▶ The Words-and-Rules Theory: Evidence from Chinese Morphology

漢語述賓式合成複合詞對詞項與規則理論的映証

THE WORDS-AND-RULES THEORY:
EVIDENCE FROM CHINESE MORPHOLOGY[*]

Yuanjian He

ABSTRACT

Chinese OVS compounds are generated by morphological rules, whereas VOS compounds are far less homogenous. The VO part is a root word if the V and O are monosyllabic, or a VP if the V and/or O are disyllabic. In the former case, the syllable count may have conditioned a pattern-associated memory that has transformed a VP into a root. In the latter case a lexico-syntactic compound has been created. These findings are in agreement with the principle of memory vs. symbol-processing proposed by Pinker (1999) and provide proof that the principle works in more areas than inflectional morphology and may indeed be universal.

1. INTRODUCTION

Some Chinese synthetic compounds involving a Verb-Object relation appear to form the paradigm in (1), which takes into account: (a) the number of syllables of the constituent stems acting as V and O, and (b) the word order:

(1)
	Monosyllable-Stemmed V&O	Disyllable-Stemmed V&O
VOS:	+	+
OVS:	-[1]	+

---

[1]   With a few exceptions, e.g., "*rou shi zhe* (肉食者), meat-eater".

V stands for verb, O for object, and S for subject. "+" indicates the existence of a relevant construction, and "-" the non-existence of a relevant construction.

There are two types of contrast, one between word orders, i.e., VOS vs. OVS, and the other between syllable counts, i.e., *disyllabic* V's and O's vs. *monosyllabic* V's and O's in VOS compounds. Examples of VOS and OVS compounds with *disyllabic* V's and O's are:

(2) VOS:

[製造謠言]者	[传播病毒]者	[盜竊國寶]犯
[V O]S	[V O]S	[V O]S
[zhizao yaoyan] zhe	*[chuanbo bingdu] zhe*	*[daoqie guobao] fan*
[make rumour] person	[spread virus] person	[steal state-treasure]
- rumour-monger	- virus-spreader	criminal
		- thief of state-treasures

OVS:

[謠言製造]者	[病毒傳播]者	[國寶盜竊]犯
[O V]S	[O V]S	[O V]S
[yaoyan zhizao] zhe	*[bingdu chuanbo] zhe*	*[guobao daoqie] fan*
[rumour make] person	[virus spread] person	[state-treasure steal]
- rumour-monger	- virus-spreader	criminal
		- thief of state-treasures

And examples of VOS compounds with *monosyllabic* V's and O's are:

(3) VOS:

[造謠]者	[畢業]生	[攝影]師
[V O] S	[V O] S	[V O] S
[zao yao] zhe	*[bi ye] sheng*	*[she ying] shi*
[make rumour] person	[finish studies] student	[take photo] master
- rumour-monger	- graduate	- photographer

[探險]家	[播音]員	[理事]長
[V O] S	[V O] S	[V O] S
[tan xian] jia	*[bo yin] yuan*	*[li shi] zhang*
[explore danger] expert	[spread sound] clerk	[execute business] chief
- explorer	- broadcaster	- chief executive

Note that any VOS compound with a *disyllabic* V and O can be reserved to OVS order, as shown in (2), but not necessarily vice versa. In fact, many OVS compounds do not have VOS counterparts, e.g., "*yizhu zhixing ren / *zhixing yizhu ren* (遺囑執行人／*執行遺囑人), will-executor"; "*feiji sheji shi / *sheji feiji shi* (飛機設計師／*設計飛機師), aircraft designer". However, such cases are not our concern here.

In addition, the syllable count which is relied upon in the Paradigm (1) applies strictly to such paradigm only. In fact, there are VO's where V is monosyllabic but O disyllabic, as in "*mai boli zhe* (賣玻璃者), glass-seller", or vice versa, as in "*langfei qian zhe* (浪費錢者), money-waster". How these variations in syllable count compare with those in Paradigm (1), which dwells on disyllabic and monosyllabic V's and O's only, requires the formation of an extended paradigm that is also beyond our concern here.

Our current concerns are, first, VOS compounds in (2) that contrast with their OVS counterparts in the ability to inflect for plurality, as in: [2]

(4) VOS: *製造謠言者們   *傳播病毒者們   *盜竊國寶犯們
       [V   O   S]-*men*   [V   O   S]-*men*   [V   O   S]-*men*
            plural   plural   plural

    OVS: 謠言製造者們   病毒傳播者們   國寶盜竊犯們
       [O   V   S]-*men*   [O   V   S]-*men*   [O   V   S]-*men*
            plural   plural   plural
       - rumour-mongers - virus-spreaders - thieves of
                                      state-treasures

VOS compounds do not, as we see, inflect for plurality, as ungrammaticality results from attaching the plural suffix "-*men*" to them. In direct contrast, OVS compounds *are* able to take the plural suffix "-*men*". We want to know why.

Second, though VOS compounds with *disyllabic* V's and O's do not inflect as (4) shows, those with *monosyllabic* V's and O's do, as shown below using the first row of items of (3) as example:

---

[2]   In Chinese, the plural suffix "-*men*" is inflected for human nouns, and is mainly for stylistic purposes, because, as is well known, Chinese nouns as a category do not differentiate singular from plural. Nonetheless, the "-*men*" plural suffix does represent inflectional morphology in Chinese.
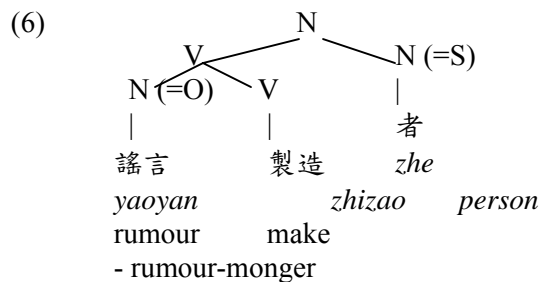
(5) VOS: 造謠者們　　　　畢業生們　　　　攝影師們
　　　　　[V O S]-*men*　　[V O S]-*men*　　[V O S]-*men*
　　　　　　　plural　　　　　　plural　　　　　　plural
　　　　　- rumour-mongers　- graduates　　　- photographers

We also want to know why this is so. In the following, I will examine the above-observed two issues one by one and draw a principled account for them based on Pinker's (1999) principle of memory vs. symbol-processing for linguistic constructs across languages. In this context, I further propose a construal of logical form for synthetic compounds in order to dissolve the seemingly disparity between synthetic compounding in Chinese where thematicity and structural integrity go hand in hand, *and* that in English (taken as representing another language in general) where thematicity and structural integrity may split. Other related empirical issues are also discussed.

