Chapter 1  Categories

We take it as our starting point that a Chinese sentence is composed of words and that words have different behaviors in a sentence. For instance, while dayan fei ‘wild.goose fly’ is an acceptable sentence, *fei dayan ‘fly wild.goose’ is not. The most obvious reason for the contrast is that dayan ‘wild.goose’ is a noun that canonically serves as the subject of the sentence and fei ‘fly’ is a verb whose canonical function is to be the predicate occurring after the subject. This means that in order to understand the syntax of Chinese, or the syntax of any language for that matter, we minimally need to understand how the words in a language are classified and how these different classes of words are put together to form sentences. In this book, word classes are referred to as lexical categories, or just categories for short, following the terminological convention of generative syntax.

While the basic distinction between nouns and verbs is universally recognized in modern literature on Chinese syntax, scholars differ, sometimes drastically, on other categories. See Chao (1968), Li and Thompson (1981), Zhu (1982), and Xing and Ma (1992) for a few examples. The differences in opinion arise partly because linguists with different theoretical backgrounds may employ different criteria for word classification, and partly because we still lack sufficient knowledge about certain words and their properties. Regardless, it is without question that the ultimate task for anyone studying lexical categories in Chinese is to identify them in such a way that they both allow an accurate description of the syntactic behaviors of the language, and provide insights into the nature of word classification.
With this goal in mind, we will introduce a theory of lexical categorization in Mandarin Chinese in this chapter. The theory consists of two intertwined parts. First, a set of categories is confirmed and examined on the basis of the syntactic behaviors of Chinese words and morphemes. Second, a decompositional theory that characterizes the intrinsic relations among these categories is defended. It is important to mention up front, however, that we do not intend to spread our discussions evenly among all issues related to lexical categorization, nor do we attempt to provide an exhaustive list of categories in the language. Rather, the chapter concerns itself primarily with where we believe new insights are available from recent research. The same approach also applies to the organization of the whole book.

1.1. Lexical categories

This section focuses on verbs (V), nouns (N), prepositions (P), and adjectives (A).

1.1.1. Verbs and Nouns – Basic distinctions

It is common wisdom in modern linguistics that N and V are two basic categories. In Chinese, the two categories can be clearly distinguished on the basis of their modifiability by the negative morpheme bu. The basic data is given in (1)-(2):

(1) Verbs
   a. bu shui ‘not sleep’
b. bu tongzhi ‘not inform’

c. bu sai-qiu ‘not play ball’

(2) Nouns

a. *bu shu ‘not tree’

b. *bu xiaoxi ‘not news’

c. *bu qiu-sai ‘not ball game’

To our knowledge, all verbs can be negated by *bu, and no noun can. It must be pointed out that *bu can also negate adjectives such as da ‘big’ and lei ‘tired’. As we will see in subsequent sections, this similarity between verbs and adjectives poses no problem for the N-V distinction.

Examples exist in modern Chinese that seem to suggest that nouns can be modified by *bu, such as *bu-ren-bu-gui ‘not-human-not-ghost’. However, there are reasons for not regarding such examples as a problem for the *bu-test of the noun/verb distinction. First, they are not formed with a productive process. A change of nouns typically results in unacceptability:


b. *bu-fan-bu-cha ‘not-food-not-tea’
Second, the nouns in these examples must be monosyllabic, even when multi-syllabic counterparts exist, further confirming that _bu_ cannot really modify a noun in modern Chinese:

\[(4)\]
\[\text{a. } \,*bu\text{-huoren-bu-sigui ‘not-live.human-not-dead.ghost’} \]
\[\text{b. } \,*bu\text{-renlei-bu-guilei ‘not-humankind-not-ghost.kind’} \]

Lastly, even with the nouns that _bu_ can accompany, a single _bu-N_ pair is not permitted, contrasting sharply with verbs in (1):

\[(5)\]
\[\text{a. } \,*bu\text{-ren ‘not-human’} \]
\[\text{b. } \,*bu\text{-gui ‘not-ghost’} \]

As a result, we regard the few exceptions not as undermining the reliability of the _bu_-test, but as idiomatic expressions not subject to the general rules we are pursing.

N and V also differ in many other ways reported in various grammar books (e.g., a subset of V allows aspectual suffixation, while no word used as N does). For the present chapter, the data below are of particular interest:

\[(6)\] Verbs
\[\text{a. meiti baodao-le na-ci shigu.} \]
\[\text{media report-LE that-CL accident} \]
\[\text{‘The media reported that accident.’} \]

Zhangsan translate-LE one-CL novel

‘Zhangsan translated a novel.’

c. laoshi piping-le zhe ji-ge yanjiusheng.

teacher criticize-LE these some-CL graduate.student

‘The teacher criticized these graduate students.’

(7) Nouns

a. meiti *(dui) na-ci shigu de baodao¹

media on that-CL accident DE report

‘the media’s report of that accident’

b. Zhangsan *(dui) yi-bu xiaoshuo de fanyi

Zhangsan on one-CL novel DE translation

‘Zhangsan’s translation of a novel’

c. laoshi *(dui) zhe ji-ge yanjiusheng de piping

teacher on these some-CL graduate.student DE criticism

‘the teacher’s criticism of these graduate students’

¹ Parentheses are another notational convention. The expression between a pair of parentheses is optional. E.g., A(B)C indicates that both AC and ABC are acceptable facts. If an asterisk "*" immediately precedes the expression inside the parentheses, as in A(*B)C, then AC is acceptable but ABC is not. If instead the asterisk immediately precedes the left parenthesis, as in A*(B)C, then ABC is acceptable but AC is not. All the examples in (7) are of this type.
The two groups of examples, though both based on *baodao, fanyi* and *fanxiu*, exhibit three differences. Take (6a) and (7a) for example. First, the semantic object occurs to the right of *baodao* in (6a) but to the left in (7a); second, a preposition *dui* is required to introduce the object only in (6a); third, the morpheme *de* is required before *baodao* in (7a). The nature of these facts will become clearer as we proceed. For now, it is sufficient to note that nouns depend on prepositions like *dui* for the grammaticality of their object whereas verbs do not. This is a very reliable test to separate N from V, with the limitation that it only applies where the semantic subject of the N/V is present.  

The fundamental distinction between N and V might be a reflection of proto-categories, a concept that traces its origin to psychological studies of human cognition. It is possible that our brain divides the world into two elementary kinds of entities: things that exist and situations that take place. Proto-N is the linguistic representation of the former kind and proto-V, that of the latter kind. All specific lexical categories are then the derivatives of these two proto-categories. Let us represent the proto-categories as two features, \([N] \text{ and } [V]\). Since a word either belongs to proto-N or does not belong to proto-N, the feature for this proto-category has two values, \([+N]\). The same logic leads to \([+V]\). These two binary-valued features yield four possible combinations: \([+N, -V]\), \([+N, +V]\), \([-N, -V]\), and \([-N, +V]\). If these feature combinations indeed correspond to lexical categories in languages, then it is obvious that nouns are \([+N, -V]\) and verbs are \([-N, +V]\). That is, a noun conforms to proto-N but not to proto-V.

