Introduction

Over the last quarter-century, there has been a surge of research on Chinese syntax. A cursory look at the programs of Chinese linguistics conferences held since 1985 shows that at least a full continuous session has been devoted to Chinese syntax throughout each day of every conference. Those who were involved in organizing such conferences can also recall the large number of syntax abstracts, routinely accounting for fifty to sixty percent of all abstracts received for review. It is also during this past quarter-century that a significant number of theoretically oriented works on Chinese syntax began to appear in major refereed academic journals published in the West. Several monographic, theoretical treatments of Chinese syntax have also appeared that distinguished themselves from earlier general descriptions or reference grammars. In the field of theoretical linguistics, more works than before make crucial reference to Chinese syntax. It is clear that research on Chinese syntax that is informed by modern linguistic theories has been productive. In turn, it is also clear that the study of Chinese syntax has played an ever-increasing role in linguists' construction of modern "mainstream" syntactic theories.

Most of these "modern syntactic theories" are in one form or another theories falling under the formal paradigm of generative grammar. Of these formal treatments, much research has been carried out in the Principles-and-Parameters (P&P) approach initiated by Chomsky and his colleagues and students around 1980, plus and minus two or three years, in its various incarnations including the so-called Government-and-Binding (GB) framework, the Barriers framework, and recent attempts at theoretical economy aimed at the ideals of the Minimalist Program (MP). The P&P approach marked the beginning of an era that distinguished itself from the first quarter-century of generative grammar (since 1957) in enabling the construction of a theory of grammar that is at once general enough to capture common properties of human language and flexible enough to account for language variations. It provided a way to make good sense of the innateness hypothesis (or "biolinguistic approach") that characterized Chomsky's approach since it was introduced twenty-five years before, a hypothesis that takes the internalized grammar of any language to be a combined product of nature and nurture.

It also allowed for the productive description of languages of various typological types and, most importantly, for the study of a variety of languages to directly contribute to the construction of general linguistic theory. The construction of the GB theory as we know it today, for example, was itself in part informed by some analyses of Chinese syntax.

The volume of research products that have appeared in this period, quite unlike the situation ever before, far exceeds the amount anyone can easily recall or enumerate when pondering over one particular topic or another. Various grammatical constructions have been given multiple different treatments. Some grammatical constructions that seemed irrelevant to generative grammar in its early periods are now actively analyzed, while objections to certain formal analyses have now lost ground. Yet we continue to hear objections and questions from scholars unfamiliar with the paradigm – either those who were educated in the pre-GB model with many assumptions that are no longer held by current generative researchers or those who are less informed about formal approaches. Part of this situation, we believe, arises from misunderstanding or lack of accessible information. The fact is that, for almost every topic of Chinese syntax, there now exists a sizable amount of generative literature within the P&P paradigm. The problem, for those who for one reason or another have not been able to follow the recent developments, comes in part from the fact that most research products come in single articles – from journals, edited volumes, and conference presentations – and there is no work as yet that attempts to take stock of the major results that have been produced and describe them in some depth – within one volume – that might serve the double purpose of informing the readers less familiar with (or less committed to) formal linguistics and the current status (in our view) of formal Chinese syntax, and of bringing further questions onto the research agenda for other scholars and students interested in the enterprise of providing rigorous analyses of Chinese linguistic facts and bringing them to bear on the construction of an optimal theory of human linguistic competence and its possible variations, as part of a theory of the "mirror of the mind."

The desire to take a first step toward filling this gap was a major motivation that led us to take up the project of writing this book. It is our hope that a volume consisting of the topics we have chosen will present a more comprehensive outlook of the syntactic system of the whole language to the reader, and that our discussion of the various analyses on each topic will help both to sum up some of the important results and to provide new points of departure for further research. It is also our intention to use this book to demonstrate, for each topic selected, how a formal generative analysis may help make sense of certain observed properties of the language, perhaps in ways better than other imagined approaches, and how it may be seen as a contribution to linguistic theory.