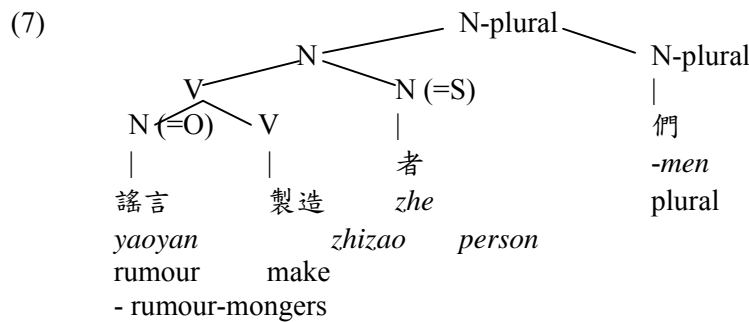## 2. LEXICAL OVS COMPOUNDS VS. LEXICO-SYNTACTIC VOS COMPOUNDS

The answer as to why OVS compounds in (4) inflect but VOS ones do not lies precisely in the run of the word order. OVS compounds reflect the truest sense of endocentric compounding, for they comply strictly with a peripheral headedness rule, in this case, the Right-Hand Head Rule (Williams 1981: 248). By the RHR, which entails the lexical rule X → Y X, the right-hand stem determines the category of the resultant constituent in endocentric compounding, and hence is the head. That is to say, the OVS compound commands the following structure:

(6)
```
                    N
            V              N (=S)
      N (=O)    V            |
        |       |            者
       謠言     製造         zhe
      yaoyan   zhizao      person
      rumour    make
      - rumour-monger
```

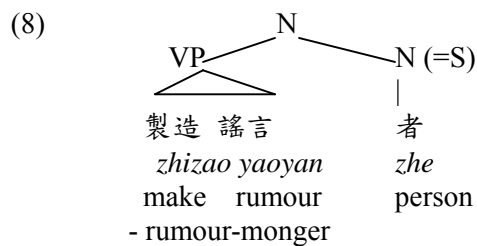*Zhe*（者）in (6) is the head of the compound. Whether it is a free or bound

item is not within the concern of the RHR, which treats free or bound items alike as long as they both have the categorial capacity to determine the category of the whole compound. (6) is thus taken as a canonical lexical structure and is, presumably, able to inflect for plurality. When this happens, the plural suffix "-*men*" will be the head of the compound, as shown in (7) below, where, according to the RHR (also see Di Sciullo and Williams 1987: 25; Katamba 1993: 313), inflectional affixes will mark the category of the whole word or compound:

(7)

```
                              N-plural
                    N                      N-plural
            V             N (=S)              |
      N (=O)    V           |                門
        |       |          者               -men
       謠言    製造        zhe              plural
      yaoyan   zhizao    person
      rumour   make
      - rumour-mongers
```

Note that the symbol "plural" is not just symbolic. It means that by virtue of the RHR, the compound is a plural noun and will get plurality interpretation when it enters in the Logical Form (LF). Without it, the compound is just an N, violating the spirit of the principle of Full Interpretation (Chomsky 1991: 441-442, 1993: 26-27, Chomsky & Lasnik 1995: 27).

VOS compounds on the other hand *cannot* be of a canonical lexical structure, because their VO sequence is of the canonical word order of a VP in Chinese, such as in:

(8)

```
                    N
         VP                N (=S)
                             |
     製造 謠言               者
    zhizao yaoyan          zhe
    make   rumour          person
      - rumour-monger
```

As a VP, the VO in (8) will differ in its grammatical properties from the OV in (6). For instance, it may take adjuncts modifying either the V or

the O:[3]

(9) a. [積極製造謠言]者
         [jiji zhizao yaoyan] zhe
         [actively make rumour] person
     b. [製造許多謠言]者
         [zhizao xuduo yaoyan] zhe
         [make many rumour] person
         a&b: wild rumour-monger

Just as a VP does in a sentence:

(10) a. 他[積極製造謠言]。
          Ta [jiji zhizao yaoyan]
          3sg [actively make rumour]
      b. 他[製造許多謠言]。
          Ta [zhizao xuduo yaoyan]
          3sg [make many rumour]
          a&b: He is a wild rumour-monger.

In contrast, the OV sequence in (6) cannot take similar adjuncts:

(11) a. *[謠言積極製造]者
          [yaoyan jiji zhizao] zhe
          [rumour make] person
      b. *[許多謠言製造]者 [4]
          [xuduo yaoyan zhizao] zhe
          [many rumour make] person

　　But how does a VP become a stem in a compound? It does so by being looped back from syntax, according to Pinker (1999: 205). The loop theory, so to speak, is based on the way lexicon and syntax are connected, which according to Kiparsky (1982) is roots→complex word formation→regular inflection→syntax.[5] Because syntactic structures

---

[3]　　In glossing Chinese, 3sg = third person singular, cl = classifier.

[4]　　It can mean "many rumour-mongers", which is, however, not what it is supposed to mean.

[5]　　The connection does not mean that every word form has to go through all the

are generated in syntax and when they show up in complex word formation, e.g., as a stem in compounding, they are taken to be looped back from syntax. Proof for such looping is that a compound that contains an XP as stem no longer inflects well (Pinker 1999: 200-208). Taking examples from English as illustration:

(12) admissions committee   *vs.*   ?admissions committees
   benefits cut   *benefits cuts
   cuts package   *cuts packages
   enemies list   *enemies lists
   grades meeting   *grades meetings
   injuries reports   *injuries-reports
   landmarks commission   *landmarks commissions
   records department   *records departments
   skills gap   *skills gaps
   twins project   *twins projects

The left stem is identified as an NP headed by a plural noun, rather than just a "root+plural", the reason being that regular inflection applies after complex word formation, as is stated above. In other words, this stem with a plural affix would have entered in syntax first, before it could be looped back to be part of a compound. Once in syntax, an item is projected into an XP, as is generally assumed, in this case an NP. A consequence of such looping is that the compound with an NP-stem may no longer inflect well, as the right column of (12) shows. Of course, not every compound may take an XP-turned stem, but for those that do, their XP-turned stem does appear to bar regular inflection applied to the whole compound.

