TOPICS ON A CATEGORIAL THEORY OF CHINESE SYNTAX

DISsertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of the Ohio State University

By

Ying-yu Sheu, B.S., M.A.

* * * * *

The Ohio State University

1990

Dissertation Committee:

David Dowty
Arnold Zwicky
James Tai
Michael Geis

Approved by

[Signature]
Advisor
Department of Linguistics
ACKNOWLEDGMENTS

First and foremost, I wish to acknowledge my deep indebtedness to my advisor, David Dowty, for his patience and invaluable help, criticism and suggestions throughout the writing of this thesis. He has influenced me greatly through his teaching and writings in my way of thinking about language. Much of the improvement in the final product over the original drafts is due to him and this would have been a better thesis if I had followed all of his suggestions and comments. This is not to say that he agrees with me on all aspects of the content here, as I remain solely responsible for all errors and infelicities.

Thanks also go to the other committee members: Arnold Zwicky, James H-Y. Tai and Michael Geis for their service in the committee. Arnold carefully read the manuscript and provided valuable criticism. Without question, I am also indebted to him for my development as a linguist during my years at Ohio State. Special thanks go to James H-Y. Tai for his intellectual guidance throughout my graduate years and during the writing of this dissertation. I have benefited from his Chinese linguistics seminars, my discussions with him on Chinese linguistics, and his ready encouragement and support.

Jack Hoeksema should also be mentioned here for introducing Categorial
Grammar to me and for inspiring me to develop the material contained here into its present form.

Acknowledgements also go to Chu-ren Huang and James C-T. Huang for their encouragement and for their generosity in sharing with me their pre-publication manuscripts. I thank Chu-ren for our modem-aided discussions across the Pacific of some of the materials here.

I am deeply grateful to the Ohio State Department of Linguistics, its faculty and chairpersons, Brian Joseph in particular, for the many privileges and opportunities provided to me, including my university fellowship, the numerous teaching and research assistantships, as well as the special travel grant which enabled me to attend the 1987 Summer Institute at Stanford, at which my flagging interest in linguistics was rekindled.

For their friendship, I thank Uma Subramanian, Rom, Joel Nevis, Yong-Kyoon No, Chu-ren Huang, Yung-O Biq, Wen-yu Chiang and Jane Smirniotopoulos. I owe special thanks to Uma for her moral support throughout the writing of this thesis. With her I also enjoyed and benefited from talking about linguistics and other things in life, whether or not we agreed on everything. Thanks are also due to Marlene Payha and Chriss Large for their very competent and ever-cheerful administrative support.

Finally, I would like to thank my husband, Richard K. Schwartz, who over the years has become more proficient than I in explaining to people what linguistics is and what exactly a linguist does for a living. I thank him for his patience, his sense of humor, and his support, emotional and otherwise, and dedicate this dissertation to him.
VITA

November 15, 1959 .......................... Born – Tainan, Taiwan

1981 ........................................... B.A., National Taiwan University, Taipei, Taiwan

1985 ........................................... M.A., Ohio State University, Columbus, Ohio

PUBLICATIONS


FIELDS OF STUDY

Major Field: Linguistics
Studies in Syntax, Semantics and Structure of Chinese
# TABLE OF CONTENTS

**ACKNOWLEDGMENTS** .................................................. ii

**VITA** ....................................................................... iv

**CHAPTER**  

<table>
<thead>
<tr>
<th>I. INTRODUCTION</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0. Statement of Goal</td>
<td>1</td>
</tr>
<tr>
<td>1.1. Categorial Grammar as a Descriptive Model for Chinese</td>
<td>2</td>
</tr>
<tr>
<td>1.1.1. Parts of Speech–Based Theory of Grammatical Categories vs. the Categorial Theory of Syntactic Categories</td>
<td>2</td>
</tr>
<tr>
<td>1.1.2. Interaction of Syntax and Morphology</td>
<td>5</td>
</tr>
<tr>
<td>1.1.3. Accommodation of Syncategorematic elements</td>
<td>6</td>
</tr>
<tr>
<td>1.2. Organization of the Dissertation</td>
<td>6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>II. THE INTERRELATIONSHIP BETWEEN SYNTAX AND MORPHOLOGY IN CHINESE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.0. Introduction</td>
<td>8</td>
</tr>
<tr>
<td>2.1. Modularity</td>
<td>9</td>
</tr>
<tr>
<td>2.2. Morphological Structure of Chinese</td>
<td>13</td>
</tr>
<tr>
<td>2.2.1. The Notion of Word in Chinese</td>
<td>15</td>
</tr>
<tr>
<td>2.3.1. The Ionization Phenomenon</td>
<td>31</td>
</tr>
<tr>
<td>2.4. A Lexicalist Account of Ionization</td>
<td>36</td>
</tr>
<tr>
<td>2.5. Summary and Final Remarks</td>
<td>53</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>III. A CATEGORIAL–PROCESSUAL APPROACH TO CHINESE SYNTAX</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.0. Introduction</td>
<td>56</td>
</tr>
<tr>
<td>3.1.1. Item–and–Arrangement vs. Item–and–Process</td>
<td>57</td>
</tr>
<tr>
<td>3.1.2. IP Approach to Syntax</td>
<td>59</td>
</tr>
<tr>
<td>3.1.3. A Categorial–Processual Approach to Syntax</td>
<td>61</td>
</tr>
<tr>
<td>3.2.1. Wrapping (Infixed) Operations in Syntax</td>
<td>63</td>
</tr>
<tr>
<td>3.2.2. The Infixed Adverbial Construction</td>
<td>64</td>
</tr>
</tbody>
</table>
3.2.3. Formalizing Wrapping ........................................ 69
3.3.4. An Account of IAC ........................................ 72
3.2.5. Word Order Principle as a Possible Motivation
for Wrapping in IAC ........................................ 78
3.3. Infixal Argument Construction ................................ 87
3.4.1. Reduplication in Syntax .................................. 94
3.4.2. The Reduplicative Interrogative Construction .......... 95

III. CATEGORY CHANGING RULES AND MARKING OF DE IN COMPLEX CN
CONSTRUCTIONS ........................................... 109

4.0. Introduction ................................................. 109
4.1. Category Change ............................................ 109
4.2. An Analysis of Category Change Marker de ............... 113
4.2.1. Adjectives in Chinese .................................... 115
4.2.2. Treatment of Adjectives .................................. 118
4.2.3. Non-intersective Adjectives in Chinese .................. 121
4.2.4. de as a Category Change Marker ......................... 125
4.2.5. Relative Clause-like Constructions ..................... 128
4.2.6. Stacking of de-marked Expressions
       at the Pre-nominal position ................................ 132
4.2.7. Possible Counterexamples ............................... 138
4.2.8. Privative Adjectives .................................... 140

IV. A TVP ACCOUNT OF THE CHINESE RESULTATIVE CONSTRUCTIONS 144

5.0. Introduction ............................................... 144
5.1. The Resultative Construction in Chinese .................. 144
5.2. The TVP Category .......................................... 147
5.3. Resultative Predicates as TVP's ........................... 149
5.4. Semantics of the RC's ...................................... 156
5.5. The Wrapped Resultative Construction ..................... 165
5.6. The Reduplicative Resultative Construction ............... 167
5.7. The Resultative Verb Compounds ............................ 171

V. CONCLUDING REMARKS .......................................... 176

6.0. Summary .................................................... 176
6.1. Separation of Rules and Operations ....................... 176
6.2. Tectogrammatics vs. Phenogrammatics ....................... 179
6.3. One-place Rules ........................................... 188
6.4. Non-lexicalist Syntax ..................................... 181

BIBLIOGRAPHY .................................................. 182
CHAPTER I

INTRODUCTION

1.0. **Statement of Goal.**

The goal of this dissertation is twofold. I hope to show that Categorial Grammar provides a most viable descriptive model for (Mandarin) Chinese, a language known to be typologically distinct from the more familiar European languages, by analyzing some of the well-known problematic constructions in the language. In the context of current dominance of a multi-stratal theory such as Government and Binding Theory, within which research on Chinese syntax is carried out almost exclusively, I believe it is important to show that Chinese data can be most successfully analyzed by a monostratal theory of syntax (with limited non-concatenative operations) which obeys strict compositionality, in particular by the theory of Categorial Grammar, the basic tenet of which is that complex expressions are characterized by a functor argument relationship between its parts.

In the process of accomplishing the above, I also hope to use data from Chinese to motivate certain extensions of Classical Categorial Grammar, namely incorporation of one-place rules (in addition to the schema of functional application
and functional composition) and of non-concatenative operations.

1.1. **Categorial Grammar as a Descriptive Model for Chinese.**

In the following I will discuss certain methodological and implementational advantages that Categorial Grammar offers as a descriptive grammatical model for Chinese.

1.1.1. **Parts of Speech–Based Theory of Grammatical Categories vs. the Categorial Theory of Syntactic Categories.**

A basic assumption in any syntactic theory is that words (and expressions) in a language fall into classes or categories which generalizations about the syntax of a language are based upon. Most modern syntactic theories are built upon some versions of X–Bar syntax, which relies upon the traditional parts–of–speech classifications of words on the basis of paradigmatic variations, with phrasal categories projected from them. According to the definition given in Hockett 1958, 'a part of speech is a form–class of stems which shows similar behavior in inflection, in syntax, or both. The part of speech system of a language is the classification of all its stems on the basis of similarities and differences of inflectional and syntactical behavior'. In other words, the assumption behind the use of morphologically defined word class systems as a basis of syntactic categories is that words in a morphological class tend to behave similarly syntactically. Such a system of syntactic categories based on sorting out the morphological parts–of–speech distinctions among lexical expressions becomes problematic when carried over to a basically isolating language such as Chinese, where word forms do not vary for any morphosyntactic categories such as case, number,
Chinese grammarians have long been agonizing over the problem of how parts-of-speech classes can be set up in a language which is devoid of almost any morphosyntactically triggered variations among word forms (e.g. Karlgren 1961, Halliday 1956, Kratochvil 1968, Simon 1937, Chou 1950, and Tai 1982 etc.)\(^1\). A constant theme in the literature of Chinese linguistics, therefore, has always been the perhaps futile debate over the parts-of-speech assignments.\(^2\)

A fundamentally different approach to syntactic categorization is taken in Categorial Grammar, where the classificatory scheme of expressions in a language is not based on the morphologically-oriented parts-of-speech classes; rather, syntactic categories directly correspond to semantic types of expressions and are defined recursively over a small set of basic semantic types. If we take the basic types to be individuals (e) and truth values (t), the following definitions recursively define a universal set of categories.

---

\(^1\) As far as Classical Chinese is concerned, Kennedy (1964: 323) concluded: 'in the final analysis word-classes cannot be defined, hence...Chinese grammar must start from different premises.'

\(^2\) To cite just a few examples, Li and Thompson 1974, Ross 1984 & 1986, Liang 1971, S.F. Huang 1974, Cartier 1970 discuss whether prepositions (traditionally called 'co-verbs') in Chinese can be distinguished from verbs, Ernst 1987 & 1988 discuss the postposition vs. noun status of a class of postposition-like expressions, Chao 1968, Tang 1979, and Ross 1983 discuss the (lack of) distinction between verbs and adjectives, and Zhu 1956 and Fu 1954 focus on how adjectives cannot be clearly distinguished from adverbs.
(1) a. e is a category.
   b. t is a category.
   c. If α is a category and β is a category then αβ (and αιβ, etc) are categories.

(2) αβ + β → α

By virtue of the functional application scheme in (2) (which is non-directional here), an expression of category αβ is something that combines with the expression of category β to form an expression of category α. Such definitions of syntactic categories are based on the combinatorial (and semantic) properties of expressions (which more or less provide direct characterizations of the notions of grammatical functions) rather than on the morphological characteristics of expressions. In this way the Categorially defined categories are universal (though different languages may make use of different sets of available categories recursively defined in (1)), whereas parts-of-speech categories are problematic in languages such as Chinese where words do not exhibit paradigmatic variations for morphosyntactic categories.

---

3 Note that the slashes are not directional here (as is the case in most versions of Categorial Grammar where the direction of the slash indicates the direction where the argument is located relative to the functor), as I am adopting a system where word order (and other possible operations) is to be stipulated by syntactic rules (cf. Pollard 1984).

4See Dowty 1989 for discussions of possible deficiencies in the Categorial theory of categories. For example, prepositions have to be assigned to multiple Categorial categories such as VP modifier (VP/VP), CN modifier (CN/CN) and S modifier (S/S), thus missing a potential generalization that in languages that have prepositions PP necessarily modifies various categories indiscriminately.
1.1.2. **Interaction of Syntax and Morphology.**

Another important typological feature of Chinese which has an effect on the characteristics of grammatical rules of the language is its monosyllabic origin (i.e., each lexeme is a monosyllable). In the modern language, polysyllabic, mostly bisyllabic, words subsequently developed (via archaic but synchronically obsolete compounding processes). Due to such typological characteristics, we find that certain word–internal structural information seems to be accessible to syntax. Such phenomenon, if proven, argues for a modification of a strictly Lexicalist–based approach, where interaction of components is forbidden. In phrase structure grammars of numerous varieties, including the program of Generalized Phrase Structure Grammar, interaction of the two components is allowed only through the medium of morphosyntactic features, which facilitate the transmission of morphosyntactic information between the components of syntax and morphology. No other lexical information can be accessed by syntax. The formalization of (some versions of) Categorial Grammar, on the other hand, is such that it allows certain interactions of the components of syntax and morphology, an issue I regard to be empirical rather than meta–theoretical. In many versions of Categorial Grammar, including the one adopted by Montague, syntactic rules directly manipulate (that is, combine, etc.) strings of expressions rather than manipulate information–encoding devices such as features (including category indexes), and thus allow certain interaction of morphology and syntax, which I hope to show is necessary in Chinese.
1.1.3. Accommodation of Syncategorematic elements.

In inflectional languages, morphosyntactic properties are realized through paradigmatic variations in word forms. It is generally known that Chinese does not utilize such a formal apparatus for this purpose. Instead, morphosyntactic properties are realized through particles called *xu zi* ('empty words') in the traditional terms, which are acategorial in nature, or even through processes such as reduplication in syntax. These syncategorematic elements can be straightforwardly accommodated in the variety of Categorial Grammar adopted by Montague, which I call the Categorial–Processual Approach. Here one does not restrict syntactic rules to functional application and operations to concatenation (as in orthodox varieties of Classical Categorial Grammar), thus allowing the possibility of syncategorematic elements, whether in the form of particles or certain processes.

1.2. Organization of the Dissertation.

The materials in this dissertation are organized as follows. Chapter two discusses the morphological characteristics of Chinese and how what has been known as the 'ionization phenomenon' exhibits syntactic accessibility of word–internal structures, therefore arguing for a theory permitting in principle certain interactions between the components of morphology and syntax. In Chapter three, I discuss how the Categorial–processual approach provides a framework within which adequate

---

5 cf. Zwicky 1988b's remarks: 'Particlexemes [particle words] serve as marks of syntactic constructions in the same way that IM [inflectional morphology] does. What one language does with IM another does with particlexemes.'
descriptions of constructions exhibiting ionization and those involving reduplications can be given. I will argue that the data motivate the addition of one-place rules and non-concatenative operations. Accounts of the constructions marked by what I call 'category-changing' de and of the Resultative Constructions, both involving one-place rules and the latter involving reduplication as well, are given within the postulated framework in Chapters four and five respectively.
CHAPTER II

THE INTERRELATIONSHIP BETWEEN SYNTAX AND MORPHOLOGY IN CHINESE

2.0. Introduction.

This chapter deals with the inter-relationship between syntax and morphology in Chinese, which is a classic example of an isolating language. In particular, I will present a group of verbal expressions from the language and show how these data present a challenge to certain claims of the Thesis of Modularity and how we need a formalization of grammar such as the one provided by Categorial Grammar in order to accommodate these data. I will begin with a brief discussion on the issue of modularity (Section 1). Before anything meaningful can be said about the interaction of syntax and morphology in Modern Chinese and its relevance to the modularity issue, we also need to know more about the morphological structure of the language\(^1\); in particular, we need to be able to determine the structural level parallel to words, an all important structural level in most Western languages, which motivates the

---

\(^1\) For references and general discussions on the morphological structure of Modern Chinese, see Chao 1968, Li and Thompson 1981, Kratochvil 1968, Norman 1988, & Kennedy 1946.
modularization of grammar (Section 2). I will then show how a group of data, which exhibit what is referred to the 'ionization' phenomenon, show that certain lexical information can be relevant to syntax, and therefore argues for a formalization of grammar which allows for sharing of information between syntactic rules and word–internal structures beyond the morphosyntactic information permitted by the Modularity Thesis and the Lexicalist Hypothesis (Section 3). I will also discuss in some detail J. Huang 1984, which assumes the universal validity of the Lexicalist Hypothesis and offers a Lexicalist account of the ionization phenomenon, which I will argue to be unsatisfactory (Section 4). A summary and final remarks are given in Section 5.

2.1. **Modularity.**

The dominant view regarding the overall architecture of the grammar of a language, a view often taken to be a meta–theoretic design feature of a grammar, is the following thesis known as the Thesis of Modularity, which at a very general level can be stated as the following.

(8)The Thesis of Modularity:
The grammar for a natural language consists of several autonomous components or domains, each with its own fundamentally different objects and rules of formation, with minimal interface between these components.

An important corollary of this thesis is what is known as the Lexicalist Hypothesis, which underlies many theorizations of how a grammar should be
organized. It is also called alternatively as the Principle of Lexical Integrity, the Thesis of the Atomicity of Words (Di Sciullo and Williams 1987), or the Principle of Morphology–Free Syntax (cf. Zwicky (to appear)). The underlying assumption here is that the structural level of word divides a grammar into two subtheories of language, namely, morphology and syntax; the kind of linguistic objects above the level of word are fundamentally distinct from those below the level of word and the nature of rules and principles governing the former are very different from those governing the latter. Minimal interaction is expected of these components of grammar. Further, words are 'syntactic atoms' in the sense that their internal structure is not visible or accessible to syntax, or in other words, that syntactic rules are blind to morphological information (cf. Jackendoff 1972, Chomsky 1970, Kiparsky 1982, Zwicky (to appear), Aronoff 1976, Di Sciullo & Williams 1987, Anderson 1982, Lapointe 1979, Scalise 1984, J. Huang 1984).

To be sure, all theories of syntax–morphology interface have to allow certain interactions of the two components, in order to deal with, for example, paradigmatic variations of word forms as determined by the syntactic configurations in which words occur in. For example, the N heading a subject NP takes a form which is covariant with the verb heading the VP in a sentence in languages that have subject–verb agreement. In the program of Generalized Phrase Structure Grammar, such interface of the components of morphology and syntax is mediated through morphosyntactic features (or 'gramcat' features, in Zwicky's 1988b terminology). The formalization of
GPSG allows the transmission of morphosyntactic information between the components of syntax and morphology by encoding the (inflectional) morphological operations that are to take place due to the configurational properties of expressions as features augmented to categories. Any other lexical information internal to a word, such as those relating to the derivational history of a word (e.g., properties of affixes occurring in words or the particular derivational rules that have applied) and the internal boundaries or constituency within words as well as the morphophonological or phonological properties of morphemes within words are opaque and inaccessible to syntax (cf. Zwicky (to appear)). For example, syntactic rules should not be able to distinguish two nouns such as *stupidity* and *foolishness* which contain different derivational affixes. As another example, syntactic rules cannot be sensitive to the fact that the morphosyntactic feature of plurality is realized differently in nouns such as *girls* and *geese*.

While there are authors that take the Principle of Morphology–free syntax to be a global design feature of grammar (cf. Di Sciullo and Williams 1987) and not subject to empirical challenge, my discussion of the Chinese data below is based on the assumption that this thesis is an empirical claim about grammar of languages and

---

\(^2\)In GB theory, on the other hand, the argument structure of lexical items also has to be made known to syntactic rules such as theta–role assignment and control (cf. Di Sciullo and Williams (1987:48)).

\(^3\)Note, however, that there are theories of morphology and syntax in which the notion of affix is available to syntax and that syntactic rules directly maneuver affixes (e.g. Botha 1980).
hence is subject to possible falsification or modification. What is to be noted here, however, is that while the main content of the Lexicalist Thesis has been about the interaction of derivational morphology and syntax, that is, about the potential syntactic relevance or accessibility of word formations, particularly those involving derivational processes, what I will be discussing in the following has little to do with derivational morphology but rather is about the syntactic relevance of a particular kind of lexical information, namely, the information regarding the internal boundaries and bracketing of words (which, as we will see soon, are not introduced by derivational morphology). This kind of information, unlike morphosyntactic information, cannot be encoded as (GPSG–style) features and is therefore beyond the realm of permissible interaction between the two components allowed by the Thesis of Modularity.

If the data presented here will indeed show that certain lexical information is available for syntactic purposes, this would mean that the boundaries of words, in Chinese at least, do not block off all information from syntactic rules as is the case in many languages. It would then call for a modification of the Lexicalist Hypothesis so as to allow for availability of a certain kind of lexical information (in addition to the usual morphosyntactic properties of words) to syntax, at least for languages such as Chinese.

Obviously, in order to substantiate the claim that the Chinese data constitute violations of this important thesis, it is necessary to show that the structural level of
word (a level between morpheme and phrase) can be motivated in the language to some extent. In the following, I will discuss the morphological structure of Modern Chinese and show that the structural level of words can be motivated and maintained.


The unique morphological characteristics of Modern Chinese are closely related to the 'monosyllabic' nature of Classical Chinese. Classical Chinese is what can be called a monosyllabic language in the sense that each syllable is generally a unit of meaning, that is, a morpheme, and further, a word, since Classical Chinese is also monomorphemic, i.e., each word consists of one morpheme.\(^4\) Therefore, the central linguistic unit intermediate between a phoneme and a sentence in Chinese is the monosyllable, which is called zi\(^5\) in Chinese (Chao 1945 & 1968, Kennedy 1951, Li and Thompson 1981, Norman 1988, etc). A zi is an invariantly monosyllabic linguistic unit, which is represented by a character in writing (that is, it is the unit of writing in the non-alphabetic writing system)\(^6\) and is associated with one or occasionally a set of

---

\(^4\) The notable exceptions of polysyllabic morphemes in Classical Chinese are mostly loan words, for example, *putao* ('grape'), *boli* ('glass') and *meigui* ('rose'). Native bisyllabic words are extremely rare, e.g. *kangkai* ('to be generous'). For more discussions of the isolating nature of Classical Chinese, see Norman 1988, Chap. 4.

\(^5\) Chao (1967:284) called it the 'word–syllable' and also the 'sociological word' on some occasions.

\(^6\) In general, each zi is represented by a unique character in the writing system, but homographs can also be found, where different zis are written with the same character.
(etymologically) related meanings\textsuperscript{7}. Throughout history, some zis have ceased to be lexemes in the modern Language, but it is still possible for speakers to assign them meanings based on their knowledge of Classical Chinese\textsuperscript{8}.

In the course of its historical development, Chinese developed an ever increasing number of polysyllabic (mostly bisyllabic) expressions by various types of compounding rules which were active at some stage of Classical Chinese, but have long ceased to be so\textsuperscript{9}. Since there is only a very small set of affixes available for derivational morphology in the language, compounding is the most important means to multiply the number of words and thus expand the vocabulary of the language. This gradual process of polysyllabicizing, i.e. the process of increasing the number of syllables in a word, is partly necessitated by the quite pervasive homophonism, particularly in Mandarin, which resulted from phonological changes that gave rise to loss of tonal distinctions and final consonants\textsuperscript{10}. The effect of these archaic

\textsuperscript{7}Sometimes the meanings that are associated with a particular zi may become synchronically unrelated. This would mean homonyms for an average speaker not well-versed in Classical Chinese.

\textsuperscript{8}In Chinese society, study of the Classical language is an essential part of the literacy process.

\textsuperscript{9}These rules are to be distinguished from true synchronic compounding rules.

\textsuperscript{10}The resulting paucity of distinct syllables puts the language at a tremendous disadvantage in that it greatly limits the number of possible morphemes (due to monosyllabicity) and hence the number of possible words (due to monomorphemic tendencies). Chao 1945 cited C.W. Luh's compilation of a list of 5000 monosyllabic morphemes in the modern Beijing dialect (the standard dialect of Modern Chinese). Since the dialect has only about 1300 different syllables, there is an average of four (continued...)
compounding (or polysyllabicizing and polymorphemicizing) processes is most prominent in modern Mandarin, which has a far greater number of polysyllabic and polymorphemic words than other dialects (cf. Kratochvil 1964 & Kennedy 1946).

2.2.1. The Notion of Word in Chinese.

The notion of word as a central structural level of the grammar of a language is rooted in the grammatical tradition of the classical languages. Its special status derives from the fact that one finds units of phonology, morphology and syntax converge to a large extent at the word level. For example, in languages such as Latin, it is the unit where accent is determined. This is also the unit that is the minimal free-standing unit (hence words are defined as free forms, as opposed to bound forms, in Bloomfield's classifications). Other properties symptomatic of words include cohesiveness (or non-interruptability), fixed ordering and non-recursiveness.\(^1\) Moreover, words can be structurally complex, i.e. analyzable into meaningful parts, called morphemes by the structuralists, the organization of which constitutes the subtheory of morphology.

From the above discussion it is easy to see why the determination of words (and morphemes) may not be a straightforward matter in modern Chinese, particularly because of the numerous polysyllabic (mostly bisyllabic) expressions (resulting from

\(^{10}\) (...continued)
morphemes for each syllable.

\(^{11}\) These criteria for wordhood are discussed in Matthews 1974.
the archaic compounding processes), which contain parts (i.e. zi's—the monosyllables) which were each units of meaning (that is, lexemes) at older stages of the language but may have ceased to be such in the synchronic grammar. There is much confusion in the literature as to whether such polysyllabic expressions should be considered phrases, words, or even morphemes. In other words, it is not straightforward whether the monosyllables making up the polysyllabic expressions—which, as explained above, may no longer be independent units of meanings in the modern language—should be counted as semantics-free syllables, morphemes, or words. Note that these bisyllabic expressions are often referred to as 'compounds' in the literature, but they are obviously to be distinguished from true compounds in the language which by definition are products of synchronic compounding rules. (This will be discussed more later.)

Part of the difficulty in classification is due to the fact that whether these polysyllabic expressions can be further analyzed into meaningful parts (whether etymologically correct or not) is partially dependent upon the degree of individual speakers' literacy in Classical Chinese. In Classical Chinese each monosyllable is a lexeme, in contrast to the modern language where some monosyllables (zi's) are no longer such (only surviving as part of bisyllabic expressions) and their etymologically correct meanings are not necessarily known to the modern speakers.

\[12\text{excluding the few polysyllabic monomorphemic words (cf. footnote 4) and true compounds in the modern language.}\]
To be sure, in all languages there is a structural level between phonemes and phrases which rules of grammar manipulate to form larger units in a compositional way. This structural level is what 'lexeme' is usually conceived to be. In languages where word form varies for morphosyntactic properties, a lexeme is an abstract unit which is instantiated by different words. (For example, die and died in English instantiate the same abstract lexeme.) In Chinese, due to lack of word form variations, the most natural way to conceive of words is to equate them with lexemes. Certainly, the substance in the postulation of the theoretic entity of word derives from the fact that this level has been shown to exhibit a cluster of presumably independent properties (much like the theoretic construct of head upon which independent properties seem to converge), such as those mentioned above. Such clustering of properties may be more prominent in some languages than others, and perhaps negligible in some others, which, as Matthews 1974 demonstrates, is the case for Turkish\textsuperscript{13}.

In the following, I will summarize how in Chinese the level of word, which, as stated above, can be identified with lexemes due to lack of paradigmatic variations, is relevant in various aspects of grammar, exhibiting certain properties or symptoms pertinent to words in other languages (though, as I will to show later, not all of the properties) and is a notion maintainable in the grammar of Chinese. (As will also be

\textsuperscript{13}which may best be described with a morpheme–based approach.
discussed later, whether or not the level of *morpheme* can be also shown to be clearly delimited as an level between *phonemes* and *words* is a separate matter).

First of all, an important aspect of wordhood is that it is a phonologically relevant unit (though the phonological correlations of word boundaries exhibited in languages differ). For example, in many languages it is the unit where stresses are determined. In Latin, words are marked off by the penultimate and antepenultimate stress patterns. In Hungarian, the initial syllable of most words is stressed. In Japanese, only syllabic nasals are allowed in word final positions. Languages such as Turkish and Hungarian are also known to exhibit word–internal phonological phenomenon such as vowel harmony (cf. Robins 1971). As Chao 1968 (Section 3.3) shows, stress and tonal patterns are not reliable in marking off words in Mandarin\(^4\), where, for example, the third tone sandhi patterns the same way (e.g. raising a third tone syllable to the second tone when it occurs before another third tone syllable) within a word and between words. However, as Chao points out, it is nonetheless possible to identify word boundaries on the basis of potential pauses; boundaries between words admit of pauses, such as hesitations or the pause particles *a(ia)*, *me*, and less commonly, *ba*, while word internal boundaries do not admit of any such pauses. When a polysyllabic word in actual speech is interrupted by pauses of any kind, the repetition of the whole word is obligatory, whereas if a phrase is interrupted

\(^4\)In Wu dialects, in contrast, word boundaries can be recognized on the basis of tone sandhi, which are different within words from sandhi between words (Chao (1968:147)).
at a word boundary, it does not have to be repeated. (See Chao (1968:151–155) for an illustration of this). That is to say, equating words with lexemes, we find that this is a phonologically significant unit in that their boundaries tolerate phonological pauses or interruptions while boundaries of units below this level do not.

A related property of word is what Bloomfield calls 'minimum) free forms', i.e. the property of occurring in isolation (though qualifications need to be made here\(^{15}\)). This would be another useful criterion in determining wordhood in Chinese. What needs to be clarified, however, is the usage of the term 'bound forms', which are forms that are phonologically bound and cannot occur in isolation, as opposed to 'free forms'. By 'bound forms' Bloomfield possibly had in mind only affixes (derivational and inflectional)\(^{16}\). If so, the term 'bound form' may not include those monosyllables in Chinese which were once lexemes but are now no longer free forms and only found as part of one or a few polysyllabic expression in the modern language. It may be helpful here to distinguish affixes, which are phonologically bound but distributionally 'free' or 'versatile' in the sense that they can be combined with other forms by rules of the grammar, from the monosyllables in question, which are not only phonologically bound but are 'distributionally bound' as well. Chao 1968 uses the terms 'versatile forms' and 'restricted forms' for the same distinctions. Since the monosyllables in

\(^{15}\)e.g. cases of 'bound words' found in various languages (cf. Nevis 1985).

