Chapter 3  The Verb Phrase

The notion of a phrase came up many times in Chapter 1. One of the main discoveries in the past thirty-plus years is that a phrase is much more than simply a group of words acting as a unit. Phrases have several important characteristic properties. One of them was introduced in Chapter 2, i.e., that words inside a phrase are combined according to a particular schema which holds across categories and, linear order aside, perhaps across all languages as well. In this chapter, we focus on the verb phrase (VP), examining the various constituents affiliated with V.

3.1.  Adjuncts and complements

Consider a typical verb phrase in the example below:

(1) ta dasheng chang minge.

he loud sing folk.song

‘He sings folk songs loudly.’

In addition to the verb chang ‘sing’ which we refer to as the head of VP, the phrase also contains the verb’s object, minge ‘folk song,’ and a modifier, dasheng ‘loud,’ that describes the manner of singing. That non-head components inside VP are divided into objects and modifiers is long-held wisdom with its basis in intuition. The object is an intrinsic participant of the event described by the verb whereas a modifier provides more
“peripheral” information about the event such as time, location and the manner in which the event is carried out. The X’-theory introduced in Chapter 2 captures the object-modifier distinction as follows (for now, we treat the subject as the Spec of VP; but see definition (20), Chapter 2, and Section 3.2 below):

\[
(2) \quad \begin{array}{c}
\text{VP} \\
\text{NP1} \\
\text{AP} \\
\text{he} \\
\text{loud} \\
\text{sing} \\
\text{folk song}
\end{array}
\]

Treating the modifier as AP, we already saw in Chapter 2 that the AP in (2) is in the adjunct position and the object, NP2, in the complement position. Note first that merging V with NP2 produces V’, which is notationally different from either V or NP2; in contrast, merging AP with V’ results in another V’. This labeling system is meant to reflect two important facts about language. First, an adjunct is only peripheral to the VP because its addition does not alter the original structure – when it attaches to V’, we still have a V’, not a node of a different nature. Second, if adjunct + V’ = V’, it follows automatically that modifier-adjunction may intrinsically happen an indefinite number of times, restricted only by other factors. It is this recursive nature of syntactic structure, not limited to adjuncts, that accounts for the ability of language to produce a potentially infinite number of sentences. Also worth observing in (2) is that adjoining AP to V’ yields the correct word order, with AP necessarily preceding the verb and its
complement. But what distinguishes complements and adjuncts is actually more subtle and interesting.

The examples in (3-4) illustrate questions about the complement and adjunct of the basic sentence in (1):

(3) a. **ni chang shenme minge?**
   you sing what folk.song
   ‘What folk songs do you sing?’

   b. **ni zenme chang minge?**
   you how sing folk.song
   ‘How do you sing folk songs?’

(4) a. **ta shuo [ ni chang shenme minge ]?**
   he say you sing what folk.song
   ‘What folk songs did he say that you sing?’

   b. **ta shuo [ ni zenme chang minge ]?**
   he say you how sing folk.song
   ‘How did he say that you sing folk songs?’ (With *how* modifying *sing*)

Each example in (3) is a simple sentence, whereas each example in (4) consists of two clauses of which the embedded one is marked with brackets. Regardless, either the object or the modifier of *chang* ‘sing’ can be questioned. It appears then that the question
expressions (e.g., *shenme minge* ‘what folk song’ and *zenme* ‘how’) can occur freely either in a simple sentence or in an embedded clause.

The generalization falls apart in several ways, however. Consider the context of indirect questions, first discovered in Huang (1982a):

(5) a. ta xiang zhidao [ shei chang minge ].

   he want know who sing folk.song

   ‘He wants to know who sings folk songs.’

b. (?)ta xiang zhidao [ shei chang shenme minge ]?

   he want know who sing what folk.song

   ‘*What song does he want to know who sings?’

1

c. *ta xiang zhidao [ shei zenme chang minge ]?

   he want know who how sing folk.song

   ‘*How does he want to know who sings folk songs?’ (*how modifying sing)

Note that (5b) and (5c) would be equally acceptable if the two question words in each sentence were both part of the indirect question; under this interpretation the examples are not understood as questions in themselves. (5b) would be translated as “he wants to know who sings what folk songs” and (5c) “he wants to know who sings folks songs in what manner.” The contrast appears when *shenme minge* ‘what folk song’ and *zenme*

1 See Huang (1982a) for a discussion on why the corresponding English example is ungrammatical. The contrast between (5b-c) may not be equally clear to every native speaker of Chinese. What is important is that if there is a contrast in acceptability, (5b) is always the better one, a generalization to which we know of no counterexamples.
‘how’ are meant to turn the whole example into a question, as indicated by the question mark in (5b-c). Under this condition, the object in the indirect question may still participate in forming an interrogative whereas the adjunct appears to resist such an interpretation.

The same differentiating pattern is found when the main verb is negated:

(6) a. ta mei gaosu dajia [ ni zheme chang nei-shou minge ].
    he not tell people you this.way sing that-CL folk.song
    ‘He didn’t tell people that you sang that folk song this way.’

b. (?)ta mei gaosu dajia [ ni zheme chang shenme minge ]?
    he not tell people you this.way sing what folk.song
    ‘What folk song(s) did he not tell people that you sing this way?’

c. *ta mei gaosu dajia [ ni zenme chang zhe-shou minge ]?
    he not tell people you how sing this-CL folk.song
    Intended reading: ‘*How did he not tell people that you sing this folk song?’
    (With how modifying sing)

The contrast between (6b) and (6c) varies among the native speakers of Chinese. For some, (6b) also sounds somewhat strange. But overall, (6c) is perceived to be significantly more difficult to interpret even though the sentence otherwise feels “grammatical,” meaning that every word seems to occur in the right spot. The following triplet, using a different main verb, confirms that the contrast is not a coincidence:
In conclusion, the adjunct inside VP is much harder to question than the complement when the main clause is negative. And judging from the English translations, the same complement-adjunct asymmetry holds in English as well.²

The same pattern is found when certain adverbs are used in the main clause:³

² This “inner island” phenomenon was first noticed in English by Ross (1983). See Rizzi (1990) for a theory of why it arises. At least in Chinese, this contrast seems to hold only of embedded clauses that would be represented as tensed clauses in other languages. If the clause is non-tensed, the asymmetry disappears:

i. ta mei ting-guo ni chang shenme minge?
he not hear-GUO you sing what folk.song
‘What folk song has he not heard you sing?’

ii. ta mei ting-guo ni zenme chang minge?
he not hear-GUO you how sing folk.song
‘*How has he not heard you sing folk songs?’
(8) a. ?ta xiaoxinyiyi de shuo [ ni chang-guo shenme minge ]?
   he cautiously DE say you sing-GUO what folk.song
   ‘What folk song did he cautiously say that you sang?’

b. *ta xiaoxinyiyi de shuo [ ni zenme chang-guo minge ]?
   he cautiously DE say you how sing-GUO folk.song
   ‘*How did he cautiously say that you sang folk songs?’ (how modifying sing)

Again, while (8a) may not be a natural question, it is much easier to understand with the intended meaning than (8b) which is essentially uninterpretable. Replacing the manner adverb xiaoxinyiyi de ‘cautiously’ with another, e.g., dasheng (de) ‘loudly’ or xinbzazaiyian de ‘absent-mindedly’ produces the same result.

In sum, the object and the adverbial modifier, while both inside VP, consistently exhibit different behaviors in syntax. In the theoretical framework in which the current

---

3 That adverbs of different types may show an interference effect when one of them is a question word has been reported by many authors. See Jackendoff (1972), Schlyter (1974), Koster (1978), Travis (1988), Alexiadou (1997), Laenzlinger (1998), Cinque (1999), Rizzi (2001), Ernst (2002), among others. The generalization that holds among all types of adverbs is reported in Li, Lin and Shields (2005):

i. Let X range over types of movement and [+X] indicate whether a given adverb (or adverb class) can undergo X-type movement, then adverb A1 prevents adverb A2 from undergoing X-type movement iff
   a. A1 e-commands A2, and
   b. A1 = [+X].

Critically and unlike what has been proposed in the literature, what turns A1 into a blocker in this case is not that A1 also needs to undergo X-type movement, but that A1 has the potential for X-type movement.
book is written, this difference is ultimately attributed to the intuition that the object holds a thematic relation with the verb that adjuncts do not. See Chapter 2 for a theory about thematic relations. For now, we focus on its significance for the structure of phrases.