  The same is true of Chinese based on the examples we have seen in this section. The fact that a VOS compound refuses the plural inflectional morphology indicates that it contains a VP-stem, as illustrated in (8). In contrast, OVS compounds are able to inflect because they are canonical lexical structures, as shown in (6). Hence this is the cause for the word order contrast between them.

---

phases. It simply means that if a complex form of a word, e.g., a compound with a regular inflectional morphology has shown up in sentences, it would have gone through all the phases. There are a number of other routes for other word forms. E.g., root words can go to syntax direct, root words with a regular inflectional morphology would bypass the phase of complex word formation, and so on.

3. VOS COMPOUNDS WITH VO AS A ROOT

Now we examine VOS compounds with *monosyllabic* V's and O's, examples of which we see in (3), and which we see are also able to inflect in (5). Question is why they can inflect, whereas their *disyllabic* (V's and O's) counterparts in (4) cannot.

As is established above, the reason for which the VOS compounds with *disyllabic* V's and O's do not inflect is because they contain a VP-stem, i.e., the VO sequence is a VP, as shown in (8). In this context, the VOS compounds with *monosyllabic* V's and O's cannot possibly contain a VP-stem, because they do inflect as shown in (5), suggesting that the VO sequence here is not a VP.

A further piece of evidence for the VO sequence in question not being a VP is in relation to its referential property. To see how, let us again look at the compounds with *disyllabic* V's and O's in (2). There, both VOS and OVS compounds may carry a specific reference, but only the OVS ones can be also generic. Compare:

(13)   a.   她是[制造谣言者]。
              Ta shi [zhizao yaoyan zhe].
              3sg be make rumour person
              - She is *the* rumour-monger.
        b.   *她是一个[制造谣言者]。
              Ta shi yi-ge [zhizao yaoyan zhe].
              3sg be one-cl make rumour person

(14)   a.   她是[谣言制造者]。
              Ta shi [yaoyan zhizao zhe].
              3sg be rumour make person
              - She is *the/a* rumour-monger.
        b.   她是一个[谣言制造者]。
              Ta shi yi-ge [zhizao yaoyan zhe].
              3sg be one-cl rumour make person
              - She is *a* rumour-monger.

"Rumour-monger" in VOS order refers to "someone who has spread a particular type of rumour", whereas that of OVS order is able also to refer to "any rumour-mongers". The ability, or non-ability, of these compounds to take a numeral-classifier modifier in front of them has made this clear.

Then further consider:

(15)   a. 她是[造谣者]。
          Ta shi [zaoyao zhe].
          3sg be make-rumour person
          - She is *the/a* rumour-monger.
    b. 她是一个[造谣者]。
          Ta shi [zaoyao zhe].
          3sg be one-cl make-rumour person
          - She is *a* rumour-monger.

The VOS compound with a *monosyllabic* V and O behaves more like the OVS compound in (14), rather than the VOS compound with a *disyllabic* V and O in (13).

The paradox is apparent. On the one hand, the *monosyllabic* VO sequence appears typical of a VP construction, with its syntax-oriented word order and unmistakable verb-object relation. On the other, it cannot be identified as a VP in a VOS compound, because the compound inflects well, as shown in (5), and can be generic. Both properties indicate that the VP in question is a root word. But how can an item of a VP-like appearance become a root word?

Again I refer to Pinker's (1999) minimalist perspective to language, one that is not dissimilar to that of Chomsky (e.g., 1995, 1998). Namely, language essentially contains a package of root words and a set of rules. When rules are applied to the roots, the process then produces complex forms of words and syntactic structures. Pinker, however, emphasizes the psycho- and neurolinguistic aspects of how the rules may or may not apply. In essence, rules stop where memory takes over. Memory is where all the roots including irregular word forms are stored as symbols, and when symbols are accessed and processed, e.g., retrieving a root and attaching an affix to it, it is rule-application and involves the computing act of the mind. As access to memory does not require as much computing as does symbol-processing, it is therefore more economical in terms of grammatical operations.

Pinker's core case is inflectional morphology where the familiar paradigm of "irregulars vs. regulars" fits well in the division of labour between memory and symbol-processing. But there are two further complications. One, the division of labour has to be evolutionary, since which linguistic forms are memory-stored and which are the products of

symbol-processing are not fixed forever. Spoken French, for instance, expresses plurality by inflecting its articles, while at the same time keeping a group of irregular plural nouns. In other words, rules apply to French articles with respect to plurality. In this sense, the so-called "regular" plural nouns in French are in fact not so regular at all, for they do not have rules applied to them, and therefore are root words accessible directly from memory, just like the irregular plural nouns. Moreover, French previously used to inflect its regular nouns for plurality, suggesting that rule-application as far as to regular plural nouns is concerned has diminished in French and memory has since taken over. The silent "-*s*" that still occurs in written forms of French regular plural nouns is testimony to that. It simply shows that a linguistic form that was a product of symbol-processing in history may have become memory-stored through time.

Another complication concerns how the division of labour between memory and symbol-processing may operate in non-inflectional morphology. For instance, "What the hell do the Chinese speakers do," to quote a critic (from Pinker 1999: 263-264), "with their morphology genes or their dedicated neural mechanisms for regulars vs. irregulars?"

I thus return to the paradigm of Chinese compounds we have been investigating. Specifically, we may call the VOS compounds with disyllable-stemmed V's and O's *Compound A*, and those with monosyllable-stemmed V's and O's *Compound B*. Compound A is a "regular", i.e., its VO sequence is generated by rules, whereas Compound B is an "irregular", i.e., its VO sequence behaves like a root word and hence is likely to be memory-stored. A possible explanation is that the monosyllable-stemmed VO form in Compound B is like a frozen syntactic-structure, probably once rule-generated in history but having evolved since to become a root, just like French regular plural nouns. It is thus another example where rules diminished and memory has taken over. Leaving historical issues aside, I will give evidence below for this monosyllable-stemmed VO form being a root word in present-day Chinese.

