

Transitivity of a Chinese Verb-Result Compound and Affected Argument of the Result Verb

You-shan Chung* and Keh-Jiann Chen*

Abstract

The Chinese verb-result compound is productive, but its meaning and syntactic behaviors have posed challenges to theoretical and automatic analyses. Theory-wise, the current study proposes that VRs have inherent affecting direction, which argument mapping principles and selectional restrictions, event structures, or the kinds of semantic/pragmatic principles or real-world knowledge proposed by previous researchers do not seem to account for. Application-wise, we predict the VR's transitivity and whether the result component is predicated of the logical subject or the logical object, based on the transitivity of individual component verbs and the selectional restrictions between component verbs and arguments. Since the transitivity property and selectional restrictions of individual verbs can be annotated in our lexicon, the rules should fare well in automatic processing. Meanwhile, as the rules have been motivated by linguistic theories and have been observed to make correct predictions in most cases, they are worthy of further large-scale testing.

Keywords: Verb-result Compound, Transitivity, Lexical-semantics, Meaning Prediction

1. Introduction

Verb-result compounds (VR¹) are compounds that are comprised of an event (denoted by the first verb, henceforth V1) and the result of the event (denoted by the second verb, henceforth V2).² For example, *da-po* 打破 'broken from being hit,' *chi-bao* 吃飽 'full from eating,' and *tu-hei* 塗黑 'paint something black' are VRs, where *da* 打 'hit,' *chi* 吃 'eat,' and *tu*

* CKIP, Institute of Information Science, Academia Sinica, Taipei, Taiwan

Tel: +886-2-27883799 ext.1502 Fax:+886-2-27881638

E-mail: {yschung; kchen}@iis.sinica.edu.tw

¹ For a complete list of the abbreviations used in this paper, please refer to Appendix I.

² Despite slight variations in terminology, VRs roughly correspond to Smith's (1994) resultative verb compounds and Li and Thompson's (1981) Result RVC.

塗 ‘paint’ lead to the results *po* 破 ‘broken,’ *bao* 飽 ‘full,’ and *hei* 黑 ‘black.’ Due to limited space, two variants of VRs, directional verb compounds and phase/completive compounds (Smith, 1994), will not be discussed in this paper.

The high productivity of VRs in Chinese makes them a worthwhile topic of investigation for NLP. Nevertheless, this verbal construction poses challenges to traditional linguistic theories of the syntax-semantics interface aimed at mapping the meaning of verbal constructions to their surface structure. This is because most generative frameworks hold that mappings between event meaning representations and syntactic structure are governed by constraints imposed by the verb (e.g. theta theory, cf. Crystal 1997 for its definition). These constraints, however, do not always seem to be applicable to VRs. Since a VR contains two verbs, it is often argued that one of them is to be identified as the “head,” which is supposedly responsible for the mapping. Nevertheless, such an assumption encounters difficulties. For one thing, the head is defined as the component that dominates the resulting meaning and syntactic behaviors, but it is controversial which component it is (i.e. V1 or V2) or even if such a component exists at all in Chinese VRs (Huang, 1998; Li, 2009). Among researchers who do not address head identification, some argue that mappings for VRs are not simply constrained by the argument structures of the component verbs (Yin, 2011; Li, 2007; Huang, 2006).

Various explanations have been proposed to accommodate VRs in a larger framework of syntax-semantics interface, including syntactic operations (Mulkeen, 2011; Shen & Lin, 2005), lexical-semantic properties of the components (Yin, 2011; Li, 2007; Huang, 2006), or a split approach to different VRs (Lin, 1998). Syntactic or lexical-semantic, these accounts all appeal either to extensions of certain underlying syntactic operations, such as movement, or to decisions on thematic role mapping. As far as automatic processing is concerned, the former means adjustments to the general parsing rules and the latter requires the incorporation of real-world knowledge.

In particular, not only do non-lexical accounts present technical challenges, but also a growing consensus in linguistic studies is that the meaning and syntactic behaviors of larger linguistic structures can be explained by the syntactic and semantic properties of the composing words. Along this lexical approach, Levin (1993) presented a heavily-cited framework to explain how the meaning and grammar that are encoded by verbs are mapped onto the syntactic arguments. Therefore, we began the analysis of the VR compound, a kind of verbal compound, with a lexical approach.

For the sake of accuracy and simplicity as a linguistic account and for ease of computer processing, our treatment of VRs is lexically-based. Unlike some lexical accounts, however, we avoid postulations of underlying representations. We address two basic syntactic and

semantic distinctions of VRs by employing information of verbs' transitivity³ and selectional restrictions encoded in our on-line lexicon. The research questions are how to predict the transitivity of a VR and how to identify V2's affected argument (i.e., the argument predicated by V2).

2. The Prediction of Transitivity and V2's Affected Argument

Transitivity and selectional restrictions are two of the most basic syntactic and semantic distinctions for verbs. A special characteristic of Chinese VRs is that V2's affected argument can either be the logical subject (i.e., subject-controlled, henceforth SC) or the logical object (i.e., object-controlled, henceforth OC) of a VR. Such a contrast can be seen in the VRs *da-po* 打破 'broken from hitting' and *du-dong* 讀懂 'read something and understand it,' which belong to the OC and SC type, respectively. For example, in *Zhangsan da-po-le boli* 張三打破了玻璃 'Zhangsan broke the glass,' the affected argument of *po* 破 'broken' is *boli* 玻璃 'glass,' V1's object, whereas in *Zhangsan du-dong-le naben shu* 張三讀懂了那本書 'Zhangsan (finally) understood the book,' it is V1's subject, *Zhangsan* 張三, that is being predicated of by V2.