It should also be noted that the data discussed so far, neat as they are, are not without apparent counterexamples. For instance, not every kind of adverb in the main clause creates the complement-adjunct asymmetry seen in (8). The examples below use *gangcai* ‘a short moment ago’ and *daochu* ‘everywhere’ to illustrate one such “counterexample.”

(9) a. ta gangcai/dao chu           shuo   [ ni chang-guo shenme minge ]?
    he just.now/everywhere say       you sing-GUO what folk.song
    ‘What folk song(s) did he say just now/everywhere that you sang?’

b. ta gangcai/dao chu           shuo   [ ni zenme chang-guo minge ]?
    he just.now/everywhere say       you how sing-GUO folk.song
    ‘How did he say just now/everywhere that you sang folk songs?’ (*how*
    modifying *sing*)

But these sentences do not falsify the asymmetry established with (8). For one thing, there is no known data showing the reversed pattern. That is, there are no examples comparable to (8) and (9) except that the sentence questioning about the embedded object is bad whereas the one questioning about the adverbial modifier is good. This rules out the possibility that the asymmetry in (8) is random. Also, there is independent evidence
from English and other languages that the difference between (8) and (9) is largely predictable once adverbs are more finely classified. See Cinque (1999) for various subclasses of adverbs. For theories on how different adverb classes interact with one another, see Ernst (2002) and Li, Lin and Shields (2005), as well as Note 3.

3.2. Postverbal constituents

In this section, we take a close look at three types of constituents occurring after the verb – double objects, V-de (cf. 2.2.3 in Chapter 2), and frequency/duration phrases – and examine their implications for the syntactic structure of language.

3.2.1. Double objects and the structure of VP

Certain verbs allow or require two objects. Throughout the recent history of syntax, double-object constructions have always posed a problem for constituency:

(10) a. ta di-gei gege yi-hu jiu.
    he pass-give brother one-CL wine
    ‘He passed his brother a jug of wine.’

b. ?ta di-gei gege yi-hu jiu, jiejie yi-pan cai.
    he pass-give brother one-CL wine sister one-CL dish
    ‘He passed his brother a jug of wine and his sister a dish.’

c. *ta di-gei de shi gege yi-hu jiu.
he pass-give DE be brother one-CL wine

‘*What he passed was his brother a jug of wine.’

While (10b) may sound somewhat strange out of the context, it is not difficult to find a colloquial context where it is perfectly acceptable. (10c), however, remains bad regardless of the context. The same pattern, though sharper, is also found in English, as seen in the translations of (10). The problem with such data is the apparent contradiction that different constituency tests create. The conjunctive construction, shown in (10b), is generally believed to require each conjunct to be a constituent. Therefore, it must be concluded that the two NP objects of the compound di-gei ‘pass-give’ form a constituent of some sort. Meanwhile, the pseudo-cleft construction in (10c) is also a well-established constituency test, with whatever is after the copula shi ‘be’ being a phrase. Then why does one test say gege yi-hu jiu is a constituent while the other says it isn’t? Different solutions have been proposed, most of which are based on Larson’s (1988) work. In this book, we adopt a variant of Larson’s theory, articulated in Chomsky (1995), which meshes well with our analysis of Chinese in other chapters of the book.

Recall from Chapter 2 that the Agent theta-role is the direct result of syntax, not of the lexical verb per se, because it represents the cause that is external to the event the verb describes. Given the universally accepted belief that all and only the theta-roles of the verb are assigned to arguments inside VP, it follows that there is a component, separate from VP, which is responsible for introducing the Agent argument. In Chomsky (1995), the Agent-introducing job is attributed to v, a soundless verbal head which is

---

4 See Pesetsky (1995) for a different approach based on a double-structure for any given sentence.
somewhat “less lexical” than V (cf. Chapter 1). If VP is taken to be the structural complement of v, then, the X’-theory applied to v and V yields (11):

```
(11)
```

![Diagram](image_url)

From the semantico-syntactic perspective, (11) is the minimal “complete functional complex” (CFC; cf. Chomsky 1986b) because it is the smallest structure in which all the external and internal participants of the given event are represented.

As an immediate consequence of adopting (11), we are provided with a solution to the dilemma in (10). Let NP be the subject and XP1 and XP2 be the two objects. For some cross-linguistic reason which we will gloss over for now, the lexical V clearly must move to v in order to yield the subject-verb-object1-object2 sequence in both Chinese and English. In the case of (10a), the structure after V-to-v movement is (12), with the generic XP objects replaced by NPs:

```
(12)
```

![Diagram](image_url)
In the tree, $t$ stands for the “trace” that marks the original position of a moved constituent. When the conjunctive construction in (10b) appears to take NP2 (i.e., $gege$ ‘brother’) and NP3 ($yi-hu jiu$ ‘a jug of wine’) as conjuncts, it really takes the whole VP. Since the $V$ position is now a phonetically null trace, NP2 and NP3 are the only constituents that can be heard. As for (10c), recall that the CFC is not VP but $vP$. It is possible then that the pseudo-cleft construction, when targeting phrases containing thematic arguments, must always apply to a CFC and never part of it. This differs from the conjunctive construction, which is more flexible on what qualifies as a conjunct (conjuncts are in bold face; the conjunctive $he$ ‘and’ may alternate with a pause):

(13) ta de qinqi (he) pengyou dou lai-le.

he DE relative and friend all come-LE

‘His relatives and friends all came.’

If the conjunctive construction can link two nominal constituents smaller than a full NP in (13), it is no surprise that part of a CFC in (10b) may serve as a conjunct as well. Also note that in (11-12), one of the objects is in fact in the Spec position of VP. Certain consequences of this configuration will become clear shortly.
3.2.2. **V-de**

Other than objects, two phrasal constituents may also occur postverbally, both of which are characterized by a morpheme *de* suffixed to the verb. Consider first the resultative:⁵

(14) a. ta zou-de qichuanxuxu.

   he walk-*DE* breathe.heavily

   ‘He walked so fast that he breathed heavily.’

b. ta qi-de wo bu xiang xie xin le.

   he annoy-*DE* me not want write letter SFP

   ‘He annoyed me so much that I didn’t want to write the letter.’

Descriptively, the semantically obscure *de* introduces a clause that describes the result of the event denoted by V. (14a) is an example with V being intransitive. Given the theta-criterion, we take the subject of *qichuanxuxu* ‘breathe heavily’ to be Pro (cf. Chapter 2, Section 2.1.3):

(15) he walk-*DE* [s Pro breathe heavily ]

---

⁵ The analysis in this subsection is an extension of Y. Li (1995) and Ting and Li (1997). Also see Huang (1988c) for arguments that the post-*de* constituent is structurally a clausal complement, and Cheng and Huang (1994) for related discussion.
When V is transitive as in (14b), the NP after it (i.e., wo ‘me’) is interpreted as both the object of V and the subject of the result clause. But in syntax, this NP can only serve either as the object of qi ‘annoy’ or as the subject of bu xiang xie xin ‘not want to write the letter’ but not as both – otherwise the NP would get two theta-roles from different sources and violate the theta-criterion. There is evidence that this NP is the object of qi, as we will soon see.

The interjection ya can be inserted between a verb and its clausal object, but never between the verb and the postverbal NP object. In each example below, the object clause is between brackets. Pro is again used in the absence of an overt subject.

(16) a. ta gaosu pengyou ya, [S Pro qu touben qinqi ].
   he tell friend YA go seek.refuge.with relative
   ‘He told his friend, um, to go to the relatives for shelter.’

b. *ta gaosu ya, pengyou [S Pro qu touben qinqi ].
   he tell YA friend go seek.refuge.with relative
   ‘*He told, um, his friend to go to the relatives for shelter.’

c. ta shuo ya, [S pengyou qu touben qinqi le ].
   he say YA friend go seek.refuge.with relative SFP
   ‘He said, um, that his wife went to the relatives for shelter.’

The unacceptability of (16b) is the result of inserting ya in front of the NP object. In contrast, the overt NP pengyou ‘friend’ after ya in (16c) is the subject of the embedded clause, making ya-insertion possible.
Applying *ya*-insertion to the resultative (14b) yields the following contrast:

(17) a. ta qi-de wo ya, bu xiang xie xin le.

he annoy-DE me YA not want write letter SFP

‘He annoyed me so much, um, that I didn’t want to write the letter.’

b. #ta qi-de ya, wo bu xiang xie xin le.

he annoy-DE YA me not want write letter SFP

= ‘#He was so annoyed that I didn’t want to write the letter.’