4. THETA-ROLE ALIGNMENT AND ABSORPTION

As is well known, the relative structural positions of nominal elements with respect to the verb in a grammatical construction are the essence of the theta theory (e.g., Chomsky 1986: 97-98; Grimshaw 1990: 8;

Williams 1994: 36). The Uniformed Theta Assignment Hypothesis (UTAH) (Baker 1988: 46, 1997: 74) makes the notion more precise by stipulating that thematic elements occupy exactly the same structural positions with respect to the verb both in syntax and in morphology. This means that irrespective of the surface word orders in syntax and morphology of a language, the underlying structural positions of thematic elements with respect to the verb are the same everywhere.

For instance, we may have a perfectly matched syntax vs. morphology paradigm as in:

(16)　　a.　[教師[指導論文]]
　　　　　　　S　　V　　O
　　　　　　jiaoshi zhidao lunwen
　　　　　　teacher supervise thesis
　　　　　　- Teachers supervise theses.
　　　　b.　[[論文指導]教師]
　　　　　　　O　　V　　S
　　　　　　lunwen zhidao jiaoshi
　　　　　　thesis supervise teacher
　　　　　　- thesis supervisor

Under UTAH, a paradigm such as (16) invokes "The Mirror Principle", whereby morphological derivations must directly reflect syntactic derivations (and vice versa) (Baker 1985). In other words, the constituency of SVO in syntax and that of OVS in morphology in (16) are inherently determined by UTAH, namely, the verb assigns the same theta role to the same structural position in syntax as well as in morphology. As a result, any alternative constituency, e.g., [論文[指導教師]] (*[lunwen [zhidao jiaoshi]],* [thesis [supervise teacher]]), is ruled out. In addition, whether it is SVO in syntax or OVS in morphology, each thematic element clearly occupies an independent structural position. I call this kind of match between a theta role and a structural position "theta-role alignment".

In other cases in morphology, however, we do not have an alignment. Take English as an example:

(17)　a.　　thesis supervisor
　　　b.　　property buyer
　　　c.　　house keeper

d.　car manufacturer

If these are compounds, then the Agent does not hold an independent structural position, because the morpheme "*-or/-er*" is part of a stem. In other words, these compounds are simply of the structure:

(18)　　　　　N
　　　　　　／＼
　　　　N　　　N
　　　　|　　　|
　　　thesis　supervisor
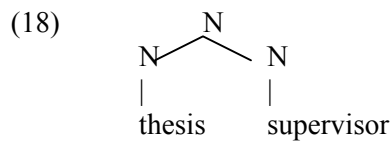
Cases like this where a thematic element is contained within a stem may be called occurrences of "theta-role absorption". The thematic element is absorbed, as it were, by the verb because it no longer needs to occupy an independent structural position, in contrast to theta-role alignment where a thematic element is not part of a stem and has to occupy an independent structural position. An example of such absorption vs. alignment in English is "*thesis supervisor*" vs. "*thesis- supervising person*".

　　Leaving aside for the moment the exact mechanism that governs theta-role absorption, to which I will return in the next section, the immediate relevance of the theta-role absorption is that when it occurs in a compound, the affected stem is a single item. In English, stems like "*supervisor*", "*buyer*", "*keeper*", and "*manufacturer*" may involve derivation. But stems absorbing a thematic element in other languages may not necessarily do so.

　　This brings us back to the Chinese VOS compounds with *monosyllabic* V's and O's. We already know that the VO sequence here is not a VP. Therefore, it has to be a root word. Question is proof, which comes in with theta-role absorption. Consider:

(19)　新聞<u>播音</u>員　　　　　電影<u>攝影</u>師
　　　[O 　<u>VO</u>　S]　　　 [O 　<u>VO</u>　S]
　　　xinwen boyin yuan　　　 *dianying sheying shi*
　　　news spread-sound person　 film take-photo master
　　　- news broadcaster　　　　 - cinema photographer

　　　電影<u>製片</u>人　　　　　電視劇<u>編劇</u>人
　　　[O 　<u>VO</u>　S]　　　 [　O 　<u>VO</u>　S]
　　　dianying zhipian ren　　 *dianshiju bianju ren*
　　　film make-film person　　 TV-drama write-play person
　　　- film producer　　　　　 - screen-writer for a TV drama series

In these OVS compounds, the V itself is part of a VO sequence with monosyllabic V's and O's. As a result, the compounds appear to have double Theme roles, which are only structurally possible if the O within the VO sequence is an absorbed thematic element, i.e., a Theme absorbed by the verb, while the left-hand O occupies an independent structural position. If this is the case, then the VO sequence is a single stem, or a root word in other words.

Further proof for the VO sequence in question being a root word with its O being absorbed by the V comes from *referential extraction*, a discoursal process where an item in either syntax or morphology is substituted for by a question word in an echo question. We find that while the VO sequence in question is referentially extractable, suggesting it be a single item, its O is not, indicating it be absorbed as predicted.

As is well known, an object in syntax is usually extractable either by movement or by substitution of a question word in a discourse. E.g.,

(20) a.　　她不播新聞。
　　　　　Ta bu bo xinwen.
　　　　　3sg not broad news
　　　　　- She does not broadcast news.
　　 b.　　新聞，她不播＿＿。
　　　　　Xinwen ta bu bo.
　　　　　news 3sg not broad
　　　　　- News, she does not broadcast.

(21) A: 她不播新聞。
      Ta bu bo xinwen.
      3sg not broad news
      - She does not broadcast news.
      B: 她不播什麼？
      Ta bu bo shenme?
      3sg not broad what
      - She does not broadcast what?

The object in B is either extracted from its original argument position after the verb, or replaced by a question word which carries a deictic co-reference from it. Hence the term *referential extraction*. Such co-reference happens only in echo questions. Compare "What does she broadcast?", which carries no such co-reference.

