

Sense and Meaning Facets in Verbal Semantics: A MARVS Perspective*

Kathleen Ahrens Chu-Ren Huang Yuan-hsun Chuang
National Taiwan University Academia Sinica National Taiwan University

This paper explores the theory of lexical knowledge set forth in the Module-Attribute Representation of Verbal Semantics (Biq 2000) and extends its scope to include issues concerning verbal polysemy. Previous versions of the theory postulated that different event structures required sense distinctions. In examining corpora data for the verbs ‘put’ and ‘set’ we argue that an additional criterion for ‘sense’ needs to be put in place, namely the standard that if two meanings can coexist in the same sentence, then their representation is at the meaning facet, and not the sense level (following Ahrens et al. 1998). In sum, this paper adds to the cross-linguistic evidence supporting MARVS as well as constrains its definition of ambiguity.

Key words: polysemy, event structure, meaning facet, lexical semantics

1. Introduction

The Module-Attribute Representation of Verbal Semantics (MARVS) is a theory of the representation of verbal semantics that is based on Mandarin Chinese data (Huang et al. 2000). This theory proposes two types of modules: event structure modules and role modules, as well as two sets of attributes: event-internal attributes and role-internal attributes which are linked to the event structure module and role module respectively. These module-attribute semantic representations have associated grammatical consequences.

Huang et al. (2000) found that both the composition of event modules and the attested lexical semantic attributes can be generalized across the natural semantic class that they belong too. For example, the contrast between *bai3* and *fang4* (verbs of putting) has to do with the fact that *bai3* has a role-internal feature of [design] attached to the location role, while *fang4* does not. Moreover, this contrast can be generalized across the semantic class of verbs that involve design on the focused location role (i.e., *hua4* ‘to paint’) and those that do not (i.e., *tu2* ‘to cover with paint, to doodle’). What

* We would like to thank participants of PACLIC-15 as well as Mei-chun Liu for their comments and suggestions on an earlier version of this paper. All errors remain our own.

we would like to determine in this paper is if similar contrasts can be found in near-synonyms of other languages. In particular, are event modules in other languages (such as English) organized along similar conceptual lines?

In this paper, we examine this theory in light of English data. In particular, we will look at the near synonym contrast of the verbs ‘put’ and ‘set’ based on data from the sampler of the British National Corpus (approximately two million words). To preview our results, we find that the event structure in English is slightly different from that of Chinese. That is, ‘put’ has the roles of Agent, Theme, and Location in its event structure, with location having a role-internal attribute of [+bounded]. ‘Set’, on the other hand, has four senses: one that involves the roles of Agent, Theme and (Place) Location with a role-internal attribute of [+position] on the Theme, a second sense which involves roles of Agent, Theme, and (Time) Location with a role-internal attribute of [+position] on the Theme, a third sense which has three roles of Agent, Theme, and Predicate, and a fourth sense which has two roles, Agent and Theme, with a role-internal attribute of [+restricted] on the theme. We will then compare the notion of ‘sense’ in MARVS with the definition of ‘sense’ and ‘meaning facets’ based on nominal data in Mandarin presented in Ahrens, Chang, Chen, and Huang (1998). We conclude that the criterion of coexistence in the same context is necessary to distinguish between sense and meaning facet and thus, is added to the MARVS theory in order to entail a unified theory of meaning representation for both nouns and verbs.

2. Methodology

Our methodology focuses on delimiting the lexical semantic distinctions between near-synonym pairs that differ slightly in both their syntactic behavior and in their semantics. However, even in cases where there is a difference in meaning, what we are looking for is the relevant differences in both syntax and semantics—that is, along what semantic lines do these two words differ, and how is this difference related to their syntactic behavior (and vice versa)?