---

2 This analysis is adopted from Y. Li (1997a). See Fu (1994) for a different treatment of the data.

3 Cf. Givon (1984) and the references cited there.
whereas a verb conforms to proto-V but not to proto-N. The hypothesis can be summarized in a feature matrix:

(8) Feature-based characterization of lexical categories (preliminary)

<table>
<thead>
<tr>
<th>Feature Category</th>
<th>N</th>
<th>?</th>
<th>?</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>[N]</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[V]</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

To avoid confusion, we make a notational clarification: [N] and [V] are categorial features which we suggested to represent proto-categories; N and V, on the other hand, are the shorthand names of the actual lexical categories that can be decomposed into combinations of categorial features. See Chomsky (1970) for the onset of this theory. Following convention (cf. Freidin 1991), the characteristic property of a noun-like category (i.e., [+N]) is its inability to take a nominal object, at least in the absence of other linguistic help, whereas [+V] is defined as the ability to function as the predicate of a standalone sentence.

A natural question arises about (8): what are the lexical categories represented by [+N, +V] and [-N, -V], which are marked with “?” in the table? Answers will be given later in this section, after a unique type of nouns is examined.

1.1.2. Localizers

---

*4 The discussion of localizers is in part based on Y. Li (1983, 2003).*
The examples below illustrate a set of words whose categorial status has always been controversial:

(9) a. wuzi li/limian
    room inside

b. chuang xia/dixia
    bed    underneath

c. da shu pang/pangbian
    big tree  side

Each expression in (9) consists of a noun followed by what Chao (1968) called a localizer (abbreviated as L). That localizers resemble nouns in syntax is widely recognized (see A. Li 1990, Y. Li 2003, and the references therein). First, to the extent that the examples in (9) are treated as phrases, which we refer to as localizer phrases (LPs) for now, they can serve as the subject and object in a sentence:

(10) a. tade chengshi/cheng wai  hen meili.
    his    city    /city    outside very beautiful
    ‘His city/The outside of the city is beautiful.’

b. wo qu-guo tade chengshi/cheng wai.
    I    go-GUO    his    city    /city    outside
    ‘I have been to his city/outside the city.’
Secondly, just as N is the last word in a noun phrase (NP), L also trails all other components in an LP, as seen in *tade chengshi* vs. *cheng wai* in (10). In syntax, this word order is referred to as “head-final,” with N and L being the “heads” of their respective NP and LP.

Another similarity between LP and NP is seen through the examples in (11):


   he P that city hold-GUO a-CL exhibition

   ‘He held an exhibition *(in) that city.’

b. *ta* *(zai)* cheng wai/li juban-guo yi-ge zhanlanhui

   he P city outside/inside hold-GUO a-CL exhibition

   ‘He held an exhibition outside/inside the city.’

For reasons to be made clear in a later chapter, NPs not functioning as the subject or the object usually need a pre/postposition (P) to occur in a sentence. As (11a) and its English translation show, this is apparently a cross-linguistic fact. In this regard, the LP in (11b) behaves exactly like NP, relying on a locative preposition to be well-formed. The same data also argues against treating L as a postposition (cf. Tai 1973, Peyraube 1980 and Ernst 1988). If L were a postposition, there would be no reason why it should not behave like one, and its presence in (11b) would be enough to introduce the nominal *cheng* ‘city’ just like *outside* does in English.
There is one property of L, however, that does distinguish it from N and make it resemble a postposition. It is the interaction between L and de, which we will turn to next. To better facilitate discussion, monosyllabic and disyllabic localizers will be examined separately.

1.1.2.1. L, de and classifiers

Starting with monosyllabic L, let us first consider the interaction between the use of de and a group of words in Chinese called classifiers (CL). In the presence of numerals and demonstrative pronouns, a Chinese noun usually needs a classifier to specify the “unit” with which the entities denoted by the noun are measured. Crucially, different nouns require different classifiers (To maintain a minimal-pair comparison between nouns and localizers, all examples in this subsection are composed of monosyllabic morphemes at the point of relevance):

(12) yi-zhang chuang, si-tiao tui, zhe-ke shu, na-pian pi
    one-CL bed       four-CL leg      this-CL tree       that-CL bark

‘a bed’ ‘four legs’ ‘this tree’ ‘that bark’

The dependency between a noun and its classifier displays interesting patterns when two nouns are concatenated with or without de:

(13) a. si-tiao chuang tui       vs.       *si-zhang chuang tui
four-CL bed leg four-CL bed leg

b. zhe-pian shu pi vs. *zhe-ke shu pi
this-CL tree bark that-CL tree bark

(14) a. (?si-tiao chuang de tui vs. si-zhang chuang de tui
 four-CL bed DE leg four-CL bed DE leg
 b. (?)zhe-pian shu de pi vs. zhe-ke shu de pi
 this-CL tree DE bark this-CL tree DE bark

In brief, *de is independently optional between two nouns; however, a CL must match the N on the right (i.e., the head N) in the absence of *de but can optionally associate with either N when *de is present.

The explanation for the pattern in (13-14) is both simple and intuitive. Suppose that the N-N cluster without *de is always a compound whereas the one with *de in between is an NP in which the noun on the left modifies the one on the right. In other words, *de necessarily and sufficiently signals a phrasal structure in the context of two consecutive nouns. Furthermore, Chinese N-N compounds are “head-final” because it is the noun on the right that determines the basic semantics of the word – a chuang-tui ‘bed leg’ is a kind of leg but not a kind of bed. As a result, only the classifier tiao, appropriate for legs but not for beds, is permitted in (13) where *de is absent. When *de is present, as in (14), the two nouns are not components of a single compound word; rather, each of them is a separate word in syntax, yielding the structure in (15) in which the pairs of brackets mark out the boundaries of noun phrases, a notational convention widely used in syntax:
The “…” in each NP is where a classifier plus a numeral/demonstrative may occur. Since there are two separate nouns, each one of them may be associated with its own classifier in syntax. In other words, a classifier at the beginning of such a string may be alternatively treated as part of NP1 (i.e., associated with leg) or part of NP2 (associated with bed). Linearly, CL occurs in the same spot; the options result from different levels of word-associations that are made possible in syntax. With this explanation, we can also understand why the first example in each pair of (14) sounds somewhat strange. In these examples, the adjacent CL-N sequence is expected to match but does not because the CL is really for the second noun, thus causing difficulty in processing. Also note that no CL is required for (15) to be acceptable (e.g., chuang de tui ‘bed DE leg’). We thus assume that an NP may be composed of a bare noun.

Turning to LPs headed by monosyllabic localizers, we note two facts: that no de is allowed between L and the preceding N, and that a CL before a N matches that N without a hitch:

(16) chuang (*de) xia,       men (*de) hou,       wu (*de) li
    bed              underneath door         behind        house         inside

(17) yi-zhang chuang xia,   zhe-shan men hou,   na-jian wu li
    one-CL bed       underneath this-CL door       behind         that-CL house       inside
(16) distinguishes L from N and underlies the proposal that L is a postposition. (17) contrasts sharply with the *de*-less N-N compounds in (13). In both (13) and (17), only monosyllabic words are used, and in both, *de* is absent. But when the rightmost morpheme is N, a classifier matching the left-hand N is totally unacceptable; when the rightmost morpheme is L, however, the otherwise identical choice of words yields 100% acceptability. Given our explanation for nominal examples in (13-14) above, it can be deduced that LP has the following structure:

\[(18) \; [LP \; [NP \; \ldots \; N] \; L]\]

Crucially, the pre-L nominal does not form a compound with L. Instead, it has its own phrase in which a classifier is permitted. Put differently, while the lack of *de* between L and the preceding N makes the cluster resemble an N-N compound in form, we have evidence now that there really is a phrasal structure.