An object is, however, not extractable from a compound by movement, owing to its lexical integrity. E.g.,

(22) A: 她不是[新聞播音員]。
                 [O   VO   S]
      Ta bu shi xinwen boyinyuan.
      3sg not be news broadcaster
      - She isn't a news broadcaster.
      B: *新聞　她不是[＿播音員]。
      Xinwen ta bu shi boyinyuan.
      news 3sg not be broadcaster

It would stretch our intuition a little for B to mean "As for news broadcasting, she isn't a broadcaster". Hence an extraction from A is not construable.

Referentially, though, the object in a compound can be extracted:

(23) A: 她是[新聞播音員]。
                 [O   VO   S]
      Ta shi xinwen boyinyuan.
      3sg be news broadcaster
      - She is a news broadcaster.

B:　甚麼播音員？
Shenme boyinyuan?
what broadcaster
- A what broadcaster?

A:　新聞播音員。
Xinwen boyinyuan.
news broadcaster
- News broadcaster.

Now consider the O in such VO sequences as in:

(24) A:　她是[播音員]。
　　　　　　[<u>VO</u> S]
Ta shi boyinyuan.
3sg be broadcaster
- She is a broadcaster.

B:　*播甚麼員？
Bo shenme yuan.
broadcast what –er [6]

The O cannot be referentially extracted. Instead, we may have:

---

[6]　　Note that it is possible to have:
A:　　她是[播音員]。
Ta shi boyinyuan.
3sg be broadcaster
- She is a broadcaster.
B:　播甚麼音？
Bo shenme yin.
broadcast what sound
- What does she broadcast?
C:　播新聞。
Bo xinwen.
broadcast news
- News.

(25) A: 她是[播音員]。
            [<u>VO</u> S]
            Ta shi boyinyuan.
            3sg be broadcaster
            - She is a broadcaster.
      B: 甚麼員？ 她是甚麼？
            Shenme yuan? *Ta shi shenme*?
            what –er   3sg be what
            - She is what?
      A: 播音員。
            Boyinyuan.
            broadcaster
            - A broadcaster.

Either the VO or the VOS (i.e., the compound) as a whole has to be referentially extracted, suggesting that either is a single item, which points to the structure [[VO]S]. Coupled with the fact that the O of the VO is not extractable on its own as shown in (24), it is logical to conclude that the O must have been absorbed by the verb. These facts, again, suggest that the VO sequence is a root word.

However, B is not referential extraction, but rather a direct question, as is made explicit by the English translation as well as by C. Technically, B is not an echo question; the "播" (broadcast) in it is the predicate verb rather than part of a word as is in A. The fact that B can be reconstructed only as in (i) below, but not (ii) or (iii), demonstrates the point:

(i)   她播甚麼音?
        Ta bo shenme yin.
        3sg broadcast what sound
        - What does she broadcast?
(ii)   *她是[播甚麼音員]?
        Ta shi bo shenme yin yuan.
        3sg be broadcast what sound person
(iii)   ?她是[播甚麼音]?
        Ta shi bo shenme yin yuan.
        3sg be broadcast what sound person
        - What is it that she broadcasts?

(i) is B, with the pro-drop subject restored. (ii) is a failed referential extraction. (iii) could be taken as B with a focused predicate, but again not as a referential extraction.

5. LOGICAL FORM FOR SYNTHETIC COMPOUNDS AND RELATED ISSUES

From the above, it appears that only thematic elements occupying independent structural positions are subject to UTAH, whereas an absorbed theta-role is not. This seems clear both with the English compounds in (17) and with the Chinese ones in (19). The status of an absorbed theta-role relates to how thematicity is reconciled with the structural integrity of a grammatical construction. Existing treatments are, for example, the proposal of argument- linking in compounds of Lieber (1992), and that of lexical decomposition of Hale and Keyser (1993). Lieber's analysis is less structurally transparent but preserves more the structural integrity of a compound. The opposite is true of the Hale and Keyser analysis. So, for example, "*thesis supervisor*" could be decomposed as *[[thesis supervis] –or]* by breaking up a stem. However, there are cases both within as well as outside English that do not demand decomposition at all. Thus, the spirit seems to be how to capture the unity of thematicity and structural integrity in some cases, and the seemingly separation of the two in others.
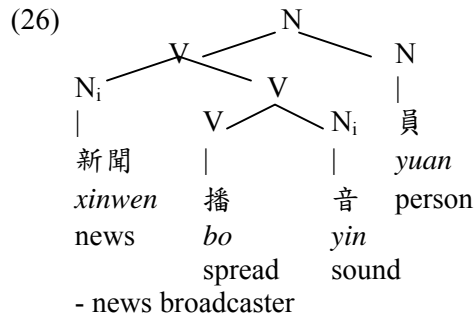
A conceivable alternative which I propose is to devise a logical form for synthetic compounds in which some kind of thematic binding operates. As we observe, the type of theta-role absorption in Chinese differs from that in English. For Chinese compounds in (19), the absorbed theta-role poses no problem for UTAH, because, as illustrated in (6), the relevant thematic elements, i.e., O and S, occupy different structural positions as required, and the absorbed theta-role has no active role to play. One way to explain this condition is to say that the absorbed theta-role has become "defective", as it were, since it is already part of a single word.

In English, however, an absorbed theta-role that becomes part of a single word is not defective at all. The absorbed theta-role in the English compounds in (17), represented in (18), does assume an agentive role, but does not occupy a structural position distinct from other theta-role(s) in the compound, thus posing a problem for UTAH.

In such a case, how to capture the seemingly opposing situation of an absorbed element across Chinese and English? It then comes to the logical form which I propose for synthetic compounds. Of the Chinese compounds in (19), the absorbed theta-role has, as it were, a thematic binder:
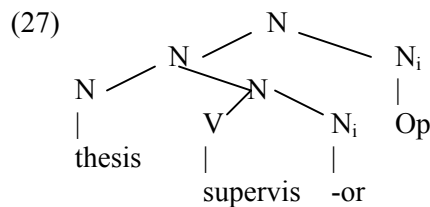
17

(26)

```
                        N
              V                   N
     N_i           V              |
     |         V       N_i        員
     新聞      |       |          yuan
     xinwen   播       音         person
     news     bo       yin
                       spread    sound
     - news broadcaster
```

The verb "播音" (*boyin*, spread-sound, i.e., *broadcast*) is decomposed for the benefit of the thematic binding, but not to the extent of breaking its or the compound's constituency. (26) is thus the lexical as well as the logical form of its own compound. Now on this account, for an absorbed theta-role in Chinese, UTAH is not a problem precisely because the absorbed element has a thematic binder.

