A Reference Grammar of

Chinese

Edited by
Chu-Ren Huang
Dingxu Shi

CAMBRIDGE

A Reference Grammar of Chinese

A Reference Grammar of Chinese is a comprehensive and up-to-date guide to the linguistic structure of Chinese, covering all of the important linguistic features of the language and incorporating insights gained from research in Chinese linguistics over the past thirty years. With contributions from twenty-two leading Chinese linguists, this authoritative guide uses large-scale corpora to provide authentic examples based on actual language use. The accompanying online example databases ensure that a wide range of exemplars are readily available and also allows for new usages to be updated. This design offers a new paradigm for a reference grammar where generalizations can be cross-checked with additional examples, and also provides resources for both linguistic studies and language learning. Featuring bilingual term lists, this reference grammar helps readers to access relevant literature in both English and Chinese and is an invaluable reference for learners, teachers, and researchers in Chinese linguistics and language processing.

CHU-REN HUANG 黃居仁 is Chair Professor of Applied Chinese Language Studies at the Hong Kong Polytechnic University and President and Fellow of the Hong Kong Academy of the Humanities.

DINGXU SHI 石定栩 is Chair Professor of Chinese Linguistics at the Department of Chinese and Bilingual Studies of the Hong Kong Polytechnic University and Fellow of the Hong Kong Academy of the Humanities.

A Reference Grammar of Chinese
中文参考语法

Edited by

CHU-REN HUANG 黄居仁
The Hong Kong Polytechnic University

DINGXU SHI 石定栩
The Hong Kong Polytechnic University

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9780521181051

First published 2016

Printed in the United Kingdom by TJ International Ltd. Padstow Cornwall

A catalogue record for this publication is available from the British Library

Library of Congress Cataloguing in Publication data
Huang, Chu-Ren, editor. Shi, Dingxu, editor.
A reference grammar of Chinese / Chu-Ren Huang and Dingxu Shi.
Cambridge : Cambridge University Press, [2016] English and Chinese.
Includes bibliographical references and index.
LCCN 2015048873 ISBN 9780521769396 (hardback)
LCSH: Chinese language – Grammar. Chinese language – Grammar, Comparative – English. Chinese language – Textbooks for foreign speakers – English.
LCC PL1099.R44 2016 DDC 495.15 – dc23
LC record available at http://lccn.loc.gov/2015048873

ISBN 978-0-521-76939-6 Hardback
ISBN 978-0-521-18105-1 Paperback

Dedicated to 献给

Y.R. Chao 赵元任　　*Zhu Dexi* 朱德熙

Giants Whose Shoulders We Stand On
高山仰止 景行行止

Contents

Figures

Tables

Contributors

Chief editors

Chu-Ren Huang, The Hong Kong Polytechnic University

Dingxu Shi, The Hong Kong Polytechnic University

Associate editors

Kathleen Ahrens, Hong Kong Baptist University

Jing Jin, The Hong Kong Institute of Education

Sophia Y. M. Lee, The Hong Kong Polytechnic University

Jingxia Lin, Nanyang Technological University

Assistant editors

Helen K. Y. Chen, Academia Sinica

Jiajuan Xiong, The Hong Kong Polytechnic University

Hongzhi Xu, The Hong Kong Polytechnic University

Chapter authors

Kathleen Ahrens, Hong Kong Baptist University

Xiaojing Bai, Tsinghua University

Yung-O Biq, National Taiwan Normal University

Shui Duen Chan, The Hong Kong Polytechnic University

Hilary Chappell, École des Hautes Études en Sciences Sociales

Chu-Ren Huang, The Hong Kong Polytechnic University

Shi-Zhe Huang, Haverford College

Yan Jiang, SOAS, University of London

Jing Jin, The Hong Kong Institute of Education

Po Lun Peppina Lee, City University of Hong Kong

Y.-H. Audrey Li, University of Southern California

Jingxia Lin, Nanyang Technological University

Stephen Matthews, The University of Hong Kong

Jerome Packard, University of Illinois at Urbana-Champaign

Haihua Pan, City University of Hong Kong

Marie-Claude Paris, Université Paris Diderot-Paris 7

Dingxu Shi, The Hong Kong Polytechnic University

Shu-ing Shyu, National Sun Yat-sen University

Chaofen Sun, Stanford University

Sze-Wing Tang, The Chinese University of Hong Kong

Virginia Yip, The Chinese University of Hong Kong

Weidong Zhan, Peking University

Preface

A reference grammar is a snapshot of language in action as well as a comprehensive description of how language is used. In addition, a reference grammar attempts to describe the range of possibilities of how words and other larger linguistic units may be put together in the language, and how they can present different meanings in different contexts. Moreover, a reference grammar aids linguists in constructing theories for one language in particular, and all languages in general. A reference grammar also provides roadmaps for both language learners and teachers to navigate the complexity of learning or teaching a new language. Lastly, a reference grammar shares with all language lovers the beauty of language as a knowledge system, including the cultural heritage of the language.