Below, we will briefly introduce how two lexical-semantic accounts (Li, 2007, Huang, 2006) and one lexical-syntactic account (Lin, 1998) address transitivity and V2's affected argument.

2.1 Review of Lexical-semantic and Lexical-syntactic Works on Chinese VRs

Lexical-semantic accounts like Li (2007) and Huang (2006) argue that event structure (and argument structure) can explain the meaning and syntactic behaviors of VRs through a set of linking rules that map event participants of various thematic roles (e.g., agent, patient, causer, causee) onto sentence positions (e.g., subject, object).

Li (1997)

Under Li's framework, transitivity of a VR compound and V2's affected argument fall out of the options in expressing the causing factor and the affected entity, assigned the thematic roles "causer" and "causee," respectively; the causee amounts to V2's affected argument in our terms. The roles are assigned by the external and internal arguments of V1 and of V2 as well

³ A transitive verb is one that has a direct object, although the object does not have to be overt (i.e., occurring in the surface structure), whereas an intransitive verb does not allow a direct object at all (overt or covert). For example, the transitive verb *eat* has a direct object, although both *He ate a sandwich* and *He ate* are possible. In contrast, as an intransitive verb, *go* usually cannot have a direct object. (cf. <http://www.glottopedia.de/index.php/Transitive>.)

as by the event structure of the V1+V2 combination.

With the causer and causee, as well as their corresponding arguments (i.e., internal/external/neither), identified, a proposed set of linking rules determine the sentence positions for the two thematic roles to be realized. Below is one of them (Li, 2007: 96).

- (1) The causer argument is realized in subject position and the causee argument in object position, when both arguments are overtly expressed by different linguistic expressions.

As can be seen, Li's reference point is the causer and the causee. Nevertheless, as far as automatic processing is concerned, the causer and causee have to be identified rather than taken as given. So, Li's rules are not readily applicable for a system that has no access to knowledge of causer and causee.

Besides methodological issues, Li does not explain his criteria for distinguishing between causer and causee, which his linking rules hinge on. As thematic roles, causer and causee should be defined on semantic grounds rather than by sentence positions. Under his analysis, *na-bao yifu* 那包衣服 'the bundle of clothes' in *Na-bao yifu xi-lei-le Zhangsan* 那包衣服洗累了張三 '(Zhangsan washed that bundle of clothes) and the clothes got Zhangsan tired' (Li, 2007: 119) and *shu* 書 'book' in *Zhangsan kan-lei-le shu* 張三看累了書 'Zhangsan read a book and as a result became tired' (Li, 2007: 115) differ in thematic roles. According to Li, in the former sentence, the causer is *na-bao yifu* 那包衣服 whereas in the latter, the causer is *Zhangsan* 張三. In terms of meaning, however, it appears that, in both sentences, it is *Zhangsan* 張三 who is the agent engaging in the activities *xi* 洗 'wash' and *kan* 看 'read,' from which he becomes tired. So, it is not clear why the causer-causee mappings in these two sentences are not the same.

Huang (2006)

There are two components in Huang's framework: event frames and linking rules. The event frames describe the obligatory and optional participants and predicates for transitive and intransitive verbs. These event participants map to certain syntactic categories and sentence positions by default. For example, the subject of a sentence is usually a noun phrase that receives an agent role.

Although Huang's account can explain mappings between thematic roles and sentence positions, his model needs to identify the transitivity of a VR first in order to determine the relevant event participants so as to discover the thematic roles of the subject and the object. Nevertheless, he does not address how to determine the transitivity of a VR compound.

Meanwhile, the issue of the definitions of causer and causee arises again. In *Baozhi kan-hua-le ta de yan* 報紙看花了他的眼 'The newspaper got his eyes blurred from reading it' (Huang, 2006: 27), Huang regards *baozhi* 報紙 'newspaper,' instead of *ta* 他 'he,' who engages in the action that leads to the result state, as the causer. It is not clear why *baozhi* 報

紙 here, *yifu* 衣服 in *Na-bao yifu xi-lei-le Zhangsan* 那包衣服洗累了張三 and *shu* 書 in *Zhangsan kan-lei-le shu* 張三看累了書 from Li (2007) differ in thematic roles. These examples suggest lack of a consistent way to identify causer and causee between researchers and within individual researcher's analyses while such criteria underpin their theories.

Lin (1998)

Lin (1998) predicts a VR's transitivity and V2's affected argument based on V1's transitivity and whether V2 is predicative of animate or inanimate entities. This approach seems promising for automatic processing since transitivity properties and selectional restrictions can be annotated.

Regarding transitivity, she argues that the transitivity value of a VR is the same as that of V1. Her prediction of the affected argument, however, is nondeterministic, as exemplified in her analysis of transitive V1s paired with intransitive V2s predicative of animated noun phrases, e.g., *zhui-lei* 追累 'chase-tired' and *chi-bao* 吃飽 'full from eating.' She argues that, in such cases, the V2 can be predicative of either the subject or the object, or both. For example, she thinks that *Zhangsan zhui-lei-le Lisi* 張三追累了李四 allows three readings, which are, with descending degrees of acceptability, 'Zhangsan chased Lisi and Lisi became tired,' 'Zhangsan chased Lisi and Zhangsan became tired,' and 'Lisi chased Zhangsan and Lisi became tired.' In contrast, *chi-bao* 吃飽 can only have an SC reading. Nevertheless, she does not explain the conditions for a reading to be available, or at least preferred.

In sum, regarding theoretical comprehensiveness, Li's and Huang's identification of causer and causee seems arbitrary. In terms of automatic processing, their predictions depend on already knowing what automatic processing needs to figure out, whereas Lin's account does not address disambiguation.