How do we determine these collateral differences? First, we examine these near-synonym pairs by first combing a corpus for all relevant examples of the words in question. These examples are then categorized according to their syntactic function. Third, each instance is classified into its argument-structure type. Fourth, the aspectual type associated with each verb is determined, and fifth, the sentential type for each verb is also determined. We find that near-synonyms usually have several cases of complementary distribution of syntactic function. It is often these cases of complementary distribution that allow us to formulate a hypothesis concerning the relevant nature of their semantic

differences.¹

In the Module-Attribute Representation of Verbal Semantics (MARVS), lexical knowledge is classified into two types: structural information is represented with the composition of atomic modules, while content information is represented with attributes attached to these modules. The overall shape of event structure is defined by the composition of five *Event Modules*. The roles that participate in the event are represented in the *Role Modules*. The semantic attributes pertaining to the whole event are called the *Event-internal Attributes* and attached to the event modules. The semantic attributes pertaining to each role are termed *Role-Internal Attributes* and attached to the appropriate role within the role module. It is important to note that the eventive information is attached to the sense of a verb. Verbs with different senses will have different eventive information. Previous versions of the theory postulated that different event structures required sense distinctions. In examining corpora data for the verbs ‘put’ and ‘set’ we argue that an additional criterion for ‘sense’ needs to be put in place, namely the standard that if two meanings can coexist in the same sentence, then their representation is at the meaning facet, and not the sense level (following Ahrens et al. 1998). In sum, this paper adds to the cross-linguistic evidence supporting MARVS as well as constrains its definition of ambiguity.

3. Analysis of ‘put’ and ‘set’

The English verbs ‘put’ and ‘set’ seem synonymous and interchangeable in most contexts (1).

- (1) a. Put/set the book on the table.
b. He set/put the pin on the cushion.

Moreover, the distributional differences that exist for *bai3* and *fang4* in Mandarin, such as cooccurrence with a progressive to describe a process (2a), taking a resultant object (3a), and being modified with an orientational adjunct (4a), do not exist in English (see examples (2b)-(4b) respectively).

- (2) a. ta zhengzai fang/*bai shu zai zhuo-shang
 s/he DURATIVE put/set book at table-top
b. She is putting/setting the book on the table.

¹ This methodology was first laid out in Tsai (1998) and elaborated in Chang (2000).

- (3) a. Ta *fang/bai-le yi zhuo cai
s/he put/set-ASPECT one table dish
b. He *put/*set a table of food. (cf. He set/put food on the table.)
- (4) a. Ta ba yizi chao dong *fang/bai
S/he BA chair face east put/set
b. She put/set the chair facing east.

Difference between ‘put’ and ‘set’ do exist, however. We found the following contrasts based on 100 randomly extracted samples of ‘put’ and ‘set’ from BNC (Ahrens and Huang 2001). An additional 300 examples of ‘set’ were also analyzed (Chuang 2003).² The following generalizations could be found.

First, in the cases of concrete objects and locations, ‘put’ often had the theme being place *into* a location (cf. examples (5)a-f with examples (6)a-b).

- (5) a. put into one prison vehicle
b. put it into a urine container
c. put their heads into a bucket of urine
d. put the Sandinistas in the dock
e. put the muzzle of his weapon into Gruson’s mouth
f. They put us back into the van
g. put more money into Hong Kong
- (6) a. put more pressure on the pedals
b. put men on the moon

It is important to note that the location in all of the cases is a contained space.

However, in the case of ‘set,’ graspable objects were set ‘alongside,’ ‘down,’ or ‘in’ (as ‘in a formation’—not inside something) as in (7).

- (7) a. set alongside his Eulangean law doctorate
b. set it down between Elisabeth and himself
c. set down Miss Danziger’s breakfast herself
d. set down a tureen for one before her
e. set them (chairs) in a semi-circle round the platform

The theme is not necessarily mentioned (but it is almost always mentioned in the ‘put’ cases), and in addition, the location is not necessarily a specific, fixed, or definite space.

² Verb phrases, such as ‘put up to’, ‘put up with’, ‘set up’, ‘set off’. They were excluded from our analysis when they had idiosyncratic meanings relating to the fixed verb-preposition form, such as ‘challenge’, ‘been bothered by’, ‘prepare’, and ‘started’, respectively.

Moreover, in cases when the theme is not graspable (as in the case of a house or garden) it is ellipsed and moreover, the location is also not a definite, containing space, as in (8).