≠ (17a).

(17a) is directly comparable to (16a), with *ya* between the postverbal NP and what we believe to be an embedded clause expressing the result. When *ya* occurs between the verb *qi* and the NP *wo*, however, the sentence still sounds grammatical but has the pragmatically strange interpretation, marked “#”, that he was so annoyed that I didn’t want to write the letter. In other words, the insertion of *ya* forces *wo* to be understood as the subject of the embedded clause because, with *ya* in between, this NP cannot be the object of *qi*. Consequently, the verb *qi* is forced to take an intransitive reading. Crucially, (17a) does not have this strange interpretation, suggesting that *qi* in it is used transitively and that *wo* is indeed the object of *qi*. Since the *ya*-less (14b) has the same basic semantics as (17a) but not as (17b), we conclude that the verb *qi* is also a transitive with *wo* as its object. The structure of (14b) is thus (18):

(18) he annoy-DE me [s Pro not want write letter ].
The next question is where exactly the result S is seated in the vP-VP configuration in (12). Recall from Chapter 2, def. (28), that a complement does not create any island effect but an adjunct does. It follows that the precise location of S may be tested: if movement out of it is good, S is in the complement position; otherwise, it must be in some kind of adjunct position. The examples below are designed for this test. (19a) involves the topicalization of the NP object inside the result S; in (19b), the same NP participates in relativization, a process which, in Chinese, moves the NP to the edge of the relative clause (see Chapter 6 for details), leaving a trace at the original site:

(19) a. na-feng xin, ta qi-de wo [S bu xiang xie t le ].
   that-CL letter he annoy-DE me not want write SFP
   Lit: ‘That letter, he annoyed me so much that I didn’t want to write.’
   b. [ ta qi-de wo [S bu xiang xie t ]] de na-feng xin.
   he annoy-DE me not want write DE that-CL letter
   Lit: ‘the letter that he annoyed me so much that I didn’t want to write.’

Other than being a little too long, these sentences show no deterioration in acceptability when compared with (14b). Thus, it is confirmed that the result S indeed is in the complement position. Putting aside certain details to which we will return later, the structure below represents the vP in (14b):

(20) 

\[ \text{vP} \]
With a careful design, the same CED test may be applied to the manner V-de as well. The examples in (21) illustrate the construction, while those in (22) involve movement out of the postverbal manner phrase:

\[(21)\]
\[\begin{array}{l}
ea. \text{ ni chang-de [}Z\text{ tebie haoting }]. \\
\text{ you sing-DE especially pleasant.to.listen.to} \\
\text{ ‘You sing especially well.’} \\
b. ?\text{ ta pao-de [}Z\text{ kuai-dao neng zhuishang tuzi }]. \\
\text{ he run-DE fast-till be.able.to catch.up.with rabbit} \\
\text{ ‘He ran fast enough to catch up with a rabbit.’} \\
\end{array}\]

\[(22)\]
\[\begin{array}{l}
a. ?\text{ na-zhi tuzi, ta pao-de [}Z\text{ kuai-dao neng zhuishang t]}. \\
\text{ that-CL rabbit he run-DE fast-till be.able.to catch.up.with} \\
\text{ Lit: ‘That rabbit, he ran fast enough to catch up with.’} \\
b. ?[\text{ ta pao-de [}Z\text{ kuai-dao neng zhuishang t }] \text{ de na-zhi tuzi}. \\
\text{ he run-DE fast-till be.able.to catch.up.with DE that-CL rabbit} \\
\text{ Lit: ‘the rabbit that he ran fast enough to catch up with.’} \\
\end{array}\]
The bracketed phrases in (21-22), marked Z, have the semantic function of a manner or degree modifier for the verb suffixed with *de*. (21b) sounds marginal because Z itself contains an embedded clause *neng zhuishang tizi* ‘Pro can catch up with a rabbit’. Topicalization and relativization are applied to (21b) to yield the examples in (22). Other than the fact that they are both long and clumsy, these examples exhibit no detectable deterioration from (21b) in acceptability. Hence, it is concluded that even the manner phrase is in fact located in a complement position.\(^6\)

Though both V-*de* constructions contain a complement (S in (19) and Z in (20)), there is a difference between them: Only the resultative V-*de* allows the object NP of the verb to occur postverbally. Compare (14b)/(18) with the examples below:

(23) a. ta chang-de tebie haoting.
   he sing-DE especially good.to.hear
   ‘He sang especially well.’

b. *ta chang-de xiaoqu tebie haoting.\(^7\)
   he sing-DE ditty especially good.to.hear

---

\(^6\) One question is what types of constituents that are not thematic objects may or must occur in the complement position and why. Huang (1988c, 1992) suggests, in the spirit of Larson (1988, 1991) and McConnell-Ginet (1982), that a postverbal manner phrase is a secondary predicate (whereas a preverbal manner phrase is an adjunct). A secondary predicate is property denoting, and it may combine first with the main verb and form a complex predicate (V’) before the thematic object is merged to Spec, VP.

\(^7\) This sentence has a fully acceptable interpretation in which *de* is not manner-denoting but signals a relative clause. I.e., (23b) can mean the folk song he sang was especially nice to listen to. This is irrelevant to our current concern.
Intended reading: ‘He sang ditties especially well.’

c. xiaoqu ta chang-de tebie haoting.
   ditty he sing-DE especially good.to.hear
   ‘Ditties, he sang especially well.’

d. ta chang-de tebie haoting de na-shou xiaoqu
   he sing-DE especially good.to.hear DE that-CL ditty
   ‘the ditty that he sang especially well’

When the verb *chang* ‘sing’ is used transitively, the object NP *xiaoqu* ‘ditty’ can either be topicalized as in (23c) or undergo relativization as in (23d). But when it stays in situ, as in (23b), the sentence becomes unacceptable. Since topicalization and relativization both leave a trace in the object position (cf. (19), (22)), the contrast between the manner V-de and the resultative V-de can be summarized as follows:

(24) A phonetically overt NP object is permitted postverbally only in the resultative V-de construction.

(24) may be linked to another fact in modern Chinese, namely that there is no verbal compound in which the morpheme on the left (V1) is modified by the one on the right (V2):

\[ ^8 \text{In order to avoid irrelevant complications, we do not distinguish A from V. See Chapter 1 for their similarities and differences.} \]
(25) fei-kuai, jing-zuo, sheng-chi, zhong-shi, nu-hou, …
fly-fast quiet-sit raw-eat heavy-view angry-shout
very fast sit quietly eat raw take seriously shout angrily

In all these examples, V1 modifies V2. *Fei-kuai* ‘fly-fast’, for instance, means only flyingly fast and never fast flying. This independent fact in Chinese morphology creates a conflict of requirements. On the one hand, the manner *de* must be suffixed to a verb. On the other hand, because it ultimately introduces a manner phrase, the only possible relation it holds with the verb is one of modification, a relation that prevents it from occurring after the verb morpheme. The only way to resolve the conflict is for *de* and the verb to be separate constituents structurally but pronounced as a unit. That is, V and *de* form a phonological word based on pure linear adjacency. Since they do not form a structural unit in the sense of lexical word-formation, the compounding pattern shown in (25) becomes irrelevant.

That phonological words may be formed without structural constituency is best illustrated with the following typical Kwak’ala example, quoted from Anderson (1992:2):

(26) nanaqχsil-ida i?gχl’wat-i χliwinuxwa-s-is mestuwi la-xa migwat-i.
guides-ART expert-DEM hunter-instru-his harpoon prep-ART seal-DEM

‘An expert hunter guides the seal with his harpoon.’
In general, the functional morphemes suffixed to each lexical stem belong structurally not to their hosts but to what follows them. For instance, the demonstrative suffix –i following the noun i?g İz ‘expert’ actually is part of the nominal phrase headed by İz liwinuxwa ‘hunter’, which in turn is followed by the instrumental suffix –s and the possessive pronoun –is ‘his’ that actually are part of the syntactic phrase with his harpoon. Given (26), we claim that the manner de forms a syntactic constituent with the subsequent AP (e.g., tebie haoting ‘especially good to listen to’ in (23a)) but suffixes to the preceding verb as part of a phonological word. This is directly comparable with what happens in Kwak’ala.