A Reference Grammar of Chinese is designed with all of the above functions in mind. We are keenly aware of the challenge and constraints of containing a grammar within the confines of the book format. Hence, we worked out a solution by constructing an example corpus in parallel while drafting our grammar. This example corpus constitutes a distinguishing feature of this book. On the one hand, all the generalizations reported in this grammar are based on expert linguists' observations of a large set of corpus data. On the other hand, when reading this reference book, a reader can refer to the example corpus (at the website http://crg.cbs.polyu.edu.hk) using chapter, section, and example numbers, so that the reader can form his/her own generalizations and/or tease out more details that could not be included in the grammar due to space considerations. Thus, the example databases allow *A Reference Grammar of Chinese* to be not only an evidence-based grammar, but also a grammar not constrained by the limitations of the pages of the book.

We are also fully aware of the fact that a reference grammar must refer to a familiar framework of grammar. As English is the academic metalanguage used for this book, and English grammar is the most familiar language grammar due to its largest number of second language learners, we choose to adopt and follow the framework of English descriptive grammar. In particular, we follow the structure and drafting guidelines of *The Cambridge Grammar of the English Language* (*CGEL*) as closely as possible. In this connection, we would like to express our gratitude

to the two authors of *CGEL*, Professors Rodney Huddleston and Geoffrey Pullum. Professor Pullum flew to Hong Kong to meet both chief editors as well as many chapter authors and to provide us with advice and encouragement as we embarked on our long and arduous journey.

Similar to the *CGEL*, *A Reference Grammar of Chinese* (*ARGC*) has chief editors who have done research on various aspects of Chinese grammar in their respective academic careers. However, even though the *ARGC* chief editors were heavily involved in drafting and revising, they are not listed as co-authors of the chapters except for the chapters in which they had a direct hand in the original drafts. It is perhaps not difficult to see that in a grammar covering such a wide range of topics, inconsistencies in descriptions among the submitted early drafts would arise and the editors would eventually need to be responsible for the final revisions to ensure a consistent voice and perspective. In such a context, the chapters were meticulously modified by varying degrees. Therefore, please bear this in mind while reading and attribute the contributions to both the chapter authors and the editors. Any remaining errors, however, are the responsibilities of the editors.

The completion of this reference grammar was the result of a long journey that we took with many friends. In addition to thanking Geoffery Pullum for sharing his wisdom and encouragement, we would like to thank Helen Barton of Cambridge University Press for her unfailing support to guide us through this journey and for helping us with various unexpected difficulties. We would also like to thank the enthusiasm, professionalism, and linguistic acumen of all the editorial team members, including the associate editors Prof. Kathleen Ahrens, Dr. Jing Jin, Dr. Sophia Y. M. Lee, and Dr. Jingxia Lin, and assistant editors Dr. Helen Kai-yun Chen, Dr. Jiajuan Xiong, and Dr. Hongzhi Xu. Many colleagues in our research group helped at various stage of preparation, especially with extraction of examples from corpora, and among the people whom we cannot list exhaustively here, we would like to thank in particular Dr. Shuang Hong, Dr. Jia-fei Hong, and Ivy Wing Shan Chan. In addition, we would like to thank Professor Qin Lu for the construction of the working database and interface for extracting and storing example sentences from corpora. Last, but not least, we would like to thank the Department of Chinese and Bilingual Studies of the Hong Kong Polytechnic University, as well as the PolyU-Peking U Research Centre on Chinese Linguistics for providing support for research that was crucial to the completion of this grammar.

Acknowledgements

The editors and contributors would like to thank the generous infrastructure support of the Hong Kong Polytechnic University-Peking University Research Centre on Chinese Linguistics including but not limited to travel grants, language resources, research assistants, and copy-editing.