Extending logical forms to synthetic compounds in English, we thus provide a solution to the problem that an absorbed theta-role poses to English synthetic compounds such as those in (17). We may assume that those compounds have a logical form in which a null thematic operator that does the binding, such as:

(27)

```
                     N
              N              N_i
     N            N          |
     |        V       N_i    Op
     thesis   |       |
              supervis -or
```

Assuming no movement in lexical structures, the thematic operator will bind the absorbed agentive morpheme further down the structure. As a result, UTAH is satisfied. In other words, the English compound has a lexical structure in (18), and a logical form in (27). Should a plural compound be represented in logical form, e.g., "*thesis supervisors*", the plural morpheme "-*s*" will appear before the thematic operator, e.g., $[_N [_{N\text{-plural}} [_N [\text{thesis}] [_N [\text{supervis}][\text{-or}_i]]] \text{-s}] Op_i]$, simply because logical form is construed after syntax (and hence after compounding and inflection in the lexicon).

The essence of the above proposal for synthetic compounds to be

assigned with logical forms is to say that within as well as across languages, UTAH is satisfied either in morphology or in LF with respect to synthetic compounds. Within English, for example, there is both "*thesis supervisor*" as well as "*thesis-supervising person*". The former satisfies UTAH in LF with a logical form in (27), and the latter satisfies UTAH in morphology, namely, *[[thesis-supervising] person]*. In Chinese, UTAH is mostly morphologically met, as we see in (6) and (26), with a few exceptions of morpho-syntactic cases as shown in (8), where the object of the verb would have to satisfy UTAH in syntax (before the VP is looped back to the lexicon for compounding) and the subject complies with UTAH in morphology. Alternatively, we may assume that all synthetic compounds have a logical form and therefore satisfy UTAH in LF. Needless to say, the issue of thematicity vs. structural integrity needs to be better understood in future studies.

Note that the logical forms which we have proposed above for Chinese synthetic compounds on the one hand and for English (here representing another language in general) on the other hand are simply reflecting respective empirical facts. In English, theta-absorption is resulted from derivation, while in Chinese, which basically lacks derivation, theta-absorption is performed by other means, such as lexicalization of syntactic constructs, a grammatical process which is conditioned by the syllable count (i.e., disyllabic) and performed probably over a long period of time, and which we believe is related to pattern-associated memory as Pinker (1999) advocates. Such is the empirical basis for the proposed logical forms and, more importantly, for the structural integrity of relevant compounds in respective languages. For example, the detailed lexical structure for English "*thesis supervisors*" is [$_N$ [thesis] [$_N$ [supervis][-or]]], where derivation joins the verb and the agent together. Its Chinese counterpart, however, has the structure [$_N$ [$_V$ [lunwen] [zhidao]] [zhe]] (*thesis-supervise person*, 論文指導者), where compounding operates to join the verb and the theme first, and then this theme-verb constituent with the agent. Again, the Chinese compound constitutes a straightforward case that satisfies UTAH, while its English counterpart resorts to logical form, as in (27), for that purpose.

One might still ask, however, whether the Chinese compound could be of the same structure as in English, after all, we may have [$_N$ [zhidao] [zhe]] (*supervise person*, 指導者) independently, which shows *zhidao-zhe* is a constituent, and we are not able to have [$_V$ [lunwen]

[zhidao]] (*thesis-supervise*, 論文指導) independently as a verb, as in
**Ta lunwen-zhidao-guo liang ci* (*她論文指導過兩次*, **She thesis-supervised twice*), which seems to show *lunwen-zhidao* is not a constituent. The issue rests, again, with the operating systems of grammar (here referring to Universal Grammar, in one or another skeletal version proposed by Chomsky, e.g., 1991, 1995) as well as Chinese language-specific features. As said in the early part of Section 4.0, the system requires across-the-board structural conformity with thematicity (as stated in UTAH) when constituent structures are formed during the computation stage, subject only to language-specific conditions. So, for Chinese [$_N$ [zhidao] [zhe]] (*supervise person*, 指導者) as well as English [$_N$ [supervis] [-or]], the sole thematic element, i.e., agent, is targeted and merged with the verb. The only difference between the two is that "*zhe*" is a bound root in Chinese but "*–or*" a suffix in English, so that the former is compounded, so to speak, to the verb, and the latter suffixed to its verb. Such language-specific conditions do not matter (in so far as this present comparison goes) until the verb subcategorizes for a second thematic element, i.e., the theme, as in [$_N$ [thesis] [$_N$ [supervis][-or]]] and [$_N$ [$_V$ [lunwen] [zhidao]] [zhe]] (*thesis-supervise person*, 論文指導者). Here, these constituencies are justified on the grounds that the Chinese structure ensues from UTAH, a UG condition, and the English one follows suffixation, a language-specific requirement, which defies UTAH in morphology, and consequently UTAH has to be satisfied in LF. Such account is, I believe, appropriate, till evidence of the contrary.

As for why [$_V$ [lunwen] [zhidao]] (*thesis-supervise*, 論文指導) cannot be used independently as a verb, it is simply not a legitimate linguistic entity. From the point of view of the generative systems, such a string is formed possibly as an error of Spell-Out, perhaps during the computation of [$_N$ [$_V$ [lunwen] [zhidao]] [zhe]] (*thesis-supervise person*, 論文指導者). Similarly, any partial string that is formed from the process of constructing a sentence would not be appropriate as a legitimate linguistic entity. However, we do say, for example, *Lunwen zhidao shi yi-xiang jianku-de gongzuo* (*Supervising theses is hard work*, 論文指導是一項艱苦的工作), in which case a nominal structure is probably in order, e.g., [$_N$ [lunwen] [zhidao]] (*thesis-supervision*, 論文指導), where either "*zhidao*" is a noun by its crossing from the verb category, or the structure is left-headed. Which is more appropriate may

ensue from further investigations.