### 1.1.2.2. L as a subclass of N

The structural analysis of the behavioral contrast between L and N in the previous subsection only serves to highlight an old question: What is the best categorial classification of L that explains its syntactic properties? Logically, there are three possible answers: L is a subclass of N, L is a postposition, or L is a separate category. In
this book, we offer a theory in which the properties of L may be understood by pursuing the first possibility.\(^5\)

As we saw earlier, L exhibits three characteristic properties of N: LP is head-final, it functions as the subject and the object in a sentence, and it needs a preposition if used as a locative modifier. L also appears like a postposition because no *de* is used to associate L with the NP before it. What remains unclear is whether this NP is the object of L in the same sense that a preposition takes an NP object, since Chinese is lacking morphological cues that would help identify the NP’s grammatical function. At least in theory, it is equally possible that the NP plays the role of a “possessor” in LP. Overall, if we are to choose a category for L between N and P, N seems more appropriate.

The question, then, is how to account for the lack of *de* if L is viewed as a type of N. It should be obvious that some stipulation is unavoidable in order to allow L to be N but still different from N. To this effect, we hypothesize that a language may allow a (natural) subclass of words in a given category X to “deviate” behaviorally from X. Meanwhile, we propose that such deviations are not random but rather the result of a predictable nature.

As the first step in our account, consider *do*-insertion in English:

\(^5\) At a more fundamental level, a categorial deviate may not be distinguishable from a new category. The more important question is whether languages may potentially allow any new category, or whether even a new one must be subject to the same principle as the core set. The second choice is obviously more restrictive and therefore is assumed here. See Y. Li (2003) for alternatively treating L as a new category “between” N and P. The basic reasoning in this section applies there too.
(19)  a. Did Sam leave?
    b. Sam did not leave.
    c. Sam left.
    d. *Sam did leave.

The semantically empty modal *do* is used in forming the interrogative in (19a) and the negative in (19b). But if *do*-insertion is a legitimate operation in English grammar, why can it not happen in the declarative (19d)? Note that (19d) would be good if *did* undertook the emphatic interpretation (plus the corresponding stress). The emphatic *did* is not the same morpheme as in (19a-b) and thus not relevant for the current discussion.

Chomsky’s (1995) answer to this question is couched in the technical terms of Minimalist Program but is, to us, intuitive in its essence. He suggests that highly language-specific operations such as *do*-insertion are more “costly” in linguistic computation because they have to be learned. In comparison, the mechanisms observed cross-linguistically are presumably hardwired in the brain and come for free. If linguistic derivations somehow try to avoid more costly operations whenever possible, then the data in (19) is easy to explain. (19c) is chosen over (19d) because it results from a “cheaper” derivation by not incurring the cost of language-specific *do*-insertion.

Independent reasons require the presence of a modal in interrogatives and negatives, so not having one necessarily leads to ungrammaticality. *Do*-insertion is justified in these cases only because there is no other grammatical way to form a question or negation. acceptable (19a-b).
Interestingly, the use of *de* is also highly language-specific. Suppose that the syntactic properties of L are decided according to (20), which in turn may be the specific manifestation of a more general principle which also produces the *do*-insertion data in (19):

(20) In deciding the properties of a categorial deviate, anything language-specific in the original category is disfavored.

Of the cluster of properties displayed by Chinese nouns, heading an NP that functions directly as the subject or object and needs a preposition otherwise is a property that is found with nouns in every language. Being head-final inside NP is also a typological phenomenon found in half of the world’s languages (Greenberg 1963, Hawkins 1983). In comparison, the use of *de* is language-specific and therefore is a more costly operation that is subject to removal if any change is to happen to the cluster. As a result, L keeps all the syntactic properties of N except *de*.

To complete the analysis, we also make explicit an assumption underlying the previous discussion:

(21) The choice of the syntactic properties of a categorial deviate X must guarantee that X is distinct from all existing categories.

The validity of (21) is self-evident. If X retained all the properties of the original category, there would be no X; if behavioral change resulted in X acting completely like another
(existing) category, then it would be a categorial conversion, not a deviation. In sum, (21) enforces a partial change in behavior while (20) dictates the exact content of the change. In the rest of this subsection, we will continue to use the term localizer for this group of words, and noun for the standard nouns, so as to facilitate discussion. It should be noted that the theory offered here does not force a language to have a deviated L, even when N has language-specific properties. It only applies if the language opts for deviation.

### 1.1.2.3. Disyllabic L

The examples below illustrate disyllabic localizers.

(22) a. chuang (de) xiamian
   bed DE underneath
   
   b. men (de) houtou
   door DE behind
   
   c. wuzi (de) libian
   house DE inside

The analysis in the previous subsection is based on the fact that monosyllabic L differs from corresponding N in not employing de. In contrast, the optional use of de with disyllabic localizers makes them resemble disyllabic nouns:

(23) a. men (de) bashou
   gate DE handle
b. wuzi (de) houmen
   house DE backdoor

If the presence of *de* signals “nounhood” as we have argued, then the localizers preceded by *de* in (22) are all nouns just like *bashou* ‘handle’ and *houmen* ‘backdoor’ in (23).

This may appear to suggest that the localizers without *de* are nouns as well. The classifier test, however, argues differently:

(24) a. na-shan damen houtou, yi-jian wuzi libian, …
   that-CL big.gate behind one-CL house inside
b. *na-shan damen bashou, *yi-jian wuzi houmen, …
   that-CL big.gate handle one-CL house backdoor

Just as we saw with monosyllabic localizers and nouns, a classifier matching the noun immediately after it is perfect with the *de*-less localizer but results in totally unacceptable examples with typical N-N compounds. Therefore, it must be concluded that even disyllabic localizers in the absence of *de* take a preceding NP as complement (cf. the structure in (18)) but do not form an N-N compound with the preceding N. As no *de* is present between this NP and the localizer, the latter must be L, a deviate of N.

This conclusion means that, while monosyllabic localizers are exclusively L, their disyllabic cousins are ambiguous between L and N. This is actually easy to understand. After all, the morphemes *mian*, *tou* and *bian* are nouns meaning, respectively, ‘face’, ‘end’ and ‘side’ when used alone. As a result, *libian* ‘inside’ can be reasonably treated as
consisting of L-N. Since the nominal compounds are head-final (see discussion following (14)), it is only natural that the L-N compound inherits its category from its N head. In other words, they are simple locative nouns. It is also a fact that mian, tou and bian have lost both their tones and their concrete semantic content when they occur as the second component of the disyllabic localizer. We suggest that they are on the verge of losing their categorial identity as well. If they are still treated as nouns, the disyllabic localizer is a noun as we saw. If their categorial content is considered lost together with the tonal and semantic content, the only morpheme in the compound that can still determine the category is the monosyllabic localizer. Hence the whole compound is treated as L. See Di Sciullo and Williams (1987) for a hypothesis on computing the lexical properties of a compound from its head.

1.1.2.4. On L as a clitic

Our analysis of L is crucially built on its interaction with de by attributing the absence of the morpheme to L deviating from N and to de being more costly to use in syntactic computation. There are alternative ways to account for the distribution of de in such contexts. Liu (1998), for instance, treats (monosyllabic) L as a clitic carrying the [+loc] feature. As such, the NP combined with L essentially behaves as a location-denoting NP like xuexiao ‘school’, with L being a “phrasal affix” that forms a phonological unit with the host NP (Klavans 1980, Zwicky 1985, Anderson 1992). This account is both simple and intuitive given the well-known fact that prosody plays an active role in the syntax.
and morphology of modern Chinese. We see no reason against calling L a clitic since it indeed tends not to stand alone, but we do not believe that doing so can either eliminate the question regarding the categorial nature of L or provide an adequate explanation for L’s behaviors.