Before we go on to present the details of what this book is about, however, we must make clear what it is not. First, it is not meant to be a reference grammar for the learner of Chinese, though it might be seen as a (somewhat biased) reference on the formal linguistic analyses of Mandarin Chinese syntax. Although we have tried to include as many references as we can to the large volume of works available, we are sure to have inadvertently missed some. Even where references are included, we do not provide a detailed discussion of all alternative analyses that are worthy of consideration, other than those closely related to our own analyses. We have also excluded most references that are explicitly non-formal. Second, it is not a comprehensive treatment of Chinese syntax. As it turns out, even within formal approaches, it is impossible to touch on all the important aspects of Chinese syntax. Rather than briefly summarizing results on a comprehensive list of topics, we have chosen to provide fairly detailed analyses and argumentation of a selected number of topics, excluding some owing to space limitation and others where we have nothing new to offer. For each topic our discussion is driven by the goal of providing one or two specific analyses and explaining the rationale behind them, with the general theory of grammar in mind. It is often said that Y.-R. Chao's (1968) *Grammar of Spoken Chinese* is a comprehensive single-volume masterpiece that represents the best of the American descriptive and structuralist tradition. No single-volume formal treatment comparable to Chao's in scope has appeared in the last several decades. The rich observations and insights contained in that volume remain unsurpassed to this day. We have not attempted a comprehensive treatment of Chinese syntax in the generative tradition. Our goals are both different and limited: the book presents grammatical analyses that cover most of the constructions of (Mandarin) Chinese that have figured in the field of theoretical linguistics in the past twenty-five years, focusing on our own analyses in most cases. It is intended to show how the facts of Chinese may be profitably understood with the tools of generative linguistics, and in turn how the analyses may help settle important issues and guide further research in linguistic theory. It is intended as a contribution to Chinese syntax as a distinct subject of Chinese studies, and also to generative grammar as a hypothesis about human linguistic competence.

The rest of this book is organized into four parts comprising nine chapters. Part I (Chapters 1–3) investigates the building blocks and "canonical" structures of sentences, including the grammatically relevant properties of words and the combinatorial algorithm by which phrases are formed. Chapter 1 presents a theory of parts of speech, which we call categories. Lack of sufficient inflectional and derivational clues has made the identification of categories difficult for Chinese. Drawing on the insights gained from other languages, we rely primarily on the syntactic behaviors of a word to determine its category. It is also shown that

a category is best viewed as a cluster of plus- or minus-valued features, which enables us not only to distinguish syntactically relevant categories but also to explain why some of them display identical properties. The analysis of localizers, one of the perpetually question-begging categories in Chinese, makes use of the notion of computational cost and opens up a new possibility to understand how categorial changes happen and a new category comes into existence.

Chapter 2 focuses on the nature of argument structure. Capitalizing on a long-known and puzzling fact, i.e., that the semantic relations between a verb and its subject or object are much less restricted in Chinese than in English, and drawing on recent works by others, we propose a theory of lexico-semantic decomposition of verbs that minimizes the amount of stipulated mechanisms and components and thereby maximizes the explanatory power of the theory. In particular, it is argued that a tiny set of event-typing elements interact with a lexical root to produce the more "rigid" argument structures found with English verbs, whereas the option of using bare roots as verbs in the absence of event-typers, aided with world knowledge, is responsible for the degree of semantic freedom in Chinese.

Chapter 3 covers a broad range of topics on the "canonical" structures of the sentence, with particular focus on the verb phrase and its components. It examines the systematic distinctions between adjuncts and complements, looks for the best structural representations of five different postverbal constituents (the double-object, two V-*de*'s, and the frequency and duration expressions), and discusses how such semantic notions as aspect and modality are handled in the syntax of Chinese. In the course of presentation, it is proposed that the behavioral disparity between the resultative V-*de* and its manner counterpart may be attributed to the superficially unrelated fact that Chinese has resultative compounds but not ones with a postverbal manner modifier. Attention is also given to constructions which appear to display syntax–semantics mismatches. What unifies this large collection of miscellaneous topics is a single phrase structure pattern whose restriction on possible syntactic analyses highlights an important characteristic of this model of linguistic theory: using the least amount of independently motivated tools to account for the maximal amount of data.

In Part II (Chapters 4 and 5), we take a closer look at argument structure and its relation with lexical semantics and its effects on syntactic structure, by focusing on two constructions that have been in the center of debate from the inception of Modern Chinese syntax as a field. Chapter 4 deals with the passive *bei* construction, which takes two forms depending on the presence or absence of an Agent phrase (the long and short passive respectively). After exhibiting the pros and cons of a movement-based approach and one based on complementation, it is argued that the Chinese passive involves both movement and complementation. The long passive is derived via clausal complementation where the embedded object is brought to

the periphery of the complement clause (a process of "operator movement") and is predicated on the main-clause subject. The short passive, on the other hand, involves verbal complementation and the object is brought to the periphery of a verb phrase (a process of "argument movement") where it is interpreted with the subject.

Building on the findings of the passive construction, Chapter 5 compares and contrasts passives and the closely related *ba* construction. *Bei* and *ba* constructions are similar in argument structures. However, they differ in the range of (un)acceptable cases, which is attributed to the different subcategorization requirements of *ba* and *bei*, reflected in the syntactic structures with which they are associated. Nonetheless, the extant literature on the *ba* construction has not been as focused on its syntactic properties as on the special meaning of this construction and how to account for it. The *ba* construction has been noted as expressing "disposal" or "affectedness." We show that the special meaning cannot be due to any thematic-assigning capabilities of *ba*. Every *ba* sentence has a non-*ba* counterpart, which points to the irrelevance of *ba* in contributing to the argument structures. In the most typical examples, *ba* seems to be related to the notion of boundedness or requires a result expression. However, the complexities of the *ba* construction require the search for further possibilities and additional mechanisms for an "affected" interpretation.