Chu-Ren Huang was supported by Hong Kong Research Grant Council GRF grants nos. 543810, 544011, and 543512 for research work leading to his contribution to Chapters 1, 2, 6, 7, and 12; as well as research contributing to the overall editorial work of the grammar.

Stephen Matthews and Virginia Yip were supported by Hong Kong Research Grant Council GRF project grant no. 14413514 for research leading to their contribution to Chapter 9.

Jingxia Lin would like to thank the support of the Hong Kong Polytechnic University (grant no. 1-ZV8E) and Nanyang Technological University (grant no. M4081117.100) for supporting the research work leading to her contribution to Chapter 13.

Weidong Zhan and Xiaojing Bai would like to thank the support of National Social Science Foundation of People's Republic of China grant no. 12BYY061 and Ministry of Education of People's Republic of China grant nos. 13JJD740001 and 12JHQ046 for partially supporting their research work leading to their contribution to Chapter 14.

Hilary Chappell's research work received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007–2013) for the ERC Advanced Grant Agreement no. 230388: "The hybrid syntactic typology of Sinitic languages" (2009–13).

Yan Jiang would like to thank the support of the Hong Kong Polytechnic University – grant no. 1-ZV8M partially supporting his research work leading to his contribution to Chapter 16.

Shu-Ing Shyu would like to thank the support of NSC-100-2410-H-110-042 partially supporting her research work leading to her contribution to Chapter 17.

English–Chinese term list

A	啊
abstract noun	抽象名词
accomplishment (verb)	完结(动词)
achievement (verb)	达成(动词)
acronym	缩略词
activity (verb)	活动(动词)
adjective	形容词
adjunct	附接语
adverb	副词
adverbial	状语, 副词短语
affirmative	肯定式
affix	语缀
agent	施事
AI	唉
anaphor	照应语, 回指语
anaphoric chain	回指链, 照应链
antecedent	先行语
appositive	同位语
argument	论元
article	冠词
aspect	体, 体貌
atelic	无终结
attributive	定语(的)
auxiliary (verb)	助动词
BA	吧; 把, 将
BALE	罢了
bare NP	光杆名词, 光杆名词短语
BEI	被
beneficiary	受益者
blending	缩合
borrowing	借词

bounded	有界
case	格
CL	classifier, 量词, 分类词
cleft	分裂句
closed class	封闭类
code	语码
common noun	普通名词
complement	补(足)语
complex sentence	复(杂)句
concessive	让步语
conditional	条件句
consonant	辅音
coordinator	并列连词
copula	系词
corpus	语料库
correlative	关联(词)
DE	地, 得, 的, 之
declarative (sentence)	陈述(句)
definiteness	有定, 定指
deictic	直示
demonstrative	指示代词
denotation	外延
deontic	义务(情态)
dependency relation	依存关系, 依赖关系
derivation	派生
determiner	限定词
deverbal	动转(的)
direct object	直接宾语, 直接受词
ditransitive	双及物
entailment	衍推
existential (sentence)	存在句
experiencer	感受者
gap	缺位
GEI	给
generic	通指
goal	目标
gradability	级差性
grammatical function	语法功能
GUAN	管

GUO	过
headless relative clause	无核关系句
illocutionary force	示意语力, 言外力量
imperative	祈使
imperfective	未完成体
indirect object	间接宾语 间接受词
inflection	屈折
instrument (case)	工具格
intensifier	程度状语
interrogative pronoun	疑问代词
intonation	语调
intransitive	不及物的
inversion	倒置
irrealis	非现实
lexeme	词位
LA	啦
LAIZHE	来着
LE	了
Lit.	literal
LI	哩
MA	吗; 嘛
mass noun	不可数名词
metalinguistic negation	元语言否定
metaphor	隐喻
metonymy	转喻
modality	情态
mood	语态
morphology	形态学; 构词, 词法
NA	拿; 哪
NE	呢
NEG	不, 不是, 未, 非, 没有, 不必, 没
nominal	名词性成分
O	喔
object	宾语, 受词
oblique	旁格, 斜格
ORDINAL	ordinal marker, 第
part of speech	词类
particle	小词, 助词
passive	被动