First, it is widely accepted in the fields of generative syntax and morphology that an affix belongs in some category. Much recent research is built on assigning syntactically distinct categories to even tense and aspect affixes. Since Liu adopts Anderson’s (1992) theory that a clitic differs from an affix only by attaching to a phrase rather than to a morpheme, it is expected that clitics fall into different categories as well.

Secondly, treating L as a clitic to NP may appear to explain why de is absent – de is another clitic and incapable of serving as host for L – but actually raises various new questions. For instance, though L typically is associated with an NP, there are cases when the NP can be missing (the glossing is only suggestive):

(25) a. X L

  xiang ‘facing’  li ‘inside’
  chao ‘facing’  wai ‘outside’
  kao ‘by the side of’  shang ‘top’
  xia ‘underneath’
  qian ‘front’
  hou ‘back’

See Feng (2000) for a theory of such syntax-prosody interactions in Chinese.
b. ta kao hou zuo-zhe.

he by back sit-ZHE

‘he sits by the backside. \(\rightarrow\) he sits on the back.’

Whether the words in the X column are verbs or prepositions may be debatable, but they are definitely not nouns. And each one of them can be productively followed directly by any monosyllabic localizers (as well as disyllabic ones and the NP-L clusters) as exemplified by (25a-b). If L is simply a locative clitic, to what phrase does L cliticize? Note that there is no sense in which the phrase headed by, say, kao ‘by the side of’ needs L to acquire the [+loc] feature, with kao itself sufficiently indicating a location:

(26) ta kao qiang zuo-zhe.

he by wall sit-ZHE

‘he sits by the wall.’

On the other hand, within the theory that treats monosyllabic localizers as a deviate of N which we call L, the examples in (25) have the following structure:

(27) by/facing \([_{LP} \ldots L]\)

where “…” is a phonetically contentless pronoun which, as we will see in later chapters, Chinese employs abundantly. Technical details need to be worked out, but (27) can easily
explain the intuition that when someone sits by the ‘backside,’ as in (25b), he sits by the backside of something obvious from the discourse. As we know, such reference is typical of pronouns. In contrast, claiming a clitic to take a phonetically empty pronoun as host would seem to run against the very notion of cliticization.

1.1.3. Adjectives

There are uncontestable reasons for distinguishing adjectives from verbs in Chinese, in spite of the fact that a Chinese adjective functions as the predicate of a clause without the help of a copular verb. In this section, we first examine the evidence for the adjectival category (shorthanded as A), then we provide a feature characterization for it.

1.1.3.1. In comparison with verbs

Some adjectives can be used transitively, in the sense that they describe situations that involve two participants rather than one. If a clause contains such an adjective, one of the participants expectedly is represented as the subject. The other, however, needs to be introduced with dui (adjectives in these examples are in bold face):

(28) a. ta dui zhe-ge jieju hen buman.

he on this-CL outcome very discontent

‘He is discontent with this outcome.’
b. ??ta hen buman zhe-ge jieju.\(^7\)
   he very discontent this-CL outcome

(29) a. wo dui tade qushi feichang shangxin.
    I on his pass.away extremely sad
    ‘I am extremely sad with his death.’

b. *wo feichang shangxin tade qushi.
    I extremely sad his pass.away

(30) a. zhe-ge gongzuo dui ni hen heshi.
    this-CL job on you very suitable
    ‘This job is suitable for you.’

b. *zhe-ge gongzuo hen heshi ni.
    this-CL job very suitable you.

\(^7\) Some such examples may be perceived to be acceptable by certain speakers, possibly due to dialectal differences. For these speakers, what is judged as an adjective in our book may have changed into a verb. Such a categorical shift is likely to happen more easily in Chinese than in, say, English for two reasons. First, Chinese has no morphological markers for categories found in European languages. Secondly, Chinese adjectives function as predicates without a copula, making them appear verb-like. On the second property, see Section 1.1.3.2 below.
To the best of our knowledge, the NP after *dui* is the semantic object of the adjective.

Such a relation between the two constituents not only is consistent with the semantics of the examples but also is supported by the comparison between (30) and (31):

(31) zhe-ge gongzuo hen *shihe* ni.
    this-CL job very suit you

   ‘This job fits you well.’

Differing only in the linear sequence of the two bound morphemes *shi* and *he*, *heshi* and *shihe* are also semantically similar, especially when each one is accompanied by two NPs as in (30) and (31). Since there is no question about treating the NP *ni* ‘you’ as the object of *shihe*, it is only reasonable to treat the same NP in (31) as the object of *heshi*.

    The different behaviors of objects in (30) and (31) are explained if the latter contains a verb while the former contains an adjective. It is established knowledge that Chinese verbs take their object directly and on the right in the unmarked context, so the linear order in (31) is expected of V. From this fact, it can be inferred that *heshi* in (30) is not a verb. In fact, as far as the use of *dui* is concerned, *heshi* patterns with nouns (cf. Section 1.1.1). Interestingly, the data here also find counterparts in English:

(32) a. She loves butterflies.
    b. her love *(of) butterflies
    c. She is fond *(of) butterflies.
In English, the object occurs consistently to the right of all categories, but the NP object must be introduced by the semantically empty preposition of when the word taking the object is a noun or adjective. Putting word order aside for now, this is exactly what happens in Chinese. For nouns and adjectives, the NP object must be introduced by dui, which not only occurs in the preverbal position typical of prepositions in the language, but also is semantically empty, as is evidenced by the fact that (30a) with dui does not mean any more than (31) without dui. In conclusion, heshi, shangxin and human in (28)-(30) are different from verbs and demonstrate typical properties of adjectives.

The use of dui as a test for distinguishing adjectives from verbs matches well with a long-known fact about reduplication patterns among disyllabic predicative words in Chinese (cf. Zhu 1982, Lü 1984, among many others):

(33) a. AB → ABAB

jiancha → jianchajiancha,  jihua → jihuajihua
examine  do a check-up of  plan  do some planning

b. AB → AABB

ganjing → ganganjingjing,  jiandan → jianjiandandan
clean   rather clean  simple   rather simple

Reduplication of this kind usually requires specific discourse contexts, which we try to take into consideration in the examples given. Various conditions apply to what disyllabic words can undergo reduplication (basically those of the “conjunctive” type), but we believe that they do not affect the validity of the analysis at hand. Also see Section 1.1.3.3 for more discussion on AABB.
The words in (33a) are established verbs and can only have the ABAB pattern. Those in (33b), on the other hand, have the AABB pattern and are regarded by many grammarians as adjectives. What is interesting is that these patterns correlate with the (non-)use of dui. The most informative evidence comes from disyllabic transitive words such as mingbai which allow both patterns of reduplication:

(34) a. ta mingbai zhe-ge daoli.
     he understand this-CL reason
     ‘He understands this reason.’

   b. ta dui zhe-ge daoli hen mingbai
     he P this-CL reason quite be.clear
     ‘He is quite clear about this reason.’