\(^{16}\)Bloomfield in fact stated that there are no bound forms in Chinese, which shows by the term 'bound form' he meant affixes. This is incidentally not true because there are derivational affixes in modern Chinese.
question are also phonologically bound (i.e. cannot occur in isolation), they are also 'bound forms' in a broader sense. To distinguish them from affixes, however, I will use the term 'distributionally bound' forms (DB forms henceforth) to refer to them. The closest example of DB forms in English may be the so-called 'cranberry morphemes', e.g. *cran-* , which occur only in a single combination in the language. Such forms are marginal in English but are abundant in Modern Chinese due to its monosyllabic origin and subsequent historical development. In including these as 'bound forms' (in the broader sense), we properly exclude them as *words*.

Furthermore, rules of a grammar can be basically divided into two kinds on the basis of their productivity and semantic predictability, namely, those which are completely productive and semantically compositional and those which are partially productive and semantically not always compositional. Dowty 1979 calls these syntactic rules and lexical rules, respectively. The output of the latter are merely possible derived expressions with the actual occurring words listed in the lexicon, possibly with a meaning deviant from the one predicted by the rule. With syntactic rules, on the other hand, there is no such distinction between 'possible' and 'actual' output. The generalization to be captured here is that formation rules internal to the word, e.g. compounding and other derivational rules, tend to be not entirely regular and not predictable semantically, i.e. 'lexical', in Dowty' usage of the term, while formation rules combining larger-than-word units tend to be entirely productive and compositional, i.e. 'syntactic'. (As will be noted below, exceptions to these tendencies
are possible). In modern Chinese, although derivational morphology is quite meager compared to many other languages, morphological processes of the derivational and compounding types do exist in Chinese. For example, the bound form jia ('-ist') is a derivational suffix (e.g., kexue-jia ('science-ist = scientist') and yundong-jia ('athletics-ist = athlete')) and the bound form ke ('-able') a derivational prefix (e.g., ke-ai ('able-love = lovable') and ke-xiao ('able-laugh = laughable')) (See Li and Thompson 1981 for a list of derivational affixes in Modern Chinese.) True compounding rules which concatenate free forms (which are to be distinguished from the aforementioned archaic compounding processes) can also be found in the synchronic grammar. Among them, N–N compounding is perhaps most common, e.g., mi-dai ('rice-bag'), wangqiu-paizi ('tennis-racket'), qiyou-zhan ('gasoline-station'), and fei-bing ('lung-disease = tuberculosis'). Verbal compounds (V–V) are also quite common. One group is called the Resultative Compounds (to be distinguished from the Resultative Constructions to be discussed in Chapter 5.)¹⁷

Thus, although morphological complexity -- of the derivational kind (there is no inflectional morphology in Chinese) -- is very limited in the language as compared

¹⁷Other types of compounding, such as A–N compounding, are also possible, e.g. da-men ('big-door = main door') and xiang-shui ('fragrant-water = perfume'). Bloomfield called compounds such as the A–N type 'syntactic' compounds, as opposed to 'asyn tactic' compounds, such as N–N compounds, which do not have parallel constructions in syntax.
to many English-like languages\textsuperscript{18}, insofar as we need to recognize the level of word in any language that has derivational morphology\textsuperscript{19}, in order to account for the tendency for the processes external to this unit to be entirely productive and the tendency for the rules internal to it to be subject to exceptions and idiosyncracies, we need to recognize the level of word in the grammatical hierarchy in Chinese in order to capture the differences between formation rules external to it and those internal to it.

In this connection, it may be useful to consider the cross-classificational distinctions drawn in Dowty 1978 & 1979 between syntactic and lexical rules on one hand and morphological and syntactic operations on the other hand. As already mentioned, a distinction is made among rules in a grammar according to their productivity and semantic predictability, on the basis of which we distinguish completely productive (or syntactic) rules from lexical rules. On the other hand, on the basis of the grammatical level of expressions involved, rules can be said to involve 'morphological operations' or 'syntactic operations'. Both syntactic and lexical rules may involve morphological or syntactic operations; the former 'operate on' (concatenate or otherwise) morphemes to form words while the latter 'operate on' words to produce phrases.\textsuperscript{20} As Dowty points out, the implementation of the

\textsuperscript{18}Recent studies have, however, shown that it is in fact richer than had been previously recognized (see Chan 1984, Packard 1989).

\textsuperscript{19}that is not entirely productive as in polysynthetic languages such as Eskimo.

\textsuperscript{20}Also, cf. Zwicky 1989, which proposed a classification where lexicon refers to the component whose objects of descriptions are words and syntax (or 'syntacticon')
distinction between morphological and syntactic operations would depend upon independent differentiation of morpheme boundaries vs. word boundaries. The claims of the modularity thesis regarding the properties of words, such as fixed ordering and uninterruptability, are then captured in this program by placing restrictions on morphological and syntactic operations as to what they can or cannot do and on what sort of information they can or cannot refer to. For English, we may require, for example, that syntactic operations not interrupt expressions produced by the morphological operations.

As already mentioned, syntactic rules tend to involve syntactic operations and lexical rules tend to involve lexical operations. But as pointed out by Dowty 1979, some syntactic rules may involve morphological operations, e.g., (1) inflectional rules and (2) completely regular derivational morphology in polysynthetic languages. Both types of rules do not seem to exist in Chinese. On the other hand, English factitive constructions and verb–particle combinations can be construed as cases of lexical rules which form units of more than one word. I am not sure at this point if there are any rules in Chinese that should be properly characterized as such, i.e., as lexical rules producing forms which can later be interrupted, although I will in fact propose later in this chapter a 'reanalysis' rule which introduces word–internal boundaries that are accessible to syntactic rules but which I distinguish from regular lexical rules (see the

\[...continued\]

refers to the component whose objects of descriptions are phrases and each component include both a regular part (= grammar) and an idiosyncratic part (= 'idiosyncraticon').
discussion on pp. 44-45). In short, it seems that in Chinese entirely productive (or syntactic) rules involve in principle syntactic operations (which combine words to form larger units) and lexical rules as a rule involve morphological operations. In other words, the unit of word in Chinese is also partly motivated on the grounds that formation rules external to it tend to be completely productive (or syntactic) and that formation rules internal to it (which, again, are not abundant but do exist in the modern language) tend to be only partially productive, or, lexical. It is important, however, to note here that an analysis of the data given below exhibiting the ionization phenomenon will be presented in the next chapter where we will see that certain syntactic rules in the language involve both syntactic and morphological operations. In particular, I will show some of the ionization phenomena are a result of the fact that the wrapping operation invoked by certain syntactic rules are simultaneously syntactic and morphological in nature, i.e. the operation accesses both phrasal boundaries and word–internal boundaries. Such an analysis cannot possibly be implemented in a strictly Lexicalist theory of grammar but can be accommodated in Dowty's program.

To summarize the foregoing discussion before we proceed to discuss the status of the bisyllabic verbal expressions that are the subject of this chapter, here we are equating words with lexemes (i.e. units of meanings which rules of syntax combine in a compositional fashion to form larger expressions) due to lack of paradigmatic variations in the language. They are 'distributionally free' forms in the sense described above. They also have the property of 'minimal free forms' in Bloomfield's sense (i.e.
occurring in isolation)\textsuperscript{21}. Finally, word-external boundaries, but not word-internal ones, allow pause particles and other forms of hesitations. In short, it is possible to a certain extent to identify and maintain in the hierarchy of grammatical units the level of \textit{words} in the language.

The identification of the structural unit of \textit{morphemes}, which is defined by the structuralists as the minimal unit that bears meanings in a language, is, on the other hand, somewhat problematic, due to the nature of the polymorphemicizing process that the language has undergone over its history. While morpheme boundaries can be clearly delimited in cases where they are introduced by synchronic derivational rules (including affixal and compounding rules), it is not a straightforward matter where morpheme boundaries are to be located in polysyllabic (mostly bisyllabic) expressions which are the products of archaic compounding rules which have long ceased to be active. Often one or both of the component \textit{zi}'s in them are no longer \textit{words} in the modern language, i.e. no longer free forms (both in the phonological and distributional/syntactic sense), but only survive as parts of (one or a few) bisyllabic expressions.

\textsuperscript{21}Exceptions, however, would be made for bound word clitics such as the class of sentential final particles. (Cf. finnish particles as discussed in Nevis 1985 and C. Huang 1985.)
To illustrate, in a bisyllabic word such as *zhi#*dao\(^{22}\) ('know—way = to know'), both zhi and dao were free forms or words in Classical Chinese but both are 'syntactically bound' forms, i.e., DB forms in Modern Chinese. However, through a speaker's knowledge of Classical Chinese, which is, of course, highly variable among speakers, it is possible for a speaker relatively literate in Classical Chinese to assign the meanings (as glossed) to the component zi's zhi and dao in the word, even though they are not lexemes/words and are 'syntactically bound', i.e., not members of any lexical category to be combined by syntactic rules in the modern language. Similarly, it is possible to assign a meaning to each component zi in a bisyllabic word such as *fou#*ze ('not—then = otherwise') through one's knowledge of Classical Chinese. Sometimes the assignment of meanings to zi's in a word is not possible for speakers that are not well-versed in Classical Chinese, e.g. *ke#*shou ('?—? = to cough'), *guo#*ran ('?—so = indeed'), and *ju#*gong ('?—? = to bow'). Thus, the ability to assign meanings to zi's in bisyllabic words is highly dependent on an individual speaker's proficiency in Classical Chinese.

---

\(^{22}\)For purpose of illustration, I will use '#' to refer to zi boundaries in the compounds and '*' to indicate 'bound forms' in the sense described earlier. Also note that the glosses given for the component zi's here are deliberately not the most erudite ones.
It is obvious, then, why delimitation of morpheme boundaries in modern Chinese is problematic: Should we count each zi in the bisyllabic words\textsuperscript{23} in the modern language as a morpheme, i.e., as each a unit of meaning? Chao 1968 seemed to answer in the positive to this question.\textsuperscript{24} In the structuralist's conception of morphemes, however, morphemes, as the minimal units of meaning in a language, are necessarily 'distributionally free' (in the sense that they can be combined by rules of the language, syntactic or lexical), whether they are phonologically free or not. But the component zi's in the polysyllabic words often are not such distributionally free items; sometimes one of the two zi's in the word is a free form and the other is a DB form (e.g. *gong#lu (‘public-road' = highway) and as we have seen from the above examples, sometimes both zi's are DB forms. Therefore, as far as the bisyllabic words that contain one or more than one DB forms (which, again, are also bound in Bloomfield's sense) are concerned, the internal boundaries within these words are not exactly morphemic in nature. That is to say, these bisyllabic words are not straightforwardly morphologically complex.\textsuperscript{25}

\textsuperscript{23}Note that there are a very limited number of bisyllabic words in Classical Chinese which are historically monomorphemic (see footnote 4). They are excluded from consideration here.

\textsuperscript{24}According to Chao, 'without completely equating morpheme to the character or to the zi, it is on the whole advisable to take the maximum analysis of literate speakers, since there will be a better chance of arriving at more convergent results than if we followed the analysis in varying degrees by speakers of varying degrees of literacy.' (1968:142).

\textsuperscript{25}Bradley 1980 (as discussed in Di Sciullo & Williams 1987) reports an experiment which can be used to discover whether word boundaries or morpheme boundaries exist within certain expressions. Results obtained through psychological
What is more difficult to classify are the bisyllabic words in which both zi's are free forms in the modern language but in which the compositional relationship between the zi's is completely opaque to a modern speaker, i.e. there are no compounding rules in the synchronic grammar that could account for the combinations in question. For example, in feng#liu ('philandering'), both feng and liu are free forms in Modern Chinese, meaning 'wind' and 'flow' respectively, but the meaning of the whole expression can not be related in any way to the meaning of the parts via a compounding rule in the synchronic grammar. The high degree of specialization of meaning makes the compound more like one unit of meaning rather two. The degree of specialization of meaning in bisyllabic words (that are not products of any lexical rules), where both component zi's are free, forms a continuum, with some of them more easily construable as extensions of the combined meanings of the component zi's than others. Examples here are bisyllabic words such as xiao#xin ('small–heart = be careful'), dan#xin ('carry–heart = be worried'), rou#ma ('flesh–numb = be disgusting/gross'), and jin#zhang ('tight–tense = be nervous'). Note that the term 'compound' has been used loosely to refer to these bisyllabic words (as well as those in

---

(...continued)

experiments of this kind may be very useful for our discussions here but unfortunately to my knowledge such experiments are yet to be carried out for Chinese.

26. The situation may be considered parallel to that of understand in English. Here one can segment the compound into two familiar morphemes but it is no longer possible for modern speakers to construe the meaning of understand as an extension of the combined meanings of the two morphemes under and stand. Therefore it may be more properly considered as a monomorpheme instead of a compound.
the preceding paragraph) within the literature. This is, however, obviously a misnomer since these are not products of synchronic compounding rules combining two words and are therefore to be distinguished from true compounds in the modern language. On the basis of the definition of morphemes, however, they probably should be considered bimorphemic. For future reference, I will call the bisyllabic words which contain at least one DB form 'Type I' and those that contain only free forms but are not true compounds 'Type II'.

For our purpose here, though, what is crucial is to distinguish all of the above, which are *words*, from bisyllabic expressions that should be properly classified as (idiomatic) phrasal expressions, such as (1) below.

(1)a. chi cu
   eat vinegar
   'be jealously of'; lit., 'to eat vinegar'

b. chi doufu
   eat Tofu
   'to flirt (with)'; lit., 'to eat Tofu'

c. dai gao mao
   wear high hat
   'to flatter/to butter up'; lit., 'to wear a tall hat.'

Expressions in (1), unlike the polysyllabic expressions shown earlier, have the important characteristics of idiomatic phrases, namely, they are syntactically well-formed, or as Zwicky 1989a put it, they instantiate (or 'are parasitic on') certain syntactic constructions in the language (with respect to both their internal make-up
and external distribution)\textsuperscript{27}. (See Zwicky 1989a, Zwicky 1989b, Hockett (1958: Ch. 36), Fillmore, Kay, & O'Connor 1988, etc). Idiomatic phrases are generally products of the ordinary syntax of a language, but, in addition to a literal meaning assigned by the compositional semantics, they also have a meaning, the idiomatic meaning, deviant from that which is predicted by the semantic rule associated with the syntactic rule which builds them up. For example, an idiomatic expression such as \textit{kick the bucket} in English instantiates the verb–object construction but, in addition to its literal meaning assigned by the compositional semantics, it also has a meaning to be learned individually. This property shared by idiomatic phrases sets the expressions in (1) apart from the bisyllabic expressions discussed above, since they instantiate regular syntactic constructions (namely the Verb–Object construction) and have a meaning predicted by the compositional semantics as well as a meaning not so predicted. The bisyllabic expressions seen earlier, on the other hand, are not 'parasitic on' any syntactic constructions of the language and, as we have seen, do not have a 'predicted' meaning but only a 'learned' or 'listed' meaning. This proper distinction between true idiomatic phrases and the bisyllabic expressions in question, which are words, will be important in our later discussions.

\textsuperscript{27}Exceptions, such as \textit{by and large} in English, which are not syntactically well-formed, exist but are extremely rare.
2.3.1. The Ionization Phenomenon.

The data which argue for syntactic accessibility of lexical information (of a particular kind) involve the group of bisyllabic expressions loosely referred to as 'V–O Compounds' in the literature, such as (2)–(3) below. (Recall that ‘#’ indicates zi boundaries and ‘*’ indicates that the zi in question is a DB form. Moreover, the glosses for DB forms, which are not lexemes/words in the modern language, are based on an average speaker's knowledge of Classical Chinese).

(2)a. shang#feng
    hurt - wind
    'to have a cold'

    b. de#zui
    obtain-guilt
    'to offend'

    c. guan#xin
    close-heart
    'to be concerned'

    d. sheng#qi
    grow - air
    'to be angry'

    e. dan#xin
    carry-heart
    'to be concerned'

(3)a. *zhu#*ce
? - book
'to register'

    b. *zhu#*yi
? - idea
'to pay attention'

    c. *ju#gong
? - ?
'to bow'
d. ti*yi
   propose-idea
   'to make a proposal'

  e. shang* dang
     up  ?
     'be fooled'

As with the other bisyllabic expressions discussed above, the use of the term
'compound' for these expressions is a non-standard one since these are not products of
the synchronically active compounding rules which combine free forms of certain
categories. The bisyllabic expressions in (3), which contain DB forms, would be Type
1 bisyllabic expressions, which, as I suggest above, contain no morpheme boundaries
(though it may be possible for a speaker somewhat literate in Classical Chinese to
supply a meaning, whether etymologically correct or not, for a DB form). Those in
(2) would be Type II expressions, since they contain only free forms (though, of
course, they are not true compounds).

These are called 'V-O compounds' in the literature because in these bisyllabic
expressions, as seen from the glosses, the first component zi can be analyzed as verbal
and the second as nouny (at least in cases where the assignments of meanings to the
component zi's are possible). Since there is no synchronic compounding rule

28 Depending on the categories of the two zi's in the compounds, other types of
bisyllabic verbal expressions in addition to the 'VO compounds' include (1)'Coordinate'
type, where a verb is juxtaposed with another verb of similar meaning and (2)'subject-
predicate' type, where the first zi is a noun and the second a verb. The archaic
compounding rules producing them were important part of the language's
polysyllabizing and polymorphemicizing process for verbal expressions (see Chao
combining V and N, the analysis of the bisyllabic expressions into V–O should more
properly be regarded as reanalysis. (This notion of reanalysis of the verbal expressions
will be discussed in greater detail later as it is important to the account of the verbal
expressions that I will be presenting contra a Lexicalist account such as J. Huang
1984).

This group of bisyllabic verbal expressions is interesting for our evaluation of
the Lexicalist Hypothesis because these expressions exhibit certain discontinuity
phenomena: many of these 'pseudocompounds' can be interrupted by other syntactic
units and become discontinuous in various constructions. This phenomenon is dubbed
Ionization by Chao (1968 & 1976), due to its similarity to the chemical phenomenon
of ionization. The following data, which do not exhaust all the possibilities of
ionization²⁹, demonstrate the phenomenon in question.

A. the 'Infixal Adverbial Construction':

(4)a. wo shang le bantian *dang cf. shang#*dang
   I ASP half-day up ?
   'I was fooled for a while.' 'to be fooled'

b. ta sheng le haojiu de qi cf. sheng#qi
   he ASP long DE grow-air
   'He was mad for a while.' 'to be angry'

B. the 'Infixal Argument Construction':

(5) ni buyong dan ta (de) xin cf. dan#xin
   you need-not him DE carry-heart
   'You don't need to worry about him.' 'be worried about'

²⁹ Also see Chao (1968:425) for other constructions where one finds ionization.
(6) wo shang le Zhengsan de *dang
I ASP Zhengsan DE
'I was fooled by Zhengsan.'

(7) ni sheng shei de qi
you who DE
'Who are you mad at?'

(8) ta *su sheme ku
he what
tell-bitterness
cf. *su#ku
'to complain'

What's he complaining about?'

C. the 'Reduplicative Interrogative Construction':

(9) ta jintian hai *ke bu *ke#*shou cf. *ke#*shou
he today still not cough
'Is he still coughing today?'
'to cough'

(10) ni *zhi bu *zhi#*dao zhe jian shi
you not know this CL matter
cf. *zhi#*dao
know tell
'Do you know about this?'
'know'

It should be obvious why these verbal expressions are relevant in our
discussions of the interaction between syntax and morphology. We see that these
'pseudocompounds' display characteristics proper only to phrasal units, specifically,
they become discontinuous or separated in various constructions exemplified in (A)–
(C). In (A), a Type 1 or Type II bisyllabic verbal expression is interrupted by a time
adverbial and becomes discontinuous. In (B), they are interrupted by an oblique
argument. In (C), the first zi in the verbal word is reduplicated while a negation
particle bu is inserted between the reduplicated zi's. As I will try to show in the
following chapter, where a detailed analysis for these constructions will be given, these
data are results of certain syntactic rules having access to the internal boundaries of
these bisyllabic verbal expressions. Note that these are not boundaries introduced by
derivational rules of the language since the bisyllabic verbal expressions in question are not products of any synchronic lexical rules (of either the affixational or the compounding kind). In other words, as already discussed above, these bisyllabic expressions are not morphologically complex in the usual sense: Type 1 expressions are not morphologically analyzable (i.e. they cannot be analyzed into combinations of morphemes (in the standard usage of the term) and Type 2, though analyzable into two free forms, are not true compounds. However, as we see in (4)–(10), certain syntactic rules seem to treat them as syntactically complex and to be able to access the internal parts (of both type I and Type II), rendering the expressions discontinuous at surface.

Detailed accounts will be given in the next chapter to the ionization phenomenon of the bisyllabic verbal expressions exhibited in the above constructions, which as we can see presents a challenge to certain claims of the Lexicalist Hypothesis. In particular, they show the syntactic accessibility of internal lexical information (namely, information about boundaries of a non-morphological nature, at least in the case of Type I expressions). To keep the Lexicalist Hypothesis intact, there has been attempt in the literature, notably by J. Huang 1984, to analyze these bisyllabic verbal expressions as idiomatic phrases, so that these data would not constitute violations of the Lexicalist Hypothesis.
2.4. A Lexicalist Account of Ionization.

According to J. Huang 1984, the grammatical status of the bisyllabic verbal expressions exhibiting the 'ionization' phenomenon is indeterminate between *words* and *phrases*, depending on the context in which they occur. In particular, their grammatical status is closely connected to the following Phrase Structure Condition (PSC), which Huang motivates independently of the ionization data.

(11) The PSC in Chinese:
'Within a given sentence in Chinese, the head (the verb or VP) may branch to the left only once, and only on the lowest level of expansion.' (p. 54)

By this is meant 'the verb of a sentence must occur near the end of the sentence, and only followed by at most one constituent, in the simplest cases' (p. 54). According to Huang, the independently motivated PSC is closely related to the *word*-vs.-*phrase* status of the verbal expressions. To illustrate, consider (12)–(13), where a bisyllabic verbal expression (e.g. *dan#xin* ('to be worried') and *ju#gong* ('to bow')) occurs as an IV\(^{30}\) and (14)–(15), where the verbal expression (e.g. *dan#xin* ('to be worried') and *guan#xin* ('to be concerned')) is a TV, followed by an object NP argument\(^{31}\).

\(^{30}\)Note that while some of these verbal 'pseudo-compounds' can be both IV and TV, some can only be IV and others only TV. The subcategorization property of these verbal expressions is a purely lexical matter and can not be predicted.

\(^{31}\)In these cases, Huang considers these verbs as transitive. But note that synonymous sentences show that when the NP precedes the verbal compound, it is obligatorily marked with a preposition which typically marks an oblique argument ((b) is an alternative way of saying (a); *wo* ('me') bears the same argument relationship to *danxin* in (a) and (b), but it follows the verb in (a) and precedes the verb in (b)).

(a) ta *dan#xin* wo

(continued...)
(12) ni bubi dan#xin cf. dan#xin  you need-not worry carry-heart 'You don't need to worry.' 'to be worried'

(13) ta *ju le bantian de *gong cf. *ju*gong he ? ASP a-while DE ? ? 'He was bowing for quite a while.' 'to bow'

(14) Wo dan#xin ta cf. dan#xin I worry him carry-heart 'I worry about him' 'to be concerned'

(15) ta hen guan#xin ni cf. guan#xin He very concerned you close-heart 'He is very concerned about you.' 'to be concerned'

Data exemplified by (12)–(15) are problematic for Huang as it seems here the proposed PSC and the Principle of Lexical Integrity cannot be simultaneously maintained. The reason is this: if we treat the bisyllabic verbal expressions as idiomatic phrases (with a V–NP structure), thus preserving the Lexicalist Hypothesis, sentences (according to Huang) such as (14) and (15) would seem to violate the PSC since in these cases the verb is followed by two NP arguments. On the other hand, if we treat the verbal expressions in question as words (which I argue in Section 2 and which I consider entirely to be independent of whether or not the PSC holds in

\[31\text{(...continued)}
he worry me
'He is worried about me.'

(b) ta wei wo dan#xin
he for me worry
'He is worried about me.'

Further investigation is needed to determine the real status of these NP arguments (cf. Dowty 1982). For purposes of the present discussion, I will simply follow Huang's usage and call the verbal expressions in (14) & (15) transitives.
Chinese), the PSC would not be violated by sentences such as (14) and (15)\textsuperscript{32}, but in that case sentences like (13), which exhibit the ionization phenomenon, would present a problem for the Principle of Lexical Integrity and for the Lexicalist Hypothesis in general.

To keep both the PSC and the Principle of Lexical Integrity intact, Huang's solution is to treat all the bisyllabic verbal expressions (that can be construed as the V-O type) as (idiomatic) phrases listed in the lexicon (whether or not the particular verbal expression contains DB forms). When such a verbal expression occurs in sentences where it is followed by an (object) NP (e.g. (14) & (15)), i.e., as a TV, the PSC then 'requires it to be a word' (p.69) --- the PSC will be violated if we regard, for example, \textit{dan\#xin} in (14) as an (idiomatic) phrase consisting of a V, \textit{dan}, followed by a nominal expression, \textit{xin}, since in that case the verb would be followed by more than one constituent. To achieve this effect, Huang proposes a rule of Lexicalization, which turns an idiomatic phrase into a word in only those contexts where the verbal

\footnotesize{\textsuperscript{32}In the case that the bisyllabic verbal expressions are being treated as idiomatic phrases, Huang does not regard sentences such as (13), where a time adverbial is inserted into the verbal expression as violating PSC. The reason is that the particle \textit{de} happens to be homonymous with the possessive particle in Chinese. Therefore Huang regards the sequence consisting of the time word followed by \textit{de} and the 'nominal' part in the verbal expression as one single constituent, under the label of NP. Such an account is problematic, since first of all the particle \textit{de} here is only optional and second of all the sequence which Huang labels as NP cannot be topicalized which should be possible with object NPs.}
expression in question is followed by an object NP. The verbal expression would remain a phrase in any other contexts (i.e., not followed by an NP). Under Huang's account, for instance, the Resultative compound da#si ('beat-dead=beat to death') is a ('compound') word in (16) but an (idiomatic) phrase in (17), the passive counterpart of (16), because da#si ('beat-dead = beat to death') is followed by an NP in (16) and not in (17).

(16) ta da#si Zhangsan  
    he beat-dead Zhangsan  
    'He beat Zhangsan to death.'

(17) Zhangsan bei ta da#si  
    Zhangsan by him beat#dead  
    'Zhangsan was beaten to death by him.'

Such an account, according to Huang, is able to offer an explanation why data such as (18) are starred, where the verbal expression dan#xin ('to be concerned') is used transitively and is separated by an adverbial expression. (Cf. (A) above.)

(18)* wo dan le bantian xin ta  
    I carry ASP a-while heart him  
    'I was worried about him for quite a while.'

The explanation is this: here the 'idiomatic VP' dan#xin ('to be concerned') is followed by an NP, therefore the rule of Lexicalization would apply and convert it to a word.

---

33 The details of this rule are not further elaborated by Huang.
The Principle of Lexical Integrity would then predict that it cannot be interrupted in syntax, hence the ill-formedness of (16).

To recapitulate Huang's account, in order to preserve the validity of the Lexical Integrity Hypothesis in the face of the ionization phenomenon of the bisyllabic verbal expressions in question, Huang classifies them as idiomatic phrases listed in the lexicon. But since this would lead in many cases, i.e., when we find the bisyllabic verbal expression in TV positions, to violations of the PSC, which is a configurational principle presumed to hold in Chinese syntax by Huang, he then proposed a rule of Lexicalization which turns a V' category into a V3 category (in syntax) when the verbal expression occurs before an object NP.

Therefore we see that whether a bisyllabic verbal expression is to be considered a word or an (idiomatic) phrase in a particular sentence is dependent upon the (in)transitivity of the verbal expression in a sentence and the content of PSC; whether the rule of Lexicalization applies or not 'is required only by the independently motivated PSC' (p. 70).

A clearly unsatisfactory feature of this account is that in order to preserve the Lexicalist Hypothesis, Huang has to classify the bisyllabic verbal expressions in question as idiomatic verb phrases in the lexicon (without giving any theory of idioms to accompany the account), despite the fact that one or both zi's in such expressions
may be DB forms and despite the fact that for these expressions, the analysis of them into [V-O] is really a case of (folk) reanalysis -- since the DB forms cannot be readily assigned a meaning and are not members of any categories in the synchronic grammar (more on this in the following). As has been pointed out (e.g., Kratochvil (1968:63), Kennedy 1964), bisyllabic words (of various lexical categories) in modern Chinese have become as prevalent as monosyllabic words. According to Huang's account, however, the so-called 'V-O compounds', as well as the Resultative compounds, which together constitute a major portion of all bisyllabic verbal expressions in modern Chinese, are listed as idiomatic phrases (rather than as words) in the lexicon. To say the least, this begs the question of why a large proportion of verbal expressions have not been part of the polysyllabicizing process the language has undergone over its history (since in Huang's classifications they are idiomatic phrases), while bisyllabic words of other categories are commonplace.

Even more implausible is the proposed rule of Lexicalization (in syntax), whereby these idiomatic phrases are turned into words when followed by an NP in a sentence but in other contexts they retain their phrasal status. What is being advanced, in effect, is the peculiar proposition that the grammatical status or level of a verbal expression must be determined by reference to a configurational condition such as PSC and the subcategorization context in which it occurs. From the acquisition point of view, this implies that, for a great majority of verbal expressions in the language, children have to learn the quite sophisticated PSC before they learn whether any given
expression is a word or a phrase in the context of a sentence. This seems to be quite implausible in light of the fact that the distinction between words and phrases is supposedly a universal or a very fundamental one while the PSC is a parochial, quite arbitrary and complex configurationally-based rule\textsuperscript{34} (if the PSC is indeed a valid principle in the language\textsuperscript{35}).

Another odd aspect of the account is that whether or not an idiomatic verb phrase is turned into a word, via a synchronic process of Lexicalization, depends on whether it is followed by an NP argument. An idiomatic verb phrase, according to the account, is turned into a V\textsubscript{0} if it is inserted into a position with an NP following (so as to prevent the PSC from ruling out the sentence). For this to work, a substantive theory of subcategorization and lexical insertion for idiomatic phrases is required, so as to, among other things, provide a mechanism for inserting an idiomatic verb phrase into a slot requiring a transitive verb for sentences like (14) & (15). Huang does not provide such a theory.