The following structure illustrates the syntactic context for (23), based on the earlier conclusion that the manner phrase is in the complement position:

(27)

The verb chang ‘sing’ raises from V to v. If it is used intransitively as in (23a), there is nothing in the position of NP2. The verb in the v position remains linearly adjacent to de, making it possible to form the phonological word chang-de. If the verb is transitive, NP2 is the object xiaoqu ‘ditty’. Once raised to v, the verb is separated from de by NP2 and no
phonological word can be formed. The “dangling” de is unacceptable for lack of a verbal host. Of course de can form a phonological word with the verb before the latter raises. But because the verb and de do not form a structural constituent, they cannot raise together to v, which chang ‘sing’ presumably must do. As a result, there is no good way to form the phonological word chang-de while still meeting all the relevant requirements. This is why (23b) is bad. As for (23c-d), both contain a moved object, leaving the

---

The example below, illustrating post-subject contrastive topicalization (or focalization), seems relevant to the proposed theory:

i. ta xiaoqu chang-de hen haoting.
   he ditty sing-DE very well
   ‘He sang folk songs well.’

On the one hand, since the object NP xiaoqu ‘ditty’ precedes the verb, (i) may be part of the pattern seen in (23c-d), which are accounted for below. On the other hand, however, the fact that this NP object follows the subject might seem to suggest that the verb does not have to raise to v after all. Suppose the verb stays in V and forms the phonological word with de. Then (i) can be generated by (27) right away, with the object between the subject and the verb. However, there is evidence that the preverbal object in (i) is in fact outside of VP rather than staying in the NP2 position in (27).

ii. ta xiaoqu mei-you chang-guo.
   he ditty not-have sing-GUO
   ‘Folk songs, he hasn’t sung.’

iii. *ta mei-you xiaoqu chang-guo.
   he not-have ditty sing-GUO

Aspect words such as you ‘have’ (as well as negation) are known to be outside VP (and indeed vP), as we will see below. The contrast between (ii) and (iii) indicates that the post-subject topic/focus has been moved out of VP to a much higher position. In other words, (i) should pattern with the examples in (23c-d) where the object has moved away.
position of NP2 occupied by a trace. As traces are phonetically empty, they do not break
the superficial adjacency between the verb and de, leaving both examples well-formed.\(^\text{10}\)

The manner de forms a contrast with the resultative de. Along the same line of
reasoning, if there is any semantic relation between V and the resultative-introducing de,
it must be comparable to the one between the two verbal morphemes in the resultative
compound such as chang-lei ‘sing-tired’ (cf. Section 2.1.2 of Chapter 2). But resultative
compounds are highly productive in Chinese. It follows that V-de can be formed with the
same lexical compounding rule that is responsible for the resultative compounds. In turn,
this means that qi-de ‘annoy-DE’ in (14b), for instance, may enter syntax as a single word.
All we need to assume is that de takes a clausal complement in itself so merging the
argument structure of qi ‘annoy’ with that of de generates a complex verb that takes a
clausal complement in addition to the arguments of qi. See Chapter 2 for how the
argument structures are treated in compounding. Since V-de is a single verb from the
perspective of syntax, it starts in the V position and raises to v as a unit, crossing the
object NP2 in the process:

\(^{10}\) The English data below are taken to indicate that a trace blocks phonological word formation:
i. I want to win.
ii. I wanna win.
iii. I want Bill to win the prize.
iv. *Who do you wanna win the prize?

In (iv), the trace of who should be between want and to. The fact that wanna cannot be formed, therefore,
follows from the trace blocking the contraction of the two otherwise adjacent morphemes. However, there
is no evidence that traces interfere with phonological word formation cross-linguistically. So (i-iv) are not
considered necessary counterexamples to the analysis proposed with the Chinese manner V-de.
3.2.3. Frequency/Duration phrases (FP/DrP)

Another type of postverbal phrase describes the frequency or duration of an event:

(29) ta chang-le wu ci / liang-ge zhongtou.

he sing-LE five time / two-CL hour

‘He sang five times/for two hours.’

In this section, we try to determine where the Frequency and Duration Phrases (FP/DrP) are located in VP and how they interact with other postverbal components.

3.2.3.1. FP/DrP as adjuncts to V’

The vP-VP structure in (11) conforms to a general perception about language: that verbs
take at most two objects, not three or six. Inside VP, there are only two positions for arguments, the specifier and the complement. If (11) accurately describes the verbal structure available to syntactic computation, it limits the maximal number of objects any verb can take. This property of (11), when combined with the conclusion in 3.2.2 that the postverbal phrase in V-*de* occupies the complement position, makes a prediction: that V-*de* doesn’t occur with two objects (Ting and Li 1997). The logic is simple: with the postverbal phrase as the complement, there is only one position left inside VP, the Spec. No more than one object can be accommodated in this one remaining position. The prediction is borne out:

(30) a. wo gei-le ta henduo liwu.
   
   *I* give-*LE* him many *gift*

   ‘I gave him many gifts.’

b. *wo gei-de ta liwu duicheng-le shan.

   *I* give-*DE* him gift pile.into-*LE* hill

   Intended reading: ‘I gave him so many gifts that they piled up like a hill.”

(31) a. ta gaosu-le renmen zhege xiaoxi.

   he tell-*LE* people this *news*

   ‘He told people this news.’

b. *ta gaosu-de renmen zhege xiaoxi jiayuhuxiao.

   he tell-*DE* people this news be.known.by.everyone

   Intended reading: ‘He told people this news so often that it was known to
There is nothing semantically wrong or pragmatically implausible about the (b) examples. Still, with two NP objects and one result clause all competing for only two syntactic positions, their unacceptability is expected.

Now consider the examples below, each with an optional and legitimate FP:

(32) a. wo shang-guo ta (liang ci) jinyinzhubao.
    I award-GUO him two time money, jewelry

    ‘I awarded him money and jewelry (twice).’

b. ta gaosu-guo wo (haoji ci) tamen bu gai jin cheng.
    he tell-GUO me several time they not should enter city

    ‘He told me (several times) that they shouldn’t go into the city.’

In the presence of two objects in each sentence, we must conclude that the FP is an adjunct and thus does not compete with the objects, unlike the result clause in the V-de construction. This is illustrated in (33):
Raising V to \( v \) yields the word order in which there are two postverbal objects with the FP between them.

Occasionally, the FP may appear after the second object, a linear arrangement that (33) is unable to generate:

(34) \( \text{wo shang-gei ta jinyinzhubao yijing liang ci le.} \)

\( \text{I award-give him money.jewelry already two time SFP} \)

‘I already awarded him money and jewelry twice.’

But there is evidence that (34) may have a different structure from (32a). Note the adverbial \( yijing \) ‘already’ immediately before the FP. This same adverbial can never precede the FP when it comes between the two objects, regardless of the specific form of the double-object verb:

(35) a. \( \text{*wo shang-guo ta yijing liang ci jinyinzhubao.} \)

\( \text{I award-EXP him already two time money.jewelry} \)

Intended reading: ‘I already awarded him money and jewelry twice.’

b. \( \text{*wo shang-gei ta yijing liang ci jinyinzhubao le.} \)

\( \text{I award-give him already two time money.jewelry SFP} \)
Intended reading: Same as (34).

One possibility investigated by A. Li (1987) is that the FP in (34) is actually the predicate of a sentence whose subject is the whole clause in front of yijing ‘already,’ which modifies the predicate FP. In any case, (34) needs to be analyzed differently from (32a) and therefore offers no obvious counterargument against treating the latter example with the tree in (33).

(33) also makes it easy to accommodate transitive verbs with only one object. The examples below illustrate the alternative word order between the NP object and FP:  

(36) a. wo da-guo liang ci qifu haizi de huaidan.
    I beat-GUO two time bully child DE bad.guy
    ‘I twice beat bad guys who bullied children.’

b. ta da-guo neixie huaidan liang ci.
    he beat-GUO those bad.guy two time
    ‘He beat those bad guys twice.’

The first example may be directly represented as (33) minus NP2. The raising of the verb da ‘beat’ from V to v yields the actual word order. (36b) is formed with the object NP in

the Spec of VP plus, again, V-to-v movement. It is not yet clear whether this NP originates as the Spec or moves into place from the complement position.\textsuperscript{12} Either way, the account for the word order variation in (36) remains intact.\textsuperscript{13}

The V’-adjunction analysis of FP can be directly carried over to the postverbal DrP:

\textsuperscript{12} Baker (1988) postulates the Uniformity of Theta-Assignment Hypothesis (UTAH), which associates each theta-role with a constant structural position. In this view, if the NP \textit{bad guys} holds the same theta-relation with \textit{beat} in both examples, then they should originate in the same position, presumably as the complement of \textit{V} (cf. Soh 1998) – if they originated in the Spec, downward movement from the specifier to the complement position would be necessary to derive (36a). This is not allowed as movement always targets c-commanding positions (cf. Chapter 2). On the other hand, Y. Li (2005) discusses various problems with the UTAH and its stipulated roles in syntax. Also see Borer (2005) for arguments against the UTAH as well as our own discussion below on the interactions between FP/DrP and the definiteness of the object.