- (8) a. It was set in a trim garden
 b. set beneath the dramatic limestone
 c. set in acres (of maintained comfort)
 d. set admist flower gardens
 e. set in acres of garden
 f. set against the cottage stone

From the constrasts drawn above with the BNC data, we can observe that ‘put’ has the roles <Agent, Theme, Location> with Location having the feature [+bounded]. This contrast can be seen in (9).

- (9) a. He put water in the bathtub.
 b. *He set water in the bathtub.

Example (9a) means that he filled the bathtub with water. Example (9b) does not allow that interpretation because the bathtub is a contained location. However, the examples in (7) indicate that although a bounded space is not required for ‘set’, a position of theme is needed. In all the examples in (7), the position of the theme in relation to an actor is being described. In (8) the position of the theme is being described in relation to the landscape. Thus, the roles for ‘set’ are <Agent, Theme, Location> and the theme role carries the attribute of [+position].

This analysis explains why, in example (10a), the message can be either sitting on top of the bulletin board, or it can be stuck to the bulletin board, or it can be on an electronic bulletin board. However, in (10b), the only reading possible is one where the note is physically sitting on top of the bulletin board.

- (10) a. She put the note on the bulletin board.
 b. She set the note on the bulletin board.

In the ‘put’ case the only requirement is that the note appear on some part of the contained space of the bulletin board (whether or not that space is physical or ephemeral). In the ‘set’ case, the position of the theme in relation to the location is critical and only the positional reading is allowed.

How do these analyses account for the abstract objects and locations? In the case

of ‘put’ the analysis holds nicely. In all the cases in (11) below, the locative role is a metaphorical extension to an abstract concept or even such as ‘quandary’ or ‘trial.’

- (11) a. put flesh and blood on the skeleton structure of a possible united Europe
- b. put them all in a real quandary
- c. put on trial
- d. put the best face on the evident lack of progress

Moreover, the locative role still has the feature of [+bounded]. In all of the above cases, the locative role has a definite boundary (i.e., the quandary is a particular situation, as is the trial, the lack of progress, and the structure of a united Europe.) Each one of these situations has a definite boundary.

The analysis of ‘set’ with abstract objects and locations must be divided up into different sections. First, in example (12), we can see that locative role has been extended to time. The theme (i.e., the event in question) must be positioned in relation to a period of time, although once again, the exact position (on a timeline, for example) is not necessary.

- (12) a. set in the early 50’s
- b. set in a historical context

Thus, the event structure that we have proposed for ‘set’ can stand, although we will return to the question of whether or not this abstract instance of ‘set’ requires a separate sense listing in the next section.

Example (13) necessarily involves only agents and themes.

- (13) a. It requires member states to set rules on mandatory bids
- b. Zambia’s first development plan set the ambitious goal of providing primary school places for all children by 1970
- c. This incident set the pattern for my relationship with Frank Dick

In the above examples, the theme is restricted. This can be seen from its collocates, such as ‘rules’ in (13a), ‘goal’ in (13b), and ‘pattern’ in (13c). Thus, we postulate that this is a different sense because it has a different argument structure of <Agent, Theme>, where the Theme has the role-internal attribute of [restricted].

Examples (14) and (15) involve agents, themes, and predicates. The predicates are instantiated as resultatives in (15).

- (14) a. set crisis management operations in train
 b. setting in motion the sequence of European conflicts
 c. set in train a flood of protest
 d. set in motion a big spy swap
- (15) a. set her pulse racing
 b. set fire to banks and offices
 c. set alight two offices
 d. set our boats on fire

In the cases in (14) and (15), moreover, there is no reason to postulate that the theme is in a particular position. It is not necessary for example that the offices in (15c) are in a particular position. However, the prepositional phrase assigns a property to the theme, i.e., there is a predicative relation. Thus, the roles in the event structure for ‘set’ with an abstract theme and location is: <Agent, Theme, Proposition>, with the predicative relation stipulated. In MARVS, this information is represented in the Role Module. MARVS makes the important assumption that verbal senses represent eventive kinds, just like nominal senses represent referential kinds (Huang et al. 2000). Since eventive information is represented by the combination of the Event Module and the Role Module, there are two ways to differentiate verbal senses in this theory. That is, two different verb senses can be assigned by virtue of either having different event-module compositions or different role-module compositions. An obvious way for two role modules to differ is that they may have different (ordered) role sets. Our analysis of ‘set’ having non-locational complements indicates that these are different senses from the locational ‘set’.