\textsuperscript{34} It is of course possible that there exist some isolated instances of some 'marginal' expressions in particular languages, whose grammatical status has to rely on language-specific syntactic or morphological principles. But in the present case, it is the grammatical status of a great portion of verbal expressions in the language that are to be determined.

\textsuperscript{35} Chen (1987) discussed major classes of exceptions to this rule, for example, ditransitive verbs.
I believe that the problem inherent in Huang's account is the view that all claims of the Lexicalist Hypothesis are presumed to be universally valid, which inexorably forces a large group of bisyllabic verbs exhibiting the ionization phenomenon to be classified as (idiomatic) phrases. As I have tried to show, they should be properly classified as words (Type 1 or Type 2) on the basis of the criteria discussed above (section 2). Their ionization properties, I believe, show certain claims of the Lexicalist Hypothesis cannot be maintained for a language such as Chinese with its unique morphological characteristics.

In the following, I will suggest an explanation for the ionization of the bisyllabic verbal expressions in question, using the Infixal Adverbial Construction ((A) above) for illustration. The rule underlying data such as (4) is in fact a more general rule, which is responsible for data like the following:\(^{36}\):

(19) wo kan le yi zheng tian (de) dianshi
    I watch ASP one whole day DE TV
    'I watched TV for the whole day.'

(20) wo deng le bantian (de) huoche
    I wait ASP half-day DE train
    'I waited for the train for a long time.'

(21) wo du le liang ge zhongtou (de) baozhi
    I read ASP two CL hour DE newspaper
    'I was reading newspapers for two hours.'

\(^{36}\) See C.-R. Huang 1988a for more discussion on this construction, which he calls the 'Possessive Object' construction.
Here we see a time or duration adverbial inserted between a verb and its object NP and at the same time the particle de optionally affixed to the adverbial expression 37, which I will analyze as a case of wrapping in the next chapter. What I propose here is that data like (22)–(23) (= (4a)–(4b)) and (24) are a result of 'reanalysis' given to bisyllabic expressions such as shang#dang ((22)), sheng#qi ((23)) and ju#gong ((24)).

(22) wo shang le ban#tian (de) *dang cf.shang#*dang I up ASP half-day DE ? up ? 'I was fooled for a while.' 'be fooled'
(23) ta sheng le hao#jiu (de) qi cf.sheng#qi I grow ASP long DE air grow air 'He was mad for a while.' 'be mad'
(24) ta zai nar *ju le bantian (de) *gong cf. *ju#*gong he at there ? ASP half-day DE ? ? ? 'He was there bowing for a long time.' 'to bow'

In particular, I propose a rule of reanalysis in the lexicon which analyzes a synchronically obscure bisyllabic verbal expression into a complex expression consisting of a verbal category and a noun category. More precisely, it supplies an analysis tree or internal bracketing of [V N] (which is syntactically accessible) for an intransitive bisyllabic verbal expression. The form and nature of the analysis rule in question are in fact not unlike those of lexical rules of Dowty 1979, which are also conceived as supplying analyses for (morphologically) complex lexical items. However, the former has to be distinguished from the latter in their function in the grammar, as the analysis rule here, unlike derivational or compounding rules, can only

37 Though the variant with de sounds better than the one without de for most speakers.
be used to supply internal structure for originally unanalyzable lexical items and cannot be used to derive new expressions, which is an important function of derivational/compounding rules in general. Furthermore, this rule is called a 'reanalysis' rule here, instead of an analysis rule, because it may assign a category to an expression that is not a member of that category (e.g. in the case of Type I bisyllabic verbal expressions). For example, in my account here, *ju*gong ('to bow') in (24) is given the structure of \([ju]_v \[gong]_n\) by the analysis rule, even though *ju and gong are not members of the V category and the N category respectively; in fact they are both DB forms and are not members of any categories in the language. Note that the analysis of the bisyllabic expressions in question (which are basic expressions in Montague's sense) as having internal structures which are accessible to syntactic rules requires that we allow basic expressions with lexically introduced but syntactically accessible bracketing, a situation not represented in Dowty 1979's program of the lexicon.

Whether a certain intransitive bisyllabic verbal expression can be given such a (re-)analysis (and is hence, in the account here, ionizable) depends a lot on whether the first zi in the verb can be given a verbal meaning and the second zi a noun-like meaning. In the case that both zi's in the verb are free forms or lexemes in the modern language, it is straightforward whether the expression can be given a V–N reanalysis. For example, the two zi's in the verb *sheng*qi ('grow–air = be angry') are easily reanalyzable as a \([V-O]\) 'pseudo-construction'. The situation is similar with verbs
such as shang#feng ('hurt-wind = to have a cold'), and guan#xin ('close-heart = to be concerned'). When the bisyllabic verb involves a DB form, its (re-)analyzability into something with a [V−N] structure depends on how easily the zi's can be given a verbal meaning and a nominal meaning respectively (for speakers with average literacy in Classical Chinese). For example, in the verb shang#*dang ('up−? = to be fooled'), even though the meaning of the second zi is a DB form\textsuperscript{38} and not a member of any categories in the modern language, it is possible to give it a [V−N] analysis on the base of the verbal-ness of the first zi.

Another factor of (re-)analyzability involves the stress pattern. As Chao (1968:160 & 431) has pointed out, since a true V−O phrase always has the stress on the object, 'there is a strong tendency to treat any iambic verbal expression as V−O and ionize it, even though it is quite something else' (p.431).

What is crucial here is that the information of the internal bracketing provided by the reanalysis rule seems to be accessible to certain syntactic rules, such as the rule for the IAC Construction ((19)−(21)), which inserts a time adverbial between a verb and its object NP. As (22)−(24) show, the rule does not seem to distinguish between the truly syntactically complex expressions (i.e. expressions that are products of syntactic rules) and the verbal expressions which are words but are complex by virtue of the structure assigned to them by the postulated reanalysis rule. That is to say, the

\textsuperscript{38}There is a lexeme, dang ('to pawn'), in the modern language, but it probably should be regarded as homonymous with the dang in shang#dang.
IAC rule treats the bracketing in the bisyllabic verbs under consideration imposed by the reanalysis rule just like true bracketing induced by syntactic rules, which results in data like (22)–(24), where the temporal adverbial is inserted between the two zi's in the verbs\(^{39}\).

Therefore, whether or not a particular (intransitive) bisyllabic verb is ionizable in the IAC (and other constructions) depends on whether it is given a [V–O] analysis. As Chao 1968 has pointed out, variability among speakers exists as to whether or not a certain item is ionizable. There are some bisyllabic verbal expressions that readily admit of ionization of all types for all speakers and there are others that are inseparable for the majority of speakers\(^{40}\). In the account here, this is directly related to whether an item has been reanalyzed. Take an intransitive bisyllabic verb such as sheng\#qi ('grow–air = to be upset') for example. This expression can be easily given the reanalysis since both zi's are free in the modern language and are members of the V category and the N category respectively. Thus we find the expression ionizable in most speakers' speech. On the other hand, consider a bisyllabic verb such as *zhi##*dao ('to know'), where both zi's in the verb are DB forms but can be respectively glossed as 'know' and 'tell' through an average speaker's knowledge of Classical

---

\(^{39}\)Note that this rule also applies in the cases of true idiomatic phrases, e.g. (a).

(a). Sanbai gen Yunniang chi le yi zheng tian de cu
Sanbai with Yunniang eat ASP one whole day DE vinegar
'Sanbai was jealous of Yunniang for the whole day.'

\(^{40}\)An exception exists for distribution in the Interrogative Reduplicative Construction, which, as I will show in the next chapter, does not depend on reanalysis.
Chinese (since these both are common lexemes in Classical Chinese). Furthermore, it has the trochaic stress pattern. Together this means it is impossible to reanalyze it into a \([V-O]\) (pseudo)construction (even though it can be used as an intransitive verb). We therefore can predict that it cannot be ionized in syntax, which is in fact the case.

(25) \(\text{wo zhidaq le}
\)
\(\text{I know ASP}
\)
\(\text{'I have known (about it).'}\)

(26)* \(\text{wo *zhi le haojih de *dao}
\)
\(\text{I ASP long-time DE}
\)
\(\text{'I have known (about it) for a while'}\)

Between those bisyllabic verbs which can be readily given a \([V-O]\) reanalysis (and therefore ionizable in everyone's speech) and those such as \(zhidao\) ('know') which can hardly be given such an analysis (and are therefore never found separated), there are many the reanalyzability and hence ionizability of which depends upon the individual speaker. This variability among speakers is due to the fact that the analyzezability of an etymologically obscure item (into a verbal part and a noun-like part) is, in addition to the stress pattern, highly dependent upon factors such as a speaker's knowledge of Classical Chinese and whether the \(zi\)'s involved are among commonly used vocabulary in Classical Chinese. Finally, whether a speaker tends to be linguistically innovative or conservative also plays a part here. Chao 1968 cites bisyllabic verbs (even some genuine monomorphemic bisyllabic verbs\(^{41}\)) which can

\(^{41}\) There are some genuine bisyllabic verbs in Classical Chinese that are not results of archaic compounding rules. Kangkai ('to be generous') is such an example, which we find ionized in (a).
hardly be given a [V-N] analysis but are found to occur ionized in some speakers' speech. For example, *ti#cao ('body-manipulate = to exercise') in (41) is etymologically a 'subject-predicate' 'pseudocompound', but a sentence like (41) would indicate that for those speakers who utter it the item ti#cao ('to exercise') has undergone the reanalysis rule I propose here. (For more such data see Chao 1968 & 1976.)

(27) tamen zai nar *ti le bantian *cao
they at there body ASP half-day manipulate(?)
'They were exercising over there for quite a while.'

The strongest data supporting the reanalysis account here are data involving genuine bisyllabic monomorphemic words in the language, which are hard to come by but nonetheless can be found. For example, the bisyllabic verbal expression youmo,42 'to joke around (with), to be funny, to tease, etc.' is the transliteration of the English word humor (e.g. (28)). For some speakers, it is even possible for it to be ionized, as seen in (29).

(28) ni bie lao gen wo youmo
you don't all-the-time with me tease
'Don't tease me (or joke around with me) all the time.'

(29) ta zai nar *you le bantian *mo,

41(...continued)
(a). ni bie laoshi *kang ta de *kai
you should-not always ? him DE ?
'You shouldn't be so generous with his money.'

42 Which has the iambic stress pattern.
(29) ta zai nar *you le bantian *mo,
    he at there? ASP half-day?
    haishi meiren li ta
    still nobody pay-attention-to him
 'He was trying to be funny for all this time, but still
    nobody paid any attention to him.'

Data such as (29) lends crucial support for the present account incorporating a rule of reanalysis for bisyllabic verbal expressions, since it indicates that for some speakers the rule has been applied to a true monomorphemic expression such as youmo,

whereby the bracketing of [V N] is supplied. (29) is the result of the syntactic rule responsible for the IAC having access to the word-internal bracketing and treating the verb as if it were syntactically complex. Under Huang's account, on the other hand, even a genuine monomorphemic item such as you-mo has to be regarded counter-factually as an idiomatic phrase in the face of data like (29).

An important motivation for Huang's account, as already mentioned, is the contrast between (30a)–(30b) and (30c) in following. (30a) shows the verb dan-xin
('carry-heart = worry') occurs transitively (followed by an NP) and (30b) shows the verb is used intransitively and interrupted by an inserted duration adverbial. (30c), where the interrupted compound is followed by an NP argument, is ill-formed.

(30)a. ta danxin zhe jian she
    he worry this CL matter
    'He was worried about this.'

b. ta dan le bantian xin
    he carry ASP half-day heart
    'He was worried for a while.'
The ill-formedness of (30c) supposedly lends support for Huang's account: when a verbal 'compound' occurs transitively, it is turned into a word by the Lexicalization rule (so as to prevent it from being ruled out by the PSC). It would then follow from the Principle of Lexical Integrity that verbal 'compounds' in this position can not be interrupted in syntax, hence the starring of (30c). Under the present account, the ill-formedness of (30c) can be easily explained for: as already mentioned, the rule of reanalysis only applies to intransitive verbs and never to transitive verbs. When a verb such as [dan#xin]$_{TV}$ ('be concerned') is inserted into the slot preceding (or combining with) an NP, there would be no internal structure available for the rule of IAC, hence the ill-formedness of (30c). While Huang relies on a configurational condition such as the PSC to somehow trigger the obligatory rule of Lexicalization for a verbal 'compound' that occurs before an object NP, hence explaining its inseparability, the account here appeals to a very natural assumption that while it is possible to reanalyze an IV as consisting of a verb and an object NP, it would be quite unnatural for speakers to reanalyze a transitive verbal expression as such.

---

43But recall that many verbal compounds have dual categorial membership in both the category of IV and TV, while some can only be TV and others can only be IV. For those compounds that can only occur as TV, the rule of reanalysis can never apply to them, which explains why they can never be found interrupted in syntax.
Furthermore, under Huang's proposal, all 'V-O compounds' (and the Resultative compounds as well) are listed in the lexicon as idiomatic phrases and they are turned into words only when they occur before an object NP, thus implying the determination of the grammatical level of these verbal expressions in a sentence is totally dependent on its (in)transitivity. This predicts that an intransitive verbal 'compound' such as chu#ban ('to be published'), being an idiomatic phrase in Huang's account, would be ionizable, which is incorrect.

(31)a. tade shu chu#ban le
    his book be-published ASP
   'His book was published.'

   b.* tade shu chu le henjiu de ban
      his book ASP long-time DE
   'His book was published for a while.'

   cf.
   c. tade shu chu#ban le henjiu le
      his book be-published ASP long-time ASP
   'His book was published for a while.'

Huang does offer an explanation of why some IV's cannot be separated in syntax by remarking that 'some items have completely been turned into words' (p.71). That is to say, some bisyllabic verbal expressions are listed as words and not idiomatic phrases and we therefore predict their inseparability in all contexts. This, however, is not a satisfactory explanation, since we have IV's such as you#mo ('to joke around') — a transliteration of the English word humor—which is clearly a word and cannot be an idiomatic phrase in anyone's account, yet it is ionizable as shown above. The analysis given here, on the other hand, classifies the bisyllabic verbal expressions as words
(excluding true idiomatic verb phrases such as *chi doufu* ('to eat Tofu' or 'to flirt with')), and, by appealing to a rule of reanalysis (for intransitive verbs), the application of which is dependent on the relative semantic transparency of component *zi*’s in the verb and factors such as stress pattern, is able to give an explanation of why some of the verbs are ionizable and some are not.

2.5. **Summary and Final Remarks.**

To summarize, under Huang 1984’s account, in order to explain the separability of bisyllabic verbal expressions, including those which are usually referred to as the 'V–O compounds' and the Resultative compounds (which I do not discuss here), he considers them as listed (idiomatic) phrases in the lexicon. Furthermore, some idiomatic phrases are turned into words via a rule of Lexicalization in syntax, if they are used transitively, so as to prevent them from being ruled out by the PSC. What I have tried to argue here is that these bisyllabic verbal expressions should be properly considered as words and not idiomatic phrases as idioms prototypically have transparent internal structure and two levels of interpretations, which is not the case with the bisyllabic verbal expressions under consideration. I suggest furthermore a rule of reanalysis which applies to members of the IV category and assigns a structure of [V–O]. The ionization phenomenon, as I demonstrate with the Infixal Adverbial Construction, is then a result of the superimposed bracketing and boundaries being accessible to syntactic rules such as the IAC, which seems not to discriminate the grammatical level of their input expressions. The variability in the ionization
phenomenon can be further explained by appealing to the fact that the application of the reanalysis rule for many etymologically obscure items is not uniform among speakers, but subject to factors such as the relative semantic transparency of the component zi's and stress patterns.

I have tried to demonstrate in Section 2 that the level of word is motivated to a great extent in Chinese and, in general, has the properties attributed to words, such as uninterruptability, fixed ordering, etc. However, the ionization phenomenon of certain bisyllabic verbs shows that word boundaries do not block all word–internal information from being available for syntactic purposes, as maintained by the Lexicalist Hypothesis. Specifically, information regarding internal bracketing and boundaries (imposed by a rule of reanalysis and not by any derivational rules in the grammar) is accessible to certain syntactic rules. (There are no syntactic rules, however, which can be shown to refer to lexical information such as the specific phonological shapes of parts of words.)

In a phrase structure theory like GPSG, where syntactic rules take the form of well-formedness conditions on the distributions of features, the only information sharing permitted between the components of syntax and morphology is through the medium of morphosyntactic features. The data presented in this chapter, which show that internal boundary information is available to syntax, would be very difficult to accommodate in such a theory, since lexical information such as word–internal
boundaries, unlike other grammatical properties of *words*, are not encodable as morphosyntactic features. The data here calls for an implementation of grammar such as that of Categorial Grammar, where rules can refer directly to expressions instead of only to categorial and other morphosyntactic features, as is assumed in a strictly modular theory such as GPSG. An analysis of these data within the Categorial framework will be given in the following chapter.
CHAPTER III

A CATEGORIAL–PROCESSUAL APPROACH TO CHINESE SYNTAX

3.0. Introduction.

In this chapter, I discuss certain constructions from Chinese to motivate what can be called the Categorial–Processual approach to syntax, which has been adopted by Montague in his PTQ and has been also argued for by authors such as Schmerling 1983a, Carlson 1983 and Hoeksema and Janda 1988. Whereas Classical Categorial Grammar (e.g., Bar–Hillel 1964, Ajdukiewicz 1935) restricts all syntactic rules to two-place rules in the form of functional application and permits only concatenation as an operation (hence having the generative power of a context–free phrase structure grammar), the Categorial–Processual approach can be considered an extension of Categorial Grammar in that it admits of certain non–concatenative combinatory methods and in that it allows one–place rules in addition to the two–place functional application (and functional compositions that are now widely–accepted). Such an approach to syntax is along the line of what C. Hockett called the 'Item and Process' grammatical model (Hockett 1954), in contrast to the 'Item and Arrangement' model which characterizes most modern syntactic theories. The goal of this chapter is to show that an extended Categorial Grammar incorporating (1) one–place rules and (2)
non-concatenative operations, such as wrapping and reduplication, provides the
descriptive tools necessary for treating certain phenomena in Chinese syntax of a quite
genuinely 'processual' nature and thus provides a descriptive model for the language
which is more viable than phrase structure grammar–based approaches.

In Section 1, the background of the IP approach to syntax will be briefly
discussed. In Section 2, I will present an account of the 'Infixal Adverbial
Construction' and argue that the phenomenon of ionization exhibited in such
construction is a genuine phenomenon of syntactic infixing, the description of which
requires wrapping. Certain word order principles in the language will be offered as an
explanation for the use of wrapping (as opposed to concatenation). Additional data
involving the 'Infixal Argument Construction', another construction exhibiting the
ionization effect, will be discussed in Section 3 to support syntactic infixing in the
language. Finally, in Section 4, I will give an account of a certain syntactic
reduplication phenomenon in Chinese, namely the 'Reduplicative Interrogative
Constructions', and will show how the account can be formalized within the present
framework.


The dominant grammatical model of the post–Bloomfieldian generative
tradition has been based on the Item–and–Arrangement (IA) model as developed by
the American structuralists. Here a grammatical analysis is concerned with how
linguistic expressions are to be sliced up into their parts. In other words, a grammatical analysis consists of the segmenting and classification of surface morphological materials and determining immediate constituency or configurational relationships among the parts. Grammatical rules are exclusively in the form of specifications of (left-to-right) combinatory possibilities of expressions (of particular categories). Meanings, either lexical or grammatical, are therefore directly associated with or assigned to particular continuous surface segments. For example, in an IA-based morphological description, the affix \textit{-ed} in English is regarded as directly associated with the grammatical meaning of Past Tense. Such is the underlying assumption behind the variety of syntactic theories based on the X-\textit{Bar} theory, e.g., the Generalized Phrase Structure Grammar.

The 'Item and Process' model, on the other hand, adopts a fundamentally different view of what a grammatical analysis ought to be. One is concerned with how expressions or categories of expressions of a language relate to other expressions via formal operations (Schmerling 1983a, Hoeksema and Janda 1988). This implies that we recognize syncategorematic elements, which are introduced by operations associated with rules and can be viewed as reflections of the applications of the rules, as opposed to lexical or categorial expressions. For example, within this approach, the past tense \textit{-ed} is a syncategorematic element which may be seen as reflecting the formal operation associated with the past tense rule in English. Here, in addition to the affixation of morphological materials, we also expect to find other types of formal
operations that can be used to derive one expression from another or to combine
several expressions into a larger one.

Everything being equal, the IA model is much more restricted in its descriptive
power and is therefore preferable on meta-theoretic grounds. The assumption in the
adoption of an IP model is therefore that we can find phenomena in natural languages
which go beyond the descriptive power of the former. In morphology, when we are
dealing with affixal morphology, these two approaches are equal in their descriptive
adequacy. The advantage of an IP approach can only be seen when we encounter what
is called process morphology, where certain functional meanings are associated with
non-concatenative processes such as infixation, reduplication, internal modification,
subtraction, etc., and therefore can not be attributed to particular surface segments
added to the base. This means that such data cannot be treated with context-free
rewriting rules of the IA model (cf. Selkirk 1982). Thanks to the widely-attested data
of process morphology, the IP approach to morphology has been embraced by writers
such as Matthews 1965 & 1972, Zwicky 1985, Anderson 1982, Janda 1983, and
Hoeksema and Janda 1988, among others.

3.1.2. IP Approach to Syntax.

In syntax, the IA model, which context-free phrase structure grammar (e.g., the
X-Bar variety of syntactic theory) is one formalization of, has clearly been the
dominant model. Although it was adopted by linguists working on native American
languages (e.g., F. Boas and E. Sapir) and has been the form of syntax used by logicians for symbolic languages, it was not until Montague's 'Universal grammar' (Montague 1970), which can be regarded as a formalization of the IP approach to syntax, that the IP model was taken seriously as a possible framework for syntax by theoretical linguists (Carlson 1983, Schmerling 1983a, & Hocksema & Janda 1988).

In UG, a grammar is defined essentially as the following:

(1)  
1. a set of expressions indexed by a set of category indices;  
2. a set of n-place operations;  
3. a set of syntactic rules, formalized as a triple:  
\(<F;c_1;c_2;...;c_k;c_{k+1}>\) where F is a k-place operation and \(c_1, ..., c_k\) are input categories and \(c_{k+1}\) is the syntactic category of the resulting expression.

The syntactic rule is interpreted in this way: if expressions \(x_1, ..., x_k\) are of category \(c_1, ..., c_k\) respectively, a k-place function F can be applied to \(x_1, ..., x_k\) to produce a new expression of category \(c_{k+1}\). In other words, a rule consists of specifications of input categories, an output category, and the operation to be applied to the input expressions. Here I follow Bach 1979 in assuming that syntactic operations are defined over labeled bracketed strings (see also Partee 1979). (This contrasts with the more restricted Head Grammar proposed in Pollard 1984, where input to head-wrapping operations are headed strings, the internal structure of which is not accessible to the operations.)
3.1.3. A Categorial–Processual Approach to Syntax.

Classical Bidirectional Categorial Grammar, although clearly based on the IA model, can be considered a formalization of the above rule schema, with the restrictions that $k=2$ and that all rules are to be semantically interpreted as functional application; furthermore, the only operations allowed are two-place directional concatenations, which are encoded in the functor categories.

Recent advances made in Categorial Grammar take two fundamentally different directions. In the work of Steedman and others (Ades and Steedman 1982, Steedman 1985, Oehrle, 1988, among others), combinators other than functional application are proposed, most notably functional composition, which has been shown to deal efficiently with, among others, long distance dependency constructions. Another way that Categorial Grammar has been generalized is exemplified in Montague’s own formalization of Categorial Grammar in PTQ, which is constructed within the general framework of UG. Rule schemas other than functional application are allowed, such as the substitution rules for treating quantification; furthermore, operations associated with rules may go beyond concatenation. A generalization of Categorial Grammar along this line has also been adopted in the work of Bach (1979, 1983b, 1984, and 1988), Dowty (1978, 1979, 1982a), Schmerling (1983a, 1983b), etc. I will call this the Categorial–Processual approach\(^1\) and will, in the following, present data from Chinese

\(^1\)Note here that the adoption of the IP model also makes a difference as to how syntactic rules are supposed to work. While Classical Categorial Grammars are primarily conceived as applying in a top-down parsing manner, the rules in an IP (continued...)
attesting to syntactic infixations (Sections 2 and 3) and reduplications (Section 4), both of which are problematic for the phrase structure theory of syntax or for the Classical variety of Categorial Grammar. What I will try to motivate is specifically (1) non-concatenative operations, such as wrapping and reduplication, and (2) one-place rules which involve formal operations on the input categories, within the Categorial-Processual approach.

\[ \ldots \text{(continued)} \]

formalization, such as those in Montague Grammar, are primarily conceived as applying in bottom-up generative manner, that is, as instructions for building up well-formed sentences of a language (see Karttunen and Zwicky 1985 for more discussion on the various interpretations of grammatical rules.). Take, for example, two expressions \( \alpha \) and \( \beta \) (of the categories of A/B and B, respectively) which occur adjacent to each other. While in the former approach, a partial analysis involving the sequence \( \alpha \beta \) may take the form of (a), the analysis tree involving the same sequence in the latter approach would look like (b):

\[ \begin{array}{c}
\alpha \\
\beta \\
\hline
A/B \\
B \\
\hline
A
\end{array} \]

\[ \begin{array}{c}
\alpha \,
\beta , A \\
/ \\
\alpha , A/B \quad \beta , B
\end{array} \]

These two conceptualizations of syntactic rules (and the different ways to represent an analysis tree) obviously make no difference if we allow only concatenations as a combinatorial device. Only the latter enables us, however, to accommodate and express operations other than concatenations.
3.2.1. **Wrapping (Infixing) Operations in Syntax.**

The best known non-concatenative operation proposed for natural language syntax within Categorial grammar\(^2\) is what has been called wrapping (Bach 1979, 1980, 1984, Dowty 1982a, Pollard 1984). It is proposed as an alternative combinatory method whereby a string of expression 'wraps around', instead of being concatenated with, another expression. This operation is primarily motivated by the discontinuous constituent (or so-called 'bounded extraction') phenomenon and has been used to treat complex TVP, e.g., *persuade to leave, & consider as foolish*, in English\(^4\); the VSO order in languages such as Irish and Breton (see Dowty 1982a); particle movement in English (see Jacobson 1987); subject–aux inversion; cross-serial dependency phenomenon in subordinate clauses in Dutch (Pollard 1984); and possibly the well-known second position (Wackernagel) phenomenon, where one finds that elements such as pronominal subject, mood and tense markers, etc., appear invariably in the second position in languages (see Hoeksema and Janda 1988). Attempts have been made in recent years to analyze data involving discontinuous constituents in a strictly concatenative syntax. For example, Zwicky 1986b proposed the Liberation Metarule, which is essentially used to flatten a structure so as to enable linearization of non-siblings (also see Uszkoreit 1987).

\(^2\)Pollard's Head Grammar (1984), which is in the phrase structure framework, also incorporates wrapping.

\(^3\)Jackobson 1987's terminology.

\(^4\)See more discussions on TVP's at Chapter 5.
Since operations such as wrapping take the power of the grammar beyond context-free grammar (see Bach 1988, Pollard 1984) while bidirectional Categorial Grammar is equivalent in its descriptive power to context-free phrase structure grammar, it would in principle be desirable to do away with them whenever possible. What I attempt to demonstrate in the following (Sections 3 and 4) is that there are data which may not be subject to alternative analyses, as they involve infixing into lexical categories.

3.2.2. The Infixal Adverbial Construction.

Consider first (2)–(4), which show that when an IV combines with a durative or frequentative adverbial, the latter is ordered after the IV.

(2) ta zou le yi ge zhongtou le
    he leave ASP one CL hour ASP
    'He was gone for an hour.'

(3) ta lai le ban nian le
    he come ASP half year AS
    'He has been here for half a year.'

Both VO and OV are possible word orders within an VP in the language, and

---

Jacobson 1987, for example, reanalyzes the primary data motivating wrapping in a phrase structure grammar augmented with features. A verb raising process is proposed to deal with discontinuous constituents such as TVP. The verb raising rule 'promotes' the verb in the TVP, and the discontinuity of the complex TVP is seen as a consequence of independently motivated word order principles in English.

VO is, however, the unmarked order. OV order is possible under certain discourse conditions and conditions involving the thematic role associated with the
when the adverbial modifies a complex IV (=VP) consisting of an NP followed by the
verb, it follows the verb (e.g., (4)-(5)).

(4) ta ba wo ma le bantian
   he BA me scold ASP half-day
   'He yelled at me for a while.'

(5) ta xin xie le liang ge zhongtou cai xiewan
    he letter write ASP two CL hour then write-finish
    'It took him two hours to write the letter.'

When the VP consists of a verb followed by an NP, the adverbials (e.g., yizhengtian
('the whole day') in (6), bantian ('half day = for a while') in (7), and liang ge zhongtou
('for two hours') in (8)) end up in a slot intervening between the V and the NP
argument; meanwhile a particle, de, is optionally affixed to the adverbial expression.

Note that the adverbials in question cannot appear at the end of the VP (e.g., (b)
sentences in (6)-(8)).

(6)a. wo kan le yizhengtian (de) dianshi
       I watch ASP the-whole-day DE TV

"(...continued)

object NP, the nature of which has yet to be clarified. For a recent discussion on this,
see Sun & Givon 1985.

Note that I am omitting data involving idiomatic phrases (e.g. chi cu ('to eat
vinegar=to be jealous') even though they can also be distributed in this construction.
The reason is that an analysis including such data would require a substantive theory
for idiomatic phrasal expressions in Chinese, which are characteristically more mobile
and less frozen than those in many other languages and the nature of which awaits
future research. (But see C. Huang 1988a for a possible treatment for idiomatic
phrases in a similar construction, i.e., the 'Possessive Object Construction' (or 'Inserted
Argument Construction'), which will be discussed later in the chapter.)
'I watched TV for the whole day.'

b.* wo kan le dianshi yizhengtian
I watch ASP TV the-whole-day
'I watched TV for the whole day.'

(7)a. wo deng le bantian (de) huoche
I wait ASP half-day DE train
'I waited for the train for a long time.'

b.* wo deng le huoche bantian
I wait ASP train half-day
'I waited for the train for a long time.'

(8)a. wo du le liang ge zhongtou (de) baozhi
I read ASP two CL hour DE newspaper
'I was reading newspapers for two hours.'

b.* wo du le baozhi liang ge zhongtou
I read ASP newspaper two CL hour
'I was reading newspapers for two hours.'