Applying different reduplication patterns to the word yields the sharp contrast below:

(35) a. ta yinggai mingbaimingbai zhe-ge daoli!
     he should get.to.understand this-CL reason
     ‘He should get to know this reason!’

   b. *ta (yinggai) mingmingbaibai zhe-ge daoli.
     he should be.rather.clear.about this-CL reason

(36) a. ta dui zhe-ge daoli mingmingbaibai.
     he P this-CL reason be.rather.clear
‘He is quite clear about this reason.”

b. *ta (yinggai) dui zhe-ge daoli mingbaimingbai
he should P this-CL reason get.to.understand

In other words, the AABB pattern is correlated with using *dui to introduce a preceding object, whereas the ABAB pattern rejects the use of *dui and retains the object-related properties of the original verb. Later, we will provide a theoretical framework for understanding why the use of *dui is associated with adjectives but not with verbs. For now, what is significant is the fact that the reduplication patterns corroborate with the verb-adjective distinction established on the basis of *dui.

It is often suggested in the literature that the adverb *hen ‘very’ can be used to define adjectives. While typical adjectives indeed associate well with *hen, the following data indicate that this test does not apply exclusively to adjectives:

(37) a. ni hen ai ta.
you very love him
‘You love him very much.’
b. ta hen ganxie dajia de bangzhu.
he very be.grateful.for people DE help
‘He is very grateful of people’s help.’
c. wo hen tongqing ni de zaoyu.
I very sympathize you DE bad.experience
‘I am sympathetic with your bad experience.’
In all these examples, the object occurs without the use of *dui* or any other grammatical help. This is typical of the object of verbs. If the goal were to formulate an isolated grammatical theory of Chinese, it might be plausible to treat *ai* ‘love, *ganxie* ‘be grateful for’ and *tongqing* ‘sympathize’ as adjectives. This is so because the test of *hen* and the test of *dui*-less object do not yield identical results, and there would be no particular reason to favor one test over the other. But we aim to capture the patterns in Chinese without losing the important fact that Chinese is also a human language. With the behaviors of adjectives in other languages taken into consideration, it is clear that the *dui* test should be favored and the words at issue in (37) are verbs, not adjectives. More generally, linguistic debates may result from lack of decisive evidence on both sides when the investigation is confined to a single language, but may be resolved, and sometimes may not even arise, if the universal aspect of language is duly considered.

### 1.1.3.2. As opposed to nouns

Given the fact that A patterns with N according to the *dui* test, it seems appropriate to assign [+N] to A. But it is equally clear that A and N have different syntactic behaviors. Most relevant to the current concern is the use of the copular *shi*. As the predicate of a clause, NP typically requires *shi*, with the option of omitting it in highly colloquial speech. The AP predicate, however, rejects *shi* completely. The contrast is shown below:

(38) a. ta shi yingxiong.
    \[
    \begin{array}{lll}
    \text{he} & \text{be} & \text{hero}
    \end{array}
    \]
‘He is a hero.’

b. ?ta yingxiong.
   he hero
   ‘He is a hero.’

(39) a. ta hen yingyong.
   he very heroic
   ‘He is heroic.’

b. *ta shi hen yingyong.\(^9\)
   he be very heroic

Intended reading: Same as (39a).

In fact, it is precisely this ability to function as the predicate without a copula that makes many researchers treat A as a subcategory of V in Chinese. In terms of categorial features, then, A is a candidate for [+V]. In sum, we see evidence that A should fill the gap of [+N, +V] in (8).\(^{10}\)

\(^9\) One needs to distinguish the copular shi from the emphatic shi, which is permitted in this example. The most salient differences between the two morphemes are that the emphatic shi must be stressed in this context while the copular shi is typically not, and that the emphatic shi, as its name implies, is used only to emphasize some constituent after it, either reflecting the assertive attitude of the speaker or bringing out a contrastive interpretation. The copular shi in (38a) does not have this semantic property at all. The emphatic use of shi and its syntactic representation will be examined in detail later in the book.

\(^{10}\) This conclusion fits seamlessly into Chomsky’s (1970) theory of categorial features, but it also raises questions about the nature of English adjectives, which crucially cannot be used predicatively in a
1.1.3.3. More on AABB

While ABAB and AABB are shown to correlate well with the different ways in which a verb and an adjective introduce their NP objects, it would be an oversimplification to match them directly with the verb-adjective distinction because certain disyllabic verbs may also be reduplicated as AABB:

(40) a. chaorang → chaochaorangrang
argue.yell
b. fengbu → fengfengbubu
sew.mend
c. laiwang → lailaiwangwang
come.go ‘mutually visit’

For these verbs, taking the AABB form does not turn them into adjectives even though, where it is possible to tell, they do lose the ability to take a postverbal object:

(41) a. ta ba na-jian yishang fengfengbubu, chuan-le hendo nian.

standalone sentence without a copula. Various possibilities arise but we will leave them aside. In essence, our analysis claims that Chinese A is the typical [+N, +V]. English A must be something else.

11 We thank the audience of Yafei Li’s Syntax of Chinese class for bringing this fact into our attention. See Lü 1984 for more examples.
He kept sewing and patching that coat and wore it for many years.

b. *ta fengfengbubu na-jian yishang, chuan-le henduo nian.
   he sew.here.mend.there that-CL coat wear-LE many year

Intended reading: Same as (41a).

The use of ba with the preverbal NP na-jian yifu ‘that coat’ in (41a) indicates that fengfengbubu is still a verb (cf. Chapter 4), but the postverbal position is no longer legitimate for the NP object, shown in (41b).

Under the assumption that the AABB pattern in both (41) for verbs and in the adjectival data in Section 1.1.3.1 is produced by the same morphological process, we conclude for now that the process itself is not category-changing, although it is necessarily correlated with the lack of a postverbal object. Many questions remain, one of which is why this happens only with reduplicating disyllabic words.

1.1.4. Prepositions

The class of prepositions is one of the most poorly defined categories in Chinese, due to the facts that the so-called prepositions in the language all have their historical origins as verbs and that Chinese has no inflectional morphology to mark verbs. In this section, we examine four classes of words that are labeled as prepositions in the literature. It is not our intention to exhaust either all preposition-like words or even all the usages of any
specific word in this group. Rather, we hope to provide an in-depth analysis of a subset of such words in an attempt to uncover some inner workings of the human linguistic faculty.

1.1.4.1. Behavioral clarifications

Some candidates for $P$ are given below:\footnote{The English glosses are only meant as approximate translations. In fact, not every word in this list is directly translatable.}

(42) a. zhiyu ‘as for’, guanyu ‘about’, …
    b. cong ‘from’
    c. gei ‘to/for’, zai ‘at’, xiang ‘toward’, …
    d. ba, bei

The words in (42) all can introduce an NP on their right without the help of any other morpheme. This property is shared by $P$ and $V$ -- both $P$ and $V$ can take a nominal object directly -- but not by $N$ or $A$, which are incapable of directly taking any nominal object. The reader is invited to test this claim with Chinese as well as with other languages. Given the definition of $[N]$ above, both $P$ and $V$ should be $[-N]$. While the same logic dictates that all the words in (42) are $[-N]$, they also differ in interesting ways, which we examine one by one. To facilitate discussion, we will temporarily call all these words prepositions until their actual categories are identified.
Though all the prepositions in (40) are typically followed by an NP, only those in (a) must occur with the NP in a pre-subject position:

(43) a. guanyu zhe-jian shi, tamen yijing taolun-guo le.
    about this-CL issue they already discuss-GUO SFP
    ‘Regarding this issue, they already discussed (it).’

b. ta cong nali dai-huilai henduo jinianpin.
    he from there bring-back many souvenir
    ‘He brought back many souvenirs from there.’

c. ta gei ban li de ren zuo-guo bushao shiqing.
    he for class inside DE people do-GUO not.few thing
    ‘He did quite a few things for the people of the class.’

d. ta ba guizhong de shoushi cang-zai waguan li.
    he BA expensive DE jewelry hide-in clay.pot inside
    ‘He hid expensive jewelry in a clay pot.’