Despite these suggested theoretical and practical inadequacies, many such lexical-semantic accounts have turned from relying solely on argument structure to what argument structure and selectional restrictions (i.e., the semantic restrictions on what arguments can be taken by a verb) cannot explain. Within the lexical tenets, the general consensus is that, besides the argument structures of individual verbs, the VR construction has its own event structure with thematic roles to assign. The roles from the two structures (i.e., argument structure and event structure) conflate to decide the resulting meaning and syntactic behaviors of the VR. Also frequently considered is the variation in the kinds of composing predicates that exist and whether they are optional/obligatory or general/specific in the event structures of VRs⁴ of different event types, consequently taking semantic and syntactic variations between different verbs and within the same verb into account.

⁴ Examples of actual representations can be found in Li (2007: 95) and Huang (2006:21).

When these interactions and variations allow ambiguities, it is suggested that semantic and pragmatic, or real-world knowledge⁵, motivations screen or rank the remaining readings. Some researchers, like Li (2009) and Huang (2002), only cite the role of real-world knowledge in determining whether a VR is possible or not but leave the question of whether an acceptable compound can occur in a sentence with certain noun phrases in a certain order or not and, if it can, what the predication relation is to syntactic, semantic, or pragmatic constraints. For example, Li proposes the Animate Subject as Agent or Experiencer Strategy (ASAES) as well as the principle of prototypicality, both of which he considers pragmatic principles. The content of the former has been made clear by its name; the latter states that the entity that carries out the V1-denoted act is less likely to be the V2-affected entity. Both principles are meant to rank the possibility of acceptable readings. Other researchers contend that real-world knowledge remains an element in the reading(s) of a legal VR. Huang (2006) incorporates the role of real-world knowledge in the event frame of VRs, which is the interface where linking rules are applied to map semantic configurations to sentence positions.

The current study suggests an even more central role for real-world knowledge in the interpretation of VRs. So far as we are aware, Huang's (2006) constructional framework is one of the few accounts that formally represents real-world knowledge, if not the only account. Nevertheless, even he thinks that real-world knowledge is only an optional participant in the event frame. In spite of this, we not only agree with Huang that real-world knowledge remains at play in deciding the reading(s) of a legal VR, but we also have found real-world knowledge to be central to the meaning and syntactic behaviors of VRs throughout its formation and interpretation. We will support this position with examples in Section 3.

2.2 The Model of the Current Study

We first propose an affecting direction principle to model the transitivity and affected argument of VR compounds. Then, we develop a set of heuristic rules to emulate the principle using only the information of individual transitivity of V1 and V2 and the selectional restrictions of V1 and V2 but not using real-world knowledge. Based on a lexical approach, we encode the semantic and syntactic properties of each word in the lexicon. Transitivity is one of the listed properties, and so are the selectional restrictions. We manually encode such information based on observations of corpus data. For example, *da* 打 'hit' requires an animate agent and a physical theme. This is shown in (2):

⁵ Some researchers use the terms "real-world knowledge" and "pragmatic constraints" interchangeably. Li (2007), however, seems to distinguish between them.

(2) *Zhangsan da Lisi* 張三打李四

Zhangsan hit Lisi

‘Zhangsan hit Lisi.’

In contrast, the verb *po* 破 ‘broken’ requires an inanimate theme (PERF=perfective marker):

(3) *Boli po le* 玻璃破了

glass broken-PERF

‘The glass is broken.’

Thus, the lexical entries of *da* 打 ‘hit’ and *po* 破 ‘broken’ would be partially represented as (4) below (Vt=transitive verb; Vi=intransitive verb; [+ANI]=animate entity; [+PHY]=physical entity; [-ANI]=inanimate entity):

(4) *da* 打 ‘hit’

Transitivity: Vt

Selectional restrictions: [+ANI] agent, [+PHY] theme

po 破 ‘broken’

Transitivity: Vi

Selectional restrictions: [-ANI] theme

With the encoded information, when the computer encounters the compound *da-po* 打破 ‘broken from being hit,’ the rules of our model can predict the VR’s transitivity and V2’s affected argument. Nevertheless, such information cannot always predict the resulting meaning and syntactic behaviors of VRs. We think this is because the fundamental factor that decides a VR’s transitivity and V2’s affected argument is the affecting direction of the compound.

By “affecting direction” we refer to native speakers’ knowledge of the range of entities that can be affected by the state resulting from an action or another state and of whether the VR can take an argument/arguments or other noun phrases or not. For instance, Mandarin Chinese speakers know that the entity affected by of the state denoted by *chi-bao* 吃飽 ‘full from eating’ is an animate being that eats and that the VR cannot be followed by noun phrases inside or outside of V1’s argument structure. In formal terms, we say that *chi-bao* 吃飽 ‘full from eating’ is intransitive and SC. Such knowledge is implicit, with linguistic and non-linguistic motivations underlying it. We will discuss the motivations in more detail in Section 3.