In Categorial Grammar, adverbial expressions denote functors which apply to properties denoted by VPs to give properties denoted by VPs and therefore are assigned the category of VP/VP or VP\VP (with slashes encoding the relative linear order of VPs and adverbials in the language). Such a category assignment for adverbials presumes that optional modifiers typically occur external to the verb and its obligatory arguments in a VP (i.e., simple concatenations (RCON or LCON) are used in combining VP and VP modifiers), as is the case in English and many other languages. But we should also expect to find languages where optional modifiers intervene between a verb and its obligatory argument, as is the case in Chinese (e.g., (6)-(8)). Such a situation, where we need to linearize daughters of a VP and a VP adjunct, is potentially problematic for Categorial Grammar, since here only sibling
categories can be ordered (via slash directions in the functor category in a directional system).

One way to account for the data in question without invoking wrapping is to consider these temporal adverbials as modifying TV's, rather than VP's (or IV's), which would make it unnecessary to have an account of discontinuous constituents. One variant of this approach is to appeal to a schema of generalized type raising within the Categorial framework which has been referred to as Geach's rule (as in (9)) (and was originally proposed in Geach 1972 to account for cross–categorial expressions such as negations (see van Bentham 1988)). A type raising rule converting VP/VP to TV/TV can then be seen as an instance of the general schema. 'Type–raising' the adverbial expressions as TV/TV would enable us to combine it with the verb first (and then with the object NP), such that the Principle of Adjacency (Ades and Steedman 1982) is adhered to.

(9) Geach's rule:
\[ X/Y \Rightarrow (X/Z)/(Y/Z) \] for arbitrary Z.

(10) VP/VP \Rightarrow (VP/NP)/(VP/NP)
semantics: \[ f \Rightarrow \lambda g \lambda x[f(g(x))] \].

Crucially, one finds data like (11a)–(13a) (which exist side by side with the expected (11b)–(13b)), where the adverbial phrase occurs to the right of the IV (as in (2)–(3)), in which the frequentative or durative adverbial is inserted inside an
intransitive verb. For these data, where the adverbial is infixed into a lexical category, in particular, into a slot between the two zi's in the verbs (such as shang#dang ('be fooled') in (11a), sheng#qi ('be mad') in (12a), and ju#gong ('to bow') in (13a)), the alternative analyses just mentioned, which obey the Adjacency Principle, are obviously not available.

(11)a. wo shang guo yici (de) *dang
    I ASP once DE 'I was fooled once.'

cf. b. wo shang#*dang guo yici
    I be-fooled ASP once 'I was fooled once.'

(12)a. ta sheng le haojiu (de) qi
    he ASP long DE 'He was angry for a while.'

cf. b. ta sheng#qi le haojiu
    he angry ASP long 'He was mad for a while.'

(13)a. ta zai nar *ju le bantian (de) *gong
    he at there ASP half-day DE 'He was there bowing for a long time.'

cf. b. ta zai nar ju#gong le bantian
    he at there bow ASP half-day 'He was there bowing for a long time.'

These data ((11a)–(13a)), where the discontinuous VP is a lexical IV, show that some kind of infixing device is indeed essential and that natural languages do employ wrapping as an alternative means to combine expressions, resulting in discontinuous

---

8See Chapter II for discussion on the grammatical status of zi's in the language.
constituents and sometimes even in the disjoining of expressions which are members of a lexical category.

3.2.3. Formalizing Wrapping.

The operation of wrapping has been proposed in the literature as a possible operation other than concatenation that syntactic rules invoke to combine expressions. There are in general two kinds of wrapping operations which have been proposed in the literature: Head Wrap, where a string is adjoined either to the right or the left of the head of another string, and a wrapping which makes reference to the margin of a constituent. The earliest formalization of wrapping was proposed by Bach (1979:516) and was called RWRAP.

(14) RWRAP:

If $\alpha$ has the form $[X W]_\rho$, then $\text{RWRAP}(\alpha, \beta)^9 = X^*\beta^*W$.

Note that this definition of wrapping presupposes some version of the X–Bar system of categories and identifies $X$ as the head of a complex phrase $XP$. The way RWRAP is formulated here, it is applicable only when the head is the left–most element of the functor expression. A functor expression such as this where an argument expression is inserted after its head is called 'circumfix' in morphology (Hoeksema & Janda 1988). It is also possible that an argument expression is wrapped

---

$^9$By convention, $\alpha$ designates the functor expression and $\beta$ designates the argument expression.
around the functor expression (i.e., the functor is infixed into the argument), e.g., in Bach 1984's treatment of the verb-raising structures in Dutch and in a possible wrapping treatment of second position clitics in general. Such wrapping can be called Right Infixing (RINFIX)—short for 'infixing to the right of head'. This can be defined as in (15); again, the way it is defined here, it is applicable only when the head in the argument expression is the right-peripheral element.

(15) RINFIX:
If $\beta$ has the form $[X \ \bar{W}]_\alpha$, then $\text{RINFIX}(\alpha, \beta) = X^\alpha \bar{W}$.

We can see that both RWRAP and RINFIX essentially allow the adjoining of something external immediately to the right of the head of a complex expression. As already mentioned, the formalization here relies on some version of the X-Bar system of categories such that the head of a complex phrase XP is the category X. In the literature of Categorial Grammar, however, the notion of head is defined derivatively in terms of the notions of functor and argument categories (cf. Vennemann and Harlow 1977, Hoeksema 1985, Hoeksema and Janda 1988, and see Dowty 1989 for some discussion on this):

(16) In $[A_\lambda B]_\alpha$ (or $[B A_\lambda B]_\alpha$), B is the head if $A = B$, $A_\lambda B$ is the head otherwise.

It is possible to reformalize RWRAP and RINFIX in a way that builds in this (theory–internal) definition of head. One such formalization of RINFIX and RWRAP
is (17) below. Both wrapping operations as defined in (17) suffix some external element to the right of the head of a complex expression, but unlike Bach 1979’s definition, (17) covers both the right-edge location and the left-edge location of the head:

(17)

**RINFIX:**

(i) If $\beta$ has the form $[A;B]_{A'}$, and if $A = B$, $\text{RINFIX}(\alpha; \beta) = \text{LCON} (\alpha; \beta) = \beta^* \alpha$; otherwise, $\text{RINFIX}(\alpha; \beta) = A;B \times \alpha^* B$.

(ii) if $\beta$ has the form $[A A;B]_{A'}$, and if $A = B$, $\text{RINFIX}(\alpha; \beta) = A \times \alpha^* A;B$; otherwise, $\text{RINFIX}(\alpha; \beta) = \text{LCON} (\alpha; \beta) = \beta^* \alpha$.

**RWRAP:**

(i) If $\alpha$ has the form $[A;B B]_{A'}$, and if $A = B$, $\text{RINFIX}(\alpha; \beta) = \text{RCON} (\alpha; \beta) = \alpha^* \beta$; otherwise, $\text{RINFIX}(\alpha; \beta) = A;B \times \beta^* B$.

(ii) if $\alpha$ has the form $[A A;B]_{A'}$, and if $A = B$, $\text{RINFIX}(\alpha; \beta) = A \times \beta^* A;B$; otherwise, $\text{RINFIX}(\alpha; \beta) = \text{RCON} (\alpha; \beta) = \alpha^* \beta$.

(As we can see, some instances of wrapping operations defined here do not result in discontinuous constituents.)

On the other hand, data found across languages show that sometimes reference to margins of expression, instead of to the head, is made by syntactic rules (see Hoeksema and Janda 1987). To allow such reference, Bach 1984 introduced the one-place analytic operation\textsuperscript{10} **FIRST** ((18)) to help define the right-edge-anchored (as

\textsuperscript{10} which corresponds to certain functions in LISP (Bach 1984:260)).
opposed to head–anchored) wrapping, which I will call RWRAP' (to be distinguished from RWRAP which, as defined above, is a head–wrap)\textsuperscript{11}. RWRAP' as defined in (19) allows the insertion of an argument expression to the right of the left–margin element of a complex phrase:

(18) Let \( x \) be the strings \( x_1, \ldots, x_n \)

\[
\begin{align*}
\text{FIRST}(x) &= x_1 \\
\text{RREST}(x) &= x_2, \ldots, x_n.
\end{align*}
\]

(19) \( \text{RWRAP}'(x,y) = \text{FIRST}(x) \rightarrow y \rightarrow \text{RREST}(x) \).

We can likewise formalize the infixing of a functor expression after the first element in the argument expression. Let's call this RINFIX', which, among others, can be used to describe Wackemagel or second position clitics found in languages.

(20)\( \text{RINFIX}'(x,y) = \text{FIRST}(y) \rightarrow x \rightarrow \text{RREST}(y) \)

3.3.4. An Account of IAC.

We can give (2)–(11), which illustrate the positioning of (durative and frequentative) adjuncts within a VP, a unified account by invoking some kind of wrapping operation in the rule. The questions here are whether a head–anchored (RINFIX) or a (left–)edge–anchored wrapping operation (RINFIX') is invoked by the

\textsuperscript{11}Note since head is often the peripheral element in a constituent (which is trivially true in binary structures), cases involving RWRAP and RINFIX can often be reanalyzed as cases of wrapping which makes reference to the margin of a constituent.
IAC rule, or whether these two are descriptively equal as far as these data are concerned. Consider (6)–(8) first, where one finds a time adverbial is adjoined to the right of the verb in a VP. Since the head (according to the above definition) in the VP is also the left–most element, it seems that both kinds of wrapping can describe the data equally. However, data in (4)–(5), where the internal order in the VP is that of OV, show that reference to the head element is necessary here, since for these VP's, the adverbial is adjoined to the right of the head verb. More of such data are given below.

(21)a. ta riwen xue le san nian le
   he Japanese learn ASP three year ASP
   'He has been learning Japanese for three years.'
   
   b.* ta riwen san nian xue le
   he Japanese three year learn ASP
   'He has been learning Japanese for three years.'

(22)a. wo yao chi le bantian le
   I medicine eat ASP half-day ASP
   tou haishi hen tong
   head still very hurt
   'I have taken the medicine for a while, but I still have a bad headache.
   
   b.* wo yao bantian chi le,...
   I medicine half-day eat ASP
   'I have taken the medicine for a while,...'

These data, where the head verb is located at the right edge of the VP, suggest that reference to the head in the wrapping operation responsible for these data is necessary. Here, the adverbial expression is adjoined to the right of the head verb, resulting in no discontinuity at surface. If the rule were to invoke an edge–anchored wrapping (such
as RINFIX' in (20)) and to insert accordingly the adverbial to the right of the left-marg in element in the VP, which is the NP, we would get the ill-formed (21b) and (22b).

Now, before we state the rule for the IAC, a note should be made of the particle *de* affixed to the inserted adverbial expression in the IAC, which is the single most prominent grammatical particle in Chinese (see C. Huang 1989). As we will see throughout this dissertation, this particle shows up in many different grammatical constructions in the language. In the approach here, we can view the operation of affixing *de* as a simplex operation in the language, which, as we will see, is part of various composite operations associated with syntactic rules in the language. Since *de* is invariably suffixal, i.e., forming a unit with the preceding expression, we can define the operation in question as:

\[(23) de\text{-Suffixation:}\]
\[
F_s(\alpha)=\alpha^*de
\]

The idea that operations associated with rules can be complex, i.e., composed of more primitive operations, is first mentioned in Partee 1976b and 1979.\(^{12}\) Zwicky 1988a's notion of 'independent mark' is essentially the same idea: certain grammatical marking can be accessed by different rules either as a simple operation or as part of a complex

\(^{12}\)A precise definition of complex operations invoked by syntactic rules is provided by Montague (1974), which he calls 'polynomial operations'.

operation associated with a rule.

Having said that, we can now state the rule of IAC covering (2)–(11) as (24), which is in the form of a triple consisting of specifications of the operation, the input category(ies) and the output category:

(24) The rule for IAC:
\[ F_{13}, \langle \text{VP}^{11}, \text{VP} \rangle; \text{VP}. \]
\[
F_{13}(\alpha, \beta) = F_3(F_2(\alpha, \beta)
\text{semantics: } \alpha'(\beta')
\]
\[(F_3 = \text{RINFIX (defined in (17))}).\]

(24) says that when a durative or frequentative adverbial expression is combined with a VP, the adverbial is suffixed with the particle *de* and is positioned to the right of the head (verb). This accounts directly for (6)–(8), where the VP is verb-initial, as well as (4)–(5) and (21a)–(22a), where the VP is verb-final (and therefore resulting in no discontinuity at surface). For (24) to cover cases such as (2)–(3), where the VP consists of a simple IV, we need to be able to identify a simple IV as the head of a verb phrase (therefore accounting for the adjunction of the adverbial to

\[\footnote{Note that since linearization is achieved through operations associated with rules (and not viewed as properties of the functor categories) under the present approach, we will use a system of non-directional categories where slashes no longer encode word order.}\]
the right of the IV). One way to achieve this (and keep the definition of head intact) is to add a clause to the definitions of head–wrapping operations. For example, for the definition of RWRAP in (14), we can add a clause that says: 'if a is simple, then RWRAP (a,b) = RCON (a,b)' (see Bach 1979). For RINFIX, which is what is referred to in (24), we can similarly add a clause which says: 'if b is simple, RINFIX (a,b) = LCON(a,b)'. By adding this stipulation, we account for (2)–(3), as well as the (b) sentences in (11)–(13), where the adverbial is adjoined to the right of the intransitive verbs such as shang#dang ('be fooled'), sheng#qi ('be angry') and ju#gong ('to bow').

What remains to be explained is the sentences exhibiting infixed into a lexical category in (11a)–(13a), which are what crucially motivate a wrapping analysis for the IAC. The explanation is readily available if we appeal to the rule of reanalysis, which, as discussed in Chapter 2, supplies internal (pseudo)structure (of [V NP]) for a large set of synchronically opaque intransitive bisyllabic verbs. (11a)–(13a), which are in variation with the expected (11b)–(13b), can then be seen as results of such internal pseudo–bracketing available to the rule in (24) (or more precisely, to the operation of RINFIX invoked in the rule). The rule treats such reanalyzed verbs as if they were truly syntactically complex and inserts the adverbial to the right of the 'verb'.

\footnote{The situation here is similar to the case of TVP's; both TVP's (such as persuade \emph{to leave}) and simple TV's can be both treated as conjunction of the object NP to the right of the head.}

\footnote{As already discussed in Chapter II, the rule of reanalysis only applies to bisyllabic verbal expressions which are IV's. This explains why the following sentences involving TV's (de–zui ('obtain–guilt = to offend') in (b) and ti=\#yi ('carry–') (continued...)}
word, the 'ionization' effect exhibited in (11a)–(13a) is a consequence of (1) the internal bracketing information in a subgroup of bisyllabic verbs (superimposed by the reanalysis rule) being transparent to the syntactic rule in (24) and of (2) the rule treating these pseudo-structures in the reanalyzed verbs and true structures generated by syntactic rules alike.

Finally, note that the coexistence of the (b) sentences in (11)–(13) and the (a) sentences in (11)–(13) shows that somehow it is possible for the rule in (24) to treat the bisyllabic verbal expressions in question both as simple (i.e., with no internal structure), resulting in the former sentences, and as complex (with an imposed 'structure' of [V NP]), resulting in the latter sentences. This state of affairs could be more easily explained if we were to find data like (11b)–(13b) only in children's speech (i.e., before the bisyllabic verbs are reanalyzed) and (11a)–(13a) only in adult speakers' speech (i.e., after the verbs are reanalyzed and given some 'structure'). Unfortunately this is not the case: (a) and (b) sentences in (11)–(13) can be found in the same speaker's speech, which suggests that even though the verbs have been

15(...continued)
proposal='to propose') in (a)) are ill-formed; they do not have the internal structure which the rule of IAC needs to refer to.

a.* Wo ti le bantian yi women xianzai jiu likai
I ASP half-day we now right-away leave
'I was proposing for a while that we set out right
away.'

b.* wo de le bantian zui ta
I ASP half-day him
'I offended him for a while.'
reanalyzed and have become 'complex', it is nonetheless possible for the syntactic rule to somewhat 'overlook' their internal 'structure' and treat them as simple.

3.2.5. Word Order Principle as a Possible Motivation for Wrapping in IAC.

It is reasonable to speculate that the use of wrapping as a combinatory method is motivated by language-specific word order principles. In English, for example, the use of wrapping in the combination of the complex transitive verb phrases (e.g., consider a fool, persuade to leave) with the object NP (hence the discontinuity of the TVP) can be attributed to a parochial ordering constraint in English which requires a direct object NP to be ordered immediately to the right of the verbal head, which is part of the more general word order principle in English: the more oblique an argument is, the further to the right of the verb it is (see Sag 1987). There is no such ordering constraint in, say, Dutch; therefore similar TVP constructions in Dutch are continuous (Hoeksema 1987). In Chinese, the use of wrapping, or more precisely RINFIX, in the IAC can be similarly seen as motivated by a parochial ordering principle. The principle in question can be stated as follows\textsuperscript{16}:

\begin{quote}
(25) optional modifiers (or adjuncts) may not immediately follow obligatory arguments.
\end{quote}

It may be emphasized that the nature of a linear principle such as (25) in the present framework is somewhat different from that of the Linear Precedence rules in

\textsuperscript{16}Such a rule, like ordering principles in general, is best looked at as a default, which can potentially be overridden under certain conditions (see also Uszkoreit 1984).
GPSG or, for instance, the word order conventions in Flynn 1985 within the Categorial framework. In GPSG, linear relationships among categories are abstracted away from the immediate dominance relationship and stated as separate rules, which hold across any dominating categories. On the other hand, in Classical Categorial Grammar, word order is deemed as a property of functor categories and encoded accordingly in functor categories (via direction of slashes); inherent in the category label A/B is the linear principle, 'A/B precedes B', which holds with respect to the dominating category A. In certain varieties of Categorial Grammar, for example Flynn 1985, word order is abstracted away from functor categories and captured in cross-categorial word order conventions, somewhat in the spirit of an ID/LP approach (also see Bach 1988). In the Categorial-Processual approach taken here, linearization is achieved via the operations associated with the rules (e.g., RCON, LCON, RWRAP, RINFIX, etc.) and the ordering principle such as (25), unlike the LP rules, does not directly determine how two categories are to be linearized. Instead, (25) is to be seen as some kind of metagrammatical word order principle which is 'built into' various rules in the language combining the categories in question, to the effect that output of various rules responsible for placements of adjuncts within VP's all conform to or obey the linear principle in question. I believe the existence of such a word order principle is what underlies the employment of wrapping, and, as to be seen later, reduplication, in certain syntactic rules responsible for the placement of adjuncts within VP's. To support this, we will in the following consider more data involving the positioning of adjunct expressions within VP's.
In Chinese each class of adverbial or adjunct expressions occurs in a fixed position relative to the verb. In the default cases\textsuperscript{17}, each class of adverbial expressions is obligatorily ordered either to the right or to the left of the verb within the VP.\textsuperscript{18} For example, manner adverbials, instrumental PP's, and purpose phrases obligatorily occur in preverbal positions (e.g., (26)-(29)).

(26)a. ta manmande zou guolai
    he slowly walk over
    'He slowly walked over.'

    b.* ta zou guolai manmande
    he walk over slowly
    'He walked over slowly.'

(27)a. ta yong kuaizi zai chi huasheng
    he with chopsticks PROGRESSIVE eat peanuts
    'He is eating peanuts with chopsticks.'

    b.* ta zai chi huasheng yong kuaizi
    he PROGRESSIVE eat peanuts with chopsticks
    'He is eating peanuts with chopsticks.'

(28)a. wo ti ta ba fangzi mai le
    I for him OBJ house sell ASP
    'I sold the house on his behalf.'

    b.* wo ba fangzi mai le ti ta
    I OBJ house sell ASP for him
    'I sold the house on his behalf.'

\textsuperscript{17} Note for example, the positioning of certain locational PP's, particularly a locational phrase headed by the preposition zai ('in'), is conditioned by certain semantic factors; namely, the pre-verbal zai–phrase 'locates the event' while the post-verbal zai–phrase 'locates the individual' (see Tai 1975 for discussion on this).

\textsuperscript{18} To compare, in English, the placement of adjunct expressions is typically more flexible and often their semantic contribution depends on their positioning in a sentence, e.g., the contrast between \textit{Kim jogged reluctantly twice a day} vs. \textit{Kim jogged twice a day reluctantly} (see Pollard and Sag 1987).
In the approach here, this situation is described by the operation of RCON associated with the syntactic rule combining the adverbials with VP's, provided we can use features such as [MANNER] and [INSTRUMENTAL] to distinguish different classes of VP!VP. Here, all output conforms to the linear principle in (25). Adverbial expressions such as duratives and frequentatives, on the other hand, must occur to the right of the verb, as shown in (2)–(8). Now the use of (head) wrapping in the rule combining such adjuncts with a VP consisting of a verb followed by an NP is explained: concatenation will potentially result in the prohibited ordering of obligatory NP arguments to the left of VP adjuncts within a VP (e.g., the ill-formed (6b)–(8b)). And as we have seen, for all the data involving durative and frequentative adverbials, a unified rule (24) appealing to RINFIX (17) would correctly place these adjuncts in all cases.

The appeal to the linear principle in (25) as an underlying motivation for the invocation of wrapping in the rule for combining (right–located) adverbials with VP's is supported by data in the following involving reduplication. Sentences (30)–(33) are synonymous with (6a)–(8a), respectively.
(30) wo kan dianshi kan le yi zheng tian
    I watch TV watch ASP one whole day
    'I watched TV for the whole day.'

(31) wo deng huoche deng le bantian
    I wait train wait ASP half-day
    'I waited for the train for a long time.'

(32) wo du baozhi du le liang ge zhongtou
    I read newspaper read ASP two CL hour
    'I was reading newspapers for two hours.'

Data such as (30)–(32) show that, interestingly, wrapping is not the only combinatory mode which results in avoiding the prohibited linear sequence in question. These data show that a reduplication operation, which achieves the same effect, can also be invoked by the syntactic rule combining a VP and a durative/frequentative adverbial.

The operation in question copies the head verb and infixes it between the NP argument and the adverbial, such that the output also conforms to the ordering principle against the left-to-right adjacency of an obligatory argument and an optional adjunct. As the mechanism for describing reduplication is discussed later in this chapter, here I will omit spelling out the operation in question, except to note that it is in alternation with F₁₃ in Rule (24). To be further noted is that, as we probably expect on the basis of earlier discussion on the group of bisyllabic verbs which have an imposed internal structure through a rule of reanalysis, this alternate operation associated with the rule combining VP's with durative/frequentative adverbials also applies to these bisyllabic verbs.
Additional data would show that in fact this is not the only construction where the use of reduplication prevents the forbidden linear sequence stated in (25). This is what is called the 'Descriptive Complement Construction' (cf. J. Huang 1988b), where a verb is followed by the particle de and a predicative complement which has the function of an adverbial.

Under the present proposal, we would expect that sentences involving the Descriptive Complement construction, such as (38a) and (39a), where a TV is followed by a NP, would be ill-formed, since they contain the prohibited linear sequence. This is indeed the case. Again, we find that reduplication is used to 'break up' the potential prohibited sequence and the resulting (40)–(41) are well-formed.
(38a.) ta qi ma de kuai
    he ride horse DE fast
    'He rides horses fast.'

b. ta ma qi de kuai
    he horse rise DE fast
    'He rides horses fast.'

(39a.) ta xie xin de man
    he write letters DE slow(ly)
    'He writes letters slowly.'

b. ta xin xie de man
    he letters writes DE slow(ly)
    'He writes letters slowly.'

(40) ta qi ma qi de kuai
    he ride horse ride DE fast
    'He rides horses fast.'

(41) ta xie xin xie de man
    he writes letters write DE slow(ly)
    'He writes letters slowly.'

Finally, I will briefly comment on the Phrase Structure Condition (PSC)
proposed in J. Huang 1984 (see also Chapter 2), which appears to be a filter of some
sort that rules out sentences in violation of it. The PSC is motivated by the same
range of data as the word order principle in (25) (which I take to be a meta-
grammatical linear principle to which output of all VP-producing rules must conform).
Specifically, data like (6)–(8) and (30)–(32), as well as (38)–(41), are among the data
Huang uses to motivate the PSC, which, to reiterate, states that 'the verb or VP may
branch to the left only once, and only on the lowest level of expansion', which is to
mean 'a verb may be followed by at most one constituent'. According to Huang, (6b)
(repeated below), for example, is ill-formed because here the PSC is violated, as the
verb is followed by more than one constituent. (6a), on the other hand, conforms to the PSC, because according to the hierarchical structure Huang assigns to (6a), the sequence *yizhentian de dianshi* ('a-whole-day DE TV') is bracketed as an NP. Such a bracketing is potentially possible because the particle *de*, which is (optionally) suffixed to the temporal adverbial in this construction, is homophonous with the possessive particle in Chinese and, as a result, the IAC construction bears a structural resemblance to the possessive NP construction (e.g., *Lisi de dianshi* ('Lisi-Poss-TV =Lisi's TV') in (42)).

(6)a. wo kan le yizhentian (de) dianshi  
I watch ASP a-whole-day DE TV  
'I watched TV for the whole day.'

b.* wo kan le dianshi yizhentian  
I watch ASP TV a-whole-day  
'I watched TV for the whole day.'

Cf. Possessive NP's:
(42) wo hua liangbai kuai mai le [Lisi de dianshi]TV  
I spend 200 dollars buy ASP Lisi POSS TV  
'I bought Lisi's TV with 200 dollars.'

As we can see, however, Huang's bracketing of the sequence *yizhentian de dianshi* ('a-whole-day DE TV') as an NP is unmotivated, first of all, on semantic grounds. Secondly, if the string in question were indeed an NP, it should be possible to topicalize it, but the following topicalized sentence is bad:

(43)* yi zheng tian de dianshi wo kan le  
one whole day DE TV I watch ASP  
'A whole day's TV I watched.'
More importantly, the presence of the particle *de*, which is what enables Huang to categorize the sequence in question as an NP, is merely optional. In sentences where *de* is absent, the construction in question no longer bears any structural similarity to the possessive NP construction, and the bracketing of a string such as *yizhengtian dianshi* ('a—whole—day TV') in (6a) as an NP becomes entirely unmotivated. Therefore, when the particle *de* is absent, the verb in (6a) is followed by two constituents (i.e., a time adverbial and an NP) — since here it is impossible to bracket the whole string (e.g., *yizhengtian dianshi* ('a—whole—day TV') in (6a)) as an NP. Thus a sentence such as (6a) without the particle *de* should be bad since it is in violation of the PSC, yet it is not.

Finally, data such as (30)–(32) and (38)–(41) which involve reduplication, are also used by Huang to support the PSC. However, within the phrase structure setting of Huang 1984 it would be very difficult to provide any motivated phrase structure for these sentences containing reduplication (such that they could be seen as conforming to the PSC).

To summarize, there have been mechanisms proposed within the framework of context—free phrase structure grammar (e.g., the Liberation metarule of Zwicky 1986b and the 'promotion' rules of Jacobson 1987) for reanalyzing the phenomenon of discontinuous constituents, which wrapping has been used to account for. In this
section, we have seen data which (1) involve infixing into lexical category and data which (2) show that what is being accounted for by wrapping is closely related to the reduplication phenomenon, in that a rule invokes alternatively reduplication and wrapping to achieve the same effect. It would be very difficult to give the first type of data an analysis within the CF phrase structure grammar (or an IA model in general including Categorial Grammar of the Steedman variety), and it is simply impossible to give such alternative analyses for any data involving reduplication. Therefore, even though it may be desirable to reanalyze some alleged wrapping phenomenon within the Phrase structure setting, I believe the data here offer strong arguments that wrapping cannot be dispensed with entirely and that natural languages do employ wrapping as a combinatory operation, on a par with processes such as reduplication.

3.3. Infixal Argument Construction.

In this section I will present another syntactic construction in the language to lend support to what I have demonstrated with the IAC, in particular, (1) the syntactic accessibility of internal bracketing and boundary information and (2) syntactic infixation (into lexical category). The construction in question is what may be called the Infixal Argument Construction (IARC)\(^{19}\), where an argument of a (bisyllabic) verb

---

\(^{19}\)This construction is also called the 'Possessive Object Construction' by C. Huang (1987 & 1988a). This is due to the fact that the particle *de*, which is suffixed to the NP argument in this construction, is homophonous with the possessive particle in Chinese. As a result, the IARC (e.g.(44b)–(50b)), just like the IAC, bears a structural resemblance to the possessive NP construction, hence the name 'Possessive Object Construction'. Note, however, the presence of the particle *de* in the IARC is only optional and can be omitted.
occurs optionally at a position internal to the verb (e.g., (44b)-(50b)) in addition to an expected external position (e.g., (44a)-(50a)). Before we examine the IARC (e.g., (44b)-(50b)), consider first the related data (44a)-(50a). Here we find (bisyllabic) verbs which are subcategorized for (optionally present) adjunct–particle–marked NP arguments\(^{20}\), which occur pre-verbally.

\[(44)\text{a. } \text{nī bū bī ti } \text{tā } \text{dàn}^\#\text{xīn} \\
\quad \text{you need-not him worry} \\
\quad '\text{You don't need to worry about him.}'\]

\[(44)\text{b. } \text{nī bū bī } \text{dàn } \text{tā } (\text{de}) \text{xīn} \quad \text{cf. } \text{dàn}^\#\text{xīn} \\
\quad \text{you need-not him DE} \\
\quad '\text{You don't need to worry about him.}' \quad '\text{to worry}'\]

\[(45)\text{a. } \text{wò dà gài hùi } \text{qu } \text{géi } \text{tā } \text{bàng}^\#\text{māng} \\
\quad \text{I likely will go for him help} \\
\quad '\text{I probably will go help him (out).}'\]

\[(45)\text{b. } \text{wò dà gài hùi } \text{qu } \text{bàng } \text{tā } (\text{de}) \text{māng} \quad \text{cf. } \text{bàng}^\#\text{māng} \\
\quad \text{I probably will go him DE} \\
\quad '\text{I probably will go help him (out).}' \quad '\text{to help}'\]

\[(46)\text{a. } \text{nī bū gāi } \text{gēn } \text{Zhāngsàn } \text{shēng}^\#\text{qi} \\
\quad \text{you not should with Zhangsan angry} \\
\quad '\text{You shouldn't be angry with Zhangsan.}'\]

\[(46)\text{b. } \text{nī bū gāi } \text{shēng } \text{Zhāngsàn } (\text{de}) \text{qi} \quad \text{cf. } \text{shēng}^\#\text{qi} \\
\quad \text{you not should Zhangsan DE} \\
\quad '\text{You shouldn't be angry with Zhangsan.}' \quad '\text{angry}'\]

\[(47)\text{a. } \text{nī bīe } \text{lāo } \text{gēn } \text{tā } \text{yòu}^\#\text{mó} \\
\quad \text{you don't always with him make-fun} \\
\quad '\text{Don't make fun of him all the time.}'\]

\[(47)\text{b. } \text{nī bīe } \text{lāo } *\text{yòu } \text{tā } (\text{de}) *\text{mó } \text{cf. } *\text{yòu}*\text{mó} \\
\quad \text{you don't always him DE} \\
\quad '\text{Don't make fun of him all the time.}' \quad '\text{to humor}'\]

\(^{20}\text{The situation is similar to verbs such as rely on, approve of, be angry with, etc in English.}\)
(48a. wo mingtian del qu jichang gen ta songxing
   I tomorrow must go airport with him send-off
   'I have to go to the airport tomorrow to see him off.'

   b. wo mingtian del qu jichang song ta (de) *xing
      I tomorrow must go airport him DE
      'I have to go the airport to see him off.'
      cf. song#*xing
      send walk
      'to see off'

(49a. wo lai ge ni *zhaoxiang
       I will for you photograph
       'Let me take a picture of you.'

   b. wo lai *zhao ni (de) xiang cf.*zhaoxiang
      I will you DE reflect-picture
      'Let me take a picture of you.' 'to photograph'

(50a. ni bubu ti ta *cao#xin
      you need-not for him be-concerned
      'You don't need to be concerned about him.'

   b. ni bubu *cao ta (de) xin
      you need-not him DE
      'You don't need to be concerned about him.'
      cf. *cao#xin
      manipulate-heart
      'to be concerned about'

The verbs in (44a)–(50a) are subcategorized for an NP marked with an adjunct
particle, e.g., gen ('with'), wei ('for'), etc., the choice of which being a purely
idiosyncratic property of the individual verbs. The situation here is not unlike the
prepositions such as on and of in rely on and approve of in English, which are used for
case–marking instead of for conveying real semantic content. (See Gazdar et al 1985's
treatment for case–marking prepositions in English (pp.131–132).) Their meaning can
therefore be regarded as denoting an identity function on NP meanings, i.e., as
semantically vacuous. One way to distinguish the PP's marked by the various
prepositions is to mark the PP category with a (semantically–null) feature encoding the various particles. That is to say, we can categorize the verbs here as VP|PP\_[X], [X] being a feature (which makes no semantic contribution and ranges, in the present case, over gen, ti, and wei, etc.) on the PP category encoding the specific adjunct particle or preposition which is required by the verb.\(^2^2\) For example, the verb sheng#qi ('be angry') in (46) will have the subcategory of VP|PP[gen]. (See Bach 1983a, 1984 for discussion on the incorporation of features in Categorial Grammar.) The following rule would then account for the verb–PP combination in question.