patient	受事
perfective	完成体
phoneme	音位
PL	们
polarity	极性
PREP	preposition, 介词
preposing	前置
presupposition	预设
pro form	指代形式
progressive aspect	进行体
proper noun	专有名词
proposition	命题
prosody	韵律
QI	起
QILAI	起来
quantifier	量化成分, 量词
recipient	接受者
reduced clause	缩减小句
reduplication	重叠
reference	指称
referent	所指
reflexive pronoun	反身代词
resultative	结果补语
resumptive pronoun	复指代词
rhetorical question	反问
scope	辖域
selectional restriction	选择限制
SHI	是
state (verb)	状态(动词)
stative	静态
subordinate	从属(成分)
suffix	后缀
SUO	所
SUOXIE	缩写
syntax	句法
telic	终结
thematic role	论旨角色
theme	主位
topic	话题

trajectory	轨迹
truth conditions	真值条件
truth value	真值
type	类
variable	变数; 变体的
vocative	呼格
voice	语态
volition	意愿
vowel	元音
WA	哇
XIAQU	下去
YA	呀
YO	哟
YOU	有
ZAI	在, 正在
ZHE	着
ZHE_NE	着呢

Chinese–English term list

啊 *a4*	A
唉 *ai0*	AI
吧 *ba0*	BA
把 *ba3*	BA
罢了 *ba4liao3*	BALE
被动 *bei4dong4*	passive
被 *bei4*	BEI
变数 *bian4shu3*	variable
变体的 *bian4ti3de0*	variable
宾语 *bin1yu3*	object
并列连词 *bing4lie4lian2ci2*	coordinator
补(足)语 bu3(*zu2*)yu3	complement
不必 *bu2bi4*	NEG
不是 *bu2shi4*	NEG
不及物的 *bu4ji2wu4de0*	intransitive
不可数名词 *bu4ke3shu3ming2ci2*	mass noun
不 *bu4*	NEG
陈述(句) *chen2shu4(ju4)*	declarative (sentence)
程度状语 *cheng2du4zhuang4yu3*	intensifier
重叠 *chong2die2*	reduplication
抽象名词 *chou1xiang4ming2ci2*	abstract noun
词法 *ci2fa3*	morphology
词类 *ci2lei4*	part of speech
词位 *ci2wei4*	lexeme
从属(成分) *cong2shu3(cheng2fen4)*	subordinate
存在句 *cun2zai4ju4*	existential (sentence)
达成(动词) *da2cheng2(dong4ci2)*	achievement (verb)
倒置 *dao4zhi4*	inversion
的 *de0*	DE
得 *de0*	DE
地 *de0*	DE

第 *di4*	ORDINAL, ordinal marker
定语(的) *ding4yu3(de0)*	attributive
定指 *ding4zhi3*	definiteness
动转(的) *dong4zhuan3(de0)*	deverbal
反身代词 *fan3shen1dai4ci2*	reflexive pronoun
反问 *fan3wen4*	rhetorical question
非现实 *fei1xian4shi2*	irrealis
非 *fei1*	NEG
分类词 *fen1lei4ci2*	CL, classifier
分裂句 *fen1lie3ju4*	cleft
封闭类 *feng1bi4lei4*	closed class
辅音 *fu3yin1*	consonant
副词短语 *fu4ci2duan3yu3*	adverbial
副词 *fu4ci2*	adverb
附接语 *fu4jie1yu3*	adjunct
复(杂)句 *fu4(za2)ju4*	complex sentence
复指代词 *fu4zhi3dai4ci2*	resumptive pronoun
感受者 *gan3shou4zhe3*	experiencer
格 *ge2*	case
给 *gei3*	GEI
工具格 *gong1ju4ge2*	instrument (case)
关联(词) *guan1lian2ci2*	correlative
管 *guan3*	GUAN
冠词 *guan4ci2*	article
光杆名词短语 *guang1gan3ming2ci2duan3yu3*	bare NP
光杆名词 *guang1gan3ming2ci2*	bare NP
轨迹 *gui3ji4*	trajectory
过 *guo4*	GUO
后缀 *hou4zhui4*	suffix
呼格 *hu1ge2*	vocative
话题 *hua4ti2*	topic
回指链 *hui2zhi3lian4*	anaphoric chain
回指语 *hui2zhi3yu3*	anaphor
活动(动词) *huo2dong4* (dong4ci2)	activity (verb)
级差性 *ji2cha1xing4*	gradability
极性 *ji2xing4*	polarity
间接宾语 *jian4jie1bin1yu3*	indirect object
间接受词 *jian4jie1shou4ci2*	indirect object
将 *jiang1*	BA