4. Representation of multiple meanings

In the preceding section, we noted that MARVS stipulates different verbal senses for different event structure representations. Thus, based on the analysis so far, ‘put’ has one sense, but it has a metaphorical extension from that sense. ‘Set’, on the other hand, has three senses. In this section we examine whether the notion of sense distinction is compatible with the MARVS representation.

We will look at a model of meaning representation that is originally based on linguistic data from nominals (Ahrens et al. 1998). The authors pointed out that when a noun occurs in different contexts, the referents change. Taken the word ‘zazhi’ (magazine) for example. In (16a), the magazine refers to the physical object, in (16b), ‘magazine’ refers to the information contained within, and in (16c), it refers to a

publishing house.³

- (16) a. 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.'
- b. 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.'
- c. 美 國 各 大 雜 誌 無 不 挖 空 心 思
meiguo ge da zazhi wubu wakong xinsi
American every big magazine do dig-empty mind
爭 取 採 訪 機 會
zhengqu caifang jihui
fight-for interview chance
'Major American magazine fight for interview opportunities.'

This complexity of meaning is described as 'latent', whereas the concept of homophonic ambiguity (as in 'port' meaning 'dock' or 'liquor') is said to be an 'active' type of complexity. The notions of 'active' and 'latent' are phrased in terms of human language processing. In other words, 'active' complexity requires immediate processing towards resolution. While 'latent' complexity can remain dormant until 'activated' by a certain context that calls for resolution. Meaning extensions that are latent involve 'meaning facets', while meaning differences that are active involve 'senses'. Meaning facets are defined by the three following criteria: 1) more than one meaning facet of a sense can coexist in the same context; 2) a meaning facet is an extension from a core sense or from other meaning facets; 3) nouns of the same semantic classes will have similar sense extensions and related meaning facets. Contrasting with meaning facets, senses are defined according to the following three criteria: 1) individual senses cannot appear in the same context (unless the complexity is active, or triggered, as in a pun); 2) there is no core sense from which it is extended; 3) no logical/conceptual mappings can be established between two senses, nor can the link between the two sense be inherited by a class of nouns (Ahrens et al. 1998). For the purposes of this paper, the criterion of whether a word can have multiple meanings within the same context (a defining characteristic of meaning facets) is crucial, especially since the issue of inheritance is rarely relevant for verbs.

³ Examples (16) and (17) are taken from Ahrens et al. (1998).

To give an example of the difference between sense and meaning facets, there are three meanings that can coexist in that context for the word ‘tian1’ in (17) below, since it may refer to ‘sky’, ‘God’, and/or ‘heaven’ (Ahrens et al. 1998).

- (17) 有 人 開始 不 敬 天 也 不 拜 天 了
 you ren kaishi bu jing tian ye bu bai tian le
 there's 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) Ahrens et al. (1998:54)

Note the crucial fact that all three meanings are present in the same context. The meaning complexity cannot be resolved even though speaker are aware of such complexity (when prompted). Hence, these three meanings cannot be senses and are considered meaning facets.

When we apply these criteria to ‘put’ we find that we can create a context where the abstract and concrete meanings of ‘put’ coexist. When the theme can be either concrete or abstract as in the case of ‘paper’ or ‘magazine’, multiple, coexisting interpretations are allowed. For example, in (18) below, ‘this paper’ refers to a printed version of the paper (on paper), the electronic version of the same paper, or the abstract content of the same paper.

- (18) Put this paper on the web.