The affecting direction principle:

Case 1: When V1 is Vt

Vt+R → VRt and SC; if the affecting direction of the VR is toward Vt's subject, e.g.,
da-ying 打贏 'fight and win,' *du-dong* 讀懂 'read something and understand it'

Vt+R → VRt and OC; if the affecting direction of the VR is not toward Vt's subject, e.g.,
jiao-hui 教會 'teach someone and make him/her understand'

Case 2: When V1 is Vi

Vi + Rt → VRt and SC, e.g., *pao-shu* 跑輸 'lose in running'

Vi + Ri → VRi and SC; if the affecting direction of the VR is toward Vi's subject, e.g.,
pao-lei 跑累 'tired from running'

Vi + Ri → VRt and OC; if the affecting direction of the VR is toward Ri's subject, e.g.,
pao-po 跑破 'broken from running'

Two points of clarification are in order here. First, there are some cases where the affecting direction seems to allow both an SC and an OC reading. Two of the oft-cited cases are *qi-lei* 騎累 'ride-tired' and *zhui-lei* 追累 'chase-tired,' as in *Zhangsan qi-lei le ma* 張三騎累了馬 and *Zhangsan zhui-lei-le Lisi* 張三追累了李四, both allowing either the "Zhangsan is tired" reading or the "horse/Lisi is tired" reading. Such VRs are believed to be ambiguous in the sense that they seem to have two readings in the same sentence. Nevertheless, we consider the two readings of these VRs as being separate lexical entries, i.e., polysemous. This is based on Huang *et al.*'s (2005) criterion for distinguishing different words from different meanings of a word, which states that senses that cannot co-exist in the same context have to be treated under different lexical entries. Accordingly, while *qi-lei* 騎累 and *zhui-lei* 追累 are possible with two readings in the above sentences, the ambiguity can be resolved with a moderate number of contextual clues.

(5) 騎累₁

Zhangsan pa qi-lei-le ai-ma, jue ding rang ta xiuxi 張三怕騎累了愛馬，決定讓牠休息

Zhangsan fear ride-tired-PERF beloved-horse, decide let it rest

'Fearing that it might be exhausted from extended rides, Zhangsan decided to let his beloved horse take a break.'

騎累₂

Zhangsan qi-lei-le ma, jue de haishi kai-che bijiao qingsong 張三騎累了馬，覺得還是開車比較輕鬆

Zhangsan ride-tired-PERF horse, feel nevertheless drive-car relatively relax

'Tired from horse-riding, Zhangsan found driving more relaxing.'

A remaining question is how to decide which *qi-lei* 騎累 or *zhui-lei* 追累 is relevant when the computer encounters such VRs. Indeed, currently, we can only identify them as transitive verbs that are polysemous between an SC and OC reading. Context is required to decide the affected argument of V2.

Second, when the affecting direction of a VR is not toward the Vt's subject, it is not necessarily toward its object.

In a few cases, it is the adjunct of the Vt that is V2's affected object. For example, *qie* 切 'cut' has an agent-denoting subject that is an animate being, e.g., *Zhangsan* 張三, a patient-denoting object that is the physical object being cut, e.g., *rou* 肉 'meat,' and an instrument-denoting adjunct that is the tool for cutting, e.g., *dao* 刀 'knife.' For the VR *qie-dun* 切鈍 'blunt from cutting,' it is the instrument-denoting adjunct instead of the patient-denoting object that is affected by *dun* 鈍 'blunt.' Nevertheless, although the instrument is not the object of the Vt, the prediction that the VR is OC is still borne out since V2's affected argument is indeed the object of the VR *qie-dun* 切鈍 because *dao* 刀 is the receiver of the action⁶ described by the compound.

When V1 is Vi, again, the noun phrase that follows the VR is not necessarily V1's object. Actually, it does not even have to be any possible argument, adjunct, or complement in the verb frame. To show this, we may have to first make clear what is meant by being SC and OC. It may be suggested that the affecting direction of the VR *pao-lei* 跑累 'tired from running' is not fixed if sentence pairs like *Zhangsan pao-lei-le* 張三跑累了 'Zhangsan ran and as a result got tired' and *Zhangsan pao-lei-le liang-tiao tui* 張三跑累了兩條腿 'Zhangsan ran, and as a result his legs got tired' are taken into account. In the former sentence, the affecting direction should be toward Zhangsan 張三, but, in the latter sentence, it seems to be either toward Zhangsan 張三 or *liang-tiao tui* 兩條腿. As a result, the latter sentence seems to have both an SC and an OC reading.

Our view is that both readings belong to the SC type. First, we find *liang-tiao tui* 兩條腿 'legs' in the sentence to be more like a complement than an object of *pao-lei* 跑累 'tired from running.' The first reason is that, for a noun phrase to be allowed to follow the VR *pao-lei* 跑累, it has to specify a body part of the subject, thus functions like a subject complement, as defined in Crystal (1997). Second, it seems that "Zhangsan's legs being tired" entails "Zhangsan being tired." To the extent that Zhangsan is also tired, the sentence can be labeled as the SC type.

Sometimes, the NP in question is not V1's argument but is still the VR's object. For example, VRs with intransitive V1s like *bing-huang* 病慌 'sick-nervous' and *ji-bing* 急病

⁶ cf. Crystal 1997 for the definition of object as "the receiver of the action."

‘worry-sick’ can have OC readings, as in *Zhangsan bing-huang-le Lisi* 張三病慌了李四 ‘Zhangsan’s illness made Lisi nervous,’ *Zhangsan ji- bing-le Lisi* 張三急病了李四 ‘Zhangsan got Lisi so worried, who became sick as a result.’ It is difficult to treat 李四 as a participant in the verb frames of *bing* 病 ‘sick’ and *ji* 急 ‘worried.’ Another probably more drastic example is *ku-zou* 哭走 ‘to cause someone to leave by crying.’ Although *ku* 哭 cannot be subcategorized for objects, the VR obligatorily take objects and is possible with an OC reading only, as in *Zhangsan ku-zou-lie Lisi* 張三哭走了李四 ‘Zhangsan cried, making Lisi leave.’ Note in passing that, in Section 2, we argued for the treatment of the OC and SC readings of *qi-lei* 騎累 and *zhui-lei* 追累 as belonging to different lexical entries. On this account, *bing-huang* 病慌 and *ji-bing* 急病 are also regarded as different VRs.