\textsuperscript{13} It should be noted that, abstracting away from details, (36b) might also be analyzed on a par with (34), with the sentence-final FP serving as the predicate of a clausal subject. The structural ambiguity results from the absence of the second object which, necessarily in the complement position of \textit{V}, would serve as the reference point for determining the nature of FP. There is still reason, however, to continue treating the FP in the specific example in (36b) as an adjunct rather than a predicate. When a bare FP occurs at the end of a double-object clause, it acquires a contrastive reading:

\textbf{i.} \textit{wo shang-gei ta jinyinzhubao liang ci.}

\begin{center}
I award-give him money.jewelry two time
\end{center}

‘I awarded him money and jewelry twice (and other things once).’

But no such contrastive interpretation is required of (36b). Since the two objects in (i) force the FP to be a predicate, there must be some difference between (i) and (36b). Why the predicative use of FP in (i) is associated with the contrastive reading is still unknown, but given the contrast, treating the sentence-final FP in (36b) as the V’-adjoined adverbial at least offers a structural base for distinguishing the two sentences.
If these examples are compared with (32a) and (34a) respectively, it is rather straightforward to see that adjoining the DrPs shi-tian ‘ten days’ and yi-nian ‘a year’ to V’ yields the correct word order.

In addition to accounting for the linear locations of FP/DrP, the structure in (33) also provides a means to accommodate the following fact:

(37) a. ta yilian jiao-le wo shi tian Henan hua.

he in.a.row teach-LE me ten day Henan dialect

‘He taught me the dialect of Henan for ten days in a row.’

b. wo mai-guo yi nian yu.

I sell-GUO one year fish

‘I sold fish for a year.’

(38) a. ta ma-le san ci ren.

he scold-LE three time person

‘He scolded people three times.’

b. zhe-ge laoshi jiao-guo shi nian xuesheng.

this-CL teacher teach-GUO ten year student

‘This teacher taught students for ten years.’

(39) a. *ta ma-le ren san ci.

he scold-LE people three time

b. *zhe-ge laoshi jiao-guo xuesheng shi nian.
In general, bare NP objects must occur after the FP/DrP adjunct while definite NPs are permitted before it.

A remarkable property of the bare NPs in (38-39) is that, unlike their definite counterparts in (40), they do not refer to individuals.\(^{14}\) Anticipating extensive discussions of the syntax and semantics of nominal phrases in Chapter 8, we consider these bare NPs non-referential and propose (41) as a hypothesis for Chinese and, perhaps, for other languages too:

(41) A non-referential constituent which bears a theta-relation with a head H should be combined with H to form the smallest possible constituent.

In the construction of VP, the smallest possible constituent is the (smallest) V’.

\(^{14}\) For a brief summary of the thoughts on bare NPs, see Longobardi (2001) and the references therein.
explains why the non-referential bare NPs must come after the FP/DrP (cf. (38)) but not before (cf. (39)) – only the former word order reflects the smallest V’ consisting of the V and the bare NP; the latter would place the bare NP in the Spec, directly under VP, a much larger constituent than V’. Definite NPs, on the other hand, are referential and thus are not subject to (41). This explains the grammaticality of (40).

As it is formulated, (41) also allows the examples below:

(42) a. ta gei-guo ren henduo ci guizhong de liwu.

he give-GUO person many time expensive DE gift

‘He gave people expensive gifts many times.’

b. zhe-ge laoshi jiao-guo xuesheng shi-nian wuli.

this-CL teacher teach-GUO student ten-year physics

‘This teacher taught students physics for ten years.’

In these sentences, the bare NPs (ren ‘person’ and xuesheng ‘student’) indeed occur before FP/DrP and, given the structure in (33), are placed in the Spec of VP rather than inside V’. However, assuming that theta-relations are intrinsically ranked so that Recipient is more prominent in the hierarchy than Theme (cf. Chapter 2 and Y. Li 2005), the VP structure conforming to the thematic hierarchy must be (43):

(43) VP
   /   
  /    
Recipient-NP V’
Given (43), the smallest possible constituent containing $V$ and the Recipient bare NP is VP. This contrasts with the earlier cases in which $V$ is a simple transitive verb. As both the complement and Spec positions are available, $V'$ becomes the only constituent satisfying (41).

### 3.2.3.2. A syntax-semantics mismatch

For all the similarities between FP and DrP, there is an intriguing difference: that $de$ may be optionally added between only DrP and the subsequent NP object without semantic change. This section examines two possible analyses of this phenomenon, starting with the basic fact:

\[(44)\]

\[a.\] ta yilian jiao-le wo shi-tian de Henan hua.

\[
\begin{array}{c}
\text{he in.a.row teach-LE me ten-day DE Henan dialect} \\
\text{‘he taught me the dialect of Henan for ten days in a row.’}
\end{array}
\]

\[b.\] wo mai-guo yi-nian de yu.

\[
\begin{array}{c}
\text{I sell-GUO one-year DE fish} \\
\text{‘I sold fish for a year.’}
\end{array}
\]
Native speakers’ intuitions are that with *de*, which characteristically introduces a modifier to the succeeding head (cf. Chapter 1, Section 1.2.2), DrP and the NP form some kind of constituent, even though DrP is still interpreted as measuring the temporal length of the event described by the verb. How is this syntax-semantics mismatch explained?

One possibility, proposed in Huang (1997, 2005b), is to hypothesize the existence of a phonetically empty verb DO, which in turn takes a nominalized clause (e.g., a gerundive clause). Example (44b), for instance, may in fact have the following structure (irrelevant details are put aside):

(45)

Moving *mai* ‘selling’ to the matrix verb DO, a type of movement we have seen many times by now, will yield the actual word order in (44b) because the trace of *mai* has no phonetic content. With (45), the syntax-semantics mismatch in (44) is no longer a real problem as DrP indeed modifies the verb *mai* ‘sell’, both semantically and structurally. The perception that DrP and NP are both part of the same constituent is also correct,
with the gerundive clause properly containing both of them plus the trace of the raised verb.

Huang further shows that the postulation of movement out of a nominalized VP readily extends to an account of mismatches of the following sorts:

(46) a. tamen ge- tamen-de -xin, women fu- women-de -gu.

They ge- their -xin we fu- our -gu

They carried out their [project of] innovation, but we went on with our restoration of ancient ways.

b. ta-de laoshi dang-de hao.

He-DE teacher do-DE well

‘He serves well as a teacher.’

In (46a), the first item of the sequence ge-xin meaning ‘to innovate’ moves out of a phrase following the possessor ‘their’; and the first item of the sequence fu-gu meaning ‘to return to the old’ moves out of its original position following ‘our’. As indicated in the translations, the possessives are understood as relating to the action denoted by ge-xin and fu-gu, not as possessives of –xin ‘new’, and –gu ‘old’. In (46b), ta-de laoshi does not denote ‘his teacher’, but is understood as ‘his service as a teacher’, being a result of the verb dang having moved out of the domain ta-de dang laoshi ‘his serving-as teacher’. The possessives are, in other words, ‘fake possessives’ of the noun that immediately follows them. (For more details and arguments, see Huang 2005b.)
Another possible approach toward understanding the syntax-semantics mismatch in (44) is to follow Dowty’s (1991) theory of Incremental Theme. Consider these examples:

(47) a. chi pingguo
    
eat apple
    ‘eat apples’

b. chi yi-ge pingguo
    
eat one-cl apple
    ‘eat an apple’

c. chi yi kuang pingguo
    
eat one basket apple
    ‘eat a basket of apples’

Dowty notes that in examples like (47b), the apple in fact measures the progress of the event of apple-eating – if half of the apple is gone, the event is also half accomplished; when the whole apple disappears, the event is completed. In other words, an apple sets boundaries for the beginning and end of the event because the apple has a physical boundary in itself; or in Tenny’s (1994) terms, this object delimits the event. In contrast, (47a) does not have such a property because pingguo ‘apple’ is (or at least can be) generic and possibly plural in meaning and therefore provides no intrinsic beginning and end for apple-eating. In principle, apple-eating may last indefinitely as long as one has the stomach and there is an ample supply of apples. To describe this semantic property,
Dowty proposes that the object NP in (47b) bears the thematic relation of Incremental Theme. The same relation also holds for a basket of apples in (47c) which again measures the progress of apple-eating through the fullness of the basket. The gist of Incremental Theme is that the boundaries of the object delimit the event described by the verb.

