures the differences between the two verbs. In this paper, the lexical conceptual **profile** refers to the most prominent or salient sub-part of the whole event. Specifically, in a group of verbs that are similar in meaning, there are different focal points in different participants or different levels of verb frames. A similar but not identical idea can be found in Goldberg [1995] and Croft [1998], in which profiling is also used to describe semantic differences among verbs.

4.1 Beneficial Role

From the evidence presented in section 3, there are at least four major differences

between *fangbian* and *bianli*. First, *bianli* never appears as a verbal modifier. Second, *bianli* occurs as a transitive verb in most cases. Third, in 90 of the instances in which *fangbian* is used as a transitive verb, it takes either a sentential or a verbal object. Fourth, *bianli* cannot be negated. To account for these variations, we propose that *fangbian* profiles the whole event, whereas *bianli* profiles the beneficial role of the event. The following pair of sentences ((12a) and (12b) repeated from (3) and (4)) illustrates this.

(12a) 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 travel abroad.'

(12b) 修改 許多 法規 便利 山民 墾植
 xiugai shuduo fagui **bianli** shanmin ken zhi
 modify many rule convenient mountain-people cultivate
 'Modifying many rules makes it convenient for the aborigines to cultivate [land].'

In sentence (12a), the main verb is *fangbian*, and the verbal meaning profiles the whole embedded event "people go abroad and visit." The syntactic evidence as shown by the constructed sentences (13a) and (13b) support this argument because in (13a), the post-verbal element, the prepositional event, can be inverted to the pre-verbal position, whereas in sentence (13b), such an inversion is not allowed.

(13a) 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 travel abroad.'

(13b)*修改 許多 法規 山民 墾植 便利
 xiugai shuduo fagui shanmin ken zhi **bianli**
 modify many rule mountain-people cultivate convenient

Furthermore, in contrast to (12a), in (12b) the main verb is *bianli*, and the verbal meaning profiles the beneficial role (the aborigines) of the embedded event (to cultivate). In other words, the focus of sentence (12b) is on the aborigines who cultivate rather than the event "to cultivate" itself. Therefore, we propose a semantic feature which shows the difference between these near synonyms to be [\pm beneficial

role]. Specifically, the beneficial role in the event structure of *bianli* is prominent. 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 convenience is on different levels. For *fangbian*, it means that the way to perform the action is convenient, whereas for *bianli*, it means that for the profiled entity, the action is convenient or beneficial to perform.

4.2 Profile on Event vs. Profile on Beneficial Role

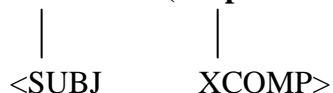
The notion that the lexical conceptual profile focuses on different sub-parts of an event also accounts for the differences between *fangbian* and *bianli*. First, we have observed that *bianli* cannot function as a verbal modifier. In other words, when people want to describe that a certain event is easily conducted, they will choose *fangbian* to express this concept. Why is this so? Since the lexical conceptual profile of *fangbian* focuses on the prepositional event, when *fangbian* modifies a verb, the eventive profile is projected to the sentential level, and semantic composition is preserved. In other words, a profile of the whole prepositional event is the inherent meaning of *fangbian*. In contrast, the lexical conceptual profile of *bianli* focuses on the beneficial role of the prepositional event; therefore, semantic compositionality cannot be maintained if *bianli* is used to modify a verb.

Second, the data from the 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, since the profile of *fangbian* focuses on the whole positional event, it can be negated like any proposition. Therefore, *fangbian* can co-occur with *bu*. In contrast, the profile of *bianli* focuses on the beneficial role rather than the whole sub-event. In order for the profile to focus on the beneficial / causee role, the whole proposition must be presupposed. Also, it is well-known that a presupposition cannot be negated/ cancelled. In addition, the semantics of the beneficial role also exclude negation since the semantics of *bianli* denote a positive meaning. It would be semantically anomalous if the predicate were negated.

4.3 Syntactic Patterns

Based on the two semantic features, the beneficial role and the lexical conceptual profile, we propose that *fangbian* and *bianli* have different event structures and argument structure frames.

(14) *fangbian* [AGENT **GOAL (Proposition)**]



(15) bianli [AGENT **BEN** GOAL (Proposition)]
 | | |
 <SUBJ OBJ XCOMP>

(14) and (15) show that *fangbian* has two roles (AGENT and GOAL), whereas *bianli* has three roles (AGENT, BEN, and GOAL). The shadowed bold text indicates the scope of the profile. That is, the profile of the event of *fangbian* focuses on the whole embedded event, whereas that of *bianli* focuses on 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 modify a verbal predicate because its semantics inherently profile 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, the difference in lexical conceptual profile also accounts for the syntactic alternation of *fangbian* and the lack of such alternation of *bianli* as shown in (5) and (6), and repeated here for convenient reference.