While the passive and the *ba* construction exemplify how modifications in lexical structure affect the syntactic relations between arguments such as subject and object, other constructions exhibit syntactic properties independent of lexical semantics. Such constructions involve operations on or beyond clauses, and often concern the logical relations between clausal peripheral elements and the clauses as a whole. Part III takes up two types of logical structure: one involving (often) overt antecedent–gap relations and the other involving, as we shall argue, covert dependency relations. The first type, dealt with in Chapter 6, is best illustrated by topic and relative clause structures, in which a clause is used to modify a head noun phrase. There have been claims that a relative construction is derived from a topic structure; however, we show that the two constructions are similar but not identical. They are alike in the set of locality conditions restricting the well-formedness of these constructions, phrased in terms of constraints on movement and rules governing the distribution of empty categories. They differ in exactly which element undergoes movement and where it lands. Variations with respect to these factors are also manifested within relative constructions in a cluster of empirical generalizations that can be traced to the absence/presence of a relative operator.

In Chapter 7, we turn to the syntax of interrogative sentences with particular attention to *wh*-questions and a special type of disjunctive question called the

"A-not-A question." After clarifying the distinctness of this question type from normal *yes-no* questions, we propose and defend a modular approach to the A-not-A questions. We discuss a number of approaches to the syntax and semantics of *wh*-questions, which exhibit covert long-distance dependencies with restrictions that, we argue, follow from an appropriately formulated theory of movement, binding, and the syntax–semantics interface.

Our presentation of the syntax of Chinese would be inappropriately incomplete without some in-depth discussion of the syntax of nominal expressions and their meanings. Part IV of this book is devoted to this area. Chapter 8 takes up the syntactic structure of nominal expressions – nouns and phrases built around nouns as their heads. We note that Chinese noun phrases, on their surface, are at once more complex and more simple than their counterparts in, say, English and other languages (for example, with respect to the requirement of numeral-classifiers, the lack of true determiners, and the occurrence of "bare" singular count nouns). We argue, however, that appearance notwithstanding, Chinese noun phrases (like those in many other languages) have more structure than meets the eye. We propose a full determiner phrase that may contain other smaller phrases headed by a numeral expression, a classifier, and a noun, and show that this allows for the derivation and explanation of certain facts of (in)definiteness, specificity, and compositional semantics.

Another important aspect of the semantics of noun phrases concerns their reference and the referential dependencies they exhibit on each other. This is the subject of the final chapter. Here our discussion addresses both the syntax and semantics of coreference and of variable binding. We show that the referential properties of nominal expressions are tied to their intrinsic properties (whether they need an antecedent or not), the syntactic position of their antecedents (if they need one), and the nature of the antecedents themselves (whether they are referential or quantificational). With respect to definite noun phrase anaphora, we devote substantial space to a discussion of the Chinese reflexive pronoun *ziji*, and show that it is both an anaphor in the sense of classical Binding Theory and a logophor within contexts of "attitudes *de se*" that describe the speech, the mental state, or the perspective of an appropriate protagonist. With respect to variable binding, we show that the crucial requirement is c-command in a proper Logical Form representation. We finish Chapter 9 with a discussion of so-called "donkey anaphora," something that has a status between definite coreference and variable binding. We present two types of "donkey sentences," each with a set of distinguishing properties, and show that a proper analysis of them helps settle an important debate between two competing theories that have figured prominently in recent treatments of indefinite noun phrases and their referential properties.

There are clearly other interesting topics of Chinese syntax that deserve coverage in a book with this title, but we have had to leave them out. Several other constructions that bear on lexical structure and syntactic projection could each deserve a chapter-length full treatment. For example, the resultative construction (both the compound and the phrasal versions), touched upon briefly in Chapter 3, has further interesting properties bearing on the structure of events and their projection in syntax. The syntax of adverbials and that of aspectual markers are two other areas that have received considerable renewed interest in recent years. Other topics falling under the area of argument structure and syntactic structure include the syntax of unaccusatives, the two types of double-object constructions, and the proper syntactic treatment of various conjunctives. With respect to logical structure and the syntax–semantics interface, we have left out much work on quantification and structures bearing on focus and presupposition. And our discussion of noun phrase anaphora also does not touch upon the distribution and reference of zero pronouns, a topic of major interest to parametric theory with implications for the interface between syntax and discourse. In selecting topics for inclusion in this work, we have used three criteria. The first is our perception of relative priority in trying to strike a balance between breadth and depth within a limited space. The second is the availability of the literature: a topic is not included when it has been extensively discussed in easily accessible monographs or journals. The third one has to do with the scope of our own research: we have left out topics on which we have not ourselves carried out sufficient research and to which we do not have something new to contribute.