In (b) through (d), the default position for the prepositions is between the subject and the main verb of the clause. Such a word order is not an alternative for the sentence in (43a). This fact indicates that the prepositions in (42a) occur in a syntactic position outside of or peripheral to the clausal structure. This contrasts with those in (42b-d), which clearly are inside the clause proper.

Furthermore, the pre-subject position is an option available to the prepositions in (42b, c) as long as they bring their nominal object along with them. This word order
alternation, viewed as *movement* away from the default position, is not available to words in (42d).

(44) a. cong nali, ta dai-huilai henduo jinianpin. (cf. (43b))
    from there he bring-back many souvenir
    ‘From there, he brought back many souvenirs.’

b. gei ban li de ren, ta zuo-guo bushao shiqing. (cf. (43c))
    for class inside DE people he do-GUO not.few thing
    ‘For the people of the class, he did quite a few things.’

c. *ba guizhong de shoushi, ta cang-zai waguan li. (cf. (43d))
    BA expensive DE jewelry he hide-in clay.pot inside
    ‘Expensive jewelry, she hid in a clay pot.’

The grammaticality of (44a, b) is expected if a preposition forms a phrase, PP, with the NP after it, the logic being that only words forming a constituent can move together.

If (44a, b) involve clause-initial PPs, the structure of (43a) can be analyzed analogously, with the preposition *guanyu* ‘about’ and the NP after it forming a PP. The difference is that the PP in (43a) does not occupy the clause-initial position through movement: where it is heard is its default position. This is why this PP cannot occur between the subject and the verb.\(^{15}\) Put differently, the PP in (43a) always modifies a

\(^{15}\) Logically, one could imagine that the PP in (44a) starts in the clause-initial position but moves to the post-subject position as an option. If this option were available, there would be no explanation for the pattern in (43). Later on, we will see that movement doesn’t happen in all directions, and that rightward
whole clause, while those in (43b, c) and (44a, b) typically modify only part of the clause, namely the phrasal predicate containing the verb. See Ernst (2002) for a comprehensive discussion on the locations of various adverbial modifiers. It should be noted that the PP formed with guanyu is also movable, as the theory expects. This can be seen in a larger syntactic environment:

(45) guanyu zhe-jian shi, wo tingshuo tamen yijing taolun-guo le.
    about this-CL issue I hear they already discuss-GUO SFP

‘Regarding this issue, I heard that they already discussed (it).’

With the interpretation that I heard that they had a discussion on this issue, the PP must have originated inside the object clause. Then its actual position in the main clause results from phrasal movement.

Given this logic, the ungrammatical status of (44c) follows if ba and the NP after it do not form a phrase – not being a constituent, ba and the NP cannot move together to the clause-initial position. The same observation is true of bei. We will elaborate on this conclusion in Chapters 4 and 5.

A frequently mentioned factor that divides the words in (42) is whether they have the option of being used alone as verbs in modern Chinese. Of the four classes, only (42c) has this option. In the examples below, the words at issue are in bold face:

Historically, *cong* in (46b) meant ‘to follow’ and *ba* in (46d) ‘to hold’, both used as verbs. But these verbal usages are no longer available in modern Chinese except inside some fossilized expressions. In contrast, it is obvious that *gei* in (46c) is used as a verb, given the aspectual markers it carries as well as the unmistakable actions it describes.

A subset of the words in class (c) of (42) may also serve productively as the second member of a complex verbal predicate, be it in a phrasal or compound form. The words of our concern are again marked in bold face:

(47) a. *ta ba* na-zhang zhaopian ji-*gei*-le wo.

he BA that-CL picture send-give-LE me
‘He mailed me that picture.’

b. ta ji-le yi-zhang zhaopian gei wo.
   he send-LE one-CL picture to me.
   ‘He mailed a picture to me.’

(48) a. ta ba xin fang-zai shu li.
   he BA letter put-at book inside
   ‘He put the letter in the book.’

b. ta fang-le yi-feng xin zai shu li.
   he put-LE one-CL letter at book inside
   ‘He put a letter in the book.’

(49) haizi-men gaoxing de chong-xiang shanding.
   child-PL happy DE dash-toward hill.top
   ‘The children happily dashed toward the top of the hill.’

To our knowledge, no preposition of other classes has this property in modern Chinese.

The significance of this correlation is explored in the next subsection.

1.1.4.2. The categorically dual status

Summarizing so far, the classes (42a, b) have the typical properties of prepositions. They form a phrase with an NP object and they cannot function as verbs, as is expected of typical prepositions and postpositions in every language. Recall that the very ability to
take an NP object without the help of any other morpheme indicates that these words are not noun-like, which is represented as [-N] in the theory based on categorial features (cf. Section 1.1.1). Now that classes (40a, b) cannot be used as verbs in the typical predicative manner, it is only natural to conclude that they are not verb-like either. Namely, they are [-V]. Therefore, a preposition is [-N, -V] in terms of categorial features. This conclusion is based on the syntactic properties of prepositions, but it also fills a gap in the feature matrix in (8). To put it differently, the theory predicts the existence of a category [-N, -V], and the prediction is borne out empirically. The only structural difference between the two classes of prepositions in Chinese is that (42a) is located somewhere outside the core structure of a clause while (42b) is inside. We return to the precise locations of these phrases later.

If prepositions are [-N, -V], then the members of class (42c) cannot be treated simply as prepositions because they can also be used as verbs, which are [+V] by definition. We believe that this class has multiple statuses. As V, the words in (42c) are [-N, +V]; and as P, they are [-N, -V]. The evidence for their verbhood is already given in (46). Now consider the argument for their prepositional use, starting with the examples below. To avoid unnecessary confusions during the discussion, gei is glossed as a theoretically non-committal GEI in the examples:

(50) a. ta gei wo zuo-le henduo shi.
    he GEI me do-LE many thing
    ‘He has done many things for me.’

    b. gei wo, ta zuo-le henduo shi.
GEI me he do-LE many thing

‘For me, he has done many things.’

c. *ta zuo-gei wo henduo shi.

he do- GEI me many thing

Intended reading: Same as (a) above.

d. *ta zuo-le henduo shi gei wo.

he do-LE many thing GEI me

Intended reading: Same as (a) above.

(50b) indicates that gei and the NP following it form a phrase. The choice of the main verb zuo ‘do’ enforces a benefactive interpretation on this phrase. Since (50c, d) are unacceptable, it must be concluded that this interpretation is not compatible with using gei as part of a complex predicate, even though gei can be so used otherwise (cf. (47)).

This apparently messy behavior of gei actually has a simple explanation. As we saw earlier, members of class (42c) can both function as standalone verbs and occur in a complex predicate. Furthermore, it is a fact that when gei is used as a verb, it does not have the benefactive interpretation. This is straightforward in (46c) and is corroborated by the recipient reading of gei in (47). Now suppose that the second part of a complex predicate must be a verb. It follows immediately that gei in (50c, d) cannot be understood benefactively. In contrast, the preverbal gei in (50a, b) is a preposition just like class (42b). Its acquisition of the benefactive meaning can be attributed to some semantic shifting and/or “bleaching”. In other words, the unacceptable (50c, d) with the intended

14 In Bantu and Iroquoian languages, a suffix to the verb root, called an applicative, often has the function
readings simply result from the independent requirement in modern Chinese that within a clause, a preposition does not ever occur after a verb. In the absence of counterexamples, we extend the same conclusion to other members of class (42c) such as zai and xiang.