It is clear that Incremental Theme, unlike the theta-roles such as Agent and Theme, is not solely determined by the verb. (47a-c) all contain the same verb, differing only in the definiteness of the NP object. In other words, calling Incremental Theme a theta-role is a misnomer; it in fact describes a semantic phenomenon which is brought into existence by the collective work of certain syntactic components.

With this caveat in mind, let us return to (44). Note first that mai yu ‘sell fish’ in itself lacks intrinsic boundaries for the event just as (47a) does. Modifying the verb with the DrP yi-nian ‘a year’ (cf. (37b)) is a straightforward way to delimit the event of fish-selling. This is accomplished structurally by adjoining DrP to V’ and semantically by restricting the event to the temporal boundaries set by DrP. The object NP yu ‘fish’ contains no constituent to delimit any boundary and no Incremental Theme is created.

Now suppose that DrP is coerced into the object NP via the use of de. This creates no structural problem because yi-nian de yu ‘a year’s fish’ is syntactically identical to the well-formed NP yi-nian de diaocha guocheng ‘a year’s investigative process’. At the semantic level, yi-nian de yu is odd and most likely uninterpretable by itself. But the DrP inside the object NP may trigger the mechanism of Incremental Theme into action, transferring the boundaries defined by the DrP to the whole event of fish-selling as Dowty has observed. This analysis explains why (44b), for instance, is semantically
identical to (37b) regardless of \textit{de}. In both constructions, DrP delimits the whole event.

Without \textit{de}, this is done directly through modification to V; with \textit{de}, the delimitation happens indirectly via Incremental Theme. It also follows that NPs such as \textit{yi-nian de yu} ‘a year’s fish’ cannot be used as anything other than the object: for independent reason, only the Theme argument may display the “incremental,” event-delimiting quality (Tenny 1994).

This alternative account of (44) does not automatically require that the examples in (46) be reanalyzed differently from a structure like (45). After all, while English has nothing like (44), it does exhibit a productive (46)-style syntax-semantics mismatch in, say, \textit{You teach your economics and I’ll hunt my coyotes. Let’s see who’ll get rich faster.} The default interpretation of such a sentence is that you do your economics teaching and I do my coyote hunting, not that economics is yours and the coyotes to be hunted are mine. The fact that the two types of mismatches do not necessarily co-occur makes it logically possible to treat them differently.

There are unresolved issues with both the gerundive-based account and the one utilizing the Incremental Theme, the most conspicuous being why the phenomenon in (44) isn’t found in, say, English. At this moment, we can only point at a possible direction of investigation: The Chinese-English contrast at issue may be correlated to another one, namely that Chinese nouns depend on classifiers for quantitative specification whereas English nouns don’t. If the use of classifiers means a noun contains no information about the unit of quantity measurement, it may become more tolerant for combining with quantitative constituents (of which DrP is one), provided that each
constituent eventually receives full interpretation (cf. Chapter 2) through legitimate linguistic mechanisms such as Incremental Theme. Similarly, on the account assuming verb-movement out of nominalized domains, the high degree of analyticity in Chinese provides for a route for syntactic V movement that results in apparent syntax-semantics mismatches. In particular, as Hale and Keyser (1993) have argued, many unergative, action verbs in English are derived via denominalization in the lexicon. It is not unnatural that the same process can occur in Chinese, but in more analytic fashion in the domain of syntax.

3.3. Preverbal constituents

Given the proposal that FP/DrP, when used as adverbials, adjoin to the left of V’, it is plausible that other types of adverbial phrases may adjoin to the left of v’. This immediately accounts for preverbal PP and ADV modifiers:

(48) a. ta cong Xi’an hui-lai-le.

he from Xi’an come-back-LE.

‘He came back from Xi’an.’

b. wo xiaoxinyiyi de xie-le yi-feng xin.

I cautious DE write-LE one-CL letter

‘I cautiously wrote a letter.’
Adjoined to \( v' \), the adverbial is to the left of \( v \), to which the lexical verb raises to generate the verb-object word order. The only detail to be added is that multiple adverbials are allowed preverbally while postverbal FP and DrP exclude each other:

\[
\begin{align*}
\text{(49) a. } & \text{ ta cong Xi’an xiaoxinyyi de hui-lai-le.} \\
& \text{he from Xi’an cautious DE come-back-LE} \\
& \text{‘He cautiously came back from Xi’an.’} \\
\text{b. } & \text{*wo mai-guo liang ci yi nian yu.} \\
& \text{I sell-GUO two time one year fish} \\
& \text{Intended reading: ‘I sold fish for a year twice.’}
\end{align*}
\]

An explanation for (49) may exist somewhere between two possibilities. First, if (49b) is taken to mean that each XP may have no more than one adjunction to \( X' \), either in the nature of the syntactic structure (see Kayne 1994 for a theory in this direction) or due to semantic reasons, then multiple preverbal adjuncts should be equated to multiple functional phrases (cf. Chapter 1), as is indeed proposed in Cinque (1999). In such a theory, (49a) is the result of having the PP and AP adjuncts each adjoin to a different functional \( X' \) at least as “high” as \( v' \) while \( V \) only raises up to \( v \). The word order follows automatically, as the reader can verify. Alternatively, it may be the case that lexical
categories allow only one adjunction in each phrase but functional categories, including $v$, allow an indefinite number of adjunctions.\footnote{A version of this idea is already proposed to handle multiple subjects in East Asian languages like Japanese, first in Fukui and Speas (1986) in the government-binding theory and later in Ura (1996) in the framework of Chomsky’s (1995) Minimalist Program.}

A few more types of constituents occur before the verb and after the subject, including the aspectual auxiliaries you (perfective as in mei-you) and zai (progressive) and modals such as neng ‘can’ and yinggai ‘should’. The traditional wisdom is that these are all part of the predicate. More recent research confirms this insight, but distinguishes the traditional sense of predicate from the structurally defined VP. In fact, there are reasons to believe that some of these constituents are outside VP. Nonetheless, we will examine certain syntactic details of these elements in the rest of this chapter because, after all, they are intrinsically associated with the verb.

3.3.1. Aspectual phrase

From the cross-linguistic perspective, it is obvious that human languages distinguish tense (T) from aspect (ASP). Briefly, tense marks the relation between the time of a described event and the time at which the description is given whereas aspect signals the speaker’s viewpoint on the progress of the event (cf. Smith 1991): the perfective aspect focuses (typically) on the final state of the event and the progressive aspect on an interval somewhere between the event’s beginning and end. This section examines the aspectual morphemes in Chinese. Given the fact that most languages have morphemes for tense, we
also assume that tense exists in Chinese, though no part of this book hinges on this assumption. For some possible motivations for syntactically represented tense in Chinese, see A. Li (1985, 1990), Simpson and Wu (2002). For alternative views see Lin (2003, 2006) and references cited.

Chinese has two systems for aspect, preverbal and suffixal, illustrated respectively in (50) and (51) with the aspectual morphemes in bold face:16

(50) a. ta **zai** chang ge.
   he at sing song
   ‘He was singing.’

b. wo **mei-you** hui jia.
   I not-have return home
   ‘I didn’t go home.’

(51) a. ta **chang-zhe** ge.
   ta sing-ZHE song
   ‘He was singing.’

b. wo hui-**le** jia.
   I return-LE home
   ‘I went home.’

c. zhe-ge ren **sha-guo** laohu.

---

16 The semantic properties of some of these aspectual morphemes are discussed below. The English translations of the examples are only approximate because the aspectual information in each Chinese sentence is often difficult to show precisely with a single English word. See Smith (1991) for details discussions and comparisons of aspectual morphemes in several languages, including Chinese and English.
This person once killed a tiger.

Anticipating a unified analysis of both systems to be introduced shortly, we adopt the following structure for aspect:

The preverbal you and zai directly fit into the ASP position, provided that NP1 eventually moves to whatever clause-initial position for the subject. The location of the suffixal -zhe, -le and -guo, however, is not as straightforward.