A word about the intended readers of this work: we prepared these chapters originally for university courses that we offer on the linguistic structure of Chinese, so the most immediate intended readers of this book are those graduate students and upper-level undergraduates who have some basic knowledge of linguistic structure. Such students, or any professional linguist of any theoretical persuasion, should find the book fully accessible, even without prior experience with the Chinese language. A student of the Chinese language may also find this work accessible with occasional reference to linguistic terminology available from syntax textbooks or linguistics glossaries. In writing the book, we have also had in mind the non-specialists who are curious about Chinese grammar and generative syntax, and have tried to briefly explain technical notions as they are first introduced. As such, we hope the book will be useful to teachers and researchers in such Chinese-related fields as language teaching, natural language processing, machine translation, language acquisition, philosophy of language, and other related areas of cognitive science.

As usual, the completion of a book of this size owes itself to the help of numerous people. It is impossible to enumerate the scholars from whom we have learned

the body of knowledge represented here. We should, however, mention a few colleagues who have collaborated with us on one topic or another with results that have been included in this work. In particular, the materials on donkey anaphora and long-distance reflexives are derived from earlier work conducted with Lisa Cheng and C.-S. Luther Liu, respectively. Our discussion of argument structure and lexical relations has also benefited from our erstwhile collaboration with Lisa Cheng and C.-C. Jane Tang. Some sections on relative constructions and *wh*-questions are incorporated from work in collaboration with Joseph Aoun. The analysis of the V-*de* constructions draws on our joint work with Jen Ting. And the discussion of the interactions among different adverb classes is a direct application of the discoveries we made together with Vivian Lin and Rebecca Shields on the intervention effects of adverbs in English and Russian. We continue to appreciate the opportunities we have had to work with them. Parts of this manuscript in one of its earlier versions have been tried out in classes and read by some of the students and faculty at Harvard University, the University of Southern California, and the University of Wisconsin–Madison, as well as the National Taiwan Normal University, Stanford University, and the University of Venice. We are gratified by the interest and support shown to us by the instructors and participants and, in some cases, for their comments and suggestions – especially those of Ressy Ai, Shengli Feng, Francesca del Gobbo, Miaoling Hsieh, So-One Hwang, Soo-Yeon Jeong, Julie Jiang, Daphne Liao, Jing Rong, Peter Sells, Yang Shen, Yuan Shen, Fuzhen Si, Jen Ting, and Yaqing Wu. In our final efforts to bring this work to fruition, we owe special gratitude to Bridget Samuels for her help in making the whole manuscript more readable than it otherwise could be. Finally, but not the least, our deep-felt thanks go to Emily, Qing, and Yu-Chin for all the best of things that life can offer; something that we have taken all these years but, probably too often, for granted.

JH, AL, & YL

1

Categories

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

While the basic distinction between nouns and verbs is universally recognized in modern literature on Chinese syntax, scholars differ, sometimes drastically, on other categories. See Chao (1968), Li and Thompson (1981), Zhu (1982), and Xing and Ma (1992) for a few examples. The differences in opinion arise partly because linguists with different theoretical backgrounds may employ different criteria for word classification, and partly because we still lack sufficient knowledge about certain words and their properties. Regardless, it is without question that the ultimate task for anyone studying lexical categories in Chinese is to identify them in such a way that they both allow an accurate description of the syntactic behaviors of the language, and provide insights into the nature of word classification.

With this goal in mind, we will introduce a theory of lexical categorization in Mandarin Chinese in this chapter. The theory consists of two intertwined parts. First, a set of categories is confirmed and examined on the basis of the syntactic behaviors of Chinese words and morphemes. Second, a decompositional theory that characterizes the intrinsic relations among these categories is defended. It is important to mention up front, however, that we do not intend to spread our discussions evenly among all issues related to lexical categorization, nor do we

attempt to provide an exhaustive list of categories in the language. Rather, the chapter concerns itself primarily with where we believe new insights are available from recent research. The same approach also applies to the organization of the whole book.

1.1 Lexical categories

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

1.1.1 Verbs and nouns – basic distinctions

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

(1) Verbs
  a. bu shui 'not sleep'
  b. bu tongzhi 'not inform'
  c. bu sai–qiu 'not play ball'

(2) Nouns
  a. *bu shu 'not tree'
  b. *bu xiaoxi 'not news'
  c. *bu qiu–sai 'not ball game'

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

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

(3) a. *bu-shu-bu-bao 'not-book-not-newspaper'
  b. *bu-fan-bu-cha 'not-food-not-tea'

Second, the nouns in these examples must be monosyllabic, even when multisyllabic counterparts exist, further confirming that *bu* cannot really modify a noun