Since VRs cannot be exhaustively listed, the affecting direction also cannot be exhaustively annotated. Thus, it has to be predicted, too. The affecting direction can generally be predicted by the transitivity properties and selectional restrictions of the two component verbs. Transitivity and selectional restrictions are as exemplified in (4) and are annotated in the lexicon. Therefore, rules based on this information are workable. Where the affecting direction cannot be predicted, however, the VR’s transitivity and V2’s affected argument become unpredictable as a result. Nevertheless, the results have seemed to be as predicted most of the time so far, suggesting that the affecting direction is usually predictable.

2.3 The Heuristics

With the possibility of unpredictability in mind, we are ready to look at the following heuristics for the automatic prediction of transitivity and V2’s affected argument.

Case 1: When V1 is Vt

- (a) $Vt + Rt \rightarrow VRt$ and SC, e.g., *da-ying* 打贏 ‘fight and win,’ *du-dong* 讀懂 ‘read something and understand it’
- (b) $Vt + Ri \rightarrow VRt$ and OC; if Vt’s object or adjunct satisfies the selectional restrictions of Ri’s subject, e.g., *da-po* 打破 ‘broken from hitting,’ *qie-dun* 切鈍 ‘blunt from cutting’
- (c) $Vt + Ri \rightarrow VRt$ and OC; if Vt’s subject and object both satisfy the selectional restrictions of Ri’s subject, e.g., *qi-lei₁* 騎累₁ ‘tired from being ridden,’ *zhui-lei₁* 追累₁ ‘tired from being chased,’ *jiao-fan₁* 教煩₁ ‘vexed from being taught’
- (d) $Vt + Ri \rightarrow VRt$ and SC; if Vt’s subject satisfies the selectional restrictions of Ri’s subject, e.g., *qi-lei₂* 騎累₂ ‘tired from engaging in the activity of riding,’ *zhui-lei₂* 追累₂ ‘tired from engaging in the activity of chasing,’ *jiao-fan₂* 教煩₂ ‘vexed from engaging in the activity of teaching’

The following rules deal with VRs where V1 is an intransitive verb.

Case 2: When V1 is Vi

- (e) Vi + Rt → VRt and SC, e.g., *pao-shu* 跑輸 ‘lose in running’
- (f) Vi + Ri → VRi and SC; if Vi’s subject satisfies the selectional restrictions of Ri’s subject, e.g., *pao-lei* 跑累 ‘tired from running’
- (g) Vi + Ri → VRt and OC; if Vi’s subject does not satisfy the selectional restrictions of Ri’s subject, e.g., *pao-po* 跑破 ‘broken from running’

Below, we show how the VR’s transitivity and V2’s affected argument might be predicted by matching a component verb’s transitivity and selectional restrictions with the other’s. Again, we *already* know the transitivity and selectional restrictions of the component verbs before the computer encounters a context that contains a VR construction. This means the judgment of the VR’s characteristics, except for the cases where both (c) and (d) apply, generally does not depend on on-line context, but we will see exceptions in the following sentences (6), (11), (12), and (15) and will address them in Section 3.

The following sentences from Li (2007) represent some of the most frequent VR patterns, where the subjects and objects are default ones that meet the selectional restrictions of the component verbs. In (6)-(12), V1 is a transitive verb. We will describe the transitivity properties before describing the matching of the selectional restrictions. The relevant rule is given at the end.

- (6) *Zhangsan zhui-lei le* 張三追累了

Zhangsan chase-tired-PERF ‘Zhangsan chased (someone) and became tired.’

Prediction: *zhui* 追 is Vt and *lei* 累 is Ri. Since *Zhangsan* 張三 can be the subject of *lei* 累, *zhui-lei* 追累 is VRi and SC. (Affecting direction, as predicted by (c) and (d), is ambiguous between VRt and OC and VRi and SC; it is disambiguated by the sentential context⁷.)

- (7) *Zhangsan ca-liang le boli* 張三擦亮了玻璃

Zhangsan wipe-shiny-PERF glass

‘Zhangsan wiped the glass shiny.’

Prediction: *ca* 擦 is Vt and *liang* 亮 is Ri. Since *Zhangsan* 張三 cannot be the subject of *liang* 亮, and the object of *ca* 擦 can be the subject of *liang* 亮, *ca-liang* 擦亮 is VRt and OC. (Rule (b))

⁷ As mentioned in Section 2.2, the VR is polysemous between an OC reading (i.e. *zhui-lei*₁ 追累₁) and an SC reading (i.e. *zhui-lei*₂ 追累₂). Nevertheless, while the heuristics are not sufficient to decide the relevant reading, the sentential context of (6) requires the SC reading. The case of (15) is similar to (6) except that, in the former, the relevant rule (i.e. rule (f)) fails to recognize ambiguities.