Conceptually, it is clearly desirable that they are affiliated with ASP. If they are also generated under ASP, the fact that they are affixed to the verb could only be because the morphemes undergo a merging process which, in syntax, means moving one of them to the other. Recall from Chapter 2 that movement must be out of a complement and target a c-commanding position. The configuration in (52) meets both requirements if V raises, via v, to ASP. There is evidence, however, that V doesn’t go out of vP.17

---

17 The argument below is based on Cheng and Li (1991), which in turn takes advantage of the insight in Pollock’s (1989) comparative study of French and English.
(53) a. ta zai dasheng chang ge.
    he at loud sing song
    ‘He was singing loudly.’
   b. *ta dasheng zai chang ge.
    he loud at sing song
    Intended reading: Same as (53a).

(54) a. wo mei-you qiaoqiao de hui jia.
    I not-have quiet DE return home
    ‘I didn’t go home stealthily.’
   b. *I qiaoqiao de mei-you hui jia.\(^{18}\)
    I quiet DE not-have return home
    Intended reading: Same as (54a).

Since the modifiers `dasheng` ‘loud’ and `qiaoqiao de` ‘quietly’ occur before the verb but after ASP, they must be adjoined to \(v\) (cf. (52)).

Now consider the linear relation between such modifiers and the aspectual suffixes:

(55) a. ta dasheng chang-zhe ge.
    he loud sing-ZHE song

\(^{18}\) This sentence is good with the reading that Zhangsan was cautious (about the trip) and therefore didn’t go home, which is irrelevant to the issue here.
‘He was singing loudly.’

b. *ta chang-zhe dasheng ge.

he sing-ZHE loud song

Intended reading: Same as (55a).

(56) a. wo qiaqiao de hui-le jia.

I quiet DE return-LE home

‘I went home stealthily.’

b. *wo hui-le qiaqiao de jia.

I return-LE quiet DE home

Intended reading: Same as (56a).

(57) a. na-ge jiahuo chishoukongquan sha-guo laohu.

that-CL guy bare-handed kill-GUO tiger

‘That guy once killed a tiger bare-handedly.’

b. *na-ge jiahuo sha-guo chishoukongquan laohu.

that-CL guy kill-GUO bare-handed tiger

Intended reading: Same as (57a).

Descriptively, the verb-suffix cluster must occur after the v’-adjoined modifiers (cf. (a) examples) and not before ((b) examples). This is not expected if –le, for instance, occupies the ASP position with the verb raising out of vP to merge with it.

The solution lies in one of the oldest ideas in linguistics combined with one of the major discoveries in modern syntax. It has long been tradition to regard a verb plus its inflectional affix as a form of the verb. Take the English verb play-s for example. While
the correct use of this inflected verb is clearly dependent on the syntactic context, play-s itself can be formed with a word-formation rule independently of syntax. By the same logic, the concatenation of hui ‘return’ with the perfective suffix –le should not rely on syntactic movement either, as long as the syntactic context that hui-le occurs in guarantees a match between the aspectual information of the clause and the suffix –le. In particular, if ASP is the syntactic node representing aspect but hui-le ‘return-Perf’ as a verb form is initially placed under V, syntax must provide a way to match the perfective suffix on the verb with whatever aspectual information coded under ASP. This is easily accomplished given the progress of our syntactic knowledge in recent years.

A prominent assumption in recent and current syntactic theory is that there exists an abstract level of syntactic derivation, called Logical Form (LF), which will be the central concern of some later chapters of this book. For now, it may be described as follows. While various constituents undergo movement in syntax, some of these movements happen “prior to” the point at which the sentence is uttered. All these pre-utterance movements are overtly reflected because the moved constituents are already in their landing sites at the point of utterance. But certain constituents move after that point, at the abstract level of LF. In these cases, the movement is not heard for the simple reason that by the time of utterance, the movement has not taken place yet. Anticipating independent evidence for LF later, we suggest that hui-le ‘return-LE’ in (56a), for instance, moves from V to v overtly but continues to move to ASP covertly at the level of LF. Since the second step of movement is covert, hui-le is heard in the v position after the adjunct phrase qiaohiao-de ‘quietly’. But because hui-le eventually lands in ASP, albeit covertly, the perfective –le does end up in ASP, thereby matching itself with the syntactic
node that carries the aspectual information. It should be pointed out that the covert movement at LF is also expected to be subject to all restrictions on movement. Given (52), the landing site, ASP, c-commands the v position from which hui-le moves, and the movement is out of vP, the complement of ASP. 19

This syntactic representation of aspect also helps us understand the contrast below:

(58) a. *ta mei-you hui-le jia.

he not-have return-LE home

Intended reading: ‘He didn’t go home.’

b. ta mei-you hui-guo jia.

19 Historically, there has been a fundamentally different approach to the one presented in the text, which, in the case of the Chinese data, takes V-Asp as a result of Affix Hopping in Phonetic Form (PF), the component of grammar that is pronounced but does not have direct consequences on the meanings of sentences. That is, we may assume that –le is an affix heading the Asp Phrase. Rather than the verb raising to Asp, the affix –le lowers to the verb. This gives the correct word order as desired, and because this “lowering” is not syntactic but phonological, it is not subject to the constraints for syntactic movement. This is in line with a number of recent treatments of English main verb morphology, and the English-French differences (Emonds 1978, Pollock 1989, Lasnik 1999) with respect to the position of the main verb. Chinese, in this respect, behaves on a par with English, in contrast to French. But unlike English (which does raise its be and auxiliary have to T), no auxiliary raising across adverbials or negation occurs in Chinese (see Huang 1994b). A variant of this idea is for both a functional morpheme such as –le and the lexical verb to stay in their respective positions and to merge into a morphological complex at PF via linear adjacency. Certain typological facts may be accounted for this way (Baker 2002, Y. Li 2005), with implications regarding more fundamental distinctions between adjuncts and other constituents in the X’-structure.
In (58a), the perfective preverbal *you* and the perfective suffixal –*le* cannot co-exist. This follows if both morphemes reflect the same aspectual information under ASP. Then it is natural that the same information under the same syntactic node does not get manifested twice. For the same reason, the co-existence of *you* and the experiential suffix –*guo* is possible because the two morphemes do not carry the identical information (cf. Smith 1991). Needless to say, each aspectual marker has its own ASPP. When *you* and –*guo* are both present, the relevant structure prior to LF movement is (59):

![Diagram]

By convention, a clause contains as many ASPPs as there are identifiable aspectual markers. In the absence of any such marker, no ASPP is present in the structure. What
remains to be answered is why the negative form of perfective aspect must choose you and not –le.²⁰

3.3.2. Modals

Examples of Chinese modals are given in (60), drawn from D. Zhu (1982):

(60) a. keneng ‘be possible’, hui ‘be likely to’, keyi ‘be permitted to’,
yinggai ‘ought to’, gai ‘ought to’, …
b. gan ‘dare’, ken ‘be willing to’, yuan yi ‘be willing to’, yao ‘want to’,
neng, ‘be able to’, nenggou ‘be able to’, keyi ‘be able to’, hui ‘be able to’, …

²⁰ There are many possible reasons for this. It may be the result of a historical accident that the Mandarin dialect of Chinese acquired two separate perfective markers along different routes of linguistic change, and opted to make use of them for separating positive and negative clauses. Or there may be a deeper reason for the distinction, considering the similarity, at least on the surface, between the Chinese perfective aspect and the English tense, both are represented with functional categories:
i. ta de-le jiang. → ta mei you de jiang.
ii. He received an award. He did not receive an award.

In both cases, a suffix is used on the lexical verb in the positive form while a totally different morpheme accompanies negation in a preverbal position. Still another possibility may be that negation (Neg) is a bound form whose morphological host must have certain yet unclear properties. In the case of perfective aspect, the host must be the Asp (because mei you kanjian ta denotes the negation of the perfectiveness).

This can be done with you, easily. But –le is too far for bu to attach to, with V-le raised to Asp only in LF, too late to support Neg.
It will become clear why we divide them into two groups and why some occur in both groups. The discussion in this section is primarily based on Lin and Tang (1995).\textsuperscript{21}

In languages like English, modals have been traditionally placed under Inflection (I), a functional head position. Though recent work (e.g., Cinque 1999) has explicitly argued that this treatment may be overly simplistic, there is still consensus that English modals belong to a functional category. (61) represents the (simplified) structure of the phrase headed by an English modal, using I merely as a convenient label (no ASPP is shown):

\begin{center}
(61)
\begin{tikzpicture}
  \node (IP) {IP};
  \node (I) [below] {I'};
  \node (vP) [right of=I'] {vP};
  \node (I') [below left of=IP] {I'};
  \draw (IP) -- (I);
  \draw (I) -- (vP);
\end{tikzpicture}
\end{center}

Such an analysis of modals cannot be directly adopted for Chinese, however. At least those in (60a) have distinctive behaviors of lexical verbs.\textsuperscript{22} Below, we examine arguments for this claim.