\(^{(51)}\)
< LCON, < VP|PP, PP, VP >

semantics: functional application

What is interesting is that for some bisyllabic verbs which are subcategorized for an adjunct–particle–marked NP argument, the NP can also occur unmarked at an internal position of the verb, with an optionally present suffixal particle de.\(^2^3\) As

\(^{2^1}\)Note that the adjunct–particle–marked NP's are optionally present, which means all the verbs that can be subcategorized as VP|PP\_[X] also have the IV subcategory. We can presumably link these two subcategories via a lexical rule.

\(^{2^2}\)In Bach 1983a's terminology, this would be considered as an 'inherent' feature, as opposed to features associated with functor categories which are built up as ordered pairs of features for the argument category and the resultant category (p.75).

\(^{2^3}\)Occasionally we find that verbs which are used transitively, such as guan#xin ('to be concerned with') (e.g., (a)), also have this subcategorization property (e.g., (b)).

\(\text{a. ni bubì guan#xin ta} \)
\(\text{you need–not concern him} \)
\(\text{'You don't need to be concerned about him.'} \)
demonstrated in (44b)–(50b), the verbs in (44)–(50) are such verbs. This is what is referred to as the Infixal Argument Construction, in which an argument NP is infixed into the verb. Examination of (bisyllabic) verbs which optionally allow the infixation of a subcategorized NP argument shows that they correspond to those verbs which allow infixed time adverbials, i.e., verbs which can be distributed in the IAC (see Section 3.2.2.).

We can accommodate this subcategorization property of the verbs in (44b)–(50b) by assigning them, in addition to the category of VP[PP[N]], a subcategory

---

b. ni bubi guan ta (de) xin you need-not him DE
   'You don't need to be concerned about him.'

Take the verb jinzhang ('be nervous') for example. It does not allow inserted adverbials, as shown in (a). (b) and (c) show that although it takes an adjunct-particle-marked NP, the verb does not allow the NP to be inserted into the verb.

a.* ta jin bantian de zhang he half-day
   'He was nervous for a while.'

b. wo ti ta jinzhang I for him nervous
   'I am nervous for him.'

c.* wo jin ta (de) zhang I him DE
   'I am nervous for him.'

However, we can occasionally find verbs, e.g., ju#gong ('to bow'), which can be distributed in the IAC ((d)) but do not tolerate well an infixed argument; (e) may not sound very good for some speakers. But such verbs are very rare.

(d) Zhangsan bu yuanyi gen Lisi ju#gong.
   Zhangsan not willing to Lisi bow
   'Zhangsan is not willing to bow to Lisi.'

(e)? Zhangsan bu yuanyi ju Lisi de gong
   Zhangsan not willing Lisi DE
   'Zhangsan is not willing to bow to Lisi.'
such as VP||NP\textsuperscript{25}, the double-slash\textsuperscript{26} being used to distinguish this category from
regular TV's. This category assignment, coupled with the rule in (52), would then
account for the IARC (e.g., (44b)–(50b).

(52) the IARC Rule:
\begin{align*}
&\langle F_{11}, <VP||NP, NP>, VP \rangle \\
&F_{11} = \text{RWRAP} \\
&\text{semantics: functional application}
\end{align*}

Again, the rule invokes a head wrapping operation, more specifically, RWRAP. As
defined in (17), RWRAP suffixes the argument expression to the right of the head of
the functor expression. As discussed in Section 3.3.4, the bisyllabic verbs in question
have the [V NP] structure (supplied by the proposed reanalysis rule), which is
available to RWRAP. Take the verb, \textit{dan#xin} ('be worried') in (44), for example. In
addition to having the VP||PP\textsubscript{gen} category (cf. (44a)), it also has the subcategory of
VP||NP, with an internal structure of [V NP] supplied by the reanalysis rule.

According to the IARC rule in (17) then, it may be combined with an argument NP via
RWRAP, which adjoins an NP, e.g., \textit{ta} ('him'), to the right of the internal head of the

\textsuperscript{25}A lexical rule may be postulated linking these two categories.

\textsuperscript{26}The use of single vs. double slashes in functor categories, sometimes referred to
as 'category splitting', was originated by Montague to distinguish between categories
which are of the same semantic category (type) but which have different syntactic
distributions, e.g., in the case of IV and CN (both being one-place predicates) and IV
modifiers and \textit{try}-class verbs (both being one-place predicate modifiers). (See
Schmerling 1983a.)

\textsuperscript{27}To account for the version of (44b)–(50b) where the particle \textit{de} is present, we
also need to associate the rule with \(F_{14}\) and \(F_{14} = \text{RWRAP}_z \) (\(\alpha, F_{5}(\beta)\)).
verb, namely, *dan*, resulting in the string *dan ta xin* ('worry about him') in (44b).

(53) \[ \text{[dan ta xin]}_{VP} \]
    \[
    \text{\textbackslash} \text{Rule (52)}
    \]
    \[ \text{[danxin]}_{VP\_NP} \quad \text{[xin]}_{NP} \]

What these data show once more is that some kind of infixed (or wrapping) operation is indeed necessary since the data here is not subject to alternate analyses adhering to the Principle of Adjacency, such as an analysis employing only functional composition. To use functional composition to deal with these data, among other things, we need to assign the component *zi*'s in the bisyllabic verb to some lexical categories with certain meanings such that they combine (in a compositional way) to form a verbal category with the meaning of the verb. This would potentially allow one of the component *zi*'s to compose with the infixed argument first when the *zi*'s in the verb are separated by the infixed argument. But this would be a highly artificial account, if not entirely impossible, since, as we have discussed in Chapter 2, many of these verbs contain 'bound forms' (in the prescribed sense), which are not members of any lexical categories in the language and do not have any isolatable meaning. Even with verbs with component *zi*'s which can be analyzed (by modern speakers) as members of certain lexical categories (e.g. *sheng\#qi* ('grow-air=to be mad at'), typically the meaning of the verbs cannot be segmented compositionally into two isolatable parts in a motivated way. Therefore any analysis along this line in an attempt to do away with wrapping has to be rejected.
3.4.1: Reduplication in Syntax.

The trend in recent syntactic theories has been toward the lexicalization of syntax, whereby syntactic information of all sorts is encoded in the lexical entries while the syntactic rules are kept to a minimum (see Karttunen 1986), e.g., Lexical–Functional Grammar, Head–Driven Phrase Structure Grammar, Unification–based Grammars and Categorial Grammars in general. In the standard varieties of Categorial Grammar, information as to the combinatorial possibilities and the direction of combination is encoded in the lexical categories, specifically, in the functor categories. Only a minimum of rules, namely functional application and perhaps functional composition, are postulated in syntax. Even with a nonconcatenative operation such as wrapping it is possible to treat the operation as a property of certain functor categories and to encode accordingly the combinatorial mode in the functor categories (similar to the way slashes are used to encode the direction of combination). Furthermore, as already mentioned, there have been alternatives proposed to reanalyze the wrapping phenomenon within an IA framework. A much more compelling type of data arguing for an IP model for a syntactic theory is the phenomenon of phrasal or syntactic reduplications, if such data can indeed be found in natural languages.\textsuperscript{28} In the following, a construction in Chinese which has come to be known as the 'A–not–A Question' will be considered, which I call the Reduplicative Interrogative Construction.

\textsuperscript{28} Such data have been extremely scarce, one exception being the phrasal reduplication in Engenni, a language of Nigeria, discussed in Carlson 1983.
(RIC) and which can be shown to involve syntactic reduplication, hence providing strong arguments for the processual approach within the Categorial framework.

3.4.2. The Reduplicative Interrogative Construction.

The RIC is one way of getting an interrogative sentence of the Yes–No type in the Chinese. It has traditionally been analyzed as involving a transformation of conjunction reduction (cf. Wang 1967, Lyu 1985, Li and Thompson 1981, Lin 1974, etc.) and was first analyzed as involving syntactic reduplication in Sheu 1987 & 1988. For arguments against the traditionally assumed conjunction reduction analysis for this construction, see J. Huang 1988a, which independently argues for a reduplication analysis.

In the RIC, a part or all of a verb phrase, when it is not too 'heavy', is reduplicated; at the same time a negative element, *bu* ('not')\(^{30}\), is inserted between the reduplicated elements. Sentences (54)–(60) represent the cases where only the verb is reduplicated. Among these, (54)–(55) involve IV's, which can be mono– or bi–sylabic, and (56)–(60) involve TV's, which, again, can be mono– or bi–sylabic.

(54) ni lai bu lai
    you come not come
    'Are you coming?'

\(^{29}\) Another way to get an interrogative sentence of the yes–no type is to add the question particle *ma* at the end of a sentence. In both patterns, no inversion is involved.

\(^{30}\) *Bu* alternates with another negative word, *mei* ('not'), depending on the aspect of the sentence (see Teng 1974).
(55) ni jin#zhang bu jin#zhang cf. jin# zhang
you nervous not nervous tight tense
'Are you nervous?'

(56) ni kan bu kan dianshi
you watch not watch TV
'Do you watch TV?'

(57) ta chi bu chi niurou
he eat not eat beef
'Does he eat beef?'

(58) ni *zhi#*dao bu *zhi#*dao zhe jian shi
you know not know this CL31 matter
'Do you know about this?'

(59) ni dan#xin bu dan#xin ta
you worry not worry him
'Do you worry about him?'

(60) ni ren#*shi bu ren#*shi Lisi cf.ren#*shi
you know not know Lisi 'recognize-know#to know'
'Do you know Lisi?'

(61)-(65) in the following represent the cases where the whole VP is reduplicated.

(61) ni kan dianshi bu kan dianshi
you watch TV not watch TV
'Do you watch TV?'

(62) ta chi niurou bu chi niurou
he eat beef not eat beef
'Does he eat beef?'

(63) ni *zhi#*dao zhe jian shi bu *zhi#*dao zhe jian shi
you know this CL matter not know this CL matter
'Did/Do you know about this?'

(64) ni dan#xin ta bu dan#xin ta
you worry him not worry him

31CL=Classifier.
'Do/did you worry about him?'

(65) ni ren#shi ta bu ren#shi ta
    you know him not know him
    'Do you know him?'

Interestingly, a third pattern for deriving the interrogative form, in addition to the copying of the verb and the whole VP, is to reduplicate just the first zi of the verb in question -- this is of course trivially true for all monosyllabic verbs. (66)-(69) exhibit the zi-copying variants of (55) & (58)-(60), respectively.

(66) ni *jin bu *jin#zhang
    you not worry
    'Are you worried?'

(67) ni *zhi bu *zhi#*dao zhe jian shi
    you not know this CL matter
    'Do/Did you know about this?'

(68) ni dan bu dan#xin ta
    you not worry him
    'Do/Did you worry about him?'

(69) ren bu ren#shi Lisi
    ni not know Lisi
    'Do you know Lisi?'

In all, there are three possible patterns of reduplication, summarized in (70). Each of these three forms of reduplication is possible for any given VP.

(70) Patterns of reduplication in the RIC:
    A. the first zi of the V (if the V is bisyllabic), e.g., (66)-(69);
    B. the whole verb, e.g., (54)-(60);
C. the whole VP including the object NP (provided it is not too heavy) (e.g. (61)-(65)).

As can be seen, the third pattern of the RIC, i.e., VP-copying, exhibits the property of unbounded syntactic string copying. The best-known data showing that natural language may not be context-free is that of the Swiss-German cross-serial verbal construction (Shieber 1985), which does not involve strict string copying but only some kind of cross-serial dependency. Thus the RIC may present a stronger case for the non-context-free-ness of natural languages. (See Radzinski 1990 for an attempt of a formal proof that (Mandarin) Chinese is non-context-free on the basis of the RIC.) This having been said, I will not pursue here further the implication of the RIC for the formal language theory of natural languages, but rather concentrate on providing an adequate analysis for the relevant data in the Categorial-Processual approach.

All three variations of reduplication can be subsumed under one reduplication pattern, namely, reduplicate the left-most constituent (or structural unit), be it a zi, word, or a phrase. That is to say, if we can make reference to the left-most constituent at different structural levels within an VP, we can generalize the reduplication operation here. As discussed in Section 3.2.3, such reference can be made via the one-place analytic operation FIRST (defined in (18)), first introduced in Bach 1984, which was originally proposed to help define edge-anchored (as opposed to head-anchored) wrapping operations.
It is implicit in Bach 1984 that the level of analysis within a constituent available to wrapping is that of word, that is, what wrapping has access to are word boundaries. But other levels of analysis available to wrapping have also been suggested (Carlson 1983, Hoeksema and Janda 1988, and Chierchia (1988:126)).

Adopting the notations in Hoeksema and Janda 1988, we can parametrize the function FIRST with respect to the level of analysis it refers to, e.g. FIRST_w(X), FIRST_p(X), etc., where w and p designate word-level and phrase-level. As discussed in Chapter 2, there is an additional level of analysis which is significant in Chinese syntax, namely, zi, the monosyllable. As (66)–(69) indicate, this level of analysis figures in the RIC. We can specify in the following the FIRST function for these three levels of analysis, namely, the phrase, the word and the zi level, which are simultaneously available to the rule for RIC.

(71) Let X=[x_1,...,x_n], where each x_i is a constituent at some structural level (and x_2,...,x_n can be empty):
FIRST_w(X) = z_i
FIRST_p(X) = w_i
FIRST_z(X) = p_i

As defined above, the function FIRST_z returns the left–most zi, or monosyllable, in a string; FIRST_w the left–most word; and FIRST_p the left–most phrasal unit. The availability of such a one–place analytic function, parametrized for
various levels, will enable us to formalize a generalized composite operation\(^3\) which the RIC invokes.

\[(72)\]
\[F_{13,i}(X) = \text{RCON}(\text{RCON}(	ext{FIRST}_i(X), F_\phi), X)\]
i ranges over z, w and p and \(F_\phi\) is a constant operation yielding the negative particle bu.

\(F_{13,i}\), as defined in (72), is in fact a schema of operations which subsumes three (composite) operations, involving, variously, \(\text{FIRST}_z\), \(\text{FIRST}_w\), or \(\text{FIRST}_p\). They are, more perspicuously:

\[(73)\]
a. \(\text{FIRST}_z(X)'bu^\ast X\)
b. \(\text{FIRST}_w(X)'bu^\ast X\)
c. \(\text{FIRST}_p(X)'bu^\ast X\)

The RIC rule can now be stated as:

\[(74)\] \(<F_{13,i}; \text{VP}; Q_i|NP>^{33}\]
i = \{z, w, p\}

\(^{32}\)called 'polynominial operation' in Montague 1974.

\(^{33}\)For a possible treatment of the semantics of yes–no question, see Karttunen 1977, where a yes–no interrogative sentence '?\(\alpha\)' is treated as denoting the unit set containing either the proposition denoted by \(\alpha\) or the proposition denoted by \(\neg\alpha\), whichever happens to be true.
F₁₃ may be (73a), (73b), or (73c) and thus for any given VP, (74) will correctly give all possible reduplicative interrogative forms, namely the zi-copying, word-copying and phrase-copying variants. Take a VP such as *zhi*dao zhe jian shi ('know this-CL matter') for illustration. The following trees represent the three reduplicative interrogative forms of the RIC construction for the VP. (Note that in each tree I only give the bracketing relevant to the particular parametrized FIRST used in the rule for the sentence in question: in (75), only zi boundaries are indicated (by '#') while word boundaries and phrasal bracketings are omitted; in (76), only word boundaries are indicated (by space); in (77) only phrasal bracketing is indicated.)

(75) ni *zhi bu *zhi*dao zhe-jian shi

you not know this-CL matter

'Do/Did you know (about) this (matter),'

ni zhi bu zhidao zhe jian shi

/ \
ni zhi# bu# zhi#dao# zhe# jian# shi

| Operation (73a)

zhi#dao# zhe#jian# shi

/ \
zhi#dao# zhe#jian# shi

/ \
zhe#jian# shi

---

³I will assume for the present purpose that demonstratives and classifiers combine to form compounds of the determiner category (NP|CN).
(76) ni *zhi*dao bu *zhi*dao zhè-jian shì

you know not know this-CL matter

'Do/Did you know (about) this (matter),'

ni zhì bu zhídào zhè-jian shì

Operation (73b)

ni zhídào zhè-jian shì

zhídào zhè-jian shì

(77) ni *zhi*dao zhè-jian shì bu *zhi*dao zhè-jian shì

you know this-CL matter not know this-CL matter

'Did/Do you know about this (matter),'

ni zhídào zhè jian shì bu zhídào zhè jian shì

[ni]_{NP} [[zhídào zhè jian shì] bu [zhídào zhè jian shì]_{Q;NP}

Operation (73c)

[[zhídào] [zhè jian shì]_{VP}

[zhídào]_{VP;NP} [[zhè-jian]] shì{NP}

[zhè-jian]_{NP;CN} [shi]_{CN}

In (77), the operation $F_{13p}$, spelled out in (73c), copies the whole VP --- this is possible because according to the way FIRST, is defined in (71) the whole VP can be analyzed as the left-most constituent (when $x_2$,...,$x_n$ are empty). Take for another example a sentence like (78a), where the verb is preceded by a PP (gen ta ('with him')); we see that both (c), where the PP is copied, and (d), where the whole VP is copied, are grammatical, since both the PP (or VP!VP) and the whole VP can be analyzed as the left most constituent at some (phrasal) levels.
(78a. ni gen ta zou
you with him go
'You go with him.'

b. ni gen bu gen ta zou
you with not with him go
'Are you going with him?'
(operation (73b))

c. ni gen ta bu gen ta zou
you with him not with him go
'Are you going with him?'

\[
\begin{align*}
&\text{ni gen ta bu gen ta zou} \\
&\text{\quad \quad ni} \\
&\quad \quad \quad \quad [[\text{gen ta} \ bu \ [\text{gen ta zou}]]_{C:NP} \\
&\quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \text{Operation (73c)} \\
&\quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad [[\text{gen ta}] \ [\text{zou}]]_{VP} \\
&\quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \text{gen ta} \\
&\quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \text{zou}_{VP} \\
&\quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \text{zou}_{VP}
\end{align*}
\]

(78d. ni gen ta zou bu gen ta zou
you with him go not with him go
'Are you going with him?'
(operation (73c))

We also predict that, in contrast to (75)-(77), which are well-formed, the
following sentence is ill-formed, since the reduplicated strings (zhidao zhe-jian 'know
this-CL' in (79a) is not a constituent (at any level):

(79)a.* ni zhidao zhe-jian bu zhidao zhe-jian shi
you know this-CL not know this-CL matter
'Do you know about this (matter).Wrap

Finally, in the RIC rule, the use of FIRST, as opposed to an operation referring
to head, is motivated by data such as (78), where the left-most element is not verbal.
In (78) a PP precedes the verb. Here the reduplication of both the preposition and the PP result in well-formed sentences, as is predicted by the RIC rule, whereas the reduplication of the verb itself results in ill-formed sentences\textsuperscript{35}. More data of this kind, demonstrating the necessity of reference to left-margin structural units, are given in the following. In (80a) an adverb, \textit{jing\#chang} ('often'), and in (81b) an auxiliary verb, \textit{keneng} ('may')\textsuperscript{36}, precede the head verb. Examples (b)–(d) give the reduplicative interrogative forms of these VP's, as predicted by the RIC rule.

\begin{enumerate}
\item[(80)a.] ta jingchang hui jia  
\hspace{1em} he often come here  
\hspace{1em} 'He goes home often.'

\item[b.] ta *jing bu jingchang hui jia  
\hspace{1em} he not often go home  
\hspace{1em} 'Does he go home often?'

\item[c.] ta jingchang bu jingchang hui jia  
\hspace{1em} he often not often go home  
\hspace{1em} 'Does he go home often?'

\item[d.] ta jingchang hui jia bu jingchang hui jia  
\hspace{1em} he often go home not often go home  
\hspace{1em} 'Does he go home often?'

\item[e.*] ta jingchang hui bu hui jia  
\hspace{1em} he often go not go home  
\hspace{1em} 'Does he go home often?'
\end{enumerate}

\begin{enumerate}
\item[(81)a.] ta keneng lai  
\hspace{1em} he may come  
\hspace{1em} 'He may come.'
\end{enumerate}

\textsuperscript{35}cf.*ni gen ta zou bu zou  
you with him go not go  
'Are you going with him?'

\textsuperscript{36}If the auxiliary verb such as \textit{keneng} ('may', 'possibly') can be shown to be the head of VP in Chinese, then data such as (81) can also be analyzed by reference to head in the RIC rule.
b. ta ke bu ke#neng lai
   he not may come
   'Might he come/Is it possible that he is coming?'

c. ta keneng bu keneng lai
   he may not may come
   'Might he come/Is it possible that is coming?'

d. ta keneng lai bu keneng lai
   he may come not may come
   'Might he come/Is it possible that he is coming?'

e.* ta keneng lai bu lai
   he may come not come
   'Is it possible that he is coming?'

Sentences (80)–(81) show that our analysis which refers to the function of FIRST makes the correct predictions regarding the interrogative forms of the VP's in question. For the VP in (80), bracketable as \[ [\text{jing#chang}]_{vp} [[\text{hui}]_{vp} [\text{jia}]_{vp}]_{vp} \] (often go home'), FIRST may return the left–most zi (i.e. #jing), the left–most lexical category (i.e. jing#chang ('often'), or the left–most phrasal constituent at the next higher level (which is the whole VP), resulting in (80b), (80c), and (80d), respectively. On the other hand, (80e), where the head verb is reduplicated, is ill–formed. Similar patterns are found in (81).\(^{37}\)

\(^{37}\)I will note here (without attempting to incorporate it into the present analysis) that the 'Descriptive Complement Construction' (e.g., (a)) and the 'Resultative Construction (e.g., (b)), both involving marked VP structure (where we find the particle de and some kind of post–verbal predicative expression), exhibit a distinct pattern of reduplication (e.g., (a') and (b')). (See Chapter 5 on the syntactic and semantic properties of the latter and J. Huang 1988b on syntactic properties of both types of constructions.)

a. ta pao de kuai
   he run DE fast
   'He runs fast.'

a'. ta pao de kuai bu kuai
As already mentioned, not all bisyllabic verbs can be distributed in the IAC and become discontinuous. However, all\(^3\) can participate in the RIC such that the verbs become discontinuous due to the zi-copying operation in (73a). For example, a bisyllabic verb, such as jin\#zhang (be nervous), has the zi-copying variant (as well as other variants --- omitted here) of the interrogative reduplicative forms ((73a)) but cannot be distributed in the IAC (e.g., (73b)).

\[(82)a. \text{ni jin bu jin\#zhang} \quad \text{you not nervous} \quad \text{'Are you nervous?'}\]

\[(82)b. \text{he run DE fast not fast} \quad \text{'Does he run fast?'}\]
\[(82)b'. \text{he run DE tired not tired} \quad \text{'Did he get tired from running?'}\]

\(^3\) Except for the Resultative compounds, which, perhaps for aspectual reasons, do not admit this pattern. What it allows is an interrogative pattern involving reduplication and infixation of the particle de and bu (not'). The traditional term for this pattern is called the 'potential' forms of the Resultative compounds (see Chao (1968:452).

\[(83)a. \text{kan-jian} \quad \text{look-see} \quad \text{'see'}\]
\[(83)b. \text{ni kan de jian kan bu jian nei dong fangzi} \quad \text{you DE not that CL house} \quad \text{'Can you see the house?'}\]
\[(83)c.* \text{ni kan bu kan jian nei dong fangzi} \quad \text{you not that CL house} \quad \text{'Can you see the house?'}\]
The contrast in the distributions of bisyllabic verbs in these two constructions is explained readily in the analysis here: the IAC invokes a head wrapping operation; therefore, only bisyllabic verbs which are (re- ) analyzed as [V NP] (hence containing a (pseudo- ) head) can be distributed in this construction. The RIC rule, on the other hand, makes reference to the left-margin unit instead of to a head element; therefore all bisyllabic verbs have the interrogative reduplicative forms (of the zi- copying variant as well as of other variants). jin- zhang ('tight- tension = be nervous') is a verb which is not reanalyzed (for almost all speakers), and therefore (82b) is bad.

To conclude the discussion on the RIC, whatever the syntactic rule is which derives (54)–(65), that rule should also account for (66)–(69)39. That is to say, the syntactic rule responsible for the RIC seems not to be sensitive to the distinction in the grammatical level of the input expression. Such a rule, however, would be impossible to state under a strictly modular view of grammar as it applies to both word- internal and word- external units. Under the strictly modular view, we would be forced to reduplicate the RIC rule: one for data such as (54)–(65) and one for data such as (66)–

39J. Huang (1984: 75 (fn.6)) dismisses this set of data as irrelevant for matters of lexical integrity in Chinese, on the grounds that the reduplication rule in question is either a phonological or morphological rule. As we can see here, however, the input of the reduplication rule can indeed be phrasal units.
(68), hence losing certain generalization proper for the phenomenon under description⁴⁰.

⁴⁰It may be noted here that this is the sort of argument that led Morris Halle to reject the phonemic/allophonic level distinction in phonology.
CHAPTER IV

CATEGORY CHANGING RULES AND MARKING OF DE IN COMPLEX CN CONSTRUCTIONS.

4.0. Introduction.

In the previous chapter, I have offered data from Chinese which I have argued should be properly analyzed with one-place rules. In some, if not all, cases, one-place rules can be construed as category changing rules. In this chapter I will present another set of data involving the marking of de in complex NP constructions, which I will argue should be properly analyzed as a syncategorematic reflex of the category changing rules deriving CN-modifying expressions.

4.1. Category Change.

A phenomenon worthy of investigation in both morphology and syntax is how the mechanisms of changing the category (or subcategory) of expression are provided in languages. For lexical categories, there is typically a part of derivational morphology which is specialized in marking the change of category of an expression

\[\text{Footnote: The relation-changing rules as applied to verbs in Dowty 1982a, which involve change of the valency of a verb or the grammatical relations associated with the arguments combined with the verb, can also be conceived as category change rules.}\]
in a language (e.g., affixes such as -ize, -ity, in English). In syntax, we also find grammatical processes cross-linguistically which serve the function of changing the category of phrasal expressions. For example, in English one finds grammatical processes such as (1) nominalization, which changes an expression of a VP category into a nominal category, e.g., gerunds, and (2) participle formation, which changes a verbal expression to a nominal-modifying expression. That is to say, affixes such as -ing can be seen as signaling or marking a change in the category (and therefore the semantic type) of expressions. The possessive marker can also be seen as such a marker, which changes an NP to a NP|CN. The 'articular infinitives' in Greek (Joseph 1983) are similar in nature to English gerunds. In Tamil, the clitic particle -a marks the category changing from a clause to a common noun modifier, and the particle -tu marks a further change from such a CN modifier to a nominal expression, which has case properties just like other nominal expressions (see Subramanian and Sheu 1989). In Korean, a similar nominalization process can be found (Lee 1967).

In Chinese, although there is a total lack of derivational morphology marking change of category in the lexicon, we find that such processes are common in the syntax. Such processes can be quite straightforwardly accommodated in the

---

2 A verb in the infinitive takes NP arguments and can be modified by adverbs but it can be combined with an article to form an NP.

3 There exists, however, a (relatively small) set of derivational affixes that changes the meaning of a word. For example, xue ('-ology') in xinlixue ('psychology') and jia ('-ist') in kexuejia ('scientist').
Categorial-processual syntax adopted here. In fact the way rule schemas are envisioned here, i.e., consisting of input category(ies), an output category and the operation invoked, all one-place rules, which are allowed under Montague's definition of syntactic rules in UG, though not adopted in standard Categorial grammars, can be construed as category-changing rules.

Before examining the category change involving the marking of *de*, let me briefly note that not all category changes in the language are grammatically encoded. The nominalization process exhibited below is one example. In (1), we see that the same form *paobu* ('run') occurs as both a predicate VP and a subject NP without any overt marking; similarly, *kan dianshi* ('watch TV') in (2) and *yonggan* ('(be) brave') in (3).