One may argue, following Pustejovsky’s (1995) Generative Lexicon, that the meaning complexity actually comes from the qualia structure of the noun, and there need not be complexity for the verb. However, envision the situation where the speaker is holding a bound version of the paper and talking to a colleague. Here, the event structure decomposition of the verb ‘put’ entails that there are two sub-events: The first is that the Agent does something to the Theme, which will be understood in this particular context as the physical paper (that the speaker is holding in hand). The second is that, as a result, the Theme beginning a new state of being at the Location. In this sub-event of ‘on the web’, the Theme clearly refers to the electronic version and/or content of the paper. With this decomposition, it becomes clear that nominal qualia structure is not the only source of the meaning complexity. The event structure encoded by the verb must also allow the latent complexity. The above example is additional evidence that the concrete and abstract meanings of ‘put’ are in fact, meaning facets, and not separate senses. This is in line with the suppositions of the MARVS theory as well.

‘Set’ also has an abstract meaning extension from the concrete locative sense

where the theme is either abstract or concrete, as in (19).

(19) Once she set the proposal on the table, people started attacking it immediately.

In (19), the proposal may be either a physical object or an abstract idea, or both. Thus, ‘set’ also contains two meaning facets for its locative sense.

However, when ‘set’ is positioned in relation to a period of time, as in ‘set in the 1950’s’, the two meanings cannot exist. In example (20), it is not the case that the novel (physical object) can be in the twenty-second century at the same time that the contents of the novel can take place during that time period.

(20) Jackson set his novel in the twenty-second century.

Thus, it is not the case that the two meanings of ‘set’ can occur simultaneously, as the definition of meaning facet above requires. However, a conceptual link can be found (i.e., a conventionalized metaphorical extension from space to time for the Locative Role). In this case, we have two criteria that conflict with each other. The coexistence criterion suggests that they are separate senses, while the conceptual mapping criterion suggests that they are instead meaning facets. In this type of case, we suggest the coexistence criterion take precedence since it is a more straightforward linguistic test to judge if two or more meanings can coexist in a single sentence, than it is to prove if a conceptual mapping exists or not.⁴

5. Representing sense and meaning facet in MARVS

We summarize in this section our analyses of the data involving ‘put’ and ‘set’, and represent our accounts in MARVS formalism. In the following representation, we propose a way to formally account for senses and meaning facets in MARVS (Huang et al. 2000).

⁴ Note that in this paper we have revised the criteria. We use ‘mapping’ instead of the original ‘conceptual link.’ The reason is because the original wording is too vague. ‘Link’ could be construed as any conceptual relations, which may well exist among senses. What we want to refer to here is a sub-lexical mapping rule that allows the different meanings to be accessed directly within the same lexical entry. Even though this criterion seems to be formally sound, it is not easy to establish empirical rules to decide whether a conceptual mapping occurs within a single lexical entry or across more than one entry. In addition, while the SPACE → TIME conceptual link is well established (Lakoff and Johnson 1980, Ahrens and Huang 2002), other conceptual links may not have the body of evidence necessary to support their existence.

MARVS theory assigns different sense entries to meanings that are distinct in their event or role modules. However, the theory was not able to clearly delineate between senses and meaning facets, which have been found to be very useful in understanding types of ambiguity in nominals (Ahrens et al. 1998). In analyzing the data from ‘set’, we concluded that the coexistence criterion took precedence over the conceptual mapping criterion because the former allows us to clearly and quickly distinguish between meaning facets and senses. Future research will focus on a statistical analysis of the ‘put’ and ‘set’ data in English as well as a statistical, corpora-based comparison with ‘*fang4*’ and ‘*bai3*’. In addition, operationally defining ‘sense’ and ‘meaning facet’ for verbs opens the way for psycholinguistic experiments on number of senses on verbs (cf. Lin and Ahrens’ (2000) experiments on nominals). A formal representation of meaning facets as conditional meaning entailments is proposed and incorporated in the framework of MARVS.

This in-depth study of bilingual near-synonym pairs, put together with the perspective provided by a comprehensive study of cross-lingual meaning equivalences such as Huang et al. (in this volume), highlights the potential of a new field of study: cross-lingual lexical semantics. On the one hand, the contrasts are fascinating and challenging for lexical semantic theories. On the other hand, for the theories of lexical semantics to have immediate and direct implications on the studies of human cognition, the most convincing proof is that they account for both sharable and contrastive features in cross-lingual lexical semantics. In short, we hope that this study leads to further understanding of the organization of multiple meanings in the mental lexicon, as well as heralds future studies in cross-lingual lexical semantics.