First, while all modals in (60) may occur between the subject and the predicative VP as expected, many of those in (60a) may also occur after a full clause, especially in colloquial Chinese:

\textsuperscript{21} Lin and Tang’s work is in turn a further development of Huang (1988a). For earlier discussion, see T. Tang (1979).

\textsuperscript{22} D. Zhu (1982) also lists modals under the large category of verbs.
(62) a. ni cizhi keyi, ta jieban bu xing!
   you resign be.permitted he take.over.one’s.position not all.right
   ‘You may resign, but he can’t be hired for your position!’

b. tamen zheyang xiang bu yinggai.
   they this.way think not ought.to
   ‘They shouldn’t think this way.’

c. zhe-ge ren shou fa gai-bu-gai?
   this-CL person receive punishment should-not-should
   ‘Should this person receive the punishment?’

d. rang wo gen ni zuodui keneng ma?
   make me with you oppose be.possible Q
   ‘Is it possible to let me be against you?’

The sentence-final location of the modals in (62) is not expected if they are under I
because vP comes after I, not before it. On the other hand, if these modals are in fact
lexical verbs that can take a clause as a subject, the word order in (62) is exactly what is
predicted. Take the first half of (62a) as an example (with V raised to v):

(63)

```
  S
    vP
      v'
        v
          ni cizhi keyi t
```
This structure is further confirmed by the parallelism between the two halves of (62a), a pattern used widely in Chinese under the condition that the two aligned constituents have the same syntactic structure. Since xing ‘be all right’ in the second half of (62a) is clearly used as the matrix verb taking a clausal subject (ta jieban ‘that he took your job’), the modal keyi in the first half must be comparably constructed. The same conclusion is also corroborated by the fact that these modals can take the pronoun zhe ‘this’ as the subject that refers to the previous sentence, illustrated schematically with (64) below:

(64) a. ni dasuan cizhi? zhe zenme keyi?
    you plan resign this how be.permitted

    ‘You plan to resign? That’s not allowed!’

b. ni touxiang-le? zhe bu keneng.
    you surrender-le this not be.possible

    ‘You surrendered? That’s impossible.’

Also common is for a modal of this group to occur sentence-initially:

(65) a. keyi ni qu, ye keyi ta qu.
    be.permitted you go or be.permitted he go

    ‘You may go or he may go.’

b. (ying)gai zanmen zhexie ren de jiang.
    ought.to we these people receive award
‘It should be that we people get an award.’

c. keneng tamen yao canjia bisai.
be.possible they will participate in competition

‘It’s possible that they will participate in the competition.’

d. hui-bu-hui²³ ta xiang chuguo?
be.likely-not-be.likely he want go.abroad

‘Is it likely that Zhangsan wants to go abroad?’

Modals of the functional type in English do not allow this word order in declaratives. The word order is indeed required in yes-no questions, but the so-called subject-auxiliary inversion applies to modals and other auxiliaries alike. This is clearly not the case for Chinese. On the other hand, if the modals in Chinese are lexical verbs, the subsequent clause can be simply treated as the object of each modal verb, directly yielding (65).

In principle, the data in (62) and (65) plus the “default” use of the modals between the subject and the predicative VP can all be subsumed under a single structure (irrelevant details are ignored):

(66)

²³ For reasons unknown to us, hui must be in A-not-A form when appearing before the subject.
The modal is the lexical verb taking a clausal complement S. If no movement takes place, the modal is the first constituent from the left, yielding (65). If the whole S raises to the matrix subject position, marked “∅”, the modal becomes the last constituent one can hear and (62) results. If, on the other hand, only the subject NP of S raises to the higher subject position ∅, the modal occurs between the raised NP and the rest of S, i.e., the predicative vP-VP. The reason for raising a complement clause to the subject position has never been fully understood. Without getting into technicalities, we note that the three optional word order arrangements with respect to the modals in (60a) are by no means a language-specific oddity. Consider these English examples with the adjective likely:

(67)  a. It is likely [ for Shawn to go abroad ].
     b. [ For Shawn to go abroad ] is likely.
     c. Shawn is likely [ to go abroad ].

The bracketed constituent is known to be an non-finite clause (cf. Chomsky 1981 and the references cited there), which stays in the complement position of likely in (67a), and raises to the subject position of the whole sentence in (67b). In (67c), only the subject of the non-finite clause raises, leaving the clause itself in situ. Except the fact that English has a semantically empty pronoun it, called an expletive, to overtly fill up the subject position in the first case, (67) patterns with the Chinese examples examined so far.

In comparison, modals in (60b) occur only between the subject and vP. Though this fact may classify such modals as functional words like can and should in English, we
would like to highlight the fact that the modals in (60b) seem to assign a theta-role to the subject NP before them. In the framework adopted in this book, this fact has non-trivial implications. For ease of discussion, we refer to the lexical modals in (64a) as *raising modals* and the group in (64b) as *control modals*, for the reason to be made clear below.

To begin with, a control modal imposes a selective restriction on the semantics of the subject NP. A raising modal has no such restriction, accepting any subject NP that is compatible with the verb after the modal. Since the specific restrictions from the control modals vary, the examples below are meant to be illustrative but not exhaustive:

(68) Raising Modals

a. wo yinggai/keneng/keyi/hui chang yi-shou xiaoqu.

   I should/will.possibly/be.allowed.to/be.likely.to sing one-CL ditty

   ‘I should/will possibly/is allowed to/is likely to sing a ditty.’

b. zhe-shou xiaoqu yinggai/keneng/keyi/hui chang yi-dian.

   this-CL ditty should/will.possibly/be.allowed.to/be.likely.to long a-bit

   ‘This ditty should/will possibly/is allowed to/is likely to be a bit longer.’

(69) Control Modals

a. wo gan/ken/neng/hui chang yi-shou xiaoqu.

   I dare/be.willing.to/be.able.to/be.able.to sing one-CL ditty

   ‘I dare/is willing to/am able to/am able to sing a ditty.’

b. *zhe-shou xiaoqu gan/ken/neng/hui chang yi-dian.

   this-CL ditty dare/be.willing.to/be.able.to/be.able.to long a-bit

   ‘*This ditty dares/is willing to/is able to/is able to be a bit longer.’
Especially worth noting is that (68b) is acceptable with *hui* only when the modal has the probability reading; the sentence becomes bad when *hui* has the ability reading in (69b).

Intuitively, this contrast between the two groups of modals is easy to understand – a control modal has the intrinsic semantics that requires the subject to have certain qualities such as sentience and free will. In modern syntax, such restrictions on arguments are typically attributed to a thematic relation, namely that the modal assigns a theta-role to the subject. Then according to the theta-criterion, the NP subject preceding the modal in, say, (69a), must not be the actual subject of the verb *chang* ‘sing’ for the simple reason that otherwise, *wo* ‘I’ would receive two theta-roles, from the modal and *chang*, simultaneously. Put differently, *wo* could not have originated as the subject of *chang* and then raised to its sentence-initial position before the modal. In fact, there is only one basic structure for (69a):

(70) \[
\text{NP, modal } [\text{X Pro, V \ldots }]
\]

where Pro is coreferential with the NP subject of the modal. In syntax, the relation between a Pro and its c-commanding antecedent is called *control*, which is why we referred to this group of modals as control modals.

### 3.4. Summary

In brief, these initial chapters provide the foundation for the following ones. Chapter 1
examined the definitions of various categories, arguing for a feature-based theory that not only is capable of accommodating enough different classes of words and morphemes for syntactic analysis but also offers a way to capture certain shared properties among such classes. Chapter 2 explored the nature of theta-roles, the type of semantic relations between a verb and its arguments that form the basis for combining words into a sentence. Chapter 3, focusing on the internal structure of VP, investigates how a verb combines with phrases of various other categories, revealing different patterns among those that receive theta-roles from the verb and those that do not. Overall, we hope to have shown that sentence-formation follows rigorous patterns, the discovery of which has allowed us to understand a wide range of linguistic facts. The subsequent chapters are built on this foundation, each of them studying a particular phenomenon in Chinese syntax in great detail.