- (8) *Zhangsan qie-dun-le dao* 張三切鈍了刀
 Zhangsan cut-blunt-PERF knife
 ‘Zhangsan cut (something) with the knife, and as a result the knife became blunt.’
Prediction: *qie* 切 is Vt and *dun* 鈍 is Ri. Since *Zhangsan* 張三 cannot be the subject of *dun* 鈍, and the adjunct of *qie* 切 can be the subject of *dun* 鈍, *qie-dun* 切鈍 is VRt and OC. (Rule (b))
- (9) *Zhangsan du-dong-le na-shou shi*. 張三讀懂了那首詩
 Zhangsan read-understand-PERF that-CL poem
 ‘Zhangsan read and understood that poem.’
Prediction: *du* 讀 is Vt and *dong* 懂 is Rt; *du-dong* 讀懂 is VRt and SC. (Rule (a))
- (10) *Zhangsan jiao-fan-le Lisi* 張三教煩了李四
 Zhangsan teach-vexed-PERF Lisi
 ‘Zhangsan taught Lisi and as a result Lisi felt vexed.’ (教煩₁, OC reading)
 ‘Zhangsan taught Lisi and as a result Zhangsan felt vexed.’ (教煩₂, SC reading)
Prediction: *jiao* 教 is Vt and *fan* 煩 is Ri. Since *Zhangsan* 張三 and *Lisi* 李四 do not differ in the possibility of being the arguments of the two verbs, *jiao-fan* 教煩 is predicted to be VRt and either OC (Rule (c)) or SC (Rule (d)). (Context beyond the sentential level is needed for disambiguation.)
- (11) *Zhangsan jiao-hui-le Lisi* 張三教會了李四
 Zhangsan teach-know-PERF Lisi
 ‘Zhangsan taught Lisi (something), and as a result Lisi learned it.’
Prediction: *jiao* 教 is Vt and *hui* 會 is Rt. Since *Zhangsan* 張三 and *Lisi* 李四 do not differ in the possibilities of being the arguments of the two verbs, *jiao-hui* 教會 is predicted to be VRt and either OC (Rule (c)) or SC (Rule (d)). Nevertheless, as the affecting direction of *jiao-hui* 教會 is toward the object of V1, *jiao-hui* 會 is VRt and OC. (The affecting direction is not predicted by rules.)
- (12) *Zhangsan e-bing-le* 張三餓病了
 Zhangsan hungry-sick-PERF
 ‘Zhangsan became sick as a result of being hungry.’
Prediction: *e* 餓 is Vt⁸ and *bing* 病 is Ri. Since *Zhangsan* 張三 can be the subject of *bing* 病, *e-bing* 餓病 is predicted to be VRi and SC. (The affecting direction, as predicted by (c) and (d), is ambiguous between a VRt and SC and a VRt and OC reading; it is disambiguated by the sentential context.)

⁸ We think that there are two polysemous *e* 餓. Besides the intransitive sense, sentences like *Baba guyi e xiaohai* 爸爸故意餓小孩 ‘The father deliberately made/makes his children starve’ suggests the existence of a transitive *e* 餓.

In sentences (13)-(15), V1 is an intransitive verb:

- (13) *Zhangsan zou-lei-le tui* 張三走累了腿

Zhangsan walk-tired-PERF leg

‘Zhang walked until his legs got tired.’

Prediction: *zou* 走 is Vi and *lei* 累 is Ri. Since *Zhangsan* 張三 can be the subject of *lei* 累, *zou-lei* 走累 is predicted to be VRi and SC (Rule (f)).

- (14) *Zhangsan ku-ya-le sangzi* 張三哭啞了嗓子

Zhangsan cry-hoarse-PERF throat

‘Zhangsan cried his throat hoarse.’

Prediction: *ku* 哭 is Vi and *ya* 啞 is Ri. Since *Zhangsan* 張三 cannot be the subject of *ya* 啞, *ku-ya* 哭啞 is VRt and OC. (Rule (g))

- (15) *Zhangsan bing-huang-le Lisi* 張三病慌了李四

Zhangsan sick-nervous-PERF Lisi

‘Zhangsan's being sick got Lisi nervous.’

Prediction: *bing* 病 is Vi and *huang* 慌 is Ri. Since *Zhangsan* 張三 can be the subject of both, *bing-huang* 病慌 is predicted to be VRi and SC. Nevertheless, the affecting direction of *bing-huang* 病慌 can be either toward the subject or the object. Therefore, *bing-huang* 病慌 can be either VRi and SC or VRt and OC. (The affecting direction is not predicted by rules; it is disambiguated by the sentential context.)

The predictions of rules (a)-(g) are borne out most of the time. Sometimes, though, these heuristics cannot account for the affecting direction. We will discuss some such cases.

3. Discussion and the Applicability of the Rules

Although language processing can benefit from the heuristics, such rules sometimes fail because they are based on lexical information, to which much of the real-world knowledge, which plays an important role in deciding the affecting direction of VRs, is invisible. As opposed to the views of Li (2007) and Huang (2002), we found real-world knowledge to be no less important in the interpretation of a VR than in its possibility of existence because both a VR's occurrence and its interpretation presume a reasonable cause-effect relationship. Notably, we suggest that the subjectivity involved in such decisions makes it only natural that definite readings are sometimes non-existent. While believing that the affecting direction is largely determined by world knowledge, we also noticed that syntactic constraints regulate how real-world knowledge can be expressed.

Recall that *jiao-hui* 教會 ‘teach someone and make him/her understand’ in (11) is

predicted by (c) and (d) to be either SC or OC, but turns out to be OC only because the affecting direction of the VR is toward the object. It might be asked if one of the verbal components has contributed to the affecting direction of the VR. As far as *jiao-hui* 教會 is concerned, the following sentences show that the affecting direction is a co-product of V1 and V2.

- (16) a. *Zhangsan jiao-fan-le Lisi* 張三教煩了李四

Zhangsan teach-vexed-PERF Lisi

‘Zhangsan taught Lisi and as a result Lisi felt vexed.’ (教煩₁, OC reading)

‘Zhangsan taught Lisi and as a result Zhangsan felt vexed.’ (教煩₂, SC reading)

- b. *Zhangsan jiao-fan-le ying-wen* 張三教煩了英文

Zhangsan teach-vexed-PERF English

‘Zhangsan taught English and as a result Zhangsan felt vexed.’

- (17) *Zhangsan xue-hui-le gangqin* 張三學會了鋼琴

Zhangsan learn-know-PERF piano

‘Zhangsan learned to play the piano.’