(1)a. wo meitian dou paobu
    I everyday all run
    'I run everyday.'

       b. meitian paobu shi hao xiquan
        everyday run is good habit
        'Running everyday is a good habit.'

(2)a. ta chang kan dianshi
    he often watch TV

---

"The category-change rules here, which affect the combinatorial properties of expressions, are to be distinguished from the type shifting mechanisms discussed in Partee (1987, 1986) and Partee and Rooth (1983). Type-shifting is a purely semantic process and refers to, for example, shifting among the various semantic types (<e>, <e,t> and <e,t,t>) associated with the syntactic category of NP's in order to account for noun phrase interpretations in various syntactic positions."
He watches TV often.'

b. kan dianshi dui yanjin bu hao
   watch TV to eye not good
   'Watching TV is not good for the eyes.'

(3)a. zhe ge ren yonggan
     this CL person brave
     'This person is brave.'

b. yonggan shi meide
   bravery be virtue
   'Bravery is a virtue.'

It is possible to hypothesize that there are category changing processes, e.g. nominalization, that are more 'natural' than others, such as the shifting between predicate expressions and noun modifying expressions to be discussed below, and that the more natural a category shift is, the more likely we find it unmarked in languages. It is worth mentioning in passing that Chierchia's IL* (1984 & 1985) provides an explanation for the naturalness of a process such as nominalization. In the semantic system of IL*, properties have 'two modes of being': the predicative mode (i.e., as propositional functions) and the individual mode, with a mapping relating properties to their individual projection in the domain of individuals. The non-encoding of nominalization in some languages may therefore be seen as a result of the naturalness of the mapping between properties and their individual projections. More likely, however, a category changing rule is grammatically encoded. Across languages we can identify particles or 'little words' which are specialized in marking these processes.

---

5 This is referred to as the 'Fregean' mapping in Chierchia 1984 and 1985.

6 This point is also made in C. Huang (1989, Chap. 3).
In the rest of this chapter, I will give an analysis of an important category change phenomenon in Chinese involving the marker *de* in complex CN constructions. It presents perhaps the most interesting case of category change in Chinese syntax, although it has not been previously recognized as such.

4.2. **An Analysis of Category Change Marker *de***

The 'category change' which will be the main focus of this chapter is that between a phrasal expression of the predicate type and that of CN|CN, which is invariably pre-nominal in the language. A predicate category in the language can be defined as expressions of the <e,t> type which combine with a subject NP to form a sentence. Throughout this thesis the category labels VP or PRED are used without differentiation to refer to this category. In Chinese, this category includes not only the verb phrases, but also, as we can see from the following data, what looks like adjective phrases, e.g., (4)–(5), and prepositional phrases, e.g., (6)–(7), as both kinds of phrases combine with subject NPs to form sentences without the copula verb *shi* ('be').

(4) zhe ge bingren tai shou
this CL patient too thin
'This patient is too thin.'

(5) zhe kuai tudi feiwo
this CL land fertile
'This piece of land is fertile.'

(6) qian zai zhuo shang
money on desk POSTP.
'The money is on the desk.'
(7) zhe dong wuzi chao nan
this CL house toward south
'This house is facing the south.'

The following data demonstrate how a predicate expression can be systematically related to a pre-nominal common noun modifier (CN-CN) in the language. Sentences (8)–(10) involve predicate expressions which are verb phrases and (11)–(12) involve those which are seemingly AP's. Sentence (13) involves those that are PPs.

(8)a. zuotian lai de(*Ø) ren ni renshe ma?
yesterday come DE person you know FP
'Did you know the man who came yesterday?'

cf. b. ta zuotian lai
he yesterday come
'He came yesterday.'

(9)a. youge dai maozi de(*Ø) ren lai zhao ni
some wear hat DE person come look-for you
'A man wearing a hat came looking for you.'

cf. b. neige ren da maozi
that-CL person wear hat.
'That man was wearing a hat.'

'On the basis of (6) and (7) only, it may be possible to analyze expressions such as zai ('at') and chao ('toward') as some kind of verbs with a prepositional meaning. Additional data such as the following, where the zai-phrase or chao-phrase are used to modify a verb, would show that these are indeed prepositions.

(6)
ea. ta zhan zai zhuo shang
he stand on desk POSTP.
'He is standing on the desk.'

cf. (7)
b. ta chao nan zou
this toward south walk
'He is walking toward the south.'
Before discussing how a category changing rule can be postulated to account for all of these data, it is necessary to discuss first (1) the status of what seems to be adjectival expressions in the language and (2) the general treatment of common noun modifying expressions within Montague Semantics.

4.2.1. Adjectives in Chinese.

First to be discussed are adjectival expressions in the language such as shou ('thin') and feiwo ('fertile') in (4) and (5). The following data show that though these expressions are translated into adjectival expressions in other languages, distributionally they pattern in exactly the same way as verbs (see Chao 1968, Ross 1983, and Tai 1982). More specifically, the following data show, in an abbreviated
way, that (1) both verbs and adjectives take degree modification ((14a) and (14b)), and (2) both take aspectual markers ((14c) and (14d)). (See Li and Thompson 1981 for more data demonstrating these patterns.)

(14)a. ta hen congming
    he very smart
    'He is very smart.'

b. ta hen dong shuxue
    he very know math
    'He knows math very well.'

c. ta shou le
    he thin ASP
    'He has become (quite) thin.'

d. wo chi le
    I eat ASP
    'I have (just) eaten.'

More significantly, as we have seen in (4)–(5) and (11)–(12), expressions denoting meanings typically associated with adjectives in other languages (1) occur at the predicate position without the copula verb and (2) cannot be used to modify common nouns. To modify common nouns, they have to be marked by the particle de.

Examples (15)–(20) below further illustrate these two patterns.

(15)a. nei duo hen hong de(*Ø) meiguihua
    that CL very red DE rose
    'That very red rose...'

cf. b. nei duo meiguihua hen hong
    that CL rose very rose
    'That rose is very red.'

(16)a. mei zhi jiankang de(*Ø) gou.
    every CL healthy DE dog
'Every healthy dog...'

cf. b. zhe zhi gou bu jiankang
this CL dog not healthy
'This dog is not healthy.'

(17)a. wo mai le yi zhang liu-chi-chang de(∅) chuan
I buy ASP one CL 6-foot-long DE bed
'I bought a 6-foot-long bed.'

cf. b. zhe zhang chuan san-chi-chang
this CL bed 6-foot-long
'This bed is 6-foot-long.'

(18)a. wo mai le yi zhi hen guì de(∅) yingwu
I buy ASP one CL very expensive DE parrot
'I bought a very expensive parrot.'

cf. b. zhe zhi yingwu tai guì
this CL parrot too expensive
'This parrot is too expensive.'

(19)a. mei ge jinzhang de(∅) qiu yuán dou likai le
every CL nervous player all leave ASP
'All the nervous players left.'

cf. b. youxie xuèshèng hén jinzhāng
some player very nervous
'Some players are very nervous.'

(20)a. ai(∅) lanchiu-duiyuan hén shào jiān
short basket-player very rare
'Short basketball players are rare.'

cf. b. zhe ge lanchiu-duiyuan hén ai
this CL basket-player very short
'This basketball player is very short.'

We see that expressions associated with adjectival meanings such as RED (hong) and
HEALTHY (jiankang), which are typically members of the class of adjectives in other
languages, behave like verbs in Chinese in that they occur at the predicate position
without a copula verb and in that they can not be used directly to modify CN's. The
(a) sentences in (15)–(20) without the presence of the marker de are ill-formed. In the
analysis here, this shows that some sort of category changing rule, to be postulated later, is obligatory.

If we take adjectives to be the class of expressions in languages which can be used to modify nouns both attributively and predicatively (cf. Schacter 1985), then Chinese does not seem to have a class of adjectives. Instead, expressions denoting meanings typologically expressed by adjectives in languages are most properly considered as a subclass of (intransitive) verbs⁸. That is to say, they are predicative, denoting sets of individuals. This explains why (1) they occur at the predicate position without the copula verb shi ('to be') and (2) they cannot modify common nouns without further marking.

4.2.2. **Treatment of Adjectives.**

Consider now the general syntactic and semantic properties of adjectives and the treatment of adjectives within Montague semantics. Distributionally, in languages that have this class of expressions, adjectives typically occur both at the predicative (or post-copula) positions and at the attributive position (e.g., pre-nominal positions in English). Semantically, adjectives seem to be a heterogeneous group; some are representable as one-place predicates, or, sets of individuals, while some others seem to have functional meanings. The former type of adjectives is often called intersective

---

⁸Therefore Chinese is what Schacter calls a 'adjectival-verb' language, as opposed to 'adjectival-noun' languages such as Hausa and Quechua.
adjectives (Bennett 1974, Siegel 1976b), since the meaning of the complex CN resulting from the combination of such an adjective and a CN is characterizable as the intersection of the set denoted by the CN and the set denoted by the adjective. These include expressions such as red, healthy, six-foot-long, four-legged, and possibly all relative clauses (Quine 1960, Keenan and Faltz (1985:122)). These adjectives can be thought of as denoting a set of entities, independently of the CN which they combine with. For example, a red rose is necessarily something that is red and is a rose. In a word, the intersective entailment is what needs to be captured somehow in the semantics of this type of CN modifying expressions.

These adjectival expressions are to be distinguished from another set of adjectives which may be called subsective adjectives. They include adjectives denoting size such as big, tall, and wide; 'value judgement' adjectives such as famous, and good; or manner adjectives such as skillful. The meanings of these adjectives are considered as not independent of the CN they combine with. For example, an adjective like big cannot be thought of as denoting a set of big entities which intersects with the set denoted by the CN to give the meaning of the complex CN, since the criteria for determining what may be, say, a big flea is quite different from those for determining what may be, say, a big elephant. In other words, these adjectives cannot be simply represented as sets (of entities). They can, however, be represented as functions mapping CN meanings to CN meanings, i.e., functions from sets (of entities) to sets (of entities). The entailment which we need to capture somewhere in semantics
for this type of adjectives is that the set denoting the resulting complex CN is always a subset of the set denoted by the CN combined with the adjective. For instance, a severe teacher is necessarily a teacher while we would not want to say a severe teacher is a severe individual.

Due to the presence of such adjectives, the general treatment of adjectives in Montague grammar, originating with Montague 1974 and the unpublished work of J.A.W. Kamp and Terence Parsons, is to assign to (attributive) adjectives the uniform category assignment of CN|CN, that is, functions from CN-meanings to CN-meanings, regardless of different entailments associated with different classes of adjectives; the different semantic entailments are considered not as matters of compositional semantics but as part of lexical semantics associated with individual lexical items and are to be captured by meaning postulates⁹ (see Dowty et al 1981 and Dowty 1985). The advantage of this approach is that we are able to give all (attributive) adjectives a uniform type/category assignment¹⁰ and to treat the divergent entailments not as differences in the type assignment in the lexical semantics.

Under this theory, where all adjectives are given the uniform category CN|CN (or type \(<e,t>,<e,t>\)), it is still possible to associate the meaning of an intersective

---

⁹Meaning postulates, first used by Carnap 1947, are employed by Montague as a device to place constraints on possible or admissible models for interpretations.

¹⁰See Dowty (1985:304).
adjective with a set of individuals, as called for by our semantic intuitions regarding such adjectives. Take the adjective *red* for example. We can use the predicate \( \text{red}'(\text{Existent}) \) (where \( \text{Existent} \) is a property true of each individual in the domain) to represent the set of things that are red in the domain. The intersective entailment associated with such adjectives can then be captured with the meaning postulate in (20).

(21) Meaning Postulate for intersective adjectives:

For all \( P \)'s and all \( x \)'s, if \( \alpha(P)(x) \) is true then \( P(x) \) & \( \alpha(\text{Existent})(x) \) is true, where \( \alpha = \text{red, four-legged, five-foot-tall,} \) and (perhaps) \( \text{healthy, etc.} \)

On the other hand, for adjectives with subsective entailments, the following meaning postulate will ensure the appropriate entailment:

(22) Meaning Postulate for subsective adjectives:

For all \( P \)'s and for all \( x \)'s, if \( \alpha(P)(x) \) is true then \( P(x) \) is true, where \( \alpha = \text{famous, good, skillful, etc.} \)

4.2.3. *Non-intersective Adjectives in Chinese.*

As already shown, however, in Chinese there does not seem to be a lexical class of CN!CN expressions, since adjectival expressions, of both the intersective and subsective kind, cannot be used to modify CN's, as witnessed by the ill-formed (15)–(20), when the marking of *de* is absent. All adjectival expressions appear to be one-
place predicates. Such a category assignment for adjectival expressions of the
intersective type is straightforward. For adjectives with subsective entailments, which
are most naturally represented as having a functional type meaning, we can say all
such adjectives in the language are systematically related to a one-place predicate
category via the rule in (23) and that only the latter category is available (to syntax):

(23) Rule for adjectives of the subsective type:

\(<F, CN|CN, VP>\)

semantics: \(\sigma'(P')\)

\((F_1\) being an identity operation)

The rule inserts a free variable \(P'\) of type \(<e,t>\), which will pick up its interpretation
from the linguistic or pragmatic context in which the adjective occurs. More
specifically, it receives as its denotation the set of individuals that is associated with
the most 'salient' or 'prominent' CN in the linguistic context. Such a CN will naturally
be the CN in the subject NP if the expression occurs at the predicate position. In the
absence of such a contextually present CN, the set of individuals to be the
interpretation of the free variable will be pragmatically determined. (An example of
this will be given below.)

Thanks to the free variable inserted by this rule, inherently functional adjectives
also come to have the type of one-place category. For example, an adjective such as
\(hao\) ('good') which has an inherent type \(CN|CN\) would be linked with \(hao'(P^*)\), which
is a one-place predicate. We can notate the latter as *hao to distinguish the two. Thus, hao, but not hao, is available to syntax. Such an analysis is able to accommodate the inherently functional nature of some adjectival expressions and to account for the predicate-like behavior of all adjectives, namely, (1) their distribution at predicate positions (without being accompanied by copular verbs) and (2) their failure to occur at pre-nominal positions as CN modifiers (e.g., a complex CN such as *gao dalou ('tall building') is ill-formed).

As just mentioned, the free variable will pick up its interpretation from the linguistic or pragmatic context. To demonstrate the latter case, consider a sentence like (24) below. The adjectival expression in the predicate expression will denote youming ('famous'), a one-place predicate (type <e,t>) which contains a free variable. The interpretation of (24) is therefore (25). (Note that hen 'very' is a predicate modifier, denoting a functor from <e,t> meanings to <e,t> meanings\textsuperscript{11}.) In the absence of a contextually present CN, certain pragmatic knowledge enables us to pick the set

\textsuperscript{11}Compare the distribution of hen in the following sentences:

\begin{enumerate}
  \item a. tian hen lan
    \begin{itemize}
      \item sky very blue
      \begin{itemize}
        \item 'The sky is very blue.'
      \end{itemize}
    \end{itemize}
  \item b. wo hen taojian Zhangsan.
    \begin{itemize}
      \item very dislike Zhangsan
      \begin{itemize}
        \item 'I dislike Zhangsan very much.'
      \end{itemize}
    \end{itemize}
  \item c. ta hen dong xinlixue
    \begin{itemize}
      \item he very know psychology
      \begin{itemize}
        \item 'He knows psychology very well.'
      \end{itemize}
    \end{itemize}
  \item d. ta hen hui shuo fawen
    \begin{itemize}
      \item he very can speak French
      \begin{itemize}
        \item 'He can speak French very well.'
      \end{itemize}
    \end{itemize}
\end{enumerate}
of cells as the interpretation for the variable P*.

(24) Mayouyou hen youming
    Yoyo Ma very famous
    'Yoyo Ma is very famous.'

(25) [hen'(youming'(P*))](m)

We therefore account for the observation that the language lacks a lexical category of CN\CN, even though some adjectives seem to have functional meanings. From the point of view of children's acquisition of language, we can perhaps describe the situation here by saying that the non-intersective adjectives are learned as one place predicates which are functors with a free variable (of type <e,t>) plugged in, the interpretation of which is to come from the (linguistic) context.

It should be noted here that Rule (23) is similar to what Flynn 1985 postulates to relate subjective adjectives (of CN\CN category) in English to a predicate type so as to account for the postcopular occurrences of such adjectives. To compare, Siegel 1976b proposes to treat such adjectives in the predicate position as having some kind of contextually supplied, but phonetically empty, CN argument, similar to his treatment for the long form adjectives in Russian in Siegel 1976a. In Russian, qualitative adjectives have the long and short forms which are systematically related to each other via their endings; the former occurs in the pre-nominal attributive and predicate positions while the latter occurs only in the predicate position. In Siegel's analysis,
while the short form has the type of a predicate, the long form has the lexical type of CN/CN. A dummy CN and a phonetically empty determiner is postulated when such long forms occur in the predicate positions. See Flynn 1985, however, for the problems associated with the postulation of phonetically empty CN's (and determiners).

In summary, then, Chinese does not seem to 'grammaticize' a lexical category of CN modifying expressions (CN|CN) in the sense that it does not have this category as a possible lexical category. No expressions can be used to modify CN's without certain extraneous marking, even though the meanings of certain expressions are best construed as CN|CN.

4.2.4. *de* as a Category Change Marker.

Having established the lack of a lexical category of CN modifying expressions, it is my proposal that all CN modifying expressions in the language are to be derived syntactically from one-place predicates through a category change rule marked by the particle *de*. That is, the particle *de* in (8)-(13) is the marker of a category changing rule deriving CN|CN expressions from predicate of type <e,t> expressions. The rule in question is formalized as (26) below. (Rule (27) combines a CN and the modifying expression.)

---

Note that *de* has an underlying neutral tone (whereas a morpheme typically carries one of the four tones), and is phonologically leaned toward the preceding morpheme, therefore a likely candidate as a 'phrasal affix' of Nevis 1986.
(26)<F₃; VP; CN|CN>
semantic operation: λQλx[α(x) & Q(x)]
(F₃ = suffixing de)

(27)<RCON; CN|CN, CN>
semantic operation α'(β')

For intersective adjectives such as hong ('red') or liu-chi-chang (6-foot-long),
the rule is straightforward. For the subjective adjectives, as already discussed, the free
variable is to receive its interpretation from the linguistic or pragmatic context. When
they occur at the pre-nominal position, as opposed to the predicate position, the
property, or set of individuals, to be the interpretation of the variable is most naturally
the property denoted by the CN which the adjective combines with, as this CN is the
most salient CN in the (linguistic) context. In the following example, P* will be the
property of Cellist'.

(28) youming de datiqinjia
    famous  DE celist
youming  famous'(P*)
youming de λQλx[famous'(P*)(x) & Q(x)]
youming de datiqinjia  λx[famous'(P*)(x) & cellist'(x)]

It might be suggested that the de-marked expressions in sentences such as
(11)–(12) and (15)–(20) are results of a lexical derivational rule. In other words, the
postulated category changing rule is a lexical derivational rule (not unlike, for
example, the ing-affixation in astonishing, amazing, etc., in English) and not a
syntactic rule. The following kind of evidence, in addition to the complete productivity and semantic predictability of the rule, indicates that the input to the rule is phrasal in nature.

(29)a. nei ge (you) congming you huopo de xiaohai that CL CONJ smart CONJ outgoing DE child 'That smart and outgoing child...'

cf. b. nei ge xiaohai (you) congming you huopo that CL child CONJ smart CONJ child 'That child is smart and outgoing.'

(30)a. neixie you qinglaq (you) nenggan de gongren those CONJ hard-working CONJ capable DE worker 'Those hard-working and capable workers...'

cf. b. neixie gongren (you) qinglaq you nenggan those workers CONJ hard-working CONJ capable 'Those workers are hard-working and capable.'

In (29a) and (30a) we find that the input to the category changing rule marked by *de* are conjoined predicate expressions, hence phrasal in nature. If the *de*-marked expressions were derived by a lexical derivational rule (with *de* as a derivational suffix), sentences such as (29a) and (30a), where *de* is missing in the first conjunct, should be ill-formed, yet they are not. This is evidence that conjoined in these sentences are two predicate phrases ((29a) and (30a)) and that it is the whole conjoined predicate phrase which undergoes the category changing rule in question.

---

13 The conjunction construction here is (you) CONJUNCT₁ you CONJUNCT₂, the first *you* ('and') being optional.
Note that it is not possible here to defend the view that the *de*-marked expressions are lexically derived by arguing that the absence of *de* after the first conjunct is the result of a conjunction reduction/deletion of some kind, because we know conjunction reduction is not supposed to delete internal parts of a morphologically complex word.\(^{14}\) Note furthermore that we should also be able to conjoin two *de*-marked expressions (which are of the CN;CN category), with each conjunct being the output of the category changing rule. Sentences like the following show that this is indeed the case:

(31) nei ge (you) congming de you huopo de xiaohai that CL and smart DE CONJ outgoing DE child 'That smart and out-going child...'

(32) neixie (you) qinglao de you nenggan de gongren those CONJ hard-working DE CONJ capable DE worker 'Those hard-working and capable workers...'

4.2.5. Relative Clause–like Constructions.

So far we have covered cases where the pre-*de* expressions, including verb phrases, adjectives, and prepositional cases (e.g., (8)–(13)), can be straightforwardly categorized as predicate expressions. Consider now the following data also exhibiting *de*-marking in complex CN constructions, which seem to involve what can be characterized as object-gapped relative clauses.

\(^{14}\) Conjunction reduction does sometimes delete the internal part of a compound, for example, *in–and out–put* in English. But we know parts of a 'derivationally' derived item can never be deleted as a result of conjunction reduction (e.g., *cheer–and beautiful*).
(33) wo xihuan de(*∅) zuojia
    I like DE author
    'Authors who I like...'

(34) Zhangsan xie de(*∅) shu
    Zhangsan write DE book
    'Books which Zhangsan wrote...'

(35) wo gai de(*∅) fangzi
    I build DE house
    'Houses which I built...'

Examples (33)–(35) exhibit what appear to be relative clauses with a missing object 
(and, in fact, the examples (8a)–(10a) previously presented may look like relative 
clauses with missing subjects). Here we see that the function of de may be somewhat 
comparable to relative pronouns in other languages. Semantically, like relative clauses 
in languages, the de–marked expressions in (8)–(10) and in (33)–(35) generate 
intersective entailments. (In a relative construction 'α that β, the resulting complex CN 
always denotes the intersection of the sets associated with α and β.)

However, it is important to note that these data cannot be treated as variable 
binding, as relativization in languages is usually treated. Treating these data with a 
variable binding rule predicts that any NP in the sentence can be relativized; but in 
these relative clause–like constructions marked by de in Chinese, the 'gap' can only be 
at either subject or object positions. The following data, (36b) and (37b), where the 
object NP of a prepositional phrase is 'relativized,' are ill–formed.
(36) a. wo fang qian zai zhuo-shang
    I put money at table-top
    'I put the money on the table.'
    b. * wo fang qian zai de zhuo-shang
    I put money at DE table-top
    'The table that I put the money on...'

(37) a. wo gen zhangsan shuohua
    I with Zhangsan speak
    'I was speaking with Zhangsan.'
    b. * wo gen shuohua de ren
    I with speak DE person
    'The person I was speaking with...'

An advantage of the present analysis of *de* as a category change marker is that it can easily accommodate such data involving the relative clause–like constructions, since these *de*–marked expressions can also be analyzed as CN|CN expressions derived from one-place predicates (type <e,t> expressions). Data involving what looks like subject relativization, such as (8)–(10), are in fact already covered by the previous rules ((26) and (27)) since the pre-*de* expressions here are clearly of the predicate category. As for what looks like object relativization ((33)–(35)), we would first need a rule, given in (38), which would allow the combination of the subject NP and the TV (= (VP|NP)|NP). In a directional system where slashes encode the direction in which a functor finds its argument, it is possible to use categories such as S\NP and S/NP to distinguish TV–object strings from subject–TV strings in, for example, English, where the former can be assigned to the category of S\NP, which is something that combines with an NP to the left (the subject NP) to give a sentence, and the latter

---

15See Morrill 1987's treatment of relative pronouns in English as (N|N)/(S\NP) for subject relativization and as (N\N)/(S/NP) for object relativization.
can be assigned to the category of S/NP, which is something that combines with an NP
to the right (the object NP) to give a sentence (Morrill 1987). In the noa-directional
system adopted here, we can use single vs. double slashes to distinguish TV-object
construction and subject-TV construction, which are both <e,t> type, the former being
labeled here as S|NP (= VP) and the latter as S||NP.

(38) the subject-TV rule:

<RCO; <S|NP||NP; NP>; S||NP>

semantics: \( \lambda x (\alpha'(x)(\beta')) \)

(note NP's are of type <e> here)

cf. the subject-predicate rule:

<LCO; <S|NP, NP>; S>

semantics: functional application

cf. the TV-object NP rule:

<RCO; <S|NP||NP, NP>; S|NP>

semantics: functional application

We can now state the 'object relativization' rule responsible for (33)–(35). As with
Rule (26), this rule also derives a CN|CN expression and is accompanied by the same
semantic operation, which captures the intersective entailment generated in the
construction in question. The rule is as follows:

(39) 'object relativization' rule

<FR; S||NP; CN|CN>

semantics: \( \lambda P \lambda x \left( (\alpha'(x)) \land P(x) \right) \)
Thus all CN modifying expressions in the language are results of category changing rules (Rule (26) and (39)) in syntax marked by *de*. That is, *de* marks rules which derive <e,t>,<e,t> type expressions from <e,t> type expressions. In a word, the particle *de* (in the data discussed in this chapter) is a CN modifier marker.

4.2.6. Stacking of *de*-marked Expressions at the Pre-nominal position.

The analysis of all *de*-marked expressions as CN|CN, along with the rule which combines a CN|CN and a CN (via RCON), predicts that these expressions can be intermingled, i.e., stacked in any order, at pre-nominal positions. This is borne out by the following data:

(40) nei tiao changfangxing de bai de ditan duoshaoqian that CL rectangular DE white DE carpet how-much 'How much is that rectangular white rug?'

(41) yi ge dai maozi de wo bu renshi de one CL wear hat DE I not know DE ren lai zhao ni person come look-for you 'A person wearing a hat who seemed a bit strange came to look for you.'

(42) wo renshi yi ge youming de dong zangwen de I know one CL famous SE know Tibetan DE hanxuejia Sinologist 'I know a famous Sinologist who knows Tibetan.'

(43) wo renshi yi ge dong zangwen de youming de I know one CL know Tibetan DE famous DE hanxuejia Sinologist

---

16This position is very close to that of Zhu 1961 (pp. 77).
'I know a famous Sinologist who knows Tibetan.'

(44) nei ge youming de de quo nuobeierjiang de that CL famous DE receive ASP Nobel-prize DE zuotian dao de wulixuejia ni jianguo le ma yesterday arrive DE physicist you see ASP Q 'Have you seen the famous physicist who has won the Nobel prize who arrived yesterday?'

Sentence (40) shows a series of adjectival expressions at the pre-nominal position; (41) shows a series of relative clause-like de-marked constructions; and (42)–(44) further show that adjectival expressions and relative clause-like expressions can be stacked (in any order) at the pre-nominal position. Such a distributional pattern of these de-marked CN modifying expressions is predicted by our analysis of all of them (e.g., all the underlined strings in (40)–(44)) as CN|CN, which combines with a CN to give another CN.¹⁷. In analyzing relative clauses headed by relative pronouns in English as CN|CN expressions, we have to somehow rule out expressions with stacked relative clauses and adjectives such as *Albanian and who Dorothy hugged student, *student Albanian and who Dorothy hugged, etc., even though they are predicted (see Keenan and Faltz 1985).

The following analysis tree demonstrates how a complex CN, such as the one in (43), where a noun is precede by a VP and an adjectival expression of the

¹⁷In comparison, it is unclear how a phrase structure syntax such as the X-Bar theory, as adopted in Jackendoff 1977, where relative clauses are daughters of N" and adjectives are daughters of N', can deal satisfactorily with the stacking effect demonstrated above in any straightforward manner.
subsective type (both categorized as predicates here), is built up and interpreted.

(45) **dong zangwen de youming de hanxuejia**
    know Tibetan DE famous DE Sinologist
    'a Sinologist who is famous and who knows Tibetan'

(English glosses are used in the translations for clarity)

[dong zangwen de youming de hanxuejia]_{CN}
'know-Tibetan-DE-famous-DE-Sinologist'
\[\lambda x[(\text{know(Tibetan)}(x) & \text{famous(P)}(x)) \& \text{sinologist}(x)]\] (after lambda conversion)

/d

[dong zangwen de]_{CN;CN}
'know-Tibetan-DE'
\[\lambda P \lambda x[(\text{know(Tibetan)}(x) & P(x))]\]

[youming de hanxuejia]_{CN}
'famous-DE-Sinologist'
\[\lambda x[(\text{famous(P)}(x) & \text{Sinologist}(x)]\]

[dong zangwen]_{VP}
'know-Tibetan'
\[\lambda x(\text{know(Tibetan)}(x))\]

[youming de]_{CN; CN}
'famous-DE'
\[\lambda Q \lambda x[(\text{famous(P)}(x) & Q(x))]\]

hanxuejia_{CN}
'Sinologist'
\[\lambda x(\text{famous(P)}(x))]\]

[youming]_{VP}
'famous'
\[\lambda x(\text{famous(P)}(x))]\]

In (45), the most 'salient' or 'prominent' CN should be the CN immediately following the adjective youming ('famous'). The free variable in the denotation of the adjective therefore picks up the set of individual as its denotation. The meaning of the whole complex CN is thus: 'Sinologist who knows Tibetan and who is famous (as a Sinologist).

Potential ambiguities arise when a (subsective) adjective such as youming ('famous') is followed by other CN modifying expressions. This is the case with the
complex CN in (46), where *youning* ('famous') is ordered before other modifying expressions. The complex CN in (46) could mean 'physicists who are famous and who had won the Nobel prize' or 'famous physicists who had won the Nobel prize'.