References

- Ahrens, Kathleen, and Chu-Ren Huang. 2001. A comparative study of English and Chinese synonym pairs: An approach based on the Module-Attribute Representation. *Proceedings of 15th Pacific Asia Conference on Language, Information and Computation*. Hong Kong: City University of Hong Kong.
- Ahrens, Kathleen, and Chu-Ren Huang. 2002. Time Passing is Motion. *Language and Linguistics* 3.3:491-519.
- Ahrens, Kathleen, Li-li Chang, Keh-jiann Chen, and Chu-Ren Huang. 1998. Meaning representation and meaning instantiation for Chinese nominals. *Computational Linguistics and Chinese Language Processing* 3.1:45-60.
- Biq, Yung-O. (ed.) 2000. Special issue on Chinese Verbal Semantics. *Computational Linguistics and Chinese Language Processing* 5.1.
- British National Corpus. <http://www.hcu.ox.ac.uk/BNC/>
- Chuang, Yuan-hsun. 2003. *Sense Distinction of Verbs in English and Mandarin Chinese: An Analysis of the Verbs "Set" and "Bai3"*. Taipei: National Taiwan University MA thesis.
- Lakoff, George, and Mark Johnson. 1980. *Metaphors We Live By*. Chicago: University of Chicago Press.
- Levin, Beth. 1993. *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago: University of Chicago Press.
- Lin, Charles, and Kathleen Ahrens. 2000. Calculating the number of senses: Implications for ambiguity advantage effect during lexical access. *Proceedings of ISCLL-7*. Chiayi: National Chung Cheng University.
- Huang, Chu-Ren, Elanna I. J. Tseng, Dylan B. S. Tsai, Brian Murphy. 2003. Cross-lingual portability of semantic relations: Bootstrapping Chinese WordNet with English WordNet relations. *Language and Linguistics* 4.3:509-532.
- Huang, Chu-Ren, Kathleen Ahrens, Li-li Chang, Keh-jiann Chen, Mei-chun Liu, Mei-chih Tsai. 2000. The Module-Attribute Representation of verbal semantics: From semantics to argument structure. *Computational Linguistics and Chinese Language Processing* 5.1:19-46.
- 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* 3.1:61-74.

[Received 6 August 2002; revised 16 May 2003; accepted 26 May 2003]

Kathleen Ahrens
Graduate Institute of Linguistics
National Taiwan University
1, Sec. 4, Roosevelt Road
Taipei 106, Taiwan
kathleenahrens@yahoo.com

Chu-Ren Huang
Institute of Linguistics, Preparatory Office
Academia Sinica
130, Sec. 2, Academia Road
Taipei 115, Taiwan
churen@gate.sinica.edu.tw

Yuan-hsun Chuang
Graduate Institute of Linguistics
National Taiwan University
1, Sec. 4, Roosevelt Road
Taipei 106, Taiwan
yuanhsun@hp.iis.sinica.edu.tw

動詞詞彙語意中的詞義與意面： 由 MARVS 理論觀點出發

安可思
國立台灣大學

黃居仁
中央研究院

莊元珣
國立台灣大學

本文旨在探討詞彙知識理論中之「模組-屬性動詞語意表達模式」(MARVS, Biq 2000)，並將此理論延伸到多義動詞的分析。承續「不同的事件結構必定有語意上的區分」的理論，並用於“put”和“set”兩動詞的語料庫為本分析。我們發現必需要有更深一層次的語意區分。換句話說，若有兩種意義 (meaning) 可同時在一語句上存在，則我們應將此兩詞義表達視為意面 (meaning facet)，而非為不同之詞義 (sense)。此理論乃承繼安可思等 (1998) 對名詞多義狀態所提的理論。總而言之，在本文中，我們利用跨語言詞彙語意之研究來支持 MARVS 理論架構以及限定動詞歧義之定義。

關鍵詞：多義詞，事件結構，意面，詞彙語意學