Return to the issue of theta-absorption. In English, as we saw earlier, derivation causes theta-absorption and therefore represents a language-specific feature in English morphology. In Chinese, theta-absorption is caused by lexicalization of syntactic constructs, which therefore represents a language-specific feature in Chinese morphology. More specifically, when a syntactic construct is lexicalized, the once-rule-generated construct is then turned into a root word, and such lexicalization process is, as is well-known, associated with syllable counts, specifically, with a disyllabic count. That is, the prospective construct must be *disyllabic as a whole*, such as the VO construction embodied by "*boyin*" (broadcast, 播音) and the like, see (3). In Pinker's (1999) terms, the specific grammatical requirement(s) for a change-of-state process, i.e., to change from rule-generated constructs to root words, are signs of patterned memories at work. In other words, for the Chinese case at hand, disyllabic counts are linguistic forms that will condition the memory system in such a "patterned" way that the relevant constructs will become root words over time. If this is true, such disyllabic-count-induced lexicalization process will have serious and worth-while implications for Chinese compounding.

Namely, not just disyllabic syntactic units (such as VO constructs) may have been lexicalized, but all other (rule-generated) disyllabic units as well, irrespective of their syntactic or morphological nature and regardless of their internal grammatical relations, be it Verb-Object, Subject-Verb, Modifier-Head, Verb-Complement, etc., might also have been lexicalized, though I will focus only on the Verb-Object relation here. Such an analysis then has two immediate consequences. One, any disyllabic VO construct (i.e., V and O being monosyllabic) is possibly just a root word, eliminating (notionally at least) the age-old problem in differentiating such constructs between words and phrases, such as "*chifan / heshui / sa-niao / lashi / fangpi / shui-jiao / zuomeng / dushu / xiezi / xiyan / hejiu / shangjie / maicai / shangban / niaotian / guohe / pashan / chengchuan / zuoche / qima*" (吃飯／喝水／撒尿／拉屎／放屁／睡覺／做夢／讀書／寫字／吸煙／喝酒／上街／買菜／上班／聊天／過河／爬山／乘船／坐車／騎馬, eat / drink water / pee / shit / fart / sleep / dream / read / write / smoke / drink / go out / buy food / go to work / chat / cross a river / hike / take a boat / take a bus / ride a horse). According to Pinker (Chapter 4, 1999), pattern-associated memory also has a generating capacity, which simply generates patterned,

memory-accessible words, such as "*ring, rang, rung; sing, sang, sung; stink, stank, stunk*" in English. This memory-based generating mechanism is believed to be more economical than the rule-based generating system. Of course, how Chinese disyllabic VO constructs behave in terms of being generated by pattern-associated memory remains to be better understood and a task which we must leave for future studies. But, if it were true that pattern-associated memory is responsible for generating these constructs in Chinese, then it would mean that these constructs are drawn as root words directly from the lexicon rather than are constructed as VPs in syntax (by using a monosyllabic verb and monosyllabic noun). One fact in support of such analysis (though a more detailed examination remains to come from further investigations) is that the so-called object in such VO constructs is basically non-referential, a situation which would be expressed simply in another language, e.g., English, by using the verb only or the verb plus a non-referential object, as is made clear by the English translations in the above. On a speculative note, the cause for such difference between Chinese and English is the lack of the article system in the former and the presence of such in the latter. The article system allows English to choose between a bare verb and a verb plus a non-referential object (i.e., *a/an* + *noun*) when constructing a VP with a non-referential object, which entails either no object or a non-referential object. It also allows English to choose between a referential and a non-referential VP by placing an article (*the* or *a/an*) where appropriate before a noun. Lacking an article system, Chinese resorts to syllable counts; so, while "*chou-yan*" (smoke, 抽烟) is non-referential, "*chou xiangyan*" (smoke cigarettes, 抽香烟) is referential where "cigarettes" is a specific and generic category. In the current context of discussion, "*chou-yan*" is a root word, but "*chou xiangyan*" a VP.

The second consequence for the notion that all (rule-generated) disyllabic constructs in Chinese might have been lexicalized as root words through pattern-associated memory is that such lexicalization process would apply across the board to syntactic and lexical units alike. Continuing to focus on Verb-Object relations, we would expect the following lexical OV constructions such as "*roushi / ke-yun / huo-yun / xiongzhao / erzhui / ersai / xiekou / menshuan*" (肉食／客運／貨運／胸罩／耳墜／耳塞／鞋扣／門閂 meat-eating / passenger-carrying / goods-hauling / breast-cupping, i.e., bra / ear-pulling, i.e., ear ring / ear plug / shoe-buckle / door-latch), to have become root words, too. These

constructs are of a lexical nature by virtue of the word order, i.e., OV as in contrast and opposed to the syntactic VO construction we have seen earlier. But both VO and OV constructs could have been lexicalized as root words directly accessible from memory instead of generated by rules, due to being disyllabic units.

If the above analysis for the lexicalization process of disyllabic constructs in Chinese is appropriate, then the crucial question to ask is how the process might have taken place. If Pinker (1999) is right in asserting a general principle of memory takes over where rules stop applying, then there must be a "route" for Chinese disyllabic constructs to evolve from rule-generated entities to words of some lexical conformity. Pinker, who did not base his study on Chinese, provides no answer. In fact, it is an age-old issue in Chinese in desperate need of fresh insights as to how disyllabic syntactic constructs might have evolved to become lexical items. As far as I am aware, studies that are up to date and may provide answers are Feng (1997a, 1997b, 2004) who has conducted much technical and insightful work on this issue and beyond from the perspective of prosodic phonology. His view, in my much too simplified version, is that the fact that a large number of Chinese words are constrained by two syllables, i.e., they may no longer be words once the syllable count exceeds two, is prosodic evidence (e.g., foot formation) for a controlled phonological process ("controlled" in the sense that disyllables are more "inducible" than multiple ones) in which prosodic requirements override other grammatical properties. How this prosodic evidence relates exactly to the change-of-state process which we have been discussing here, i.e., how rule-generated constructs may be "transformed" to word status, certainly calls for continuing investigations.