(46) youming de de guo nuobeierjiang de wulixuejia famous DE win ASP Nobel-Prize DE physicist

'physicists who are famous and who had won the Nobel Prize' or 'Nobel-prize-winning physicists that are famous'.

\[
\lambda x [\text{famous}(P^*)(x) \land \text{win}(n)(x) \land \text{physicist}(x)] \quad \text{(after lambda conversion)}
\]

\[
\lambda Q \lambda x [(\text{famous}(P^*)(x) \land Q(x))]
\]

\[
\lambda x \text{famous}(P^*)(x)
\]

\[
\lambda x \text{win}(n)(x) \land \text{physicist}(x)
\]

\[
\lambda x \text{win}(n)(x)
\]

\[
\lambda x \text{nuobeierjiang}(x)
\]

It is possible to account for the ambiguity here by appealing to the two available choices which we have for assigning interpretation to the free variable in the adjective *youning* ('famous'): both the set of individuals denoted by the CN *wulixuejia* ('physicists') and the set denoted by the CN *de guo nuobeierjiang de wulixuejia*
('physicists who had won the Nobel Prize') can be what the free variable picks up as its interpretation, since both CN's can be construed as 'salient' in the context: they both follow the adjective in question. Hence the ambiguities: if the former is picked, we get the meaning 'physicists who are famous and who had won the Nobel Prize' and if the latter is picked, we get the meaning 'Nobel-prize-winning physicists that are famous'. Compare (46) with (47) in the following, where the adjective youming ('famous') is ordered immediately before the CN wulixuejia ('physicists'). No ambiguities are generated here, as there is only one likely CN the denotation of which is what the free variable (in the denotation of the adjective) picks up as its interpretation, namely, the simple CN immediately following the adjective, i.e., wulixuejia ('physicists').
It should be mentioned here that C. Huang 1989 represents the latest and most detailed analysis of *de* in complex CN constructions\(^{18}\), in which Huang attempts to give a unified analysis to all NP constructions marked by *de*, including the Possessive construction and the Appositional construction, which the present analysis does not incorporate. In his analysis, *de* is treated directly as an intersection or meet operator.\(^{19}\)

\(^{18}\)Other work on this topic includes Zhu 1961, Chao 1968, Li and Thompson 1981, and Chu 1983, etc.

\(^{19}\)As in the analysis here, Huang also assumes that the denotation of subsective adjectives contains a contextually specified variable (of the property type). He further assumes that a rule which may be of a semantic or pragmatic nature can be postulated to the effect that the free variable will have as its value the property that is the denotation of the head noun in the complex CN.
while in the approach here *de* is the marker of category changing rules deriving the
category of CN|CN. Both approaches in fact accommodate the semantic properties of
*de*-marked expressions in quite similar ways. However, an important advantage of the
present approach is that by assigning the *de*-marked expressions the category of
CN|CN, the stacking of *de*-marked expressions in the pre-nominal position (e.g.,
((40)--(44) above) is readily predicted, whereas such data are not obviously predicted
by Huang's analysis.

4.2.7. Possible Counterexamples.

Above I have established that the language lacks a lexical CN|CN and that all
CN modifying expressions are derived via syntactic category changing rules marked by
*de*. The following data, where *de* is omitted in complex CN's, would seem to
constitute counterexamples since under the analysis here they should be ill-formed yet
they are not.

(48)a. yi ge congming ren
     a CL smart person
     'a smart person'

(49)a. hao shu
good book
 'good book(s)'

(50)a. lan tian
     blue sky
     'blue sky'
As already pointed out in the literature (Chao (1968:285), Norman (1988:161) and C. Huang (1989:83 & 102)), these should be regarded as lexicalized compounds and not as syntactically derived phrases. The first reason is that (1) such combinations are idiosyncratic in nature\(^{20}\), as demonstrated in the following:

\[\begin{align*}
(51)\text{a.} & \quad \text{yi ge yuben ren} & \text{cf.} \quad \text{yi ge yuben de ren} \\
& \quad \text{a CL stupid person} & \text{a CL stupid DE person} \\
& \quad '\text{a stupid person}' & \quad '\text{a stupid person}' \\
(51)b. & \quad \text{yi ge congming xiaohai} & \text{cf.} \quad \text{yi ge congming de xiaohai} \\
& \quad \text{a CL smart child} & \text{a CL smart DE child} \\
& \quad '\text{a smart child}' & \quad '\text{a smart child}'
\end{align*}\]

\[\begin{align*}
(52)\ast & \quad \text{gui shu} & \text{cf.} \quad \text{gui de shu} \\
& \quad \text{expensive book} & \text{expensive DE book} \\
& \quad '\text{expensive book(s)}' & \quad '\text{expensive book(s)}'
\end{align*}\]

\[\begin{align*}
(53)\text{a.} & \quad \text{lan tiankong}\^{21} & \text{cf.} \quad \text{lan de tiankong} \\
& \quad \text{blue sky} & \text{blue DE sky} \\
& \quad '\text{blue sky}' & \quad '\text{blue sky}'
\end{align*}\]

Furthermore, (2) when the pre-nominal adjective is modified by an adverb, the presence of *de* becomes obligatory (e.g. (51)–(52)); if the adjective–CN combinations in (51a) and (52a) are syntactically derived, such modification should be possible.

\[\begin{align*}
(54)\text{a.} & \quad \text{hen hao shu} \\
& \quad \text{very good book}
\end{align*}\]

\(^{20}\)The combinability of an adjective and a noun into a such a compound sometimes depends on the number of syllables in the adjective and noun in question. Further research, however, is needed to determine the exact nature of this phonological factor.

\(^{21}\)There is no meaning difference between the monosyllabic *tian* and bisyllabic *tiankong*, which are both glossed as 'sky'.
'very good book(s)'

cf. b. han hao de shu
    very good DE book
    'very good book(s)'

(55)a. * bijiao congming ren
    more smart person
    'smarter people'

cf. b. bijiao congming de ren
    more smart DE person
    'smarter people'

4.2.8. Privative Adjectives.

There is a third class of CN modifying expressions which generate neither
intersective nor subsective entailments which have not been mentioned thus far. In
English, these include what may be called 'negative' adjectives, such as fake or false,
or 'conjectural' adjectives, such as alleged or ostensible, as well as expressions such as
former. For example, a fake diamond can not be said to be in the intersection of the
set of things which are fake and the set of things which are diamonds. Nor can we
infer that it is even a diamond at all. Similar properties apply to an expression such as
a former senator. These adjectives when combined with CN's give rise to one of the
intensional constructions. To know the meaning of say, former senator (at a given
index), requires knowing the intension of the set denoted by the noun, that is, the
function from indexes (or time-world coordinates) to sets of individuals (=
extensions), particularly, the denotation of senator at indices with earlier time
coordinates.
Adjectives of this variety, as shown in (56)-(59), also occur in Chinese and they are marked by de as well.

(56) yi-zhi jia de (*Ø) laolishi
    one-CL false DE Rolex
    'a false Rolex'

(57) weizao de (*Ø) caopiao
    fake/counterfeit DE bill
    'counterfeit bills'

(58) qianren de (*Ø) xizhuren
    former DE chairman
    'former chairman'

(59) suowei de (*Ø) aigufenzi
    alleged DE patriots
    'alleged patriots'

These adjectives present a problem for the present analysis, where de is treated as the marker of category changing rules deriving CN|CN expressions with intersective or subsective entailments, depending on the pre-de expression, as well as for C. Huang 1989's analysis, where de is a meet operator. Huang, however, does not acknowledge the existence of data such as (56)-(59) and instead cites the ill-formed expressions in (60)-(61) to support his analysis of de as a meet operator. According to him, (60)-(61) are ruled out because they are 'semantically anomalous'; the presence of de will yield false intersective inference (pp. 84).

(60)* wei de caopiao
    fake DE bill
    'fake bills'

(61)* qian de guohui-yiyuan
former DE congressman
'former congressman'

The unacceptability of (60) and (61), under closer examination, however, turns out to have nothing to with any semantic factors. They are unacceptable because the pre–de element in them, wei and qian, are bound forms in the language. That is, they can only occur as part of a (multi–syllabic) word and cannot occur alone. (See the discussions in Chapter 2 regarding bound forms.) This is clearly evident by contrasting the unacceptable (60) and (61) and the perfectly well–formed (57) and (58). In the latter examples, the bound forms wei and qian occur as a proper subpart of the related compounds weizao ('fake–made = counterfeit') and qianren ('former–term =former'), respectively.22

The existence of such expressions presents a problem for any attempt to give a unified analysis of the pre–nominal de–marked expressions. Their distribution is such

---

22Huang does recognize explicitly the existence of expressions such as qianren de laoshi 'former teacher', or qianren de guohuiyiyuan 'former congressman' (pp. 104). However, he appeals to some sociological and cultural factors for explanations, citing as evidence the possibility for a Chinese speaker to address a former teacher as Teacher and a former president as President to show a Chinese speaker may assign a different meaning to an expression like qianren ('former') from what is assigned to an expression like former in English (194). However, it is difficult to see how the fact that sometimes a person may be addressed with the title he or she formerly held (both in Chinese and English) can have any bearing on the meaning of expressions like former, and given the data there is no reason to believe any alleged meaning differences exist between qianren ('former') in Chinese and former in English. More importantly, no such explanations are available for other privative adjectives such as weizao ('counterfeit') and jia ('false').
that they should be assigned the category of CN\textsuperscript{\textdagger}CN, which is consistent with the account here, since all \textit{de}-marked expressions in complex CN’s should have this category in the present analysis. However, their semantic property is such that they cannot be the product of the category changing rule postulated in (26). Unfortunately, I do not have a satisfactory explanation for why these expressions are obligatorily marked by \textit{de}, except to say that there are only a small number of such adjectival expressions in Chinese and therefore we may simply treat them as idiomatic lexical expressions, the meanings of which are functions from intension of sets of individuals (i.e. functions from indexes to sets of individuals) to sets of individuals, that is, of type \langle\langle s,\langle e, t\rangle\rangle, \langle e, t\rangle\rangle.
CHAPTER V
A TVP ACCOUNT OF THE CHINESE RESULTATIVE CONSTRUCTIONS

5.0. Introduction.
In this chapter, I will give an analysis of a set of data from Chinese loosely known as the Resultative Construction (RC's hereinafter), which is a class of constructions involving 'missing' noun phrases in predicate structures which are marked by the Resultative particle de. These constructions have been subject to different analyses in various pre-theoretical and theoretical contexts (by Tang 1977, P. Wang 1970, and most recently and notably in J. Huang 1989 within the Government Binding framework). I propose to analyze RCs along the line of the Bach-Dowty analysis of TVP (Transitive Verb Phrase) (Bach 1979, Dowty 1982a). It will hopefully be shown that RCs can be successfully analyzed in Categorial Grammar and that these data lend important support to the TVP category in the Universal Grammar.

5.1. The Resultative Construction in Chinese.
The Resultative Construction in Chinese refers to a group of constructions which are complex predicative structures involving a predicative complement with a missing subject NP. RCs therefore look like control verb constructions in English and other languages. Moreover, unlike the latter, they are obligatorily marked by a
particle, namely, *de*, which, as we have seen, also figures in many other constructions in Chinese syntax. We may call this instance of *de* the Resultative *de*. Items (1)–(8) below exemplify three different varieties of the RC. Consider the pattern of RC exhibited in (1)–(3) first, the bold-faced strings as the predicative structures in question:

(1) Zhangsan *ku de hen shangxin*
    Zhangsan cry DE very upset
    'Zhangsan cried such that he became very upset.'

(2) Zhangsan *gaoxing de ku le*
    Zhangsan happy DE cry ASP
    'Zhangsan was so happy that he cried.'

(3) Zhangsan *qi de ti ziji yi-jiao*
    Zhangsan angry DE kick self one-leg
    'Zhangsan was (so) angry that he kicked himself.'

In (1)–(3) we have an IV followed by the particle *de*, the presence of which is obligatory, and a predicative complement. The complement is controlled by the subject of the matrix verb in the sense that (1) the 'understood subject' of the complement is co-referential with the subject of the matrix verb and (2) the subject NP triggers the reflexive within the complement (e.g., (3)). It should be noted that these two sets of facts are the only manifestations of the controller relationship

---

1. ASP = aspect marker.

2. Note that the predicative complement is in its base form and not in a derived form such as Gerund or Infinitive.

3. Note in Chinese there is no number or gender agreement between the NP that triggers the reflexive pronoun and the reflexive pronoun.
between the matrix subject and the complement as there is no agreement of any kind in Chinese.

Sentences (4) and (5) illustrate a second pattern of the RC, with an extra NP marked by the object case particle ba. (6)–(7) illustrate yet another pattern, where an NP is also present in the upper clause and the matrix verb is a transitive verb instead of an intransitive verb as in (1)–(5). In both sets of examples, the extra NP is crucially marked by the (object) case particle ba, which will be discussed later. Semantically, it is this NP which is the controller of the predicative complement in question.

(4) Zhangsan ba Lisi ku de hen shangxin
Zhangsan BA Lisi cry DE very upset
'Zhangsan cried such that Lisi became very upset.'

(5) Zhangsan ba shoupia ku de shidadade le
Zhangsan BA handkerchief cry DE wet ASP
'Zhangsan cried and got the handkerchief all wet.'

(6) Zhangsan ba Lisi ma de ku le
Zhangsan BA Lisi scold DE cry ASP
'Zhangsan scolded Lisi such that Lisi cried.'

(7) Zhangsan ba ma qi de hen lei
Zhangsan BA horse ride DE very tired
'Zhangsan rode the horse and the horse got very tired.'

(8) Zhangsan ba Lisi da de liuxie
Zhangsan BA Lisi beat DE bleed
'Zhangsan beat Lisi such that Lisi bled.'
5.2. The TVP Category.

The category of Transitive Verb Phrase (TVP) (which is a complex transitive verb and not a phrase containing a transitive verb) has been proposed for expressions in English such as those listed in (9), which contrast in their syntactic behaviors with expressions such as those listed in (10), which have the category of TVP (see Partee 1976a, Bach 1979, Dowty 1982a, Dowty 1982b, and Jacobson 1987).

(9) persuade to leave
    paint white
    put into a box
    consider a friend

(10) promise Mary
    strike me

A TVP has the same combinatorial properties as a TV; syntactically it combines with an NP to form an VP. The TVP analysis of the expressions in (9) is motivated by the Categorial theory of grammatical relations (Dowty 1982a, 1982b and Bach 1979, 1980). Here the (direct) object is the NP that a transitive verb combines with to form a VP. By categorizing expressions as TVP, we identify the NP to be combined with it (via the operation of wrapping) as the direct object. An important advantage of this approach is that we are then able to account for the passive properties of these verbal expressions, as opposed to expressions in (10).

(11) Mary was persuaded to leave.
(12) John was considered a friend.
(13)* Mary was promised to leave.
(14) I was struck (by John) as unhappy.

Furthermore, the TVP analysis enables us to explain the different behaviors between verbs such as persuade, which in the TVP analysis has the category of TV/VP, and promise, which is assigned to the category of (VP/VP)/NP, in particular, the control properties of verb-infinitival complement constructions involving these verbs. The control properties of these verbs can be characterized by what can be called Bach's Principle (Bach 1979, Dowty 1982a). This principle states that the controller of the infinitival complement --- i.e., the NP that triggers the reflexive agreement or number agreement within the complement and which is also involved in semantic entailments with respect to the infinitival complement --- is the next higher argument of the verb after the verb is combined with the complement. In other words, the predicative complement is controlled by the object of a TVP, since it is the next argument to be combined with the TVP, and by the subject of the IVP (=VP), since it is the next argument to be combined with the complex predicate in question. For example, this correctly identifies the controlling NP's in the sentences in (15) and (16). (17) and (18) represent the argument structures of (15a) and (16)a, respectively.

(15)a. John persuaded Mary to leave.
   b. I consider John as happy with himself.
(16)a. John promised Mary to leave.
   b. John strikes me as very proud of himself.

(17) The argument structure for (15)a:
    [John persuaded Mary to leave] s
        /          \
      John  [persuaded Mary to leave] iv
(18) The argument structure for (16)a:

\[ \text{[persuaded to leave]}_{TV} \quad \text{Mary}_{NP} \]
\[ \text{persuade}_{TV/IV} \quad \text{leave}_{IV} \]

(The analysis of the verb–complement structure in sentences like (11) and (12) as a TVP, which is discontinuous at surface, therefore offers explanations for what has come to be known as Visser's Generalization (Visser 1973): object control verbs have corresponding passives and subject control verbs do not. The object control verbs are those that combine with an infinitive complement to form a TVP, which is predicted to have a corresponding passive.

5.3. Resultative Predicates as TVPs.

It is my proposal here that the complex predicative structure (i.e., the \([V \ de \ VP]\) sequence) demonstrated in (4)–(8) is an instance of the category of TVP, while the complex predicative structure in (1)–(3) is an instance of IVP (or VP). In the following, I will motivate the transitivity of the complex predicative structures in (4)–(8). Before proceeding, however, I should point out that J. Huang 1989's account of the various Resultative Constructions also relies on the same crucial distinction between what he calls 'intransitive and transitive V' phrases' to account for the subject–
control sentences (e.g., (1)–(3)) and the object–control sentences (e.g., (4)–(8)). However, his account is carried out within the Government and Binding framework and therefore markedly differs from the present account, which is within context of Categorial Grammar and Montague Semantics, not only with respect to the mechanisms and theory–specific principles that are available (e.g., Huang's account relies on Case theory and on mechanisms such as head movement while the present account relies heavily on proper categorizations of complex predicates and their translations into lambda expressions), but also with respect to the representations of structures of sentences and to the nature of syntactic categories. As the object here is not a cross-theoretic study, though it would be a worthy project, I will not attempt to compare Huang's analysis with the analysis given here.

The transitivity of the complex predicate structure (i.e., the [V–de–VP] sequences) in (4)–(8) can be motivated on the grounds that it exhibits all the properties of simple TV's and that the extra NP argument (e.g. Lisi in (4) and shoupa (‘handkerchief’) in (5)) exhibits characteristics of an object NP argument.

The most important evidence of the transitive nature of the [V–de–VP] sequence in (4)–(8) comes from the presence of the particle ba, which is a particle marking the (direct) object NP (L. Wang 1958, Thompson 1973, and Li and Thompson 1981, etc). In general, the word order of a TV and an object NP in Chinese may be
either VO or OV⁴. The particle ba obligatorily marks the object NP when (1) the object NP precedes the verb⁵ and (2) when it is associated with the thematic role of Patient or with an 'Affected' Theme.⁶

⁴It is still controversial as to what are the factors determining whether the correct order is OV or VO in any given case; there are classes of verbs which require VO order and others which require OV order and often either order is possible. Furthermore, discourse notions such as definiteness and indefiniteness of the NP's involved also play a role. I will not pursue this topic here.

⁵When an object NP follows the verb, it must occur in its bare form.

a. ta da le (*ba) Zhangsan yi-dun
   he beat ASP BA Zhangsan one-CL(assifier)
   'He beat up Zhangsan.'

b. wo ba ta kanjian le
   I BA he see ASP
   'I saw him.'

⁶This observation that a ba-marked NP must be associated with the Patient or 'Affected' Theme role is often made in the literature, e.g. Wang 1958, Thompson 1973, Teng 1975, & J. Huang 1989, on the basis of data such as the following:

a.* Zhangsan ba Lisi xihuan
   Zhangsan BA Lisi like
   'Zhangsan likes Lisi.'

b.* wo ba ta kanjian le
   I BA he see ASP
   'I saw him.'

I will adopt this view here, though I believe further research needs to be done to support it since we also have data like the following where the ba-marked object NP is not associated with Patient or Affected Theme.

c. wo ba tade mingzi wangji le
   I BA his name forget ASP
   'I forgot his name.'

d. wo ba zhexie haomia dou bei qilai le
   I BA these number all memorize up ASP
   'I have memorized (all) these numbers.'

Moreover, the reason that (a) and (b) are ill-formed may not be due to the marking of ba, but simply due to the fact that these verbs require the object NP to occur to the left of the verb, as (a) and (b) are bad even when the object NP is not marked by ba.

(continued...)
(19) wo ba dianshi mai le
    I BA TV sell ASP
    'I sold the TV.'

(20) wo ba cha he le
    I BA tea drink ASP
    'I drank the tea.'

(21) ta ba Zhangsan da le yi-dun
    he BA Zhangsan beat ASP one-CL
    'He beat up Zhangsan.'

In a word, *ba* is an obligatory marker of pre-verbal object NP's which are
associated with the thematic role of a patient or an affected theme. Therefore, the
analysis of the complex predicate as TVP will predict that the pre-verbal NP
arguments in (4)–(8), which, as we can see, are all associated with a Patient or
Affected Theme role, have to be marked with *ba*. This is indeed the case, as
evidenced by the ungrammaticality of the sentences where the NP is not so marked
((22)–(26)) in contrast to the grammaticality of sentences in (4)–(8) where the NP is

...(continued)

a'. * Zhangsan Lisi xihuan
    Zhangsan Lisi like
    'Zhangsan likes Lisi.'

b'. * wo ta kanjian le
    I he see ASP
    'I saw him.'

cf.

a". Zhangsan xihuan Lisi
    Zhangsan like Lisi
    'Zhangsan likes Lisi.'

b". wo kanjian ta le
    I see him ASP
    'I saw him.'
marked by *ba*.

(22)* Zhangsan Lisi *ku de hen shangxin*
   Zhangsan Lisi cry DE very upset
   'Zhangsan cried such that Lisi became very upset.'

(23)* Zhangsan *shoupa ku de shidadade le*
   Zhangsan handkerchief cry DE wet ASP
   'Zhangsan cried and got the handkerchief all wet.'

(24)* Zhangsan Lisi *ma de ku le*
   Zhangsan Lisi scold DE cry ASP
   'Zhangsan scolded Lisi such that Lisi cried.'

(25)* Zhangsan ma *qi de hen lei*
   Zhangsan horse ride DE very tired
   'Zhangsan rode the horse and the horse got very
tired.'

(26)* Zhangsan Lisi *da de liuxie*
   Zhangsan Lisi beat DE bleed
   'Zhangsan beat Lisi such that Lisi bled.'

Note that it is still possible here to argue that *ba* is used to mark the NP as the object of the matrix verb (e.g., *ku* ('cry') in (4) and (5)) and not as the object of the whole complex predicate. What is significant is that in sentences like (4)–(5), the matrix verb is an intransitive verb. Therefore, *ba* cannot possibly be used to mark the NP as the object of the matrix verb, and therefore can only be viewed as marking the NP as the object of the whole complex Resultative predicate. This crucially validates the analysis of the preverbal NP in these data as the object of the whole complex predicate, that is, the TVP.
As already mentioned, the analysis of expressions such as *persuade to leave* or *regard as a friend* in English as a TVP is supported by the fact that such verb phrases have corresponding passives. A TVP analysis of the predicate phrases in (4)–(8) similarly predicts that they have corresponding passives, which is borne out by the following data: (27)–(29) are the passives of (4)–(6), respectively. Note that Passive in the language is formed by 'moving' the object NP in the active sentence to the subject position and meanwhile either (1) marking the 'original' subject NP with the passive particle *bei* ('by'), for agentive Passive, or (2) marking the verb with the passive particle, for agentless Passive. The TVP's in question, as shown in (27)–(29), have both kinds of Passives.

(27)a. Lisi bei Zhangsan ku de hen shangxin
Lisi by Zhangsan cry DE very upset
'Lisi was made very upset by Zhangsan's crying.'

b. Lisi bei ku de hen shangxin
Lisi by cry DE very upset
'Lisi was made very upset by someone's crying.'

(28)a. shoupa bei Zhangsan ku de shidadade
handkerchief by Zhangsan cry DE wet
'The handkerchief was made all wet by Zhangsan's crying.'

b. shoupa bei ku de shidadade
handkerchief by cry DE wet
'The handkerchief was made all wet by Zhangsan's crying.'

(29)a. Lisi bei Zhangsan ma de ku le
Lisi by Zhangsan scold DE cry ASP
Lisi, was scolded by Zhangsan such that he, cried.'

b. Lisi bei ma de ku le
Lisi by scold DE cry ASP
Lisi, was scolded by Zhangsan such that he, cried.'
Furthermore, an object NP in Chinese can be topicalized. Thus the following topicalization data also support the analysis of the NP in question as the object of the complex predicates.

(30) nei tiao shoupa, Zhangsan ku de shidadade that CL handkerchief Zhangsan cry DE wet
    'The handkerchief got all wet because of Zhangsan's crying.'

(31) nei ge xiaohai, Zhangsan ma de ku le that CL child Zhangsan scold DE cry ASP
    'That child, Zhangsan scolded and got him to cry.'

Finally, note that in Chinese the category of TVP is continuous whereas in English TVP's such as persuade to leave or regard as a friend are always discontinuous at surface, though showing up as a continuous constituent in passives, hence the incorporation in Categorial Grammar of wrapping as an operation side by side with concatenations whereby expressions are combined together. The Chinese data show that in some languages concatenation instead of wrapping is used in putting together these TVP's with their object NP's and that it is merely accidental that TVP's in English show up only in discontinuous forms.

---

2=classifier.

To maintain a TVP analysis without invoking wrapping, Jacobson 1987 proposes within the generalized phrase structure grammar framework a verb raising process which promotes the verb in the TVP to become its sister.
5.4. **Semantics of the RCs.**

After motivating the transitivity of the complex Resultative predicates in (4)–(8) (as opposed to those in (1)–(3), which are IVP's), we now have to account for the semantic inference associated with these structures, which is captured in Bach's Generalization discussed above, namely, that the 'understood' subject of the verbal complements missing in these predicates is the object NP marked by *ba*, in contrast to the control pattern in (1)–(3), where the understood subject is the subject of the matrix verb$^9$.

First of all, as can be seen by the above categories assigned to the *de*-marked verbs, we are assuming that these subjectless predicative complements are VPs and non-clausal, an assumption mandated by a non-abstract syntactic theory. Under this assumption, there are still two ways to deal with the control properties (or the 'understood subject' effect) exhibited in these verb complement structures. One is to directly represent them in the compositional semantics. The other is to account for them as differences in the lexical meaning of the various verbs involved. The first approach, where we build some device into the interpretive procedure such that the 'missing' subject of the complement is plugged in at some point, is advanced in, for example, Klein and Sag 1982, where they proposed some special combinators to combine the control verb and its complement so as to reconcile the differences

---

$^9$This is the generalization regarding the RCs stated in J. Huang (1984:6): 'The missing subject of a resultative complement must refer to a *ba*-NP in the main clause if there is a *ba*-NP. Otherwise, it must refer to the main clause subject.'
between the desirable semantic type of the verb and the syntactic or distributional properties of the verb. Such an approach directly captures the differences in the meanings and control properties of the following verbs at the level of semantic structures as induced by compositional semantic rules.

(32) John wishes to leave.
    wish'(leave'(j))(j).
    John tends to win.
    tend'(win'(j))
    John persuaded Mary to leave.
    persuade'(leave'(m))(m)(j)
    John promised Mary to leave
    promise'(leave'(j))(m)(j)

Another way to capture the relevant entailments is to treat them as part of the lexical meanings of the individual verbs exhibiting different control patterns, via meaning postulates schemata (Bach 1979, Dowty 1982a, Chierchia 1984, 1985). To capture the meaning of, for example, try to leave, we can have a meaning postulate to the effect that for every \( \alpha \) and for every property \( P \), if \( \alpha \) tries \( P \), then \( \alpha \) tries to bring about \( P(\alpha) \). The contrast between, for example, persuade and promise, is then accounted for by the meaning postulates associated with these two verbs, to the effect that the predicative complement will be linked with the object of the TVP for the former and with the subject of the IVP for the latter; the differences between these control verbs are no longer represented in their compositional semantic structure.
While Dowty 1982b convincingly argues that the assumption about the difference in the semantic structures associated with different control (and equi) verbs may be gratuitous, the very different nature of the Resultative constructions, as compared to the control verb constructions in English, calls for an approach where the relevant semantic inferences are captured at the level of compositional semantics. While the different control patterns in the verb–complement constructions in English depend on the particular verbs involved, hence construable as part of lexical properties of particular verbs, the three Resultative constructions demonstrated in (1)–(8) (and other patterns of it which will be discussed later) are not item–specific and are completely productive. Any IV or TV can be freely distributed in the RCs. We can put any IV (in the variety of RCs shown in (1)–(5)) or any TV (in the variety of RCs shown in (6)–(8)) marked with the particle *de* together with a predicative complement to derive a Resultative predicate (which may be an IVP, e.g. (1)–(3), or a TVP, e.g. (4)–(10)) in a completely productive way.10

In other words, the differences in the control properties in the various RCs are not dependent on the individual verb. They are therefore not lexical in nature and should not be treated as part of lexical properties of specific (classes of) verbs via meaning postulates associated with particular verbs. Furthermore, these constructions

---

10Note that if we treat the second VP in the RCs as adjuncts, which is a possible analysis of these data (though not adopted here — more discussions on this later), we have another good reason not to treat the control properties of the second VP as part of lexical meanings of the verbs, as adjuncts are simply optional elements.
are obligatorily marked by *de*. This suggests that the control properties are in a sense built into these *de*-marked verbal constructions and are not part of the lexical meanings of individual verbs (in the form of meaning postulates). In other words, we should want to capture the relevant semantic entailments directly in the compositional rules responsible for these *de*-marked constructions.

There are two ways to implement this. Take the RC in (4)-(5) for example. We can either (1) associate an appropriate semantic operation with a syntactic rule (which invokes the operation of affixing *de*) whereby an IV and a VP are combined to form a TVP, or (2) associate the semantic operation in question with a one-place syntactic rule invoking the same operation, which turns an IV to a TVP|VP. The differences between the two may not be crucial, but I will adopt the latter approach here for the simple reason that the latter is more in line with Categorial Grammar by virtue of the fact that in this approach combinatory properties are encoded in category labels. In the former approach we would have to postulate a two-place rule which specifies that an IV and a VP combine to form a TVP, weakening a strong claim of the theory, namely, that all two place rules are in the form of functional application (or functional composition).

\[\text{Even though in the system here we will still need syntactic rules to spell out the operations invoked to combine the categories (e.g. RCON, WRAP, etc).}\]
We can now move on to state the one-place rules in question which affix *de* syncategorematically to an IV or TV. These one-place rules are accompanied by appropriate semantic operations which directly capture the different control properties in the different patterns of the RC's all marked by *de*. As already discussed, there are three patterns of RC's exhibited in (1)-(8). This means there are three such rules deriving the *de*-marked verbal expressions. For the RC's in (1)-(3), an IV occurs as the matrix verb in the [V-*de*-VP] sequence, which is an IVP, therefore the *de*-marked verbal expression should have the category of VP"/VP (double slash is used here to distinguish the category from VP modifiers) and the appropriate one-place rule should look like (33a). The accompanying semantic operation ensures the proper control effect. (To represent the Resultative meaning in a simplified manner, I will employ without further discussion an operator RESULT-IN, the nature of which is similar to the logical operators CAUSE and BECOME discussed in detail in Dowty 1979\(^{12}\).)