That *jiao-hui* 教會 ‘teach someone and make him/her understand’ and *jiao-fan* 教煩 ‘teach and got (someone) vexed’ differ in the affected arguments indicates that the affecting direction at least is not determined by V1 in all instances. Likewise, a comparison between *jiao-hui* 教會 and *xue-hui* 學會 ‘learned something’ shows that the affecting direction of the VR in question is not determined by V2 alone, either. Rather, it appears that real-world knowledge as a result of the *composition* of V1 and V2 deems it unlikely to teach oneself and make oneself understand, hence the inherent OC typing of *jiao-hui* 教會.

Li’s principle of prototypicality states that the one who acts usually is not the one that is affected. While he calls it a “pragmatic principle,” we think it is still derived from experience living in the world. To the extent that the principle of prototypicality is derived from real-world knowledge is true, it is still an abstraction that competes with other kinds of real-world knowledge. Consider the following four sentences where our heuristics make nondeterministic (i.e., (18) and (20)) or wrong predictions (i.e., (19) and (21)).

- (18) *Zhangsan zhui-lei le* 張三追累了

Zhangsan chase-tired-PERF

‘Zhangsan chased (someone) and became tired.’

Prediction: *zhui* 追 is Vt and *lei* 累 is Ri. Since *Zhangsan* 張三 can be the subject of *lei* 累, *zhui-lei* 追累 is VRi and SC. (The affecting direction, as predicted by (c) and (d), is ambiguous between VRt and OC and VRt and SC; it is disambiguated by the sentential context.)

- (19) *Zhangsan jiao-hui-le Lisi* 張三教會了李四

Zhangsan teach-know-PERF Lisi

‘Zhangsan taught Lisi (something), and as a result Lisi learned it.’

Prediction: *jiao* 教 is Vt and *hui* 會 is Rt. Since *Zhangsan* 張三 and *Lisi* 李四 do not differ in the possibility of being the arguments of the two verbs, *jiao-hui* 教會 is predicted to be VRt and either OC (Rule (c)) or SC (Rule (d)). Nevertheless, as the affecting direction of *jiao-hui* 教會 is toward the object of V1, *jiao-hui* 會 is VRt and OC. (The affecting direction is not predicted by rules.)

- (20) *Zhangsan e-bing-le* 張三餓病了

Zhangsan hungry-sick-PERF

‘Zhangsan became sick as a result of being hungry.’

Prediction: *e* 餓 is Vt and *bing* 病 is Ri. Since *Zhangsan* 張三 can be the subject of *bing* 病, *e-bing* 餓病 is predicted to be VRi and SC. (The affecting direction, as predicted by (c) and (d), is ambiguous between a VRt and SC and a VRt and OC reading; it is disambiguated by the sentential context.)

- (21) *Zhangsan bing-huang-le Lisi* 張三病慌了李四

Zhangsan sick-nervous-PERF Lisi

‘Zhangsan’s being sick got Lisi nervous.’

Prediction: *bing* 病 is Vi and *huang* 慌 is Ri. Since *Zhangsan* 張三 can be the subject of both, *bing-huang* 病慌 is predicted to be VRi and SC. Nevertheless, the affecting direction of *bing-huang* 病慌 can be either toward the subject or the object. Therefore, *bing-huang* 病慌 can be either VRi and SC or VRt and OC. (The affecting direction is not-predicted by rules; it is disambiguated by the sentential context.)

As far as Li’s (2007) two principles for ranking possible readings are concerned, (19)-(22) all satisfy ASAES (Animate Subject as Agent or Experiencer Strategy) because this paper only deals with such VR patterns. The other principle, i.e., prototypicality, makes SC readings less likely. Nevertheless, SC readings are not only preferred but are the only possibility in (19) and (21). Therefore, the knowledge that “the one who acts is unlikely to be the one that is affected” is not only insufficient to rank readings but actually allows impossible readings. That (19) and (21) allow SC readings seems to come from real-world knowledge that verifies the possibility for the one who acts (i.e., *Zhangsan* 張三) to be the affected one.

A comparison between (21) and (22) suggests that the ultimate difficulty in pinpointing the most possible reading of a VR lies in the fluidity of real-world knowledge. Such fluidity sometimes comes from subjectivity. Take the readings of *e-bing* 餓病 ‘hungry-sick’ and *bing-huang* 病慌 ‘sick-nervous’ for example. It seems that, for a VR composed of two Vis,

like *bing-huang* 病慌, to have an OC reading, it must have a causative reading. As for the Vt+Vi combination *e-bing* 餓病, the OC reading would be ‘someone makes another person starve so that this starving person becomes sick.’ While both assuming a causative interpretation under the OC readings, the two VRs contrast in two ways. First, in *e-bing* 餓病, both composing verbs take on causative readings, whereas in *bing-huang* 病慌, only *huang* 慌 has a causative reading. Second, in the absence of context, the SC reading seems more natural than an OC reading for *e-bing* 餓病. Nevertheless, if such a bias is present at all, it is not as pronounced in *bing-huang* 病慌. This is unexpected in structural terms because, while both VRs have V2s that are Vis, the V1 *e* 餓 in *e-bing* 餓病 is a transitive verb, which supposedly is more likely to affect an entity other than the one who acts (thus, have an OC reading) than the intransitive V1 *bing* 病 in *bing-huang* 病慌 does. We are led to the hypothesis that the contrasts could result from judgment of plausibility based on real-world knowledge: *E-bing* 餓病 is less natural with a causative reading because people generally do not intentionally make others suffer; it is unusual to starve people. On the other hand, in the case of *bing-huang* 病慌, *bing* 病 is involuntary and cannot be put in intentional terms, while it is human to find another person’s suffering disturbing.