6. CONCLUDING REMARKS

The paradigm in (1) has revealed complex modes of grammatical operations. The *monosyllabic* VO's are a root word, whereas the *disyllabic* VO's are VP's, and the *disyllabic* OV's are lexical structures. The first part of the findings is not entirely new, as other authors, e.g., Chao (1968: 159-160), have said that the *monosyllabic* VO's are of integral lexical items on their own, i.e., they are lexicalised. But the approach I adopt here is a fresh one and it tries to answer the deeper question of what is meant by lexicalisation. In Pinker's (1999) terms,

lexicalization amounts to memorisation. Thus, grammatical operations are given a background of psycho- and neurolinguistic properties.

The paradigm of *monosyllabic* VO's vs. *disyllabic* VO's as well as OV's in Chinese compounds thus echoes that of the "irregulars vs. regulars" of inflectional morphology in terms of comparable modes of grammatical operations. According to Pinker as mentioned earlier, irregulars are root words, whereas regulars are generated by rules, and while root words are accessed by memory, rules are activated by the computing mind. Hence the operational principle of memory vs. symbol-processing, which Pinker believes is behind all empirical paradigms, including inflectional morphology and beyond. The findings of this paper have thus strengthened that belief in a small way, though it remains to be seen whether they will be worth further research than is appropriate here.

As a last word, I am concerned mainly with the linguistic contrasts of the compounds in question, and have left any psycholinguistic implications of these contrasts unexplored. For instance, the monosyllabic VO construction is productive in Chinese, and this implies that some pattern-associated memory is at work. If so, how this construction (with its memory-based productivity) may play out in acquisition of Chinese compounds will be among issues for useful future research.

REFERENCES

Baker, M. (1985) The Mirror Principle and Morphosyntactic Explanation. *Linguistic Inquiry* 16:373-416

Baker, M. (1988) *Incorporation*. Chicago: University of Chicago Press

Baker, M. (1997) Thematic Roles and Syntactic Structure. In L. Haegeman (ed.) *Elements of Grammar: Handbook of Generative Grammar*. 73-138. Dordrecht: Kluwer

Chao, Y-R (1968) *A Grammar of Spoken Chinese*. Berkeley: University of California Press

Chomsky, N. (1986) *Knowledge of Language: Its Nature, Origin and Use*. New York: Praeger

Chomsky, N. (1991) Some Notes on Economy of Derivation and Representation. In R. Freidin (ed.) *Principles and Parameters in Comparative Grammar*. 417-454. Cambridge, MA: MIT Press

Chomsky, N. (1993) A Minimalist Program for Linguistic Theory. In K. Hale and S. J. Keyser (eds.) *The View from the Building 20: Essays in Linguistics in Honour of*

  Sylvain Bromberger. 1-52. Cambridge, MA: MIT Press

Chomsky, N. (1995) *The Minimalist Program*. Cambridge, MA: MIT Press

Chomsky, N. (1998) *Minimalist Inquiry: The Framework*. ms.

Chomsky, N. & Lasnik, H. (1995) The Theory of Principles and Parameters. In N. Chomsky *The Minimalist Program*. 13-128. Cambridge, MA: MIT Press

Di Sciullo, A. M. & Williams, E. (1987) *On Defending the Word*. Cambridge, MA: MIT Press

Grimshaw, J. (1990) *Argument Structure*. Cambridge, MA: MIT Press

Feng, S. (1997a) Prosodic Structure and Compound Words in Classical Chinese. In J. L. Packard (ed.) *New Approaches to Chinese Word Formation: Morphology, Phonology and the Lexicon in Modern and Ancient Chinese*. 197-260. Berlin & New York: Mouton de Gruyter

Feng, S. (1997b) *Prosodic Morphology and Syntax in Chinese*. (漢語的韻律、詞法與句法) Beijing: Peking University Press

Feng, S. (2004) Verb-Object Inversion and Prosodic Morphology. (動賓倒置與韻律構詞法) *Linguistic Science* 3:12-20

Hale, K. & Keyser, S. J. (1993) On Argument Structure and the Lexical Expression of Syntactic Relations. In K. Hale and S. J. Keyser (eds.) *The View from the Building 20: Essays in Linguistics in Honour of Sylvain Bromberger*. 53-110. Cambridge, MA: MIT Press

Katamba, F. (1993) *Morphology*. London: The MacMillan Press

Kiparsky, P. (1982) Lexical Phonology and Morphology. In I. S. Yang (ed.) *Linguistics in the Morning Calm*. Seoul: Hansin

Lieber, R. (1992) *Deconstructing Morphology: Word Formation in Syntactic Theory*. Chicago: University of Chicago Press

Pinker, S. (1999) *Words and Rules: The Ingredients of Language*. London: Phoenix

Williams, E. (1981) On the Notions "Lexically Related" and "Head of a Word", *Linguistic Inquiry* 12:234-74

Williams, E. (1994) *Thematic Structures*. Cambridge, MA: MIT Press

Author's address:
Department of Translation
The Chinese University of Hong Kong
Shatin, N.T.
Hong Kong
yuanjianhe@cuhk.edu.hk

<div align="center">

漢語述賓式合成複合詞對詞項與規則理論的映証

何元建

香港中文大學

</div>

帶施事題元的漢語述賓式合成複合詞有兩種：OVS 型和 VOS 型。前者是純粹的向心型詞結構，而後者卻不是如此。在 VOS 型當中，如果 V 和 O 分別是單音節的，那麼 VO 是 VP 經模式聯體記憶而轉化過來的詞根，跟 S 再組成複合詞。如果 V 和 O 分別是雙音節的，那麼 VO 是 VP 直接進入複合詞，即所謂"短語入詞"現象。模式聯體記憶和運用規則之間的相互補充反映了語法運作的經濟原則。漢、英兩種語言中的合成複合詞都為這一原則提供了證據。不同的是，漢語合成複合詞中的施事題元是一個粘著詞根，而英語合成複合詞中的施事題元卻可以是一個動詞後綴。這樣，漢語合成複合詞在成分結構中直接就滿足了"題元指派統一論"的要求，而英語合成複合詞（如果施事題元是一個動詞後綴的話）就需要在邏輯形式中去完成相同的要求。