(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 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 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 xuan-gou
 have various product convenient consumer choose-buy
 The variety of products makes selection convenient for consumers.'

(6b)*有 各種 產品， 消費者 選購 便利

you gezhong changpin xiaofeizhe xuan-gou bianli
have various product convenient consumer choose-buy

Sentences (5)-(6) demonstrate that post-verbal elements of *fangbian* can undergo inversion whereas those of *bianli* cannot. Since *bianli* has two *postverbal* elements, one of the grammatical functions cannot be inverted by itself. On the contrary, *fangbian* has only one post-verbal element⁵. In brief, the syntactic profile cannot contradict the lexical conceptual profile.

4.4 An Additional Perspective

An additional possibility is that the distinction between this pair of synonyms might have to do with the distinction between the type and token of a certain event. Since *fangbian* profiles the whole proposition event and *bianli* profiles the beneficial role of the event, *fangbian* tends to be used to describe a generic event while *bianli* tends to be used to describe the specific event. The profile of the event of *bianli* focuses 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* comments on the 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 a specific event. In conclusion, we suggest that the type and token are also the potential distinctions between *fangbian* and *bianli*. *Fangbian* refers to a group of events, that is, the type of event. *Bianli* refers to a single event, that is, the token of the event.

4.5 Summary

From distributional syntactic differences, we have discovered differences between *fangbian* and *bianli* that are not easily determined solely by means of intuition. We assert that two semantic factors determine the relevant syntactic behaviors of these 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. Finally, the distinction between event type and event token also contributes to the distributions of these synonyms.

5. What Can Near Synonyms Tell Us

⁵ 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 one lexical entry.

The hypothesis that the syntactic behaviors of verbs are semantically determined has been supported by a series of studies which have compared near synonyms. The present study can be viewed as one of the building blocks contributing to the study of Mandarin Chinese lexical semantics, based on the framework proposed by Huang and Tsai [1997]. The semantic features proposed in this paper to distinguish between the relevant syntactic behaviors of the near synonyms *bianli* and *fangbian* are lexical conceptual profile and beneficial role. The lexical conceptual profile determines both the syntactic function that a word can have and also the scope of negation in embedded predicates. The presence or absence of a beneficial role predicts the relevant syntactic alternation.

So far, this series of studies [Tsai et al. 1998 & 1999 as well as Huang et al. 2000, Chang et al 2000] has proposed several semantic features that explain syntactic differences and predict syntactic behaviors. If semantics can properly predict syntactic behaviors, then pairs of words that have exactly the same meaning should have exactly the same syntactic behaviors. Therefore, the syntactic differences between near synonyms indicate the existence of subtle semantic differences. However, these syntactic differences are not easily discovered solely by means of intuition. In the present study, we used corpus data to find differences, and we then looked for semantic explanations for the relevant syntactic behaviors. In conclusion, this approach, which is based on comparing synonyms and is aided by corpus studies, provides a new way to understand the interaction between syntax and semantics in Mandarin Chinese.

References

- Chang, Li-li, Keh-jian Chen, Chu-Ren Huang. 2000. "Alternation Across Semantic Fields: A Study on Mandarin Verbs of Emotion." *International Journal of Computational Linguistics & Chinese Language Processing*. Vol.5 no.1 pp.61-80
- Croft, William. 1998. "Event Structure in Argument Linking." In *The Projection of Arguments Lexical and Compositional Factors*. Ed. by Miriam Butt and Wilhelm Geuder. pp. 21-65 Stanford: CSLI Publications.
- Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: The University of Chicago Press.
- Huang, Chu-Ren, Kathleen Ahrens, Li-li Chang, Keh-jian Chen, Mei-Chun Liu and Mei-Chi Tsai. 2000. "The Module-Attribute Representation of Verbal Semantics: From Semantics to Argument Structure." *International Journal of Computational Linguistics & Chinese Language Processing*. Vol.5 no.1 pp. 19-46

- Huang, Chu-Ren and Mei-Chih 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 University, Nov. 25.
- Liu, Meichun. 1997. "Lexical Meaning and Discourse Patterning - The three Mandarin cases of "build1." Paper presented at the *Third Conference on Conceptual Structure, Discourse, and Language*. Boulder, Colorado.
- Liu, Meichun, Chu-Ren Huang, and Ching-Yi 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*. Stanford University, June 16-18.
- Tsai, Mei-Chih, Chu-Ren Huang, and Keh-jiann Chen. 1999. "You jinyici bianyi biao zhun kan yuyi jufa zhi hudong (From near-synonyms to the interaction between syntax and semantics)." *Chinese Languages and Linguistics*, 5:439-459.
- Tsai, Mei-Chih, Chu-Ren Huang, Keh-jiann Chen, and Kathleen Ahrens. 1998. "Towards a Representation of Verbal Semantics - An Approach Based on Near Synonyms." *Computational Linguistics and Chinese Language Processing*, 3(1), pp. 61-74.