It must be pointed out that this argument for treating gei as P in (50) does not necessarily exclude the prepositional gei from having the recipient interpretation. In fact, the sentence below is potentially ambiguous:

(51) ta gei wo ji-le yi-zhang zhaopian.
    she GEI me send-LE one-CL picture

a. ‘She sent a picture to me.’ or
b. ‘She sent a picture for me.’

We see no reason to treat the recipient gei, reflected in (51a), as a verb when the benefactive gei in (51b) is a preposition. To be sure, if we ignore semantic subtleties, (51a) is synonymous to (47), where gei is argued to be a verb. But this is not a problem considering how the semantics of the complex predicate is computed, a topic which we return to in Chapter 2.

With this discussion of prepositions, we are also ready to be more precise about the nature of [N]. Recall that N and A are [+N] because they need a preposition, of in English and dui in Chinese, to help introduce the object NP, and that such prepositions are of introducing either the goal or benefactive NP, depending on the context. Though a current analysis of applicatives treats these affixes as V (Baker 1996, Y. Li 2005), it is nonetheless worth noting that it may not be an accident that a semantically “bleached” gei also acquires similar semantic functions.
semantically empty. These facts only suggest one thing: that there is a pure grammatical demand for *of* and *dui* in such contexts. In the framework used in this book, the grammatical demand is to provide a Case for a nominal phrase under the following hypothesis, referred to as the Case Filter:

\[(52) \text{Every NP must have a Case.}\]

In syntax, a Case is an NP’s “certificate” to function as the object or the subject of a sentence. Languages like Russian and Korean choose to manifest Cases through morphological cases such as the nominative and accusative, but every language presumably employs the certificate system regardless of whether or not it morphologically marks arguments. In this theory, the fact that A and N are incapable of directly taking an NP object is attributed to their inability to provide a Case to the latter, and this inability is formally defined as [+N]. It follows that V and P are [-N] for being able to provide Cases to their objects. In the presence of an NP object for A or N, then, languages enlist a semantically empty P to provide the needed Case without altering the meaning of the expression. Hence the properties of *of* and *dui* are explained.

1.2. Functional categories

The advantage of the feature-based theory of categories in (8) is its ability to capture shared syntactic properties of certain categories, e.g., that N and A both need a P to introduce their NP object. Such behavioral similarities would be lost if each category
were treated as a non-decomposable entity in language. The disadvantage is that two
categorial features maximally produce only four categories – V, N, A, P according to the
previous section, a clearly insufficient outcome. As an example, the discussion of L in
1.1.2 critically relied on classifiers (CL) and de, neither of which can be obviously
accommodated by (8). This section examines how (8) may be revised to allow more
categories while remaining restrictive enough to be empirically insightful.

1.2.1. \([Fn], n \geq 0\)

In the tradition of Chinese linguistics, there is a widely accepted distinction between
\(shi-ci\) ‘substantive word’ and \(xu-ci\) ‘empty word’. N, V, A are classified as the former and
P usually as the latter. The \(shi-xu\) distinction corresponds to the dichotomy of lexical
categories and functional categories in the theoretical framework adopted in this book,
except that P is treated as a lexical category here, not as a functional one. Regardless of
where P belongs in the dichotomy, however, it is clear that languages in general, Chinese
being no exception, make use of functional words in syntax.\(^{15}\) In fact, one of the major
theoretical claims in the past three decades is that there are more functional categories
than lexical categories.

First consider classifiers (CL) introduced in 1.1.2. The nominal origin of CL is
widely recognized. Below is an example to highlight the relationship between CL and N:

\[\]

\(^{15}\) Jackendoff (2002) argues that in the course of evolution, the advent of functional words is a major
marker for the critical transition from some proto-communicational system to modern language.
(53) gan

a. As N: qiang-gan
   gun-barrel
   ‘gun barrel’

b. As Cl: yi gan qiang
   a CL gun
   ‘a gun’

Equally obvious is the “bleached” semantic content of classifiers. In (53), for instance, the classifier use of gan no longer refers to any specific part of the gun, but rather to the class of objects with the general shape and texture of a thin shaft. Outside the theoretical framework of this book, the transition from N to CL is sometimes referred to as grammaticalization. Independently of terminology, however, the fact remains that a classifier does not serve as a lexical noun but rather as a “functional” one whose role in syntax is semantically less concrete.

In addition to CL, two other classes of words may also occur inside a nominal phrase:

(54) na yi gan qiang
    that one CL gun
    Lit: ‘that one gun’
The numeral (Num) 一 ‘one’ typically occurs with CL. There is also evidence that the demonstrative pronoun na ‘that’ belongs in the category of determiners (D) that is separate from N, despite the fact that the two are often placed in the same category in grammar books. Num and D will be examined in detail in Chapter 8. For now, we note that for the lexical category N, there are at least three more or less “noun-like” categories, CL, Num and D, associated with it.

It is based on considerations of the similar kind that Grimshaw (1991, 2000) proposes to add to the existing feature set [N] and [V] another feature [F\textsubscript{n}], where F stands for functional and n’s value is equal to or larger than 0. [F0] indicates a lexical (i.e., non-functional) category; [F1] is the functional category structurally closest to [F0] in the given phrase; [F2] is farther from [F0] than [F1], etc. In this enriched theory of categorial features, the four nominal categories just discussed are described in (55):

\begin{align*}
N & = [F0, +N, -V] \\
CL & = [F1, +N, -V] \\
Num & = [F2, +N, -V] \\
D & = [F3, +N, -V]
\end{align*}

The precise structural relations among these categories will be discussed in detail in Chapter 8.

A similar situation exists for V. Aspectual morphemes such as le, zhe and guo were historically verbs, with guo still capable of functioning as a standalone verb meaning ‘to pass’ in modern Chinese. On the other hand, they are clearly not lexical verbs, both
because they only express various kinds of aspectuality (i.e., the developmental status of an event) without changing the basic semantics of the verbs to which they affix, and because in this usage they cannot be used alone as the predicate of a clause. As a result, they seem best described as being verb-like but functional. Following the notational convention, this category is abbreviated as ASP and is defined, for now, as \([F1, -N, +V]\), namely the closest functional word to \(V\) (= \([F0, -N, +V]\)).

The English examples below illustrate another functional category which Grimshaw places under the verbal system:

(56) a. Pat thinks **that** the moon is made of Wisconsin cheese.
    b. Pat asks **if** the moon is made of Wisconsin cheese.