As can be seen, the above judgment is a reflection of the perception of the world. Thus, it is expected that the perceived probability of OC and SC readings will be as varied as people’s perception of the world. A definite OC or SC judgment sometimes could be non-existent.

Finally, we will discuss two cases that seem to be consistent with real-world knowledge but are either downright unacceptable or odd. Sentence (22) presents a reading that is a possible real-world scenario but linguistically unacceptable:

(22) *Zhangsan zhui-lei-le Lisi* 張三追累了李四

Zhangsan chase-tired-PERF Lisi

‘Zhangsan chased Lisi and as a result Zhangsan got tired.’

In terms of real-world knowledge, it is possible that Zhangsan chased Lisi and as a result Zhangsan got tired, but such a reading is not as natural as one where Lisi is tired.

Another example is the intransitive-and-SC-only reading of *chi-bao* 吃飽 ‘full from eating.’ Consider the following sentence:

(23) *?Zhangsan chi-bao-le ji* 張三吃飽了雞

Zhangsan eat-full-PERF chicken

‘*Zhangsan ate (something) and the chicken became full.’ (OC reading)

‘?Zhangsan ate the chicken and became full.’ (SC reading)

Since both *chi* 吃 ‘eat’ and *bao* 飽 ‘full (from eating)’ can take animate subjects, the VR is predicted to be transitive and either SC or OC. Nevertheless, the impossibility of

becoming full from someone else's eating rules out the transitive-and-OC reading. As for why the transitive-and-SC reading is also impossible, however, real-world knowledge does not seem to be a factor, as there is nothing unusual about eating something and becoming full as a result.

For now, we have not been able to explain the underlying motivations for every (im)possible reading. Nevertheless, the proposed rules so far have made the correct prediction most of the time in the VRs we encountered.

4. Conclusion

In the current study, we explain how we predict a VR's transitivity and identify the argument predicated by the second verb. We propose a set of rules that are motivated by a lexical-semantic analysis. Although these rules have not been tested against a corpus, they are testable and are worth testing. This is because linguistic analyses suggest their theoretical credibility and because these rules have a uniform formalism and are modular in the sense that they use two kinds of formally-represented information, i.e., 1) the transitivity of V1 and V2 and 2) whether the arguments can be the subject and object of the two component verbs or not. If these rules prove to achieve high prediction rates, they can be readily applied.

References

- Chen, K.-J., Huang, S.-L., Shih, Y.-Y., & Chen, Y.-J. (2005). Extended-HowNet-A representational framework for concepts. In *Proceedings of OntoLex 2005*, 1-6.
- Crystal, D. (1997). *Dictionary of phonetics and linguistics*. Oxford: Wiley-Blackwell.
- Huang, C.-T. J. (2002). Resultatives and unaccusatives: a parametric view. *Bulletin of the Chinese Linguistic Society of Japan*, 253, 1-43.
- Huang, H.-C. (2006). A constructional approach to argument realization of Chinese resultatives. *UST Working Papers in Linguistics*, 2, 13-31.
- Huang, S. (1998). Chinese as a headless language in compounding, In J. L. Packard (Ed.), *New approaches to Chinese word formation: Morphology, phonology and the lexicon in modern and ancient Chinese*, 261-284. Berlin: Mouton De Gruyter.
- Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. Chicago: University of Chicago Press.
- Li, C. (2009). On the headedness of Mandarin resultative verb compounds. *Concentric: Studies in Linguistics*, 35(1), 27-63.
- Li, C. (2007). *Mandarin resultative verb compounds: Where syntax, semantics, and pragmatics meet*. Ph. D [Dissertation]. Yale University, New Haven, USA. [Online]. Available: ProQuest. [Accessed: 2011/11/13].

- Li, C. N. & Thompson, S. A. (1981). *Mandarin Chinese: A functional reference grammar*. Taipei: Crane.
- Lin, H.-L. (1998). *The syntax-morphology interface of verb-complement compounds in Mandarin Chinese*. Ph. D [Dissertation]. University of Illinois at Urbana-Champaign, Urbana-Champaign, Illinois, USA. [Online]. Available: ProQuest. [Accessed: 2011/7/13].
- Mulkeen, S. (2011). A conflation account of mono-clausal resultatives in English and Chinese. MA thesis. National Cheng-Kung University. [Online]. Available: <http://www.airitilibrary.com/searchdetail.aspx?DocIDs=U0026-0802201119204700> [Accessed:2012/4/3]
- Shen, L., & Lin, T.-H. J. (2005). Agentivity agreement and lexicalization in Resultative Verbal Compounding, *Paper Space*, 2005. [Online]. Available: http://ling.nthu.edu.tw/faculty/thlin/pdf/Shen_Lin.pdf. [Accessed: 2011/7/9].
- Smith, C. C. (1994). *The parameter of aspect*. New York: Kluwer Academic Press.
- Yin, H. (2011). The semantic structures of Chinese verb-resultatives. *International Journal of English Linguistics*, 1(2), 126-133.
- 黃居仁 (2005) 主編。《意義與詞義》系列。中文的意義與詞義。中央研究院語言所文獻語料庫與資訊所中文詞知識庫小組技術報告 05-01。
- 詞庫小組。廣義知網知識本體架構線上瀏覽系統。 [Online]. Available: <http://ehownet.iis.sinica.edu.tw> [Accessed: 2012/4/3]

Appendix I

Abbreviations

[+ANI]=animate entity

[-ANI]=inanimate entity

CL=classifier

LOC=locative preposition

[+PHY]=physical entity

OC=object-controlled

PERF=perfective marker

R=result

Rt= transitive verb that denotes a result

SC=subject-controlled

Vi=intransitive verb

VRi= VR is an intransitive verb

VRt= VR is a transitive verb

Vt=transitive verb