(33)a. Resultative Rule I: (e.g. (1)-(3))
\[ \langle F_s; VP; VP"/VP \rangle \]
Semantic operation: \( \lambda P \alpha x \left[ a'(x) \text{ RESULT-IN } P(x) \right] \)
\( F_s = \text{suffixing } de \)

\(^{12}\)At this point it is not clear to me what the relationships are between the different aspektual verb classes in Chinese and the Resultative Construction. Since the main emphasis here is on demonstrating how a TVP analysis accounts for the control properties of the various patterns of the Resultative Construction and not on representing the meaning of the Resultative Construction in the fashion of Dowty 1979, I will await for future research for discussions on the latter topic.
To illustrate, *ku-de* ('cry-DE') in (1), according to (33a), is a VPII+VP denoting:

\[ \lambda P \lambda x \text{Cry}'(x) \text{RESULT-IN P}(x) \text{]. (For the sake of clarity, I am using English glosses for all the predicates.) The complex Resultative predicate ku de hen shangxin ('cry-DE-very-upset') will then be of the VP (or IVP) category and will denote the predicate: } \lambda x \text{Cry}'(x) \text{RESULT-IN Very-Upset}'(x) \text{]. When it is combined with the subject NP, we get } \text{[Cry}'(z) \text{RESULT-IN Very-Upset}(z) \text{]} \text{ for (1). The analysis tree for (1) is given in the following.}

(33)b. Zhangsan ku de hen shangxin
Zhangsan cried DE very upset
'Zhangsan cried such that he became very upset.'

\[ [\text{Zhangsan ku } \text{de } \text{hen } \text{shangxin}]_s \text{ [Cry}'(z) \text{RESULT-IN Very-Upset}(z) \text{]} \\
/ \backslash \\
\text{Zhangsan}_{np} \quad [\text{ku de hen shangxin}]_{vp} \\
/ \backslash \\
[ku de]_{vp};[vp] \quad [\text{hen shangxin}]_{vp} \\
\lambda P \lambda x \text{Cry}'(x) \text{RESULT-IN P}(x) \quad \lambda x [\text{Very-Upset}'(x)] \\
\text{ku}_{vp} \]

An important issue needs to be discussed before we proceed to provide additional rules for the other two patterns of RC given above. The way the RC's are analyzed here, the first verb in (1)-(3) undergoes a one-place (or category changing) rule, namely the rule in (32a), which is accompanied by *de*-suffixation, such that it comes to have the category of VP*I*VP, a category which combines with a predicative complement to form a VP (the double slash here is used to distinguish the category from a true VP modifier). The category change rule in question thus can be
considered as a relation-expanding rule of some sort (see Dowty 1982a), as it adds a
verbal complement to be the argument of the verb. It should be noted though, that it
is also conceivable to analyze the first verb and the second verb in the RC as in a
relationship other than that of verb–complement (as I have done here). For example, it
is possible to treat $V_1$ and $V_2$ as in a verb–adjunct relationship. In fact, this is how the
RC’s are generally regarded in the literature. The debates arising from this analysis
have revolved around which verb should be treated as the main predicate and which
the adjunct. The treatment of $V_2$ as the main predicate and $V_1$ as an adverbial adjunct
is called the 'Primary Predication' analysis (C. Huang and Mangione 1985) and the
treatment of $V_1$ as the main predicate and $V_2$ as an adverbial adjunct is called the
'Secondary Predication' analysis (J. Huang 1988b). There seems to be equal evidence
supporting each position (see J. Huang 1988b for a review of the relevant arguments).
Instead of participating in this debate, I have opted to treat the RC’s as exhibiting the
verb–predicative complement structure, leaving a detailed investigation of the possible
empirical and non–theory–dependent differences of these three possible analyses for
further research.

Proceeding now to the second pattern of the RC (e.g. (4)–(5)), we find an IV
marked by *de* combines with a predicative complement to form the category of TVP,
which means the *de*-marked verbal category should be a TVP|VP. The following rule
turns an IV into a TVP|VP, with the semantic operation directly guarantying that the
predicative complement is linked with the object of the TVP:
(34)a. Resultative Rule II: (e.g. (4)–(5))
<Fr: IV; TVP|VP>
Semantic operation: λPλxλy [a'(y) RESULT-IN P(x)]

For example, ku–de ('cry–DE') in (4) is of the category TVP|VP and is
translated into: λPλxλy [Cry'(y) RESULT-IN P(x)] and the complex Resultative
predicate, ku–de hen shangxin ('cry–DE–very–upset'), a TVP under the analysis here,
will then have the translation: λxλy [Cry'(y) RESULT-IN Very-Upset'(x)].

Plugging in the object NP Lisi, we get: λy [Cry'(y) RESULT-IN Very-Upset'(l)].
Finally, the VP in question combines with the subject NP. The whole sentence then
has the meaning: [Cry'(z) RESULT-IN Very-Upset'(l)]. The analysis tree for (4) is
set out in (34b). Note that the two Resultative rules in (33a) and (34a) can be both
applied to an IV and therefore a de–marked verbal expression such as ku de ('cry–DE')
can be either VP|VP or TVP|VP, which have different combinatorial and semantic
properties, as can be seen in (1) and (4).
(34) b. Zhangsan ba Lisi ku de hen shangxin
    Zhangsan-BA-Lisi cry DE very upset
    'Zhangsan cried such that Lisi became very upset.'

\[ [\text{Zhangsan ba Lisi ku de hen shangxin}]_s \]
\[ [\text{Cry}'(z) \text{ RESULT-IN Very-Upset}'(l)] \]
\[ / \]
\[ \text{Zhangsan} \]
\[ [\text{ba Lisi ku de hen shangxin}]_{vp} \]
\[ \lambda y [\text{Cry}'(y) \text{ RESULT-IN Very-Upset}'(0)] \]
\[ / \]
\[ \text{Lisi}_{np} \]
\[ \text{[ku de hen shangxin]}_{tp; vp} \]
\[ \lambda x \lambda y [\text{Cry}'(y) \text{ RESULT-IN Very-Upset}'(x)] \]
\[ / \]
\[ \text{[ku de]}_{tp; vp} \]
\[ \lambda x \lambda y [\text{Cry}'(y) \text{ RESULT-IN P(x)}] \]
\[ / \]
\[ \text{[hen shangxin]}_{vp} \]
\[ \lambda y [\text{Cry}'(y) \text{ RESULT-IN P(x)}] \]
\[ / \]
\[ \text{ku}_{tv} \]

In (6)–(8), we find another pattern of the RCs: a transitive verb marked by de combines with a predicative complement to form a TVP, which means it should also have the category of TVP;VP. The one-place rule deriving the de-marked verbal category here is then (35).

(35) Resultative Rule III: (e.g. (6)–(8))
\[ <F_s; TV; TVP; VP> \]
Semantic operation: \[ \lambda P \lambda x \lambda y [(\alpha'(x))(y) \text{ RESULT-IN P(x)}] \]

For example, da de ('beat-DE') in (8) has the TVP;VP category and denotes:

\[ \lambda P \lambda x \lambda y [(\text{beat}'(x))(y) \text{ RESULT-IN P(x)}]. \]
The Resultative predicate da de liuxie ('beat-DE-bleed'), which is a TVP, is then translated into: \[ \lambda x \lambda y [(\text{beat}'(x))(y) \text{ RESULT-IN bleed}'(x)]. \]
Plugging in the object NP Lisi, we get: \[ \lambda y [(\text{beat}'(l))(y) \text{ RESULT-IN bleed}'(l)]. \]
The meaning of the whole sentence will be \[ [\text{beat}'(l)(z) \text{ RESULT-IN bleed}'(l)]. \]
RESULT–IN bleed'(1)]. (The analysis tree for (8) is omitted here.) Note that the matrix verb in the Resultative predicate phrase is an IV in (1)–(5) (e.g., ku ('cry')) while a TV is involved in (6)–(8) (e.g., ma ('scold'), qi ('ride') and da ('beat')). When the matrix verb is a TV, we see that the ba–marked object NP serves not only as the understood subject of the post–de predicative complement, but also as the understood object of the matrix TV. This effect is directly achieved by the semantic operation associated with the Resultative rule in (35).

While syntactically we treat the predicative complements as VP's, and thus non–clausal, semantically the 'missing' subject is provided for the propositional function corresponding to the predicative complement, that is, the property associated with the predicative complement is saturated by appropriate NP's without an accompanying syntactic rule providing a (morphologically empty) subject argument.

5.5. The Wrapped Resultative Construction.

Consider now (36)–(39), which exhibit an important pattern of the RC's not discussed thus far:

(36) Zhangsan ku de Lisi hen shangxin
Zhangsan cry DE Lisi very upset
'Zhangsan cried such that Lisi became very upset.'

(37) Zhangsan ku de shoupa shidadade
Zhangsan cry DE handkerchief wet
'Zhangsan cried such that the handkerchief was all wet.'
(38) Zhangsan *ma* de Lisi *ku le*
Zhangsan scold DE Lisi cry ASP
'Zhangsan scolded such that Lisi was crying.'

(39) Zhangsan *qi de ma hen lei*
Zhangsan ride DE horse very tired
'Zhangsan rode the horse such that horse got very tired.'

Items (36)–(39) are synonymous with (4)–(8), respectively. Here the extra NP is not marked by the particle *ba* and, instead of being ordered to the left of the complex predicate phrase, as we have seen so far, is ordered between the *de*-marked verb and the predicative complement. Furthermore, the control situation in (36)–(39) is the same as in (4)–(8); the predicative complement is controlled by the non-subject NP. Moreover, (36)–(39) have corresponding passives and topicalized sentences, which are identical to the passives and topicalized sentences corresponding to (4)–(8). The only difference between the pattern of the RC's in (4)–(8) and that in (36)–(39) is, to reiterate, that in the former the additional NP argument precedes the V-*de* unit and is marked by the object case particle *ba* (the presence of which, as has already been discussed above, is obligatory when the object NP precedes the verb and is an Affected Theme or a Patient), and in the latter, the additional NP argument, unmarked by *ba*, follows the V-*de* unit and precedes the predicative complement.

Interestingly, everything about (36)–(39) will follow straightforwardly from what has been said about the RC's in (4)–(8), if we analyze the TVP in (36)–(39) to be cases of discontinuous constituents, where it is interrupted by the object NP, not
unlike the persuade cases in English. In other words, all that is needed to account for (36)–(39) is to add a rule combining TVP (=VP|NP) and an object NP which invokes wrapping instead of concatenation. The wrapping operation inserts the NP after the de-marked verb. Thus in (36), for example, the TVP ku de hen shangxin (cry–DE–very–upset) wraps around the NP Lisi, resulting in the discontinuity of the TVP. The analysis tree for (36) is given in the following.

(40) cf. (34b)

\[
\begin{align*}
\text{Zhangsan} & \quad \text{ku} \quad \text{de} \quad \text{Lisi} \quad \text{hen} \quad \text{shangxin} \\
\text{Zhangsan} & \quad \text{cry} \quad \text{DE} \quad \text{Lisi} \quad \text{very} \quad \text{upset} \\
& \quad \text{'}\text{Zhangsan cried such that Lisi became very upset.}'
\end{align*}
\]

\[
\begin{align*}
\text{[Zhangsan} & \quad \text{ku} \quad \text{de} \quad \text{Lisi} \quad \text{hen} \quad \text{shangxin}]_3 \\
& \quad \downarrow \\
\text{Zhangsan}_{\text{NP}} & \quad \text{[ku de Lisi hen shangxin]}_\text{VP} \\
& \quad \downarrow \\
\text{Lisi}_{\text{NP}} & \quad \text{[ku de hen shangxin]}_\text{VP} \\
& \quad \downarrow \\
\text{[ku de]}_\text{TVP|VP} & \quad \text{[hen shangxin]}_\text{VP} \\
& \quad \downarrow \\
\text{ku}_\text{TV}
\end{align*}
\]

5.6. The Reduplicative Resultative Construction.

Another variant of the RC involves reduplication. Consider the following data where the complex Resultative predicate consists of the \( V_i \text{ NP } V_i \text{ de } VP \) sequence.

(Note that this pattern is possible only when the matrix verb is a TV.)

(41) \text{Zhangsan} \quad \text{ma} \quad \text{Lisi} \quad \text{ma} \quad \text{de} \quad \text{ku} \quad \text{le} \\
\text{Zhangsan} \quad \text{scold} \quad \text{Lisi} \quad \text{scold} \quad \text{DE} \quad \text{cry} \quad \text{ASP} \\
\text{'}\text{Zhangsan, scolded Lisi, such that he \text{'}i, cried.'}

(42) \text{Zhangsan} \quad \text{qi} \quad \text{ma} \quad \text{qi} \quad \text{de} \quad \text{hen} \quad \text{lei} \\
\text{Zhangsan} \quad \text{ride} \quad \text{horse} \quad \text{ride} \quad \text{DE} \quad \text{very tired}
'Zhangsan rode the horse and got very tired.'

(43) Zhangsan chi fan chi de hen bao
Zhangsan eat meal eat DE very full
'Zhangsan ate (a meal) and got very full.'

Sentences (41)–(43) exhibit a rather unexpected control pattern. Here the subject NP, rather than the additional NP, is the controller of the predicative complement. For example, in (41) it is the subject NP Zhangsan, rather than the NP Lisi that the complement ku le ('cry–ASP') is predicated of. This contrasts with (4) and (38) (repeated below as (44) and (45)), where a similar complement is predicated of the object argument NP of the complex Resultative predicate.

(44) Zhangsan ba Lisi ma de ku le
Zhangsan BA Lisi scold DE cry ASP
'Zhangsan scolded Lisi such that he cried.'

(45) Zhangsan ma de Lisi ku le
Zhangsan scold DE Lisi cry ASP
'Zhangsan scolded Lisi such that he cried.'

This might appear to be a problem for Bach's Generalization because in (41)–(43), it is the subject NP, not the additional NP, that is the understood subject of the predicative complement in the complex Resultative predicate, contrary to what we know about the control property of TVP's, if a TVP were indeed involved here. Under closer examination, Bach's Generalization can be maintained if we analyze the reduplicated sequence of [V₁ NP V₁ dc] as VP₁||VP, which combines with a predicative complement to form a VP. Since the subject NP is the next NP to be combined with
the complex predicate containing the predicative complement, it is predicted to be the 'understood' subject of the complement. The analysis tree for a sentence with the Reduplicative Resultative construction such as (41) would look like the following:

(46) Zhangsan ma. Lisi ma de ku le
Zhangsan scold Lisi scold DE cry ASP
'Zhangsan_1 scolded Lisi_ such that he_{i/\_j} cried.'

\[ \text{Zhangsan} \quad \text{Lisi} \quad \text{de} \quad \text{ku} \quad \text{le} \quad \text{S} \]

Passive facts support this analysis of the Reduplicative Resultative Construction. If these sentences were to contain a TVP (cf., (4)–(8)), they should have corresponding passives. But since they do not involve the TVP category (as shown in (46)), we expect the lack of passives, which is indeed the case. Sentences (47)–(48) are the passive counterparts of (41)–(42), respectively, but they are ill-formed.\(^{13}\)

\(^{13}\)Note that (46) and (47) can be well-formed if there is a phonological pause between the reduplicated verb. However, the meaning of these sentences is such that they must have come from a different source and cannot be the passive counterparts of the sentences in question; for example, (47), when said with the pause, means that the horse was ridden by Zhangsan such that the horse, and not Zhangsan, became very tired.
(47)* Lisi bei Zhangsan ma ma de ku le.
    Lisi by Zhangsan scold scold DE cry ASP
    'Lisi was scolded by Zhangsan such that he cried.'

(48)* Ma bei Zhangsan qi qi de hen le
    horse by Zhangsan ride ride DE very tired
    'The horse was ridden by Zhangsan such that Zhangsan got very tired.'

The one place rule deriving the category $(VP_{II}^n\backslash VP) \\backslash NP$ from a TV and the appropriate semantic operation (which captures the association between the subject NP and the complement) can be postulated as in (49). The complex operation involves reduplication and $de$-suffixation.

(49) The Reduplicative Resultative Rule:
    $<F_{II}^n, TV, (VP_{II}^n\backslash VP) \\backslash NP>$
    Semantic Operation: $\lambda x \lambda y [\alpha'(x)(y) \text{ RESULT-IN } P(y)]$
    $F_{II}(\alpha) = F_3(\text{RED}(\alpha))$.

Finally, we have to specify that the rule combining $(VP_{II}^n\backslash VP) \\backslash NP$ and NP employs wrapping, to the effect that the NP will be inserted after the first word of the verbal category (e.g., $ma$ Lisi $ma$ de ('scold-Lisi-scold-DE') in (46)).

The analysis for the Reduplicative Resultative given in (46) and (49) thus offers an adequate account for the contrast between (6)–(8) and (38)–(40), on the one hand, and (41)–(43), on the other hand, where we find distinct control relationships between the NP's in the sentence and the predicative complement within the Resultative predicate.
5.7. The Resultative Verb Compounds.

The case for analyzing some Resultative Constructions as TVP's is further strengthened by the existence of a group of what is called the Resultative verb compounds in Chinese. These compounds pattern very much like the various Resultative Constructions. As shown in (50)–(55), they consist of a verb directly followed by a verbal complement. Again, the same distinction between transitives and intransitives which we found among the Resultative Constructions applies to the compounds\(^\text{14}\). In (50)–(51), the intransitive Resultative compounds are illustrated, where the subject NP is the understood subject of the complement within the Resultative Compound.

(50) xiaohai ku-xing\(^\text{15}\) le
  child cry-awake ASP
  'The child cried herself awake.'

(51) wo chi-bao le.
  I eat full ASP
  'I ate and was full (as a result).'

In (52)–(55) examples of transitive Resultative compounds are shown. In (52) the main verb in the compound is an IV and, in (53)–(55), the main verb in the compound is a TV. These transitive compounds behave like regular TV's; they can

\(^{14}\)This is already pointed out by Huang 1989.

\(^{15}\) ' is used here to designate morpheme boundaries within the compound.
either (1) precede the object NP, as shown by the (b) sentences in (52)–(55), in which case, the NP is not marked by ba, or (2) occur after the object NP, in which case the NP is marked by ba (e.g., (a) sentences in (52)–(55)). Finally, the (c) sentences in (52)–(54) show that unlike the resultative predicate phrases, these compounds cannot be interrupted\(^\text{16}\) (cf. (36)–(40)), where the object NP is inserted after \(dc\).

(52)a. ta ba shoupa ku-shi le
    she BA handkerchief cry-wet ASP
    'She cried such that the handkerchief became all wet.'

b. ta ku-shi le shoupa
    she cry-wet ASP handkerchief
    'She cried such that the handkerchief became all wet.'

c.* ta ku shoupa shi le
    she cry handkerchief wet ASP
    'She cried such that the handkerchief became all wet.'

(53)a. Zhangsan ba Lisi ma-ku le
    Zhangsan BA Lisi scold-cry ASP
    'Zhangsan scolded Lisi, such that he\(_i\) started to cry.'

b. Zhangsan ma-ku le Lisi.
    Zhangsan scold-cry ASP Lisi ASP
    'Zhangsan scolded Lisi\(_i\), such that he\(_i\) started to cry.'

c.* Zhangsan ma Lisi ku le
    Zhangsan scold Lisi cry ASP
    'Zhangsan scolded Lisi\(_i\), such that he\(_i\) started to cry.'

(54)a. Zhangsan ba Lisi da-si le
    Zhangsan BA Lisi beat-die ASP
    'Zhangsan beat Lisi to death.'

\(^{16}\)The continuity property of these compounds is to be contrasted with the interruptability of the bisyllabic verbal expressions discussed in Chapter 2 and 3.
b. Zhangsan da-si le Lisi
   Zhangsan beat-die ASP Lisi
   'Zhangsan beat Lisi to death.'

c.* Zhangsan da Lisi si
   Zhangsan beat Lisi die
   'Zhangsan beat Lisi to death.'

(55)a. Zhangsan ba Lisi ti-shang le
   Zhangsan BA Lisi kick-wound ASP
   'Zhangsan kicked Lisi such that Lisi was wounded.'

b. Zhangsan ti-shang le Lisi
   Zhangsan kick-wound ASP Lisi
   'Zhangsan kicked Lisi such that Lisi was wounded.'

c.* Zhangsan ti Lisi shang le
   Zhangsan kick Lisi wound ASP
   'Zhangsan kicked Lisi such that Lisi was wounded.'

As transitive verbs, these verbal compounds have corresponding passives. The passive for (53), for example, is (56).

(56) Lisi bei Zhangsan da-si le
     Lisi by Zhangsan beat-die ASP
     'Lisi was beaten to death by Zhangsan.'

The transitive Resultative compounds exhibit the same control properties as the transitive Resultative predicates in (4)–(8); the object NP is the understood subject of the complement in the compound while serving as the object of the main verb in the compound. This provides strong evidence for the analysis of the Resultative predicate in (4)–(8) as a verbal unit, namely, as a complex transitive verb, or TVP, since these TVP's turn out to be patterned the same way as the lexical (transitive) verb
compounds in (52)–(55). The evidence that these compounds are indeed lexical units, i.e., lexically derived, comes from the fact that the combination of a verb and a verbal complement into a Resultative compound is restricted to cases where both the verb and the complement are monosyllabic. Furthermore, similar to the lexicalized verb–adjective factitive construction (such as the English verb–adjective combinations like hammer flat (Bolinger 1971, Dowty 1979)), the combination is not productive and is subject to idiosyncracies of an unclear nature.

In our analysis of the Resultative Constructions, the semantic properties are captured by the semantic operations associated with the rules deriving the de–marked verbs, which combine with predicative complements to form the complex Resultative predicates. With Resultative compounds, the relevant semantic properties are captured by the semantic operations associated with the lexical rules building the compounds.

The following are the three Resultative Compound Rules — in the tradition of Dowty 1979's lexical rules — required for the three types of Resultative compounds exhibited in (50)–(51), (52), and (53)–(55), respectively:

(57)
the Resultative Compound Rule 1:
If α and β are IV's and monosyllabic, then F(α,β) is an IV, and F(α,β)= α'β.
Translation: λx [α'(x) RESULT–IN β'(x)].
(e.g., (50)–(51))

17Occasionally, it is possible to have a Resultative compound where both the verb and the complement are bi–syllabic, for example, dasao#ganjing ('sweep–clean'). This was pointed out to me by C. Huang (pc).
the Resultative Compound Rule 2:
If $\alpha$ and $\beta$ are IV's and mono-syllabic, then $F(\alpha, \beta)$ is an TV, and $F(\alpha, \beta) = \alpha'\beta$.
Translation: $\lambda x \lambda y \ [\alpha'(y) \ \text{RESULT-}\text{-IN} \ \beta'(x)]$.
(e.g., (52))

the Resultative Compound Rule 3:
If $\alpha$ is a TV and monosyllabic and $\beta$ is an IV and mono-syllabic, then $F(\alpha, \beta)$ is a TV, and $F(\alpha, \beta) = \alpha'\beta$.
Translation: $\lambda x \lambda y \ [(\alpha'(x))(y) \ \text{RESULT-}\text{-IN} \ \beta'(x)]$.
(e.g., (53)–(55))
CHAPTER VI
CONCLUDING REMARKS

6.0. Summary.

Hopefully, the data discussed in this thesis may contribute toward motivating a
Categorial-Processual approach to Chinese syntax incorporating (1) non-concatenative
operations, such as syntactic infixation and reduplication, and (2) one-place rules
within Categorial Grammar (in addition to functional application and functional
composition). To conclude, I will summarize some relevant issues within such a
process-based approach.


In Montague's schematization of syntactic rules\(^1\), a rule consists of
specifications of input category(ies) and output categories and the operation to be
applied to the input expressions. Each rule is further accompanied by a semantic rule.
Other types of operations combining expressions in addition to concatenation are
allowed. (This drastically increases the expressive power of the grammar and
obviously needs to be constrained). Under such a view, it is possible to talk about the

\(^1\)They include lexical rules as well.

176
phenomenon of multiple marking and multiple use of a grammatical marking in syntax, which is common in morphology. In the data discussed, we can find not only cases where a rule is associated with a complex operation composed of simple operations (e.g., the rule for the IAC is associated with a complex operation which consists of \(de\)-suffixation and RINFIX) but also cases where an operation is invoked by different rules.

The latter phenomenon seems to be quite pronounced in Chinese syntax. As Zwicky 1988a suggested (also cf. Janda 1982), there seems to be parochial bias in languages' use of grammatical marking; individual languages seem to specialize in particular grammatical markings. For example, in English, the affixes \(-s\) and \(-er\) seem to be the choice of marking in quite a few morphological rules. In Modern High German, the operation of Umlaut is associated with over sixty rules in the language. Under the IA view, the phenomenon of multiple use of a grammatical marking, at least when affixal marking is concerned, is seen as homophonism where multiple (grammatical) meanings are associated with a particular morphological segment. In Chinese, the particle \(de\) is known to figure prominently in the syntax (see C. Huang 1989). Its prominence in Chinese syntax is almost comparable to Umlaut in German morphology/phonology. For example, we have seen that \(de\)-suffixation (i.e., \(F_2\)) is involved in the rules for IAC, IARC, and the Descriptive Complement Construction, and we have also seen that it figures in the Resultative Constructions, discussed in Chapter 5, and in certain category changing rules, discussed in Chapter 4, among
others. Under an IA view, we have to say that *de* is associated with a huge number of functional meanings (hence massive homonymy) and sometimes no meanings at all. For example, in both the IAC and IARC, *de*-suffixation is part of a complex operation which puts together two expressions (comparable to regular concatenation) and does not contribute any meaning. Similarly, reduplication has been demonstrated to be employed in various rules.

Under the IA conceptualization, the phenomenon of the multiple use of grammatical markings is viewed as homonymy, which begs the question of why languages would economize in the means of grammatical marking at the expense of generating seemingly unnecessary homonymy. However, under an IP conceptualization of grammar, one regards rules, not the particular morphological segments or the operations associated with rules, as units of description; even though the same operation may be accessed by different rules, it is always the case that each rule, accompanied by a semantic rule, uniquely links one or more input category with a particular output category via a certain formal operation, simple or complex. For example, in a morpheme-based view, the suffix *-s* in English is regarded as homonymous. Under the present IP view, however, the association of the operation of *s*-suffixation with various rules does not involve ambiguities; the operation is a reflex of the application of the plural rule when the category of the base is (count) Noun, and it is a reflex of the third-person-singular-present rule when the base category is Verb.

\(^2\text{cf. Zwicky 1988a specifically makes this point for process morphology.}\)
Thus, this way of conceptualizing rules and operations where operations are not viewed as directly associated with meanings provides kinds of explanations not available in an IA model for languages' tendency to minimize the stock of grammatical markings available in the language. Each language can be viewed as having a set of semantics-free formal operations invoked by the syntactic rules in the language, with different rules possibly accessing the same operations. Part of the goal of syntactic analysis would then be to provide an inventory of such formal operations, or what Zwicky 1988b calls 'formal concomitants' of rules, available for syntax (see also Schmerling 1983b).

6.2. Tectogrammatics vs. Phenogrammatics.

The distinction between syntactic rules and operations may also be viewed as a way of realizing the distinction made between what Curry 1961 calls tectogrammatics and phenogrammatics. (See also Dahl 1977, Dowty 1982a, and Huck & Ojeda 1987.) Tectogrammatic rules correspond to the specifications of input and output categories in a rule which is in one to one correspondence with a semantic rule. Phenogrammatic rules can be conceived as 'spelling out' the tectogrammatic structure, which is exactly what operations do. In this connection, it may be noted that the abstracting of linear precedence relationships away from the rules stating the immediate dominance relationship among categories in GPSG (Gazdar et al., 1985) can be seen as representing a step in the direction of recognizing the need to separate tectogrammatic rules, which are provided in the Universal Grammar, from phenogrammatic rules,
which are language-specific. In GPSG, however, phonogrammatic rules consist only of principles of linear precedence, that is, of rules regulating relative ordering of adjacent sister elements, thus making the grammar context-free. In Montague's program of syntax, phonogrammatic rules are not restricted to rules of concatenations, perhaps due to the recognition that, as Bach described it, 'language just isn't left-to-right compositional'. (Bach 1984)

6.3. One-place Rules.

The rule for the RIC discussed in Chapter 3 as well as the category changing rules marked by de in Chapter 4 are unconventional in the context of Categorial Grammar in that they are one-place rules, whereas Classical Categorial Grammar assumes function argument structure for all phrasal expressions and therefore only allows (two-place) functional application (and composition) rules. However, as Bach 1988 has pointed out, 'from the point of view of Montague's UG, it would be an artificial restriction' to disallow such rules (also see Bach 1983b). Obviously, postulation of such non-canonical rules needs to be well-motivated, because, as also indicated by Bach, without any constraints, 'any arbitrary element could be removed from the stock of 'items' and reintroduced by a suitable operation' (1988:28). The RIC rule which involves reduplication constitutes a case where an one-place rule is

\[^3\text{Note that type raising (Partee and Rooth 1983, Dowty 1988, etc), which is a one-place rule, has also been proposed in Categorial Grammar. It is, however motivated, entirely for semantic purposes and involves only a change in the type of an expression but not in its category.}\]
essential, since no non-arbitrary function argument structure can possibly be postulated for these reduplicated strings.

6.4. Non-lexicalist Syntax.

In most modern syntactic theories (e.g., Lexical Functional Grammar, Classical Categorial Grammar and the Head-driven Phrase Structure Grammar of Pollard and Sag 1988), syntax has become increasingly lexicalized and syntactic rules have increasingly been reduced to a minimum set of rule schemas, while lexical items carry most of the information regarding properties of syntactic constructions, such as the combinatorial properties of expressions in the constructions. In chapter 3, I have discussed (1) data which show discontinuity of the lexical category as a result of syntactic infixation and (2) data which show phrasal reduplication, to motivate a framework where syntactic rules (which have access to a set of formal operations) and not lexical items carry the information regarding the combinatorial and distributional properties of expressions, an approach which may seem a bit anachronistic in light of the latest syntactic theorizations. In this respect, the present approach may be seen as more in the line of the theoretic view of Zwicky 1985 (also Hoeksema 1987), which calls for a revival of the notion of constructions and, therefore, of the role of syntactic rules (as opposed to lexical entries) in descriptions of syntax.
BIBLIOGRAPHY


Conference on Chinese Linguistics, 201–214.


Thompson, Sandra. 1973. Transitivity and Some Problems with the *Ba* Construction in


Zwicky, Arnold. 1988b. Inflectional Morphology as a (Sub)component of Grammar. paper presented at the Third International Morphology Meeting.