The boldface words are complementizers (C) which have two functions, introducing an embedded clause in a bigger context and marking the type of the clause. In (54), both **that** and **if** introduce an embedded (object) clause, but the former indicates the clause to be declarative while the latter marks out the interrogative. This view, however, proves to be oversimplified when East Asian languages are taken into consideration. The Korean examples below are quoted from Y. Li (2005):

    John-TOP Mary-ACC classical music-ACC like-PRES-DEC-C believe-PRES-DEC
    ‘John believes that Mary loves classical music.’
In each example, the verb of the embedded clause is suffixed with two morphemes, \textit{ko} for introducing the embedded clause, and the other for “typing” it, with \textit{ta} for the declarative (glossed as ‘DEC’) and \textit{nya} for the interrogative (glossed as ‘Q’). In addition, only the clause-typing morpheme \textit{ta} occurs with the matrix verb. This makes perfect sense. After all, the matrix clause is itself the largest syntactic construction. As it is not embedded, there is no need for \textit{ko}. In conclusion, to the extent that we accept Grimshaw’s view that there are functional categories associated with V, analogous to the functional categories associated with N in (55), it is necessary to distinguish two more categories of morphemes, those that introduce embedded clauses, for which the name complementizer (C) is kept, and those that signal the types of clauses, which we call clause-typers (CT).

In (57), C is immediately to the right of CT and thus is farther from V. So if CT = [Fi] for any value of \( i \) greater than 1 (ASP = F1), then C = [Fi+1]. As for English, that and if must be the result of merging both CT and C into a single morpheme, a phenomenon not surprising for European languages where two conceptually separate pieces of information, such as agreement and tense, are characteristically represented as a single morpheme.

Turning back to Chinese, consider the following examples:

(58) a. ni-men zou ba.

\text{you-PL leave SFP}

‘You can leave (now).’
b. ta qu-guo ma?
    he go-GUO Q
    ‘Was he there (before)?’

b. shei xie zhe yi zhang ne?
who write this one chapter Q
    ‘Who will write this chapter?’

Each of the three morphemes, *ba*, *ma* and *ne*, signals a particular clause type. *Ba* is for imperatives, *ma* for yes-no interrogatives, and *ne* for interrogatives containing question phrases such as *shei* ‘who’ and *shenmo shihou* ‘what time’. Furthermore, these morphemes occur at a position peripheral to the clause, just like CT in Korean matrix clauses. We thus propose to treat clause-final morphemes such as *ma*, *ba* and *ne* as CT. What remains unclear is why CT in Chinese never occurs with embedded clauses. Possibly, there are unidentified discourse functions that *ma*, *ba*, and *ne* perform that are associated only with matrix clauses. See Cinque (1999) for a list of functional categories in the typical clausal structure. For the purpose of this book, these morphemes can be effectively treated as C, i.e., *[F n, +V, -N]*, where *n* is a number sufficiently large to distinguish itself from the values of those functional categories more closely associated with the lexical verb.

To summarize so far, N and V each have a set of related functional categories, which are distinguished through incremental values of *[F]*. This analysis also suggests a route for diachronic changes. In the literature, both the shift from V to ASP and from N to CL has sometimes been called grammaticalization or *xu-hua*, meaning that a lexical
morpheme adopts a more abstract meaning and starts to perform grammatical functions.
One way to grammaticalize, then, is to shift from [F0] to [Fi], i > 0, while all other
categorial values remain intact.

1.2.2. [F] and the modifier-introducing de

That de belongs to a functional category for its lack of tangible semantic content is the
conventional wisdom and will be adopted here as well. The question is where it stands in
the [±N, ±V] system. Descriptively, de occurs in the syntactic context of [ X de Y ]. If Y
= N, X can be an NP, an AP, a PP, or a full clause, as shown in (59):

(59) a. zhe-wei xuezhe de guandian
    this-CL scholar DE opinion
    ‘this scholar’s opinion’

b. shifen youren de tiaojian
    very enticing DE term
    ‘very enticing term’

c. guanyu zhanzheng de chuanyan16
    about war DE gossip

16 Note that (59c), as well as the good examples in (7), determines that de cannot be a morphological
marker of Case, for the simple reason that PPs (guanyu zhanzheng ‘about war’) don’t need any Case (cf.
(52)). The same logic applies to no in Japanese, which is sometimes treated as the Genitive Case marker but
is actually suffixed to either an NP or a PP.
‘gossip about war’

d. wo qu guowai de liyou
   I go abroad DE reason

   ‘the reason for my going abroad’

In contrast, X is largely restricted to AP when Y = V, with de being optional even then.

Pending a better understanding of this morpheme, two categorial characterizations seem plausible to explore. First, de may be \([Fi, +N, +V]\), where \(i > 0\). Under this interpretation, de is an adjectival functional word which turns a phrase inside a larger NP into a modifier. Functional morphemes that may alter the category of phrases are found in other languages as well. Consider the English example below:

(60) Beth is proud of Christine’s winning the prize twice.

On the one hand, winning functions as a verb because it can assign a Case to the object NP the prize. On the other hand, it also displays two nominal properties: providing the Genitive Case for the semantic subject Christine and needing of to be the legitimate object of the adjective proud. As we saw in 1.1.3.1, of is required in this context only when the object is nominal and thus needs to satisfy the Case Filter in (50). This mixture of verbal and nominal properties follows if –ing in (60) is \([Fj, +N, -V]\) \((j > 0)\), which nominalizes the whole VP rather than just the verb win. Since win remains V inside VP, its ability to assign Case to the object remains intact; the whole VP, however, is nominalized by –ing and therefore needs a Case for itself, just like any other nominal
phrase. For the same reason, the semantic subject in the nominal phrase receives the Genitive Case.\footnote{Also see Huang (1994b) for a related treatment of the gerundive construction.}

One may wonder why \textit{de} is needed even for an AP in (59b). A possible answer relates this to another property of A. Unlike their English counterparts, Chinese adjectives play the role of a predicate directly, without any copula (cf. 1.1.3.2). In this use, AP behaves just like VP. It follows that an AP modifier may in fact be a relative clause, which in turn is “adjectivized” by \textit{de}. The same analysis may even apply to PP modifiers, given the well-known fact that Chinese prepositions originated as verbs. Caution is called for, however, as there is no evidence that \textit{guanyu} ‘about’ has any verbal property in modern Chinese. Another area of concern is why \textit{de} is not as widely used with modifiers for V.

\textit{De} may also be analyzed as [Fx, N, V], i.e., a word without any value for the categorical features [N] or [V] nor for the [F] feature. According to common assumption (cf. Di Sciullo and Williams 1987), the constituent composed of such an unspecified \textit{de} and a pre-\textit{de} phrase of category X (i.e. [ XP \textit{de} ]) inherits X as its category; that is, X provides the values for [F], [V] and [N]. In effect, then, a noun allows modification by NP, AP, PP, and clauses, as is the case in many other languages, while \textit{de} serves merely as a morphological linker. For practical purposes, this analysis treats \textit{de} on a par with a subordinate conjunctive in the conventional sense, linking a modifier phrase to the modified. See Aoun and Li (2003) for a proposal in this direction. Again, complications arise which are yet to be resolved, one of which is why \textit{de} is even needed in the first
place. We suspect that the ultimate account of *de* in modern Chinese depends partially on understanding its predecessors at various historical stages of the language. Still, we have seen the potential for the [F, N, V] system to accommodate this morpheme, and equally importantly, *de* in either analysis is of a clearly language-specific nature and thus is disfavored in the case of categorial deviation (cf. 1.1.2.2).

---

18 *De* results from two separate morphemes, *di* and *di*, distinguished through tones, in an earlier stage (Lü 1984: 130-131). Of the two *di*’s, one was used to introduce a “descriptive” phrase for N whereas the other, judging from the examples in Lü’s work, was limited to introducing AP modifiers that are “qualitative” (p. 126). Lü’s classification of NP-internal modifiers into descriptive and qualitative might correspond to the distinction between individual-level and stage-level predicates (Carlson 1977).