接受者 *jie1shou4zhe3*	recipient
结果补语 *jie2guo3bu3yu3*	resultative
介词 *jie4ci2*	PREP, preposition
借词 *jie4ci2*	borrowing
进行体 *jin4xing2ti3*	progressive aspect
静态 *jing4tai4*	stative
句法 *ju4fa3*	syntax
肯定式 *ken3ding4shi4*	affirmative
啦 *la1*	LA
来着 *lai2zhe0*	LAIZHE
类 lei4	type
哩 li0	LI
量词 *liang4ci2*	CL, classifier; quantifier
量化成分 *liang4hua4cheng2fen4*	quantifier
了 *liao3*	LE
论元 *lun4yuan2*	argument
论旨角色 *lun4zhi3jue2se4*	thematic role
吗 *ma0*	MA
嘛 *ma0*	MA
没有 *mei2you3*	NEG
没 *mei2*	NEG
们 *men0*	PL, plural marker
名词性成分 *ming2ci2xing4cheng2fen4*	nominal
命题 *ming4ti2*	proposition
目标 *mu4biao1*	goal
拿 *na2*	NA
哪 *na3*	NA
呢 *ne0*	NE
喔 *o0*	O
派生 *pai4sheng1*	derivation
旁格 *pang2ge2*	oblique
普通名词 *pu3tong1ming2ci2*	common noun
祈使 *qi2shi3*	imperative
起来 *qi3lai2*	QILAI
起 *qi3*	QI
前置 *qian2zhi4*	preposing
情态 *qing2tai4*	modality
屈折 *qu1zhe2*	inflection
缺位 *que1wei4*	gap

让步语 *rang4bu4yu3*	concessive
施事 *shi1shi4*	agent
示意语力 *shi4yi4yu3li4*	illocutionary force
是 *shi4*	SHI
受词 *shou4ci2*	object
受事 *shou4shi4*	patient
受益者 *shou4yi4zhe3*	beneficiary
双及物 *shuang1ji2wu4*	ditransitive
缩合 *suo1he2*	blending
缩减小句 *suo1jian3xiao3ju4*	reduced clause
缩略词 *suo1lüe4ci2*	acronym
缩写 *suo1xie3*	SUOXIE
所指 *suo3zhi3*	referent
所 *suo3*	SUO
体貌 *ti3mao4*	aspect
体 *ti3*	aspect
条件句 *tiao2jian4ju4*	conditional
通指 *tong1zhi3*	generic
同位语 *tong2wei4yu3*	appositive
哇 *wa1*	WA
外延 *wai4yan2*	denotation
完成体 *wan2cheng2ti3*	perfective
完结(动词) *wan2jie2* (dong4ci2)	accomplishment (verb)
未完成体 *wei4wan2cheng2ti3*	imperfective
未 *wei4*	NEG
无核关系句 *wu2he2guan1xi4ju4*	headless relative clause
无终结 *wu2zhong1jie2*	atelic
系词 *xi4ci2*	copula
辖域 *xia2yu4*	scope
下去 *xia4qu4*	XIAQU
先行语 *xian1xing2yu3*	antecedent
限定词 *xian4ding4ci2*	determiner
小词 *xiao3ci2*	particle
斜格 *xie2ge2*	oblique
形容词 *xing2rong2ci2*	adjective
形态学 *xing2tai4xue2*	morphology
构词 *gou4ci2*	morphology
选择限制 *xuan3ze2xian4zhi4*	selectional restriction
呀 *ya0*	YA

言外力量 *yan2wai4li4liang4*	illocutionary force
衍推 *yan3tui1*	entailment
依存关系 *yi1cun2guan1xi4*	dependency relation
依赖关系 *yi1lai4guan1xi4*	dependency relation
疑问代词 *yi2wen4dai4ci2*	interrogative pronoun
义务(情态) *yi4wu4(qing2tai4)*	deontic
意愿 *yi4yuan4*	volition
音位 *yin1wei4*	phoneme
隐喻 *yin3yu4*	metaphor
哟 *yo1*	YO
有定 *you3ding4*	definiteness
有界 *you3jie4*	bounded
有 *you3*	YOU
语调 *yu3diao4*	intonation
语法功能 *yu3fa3gong1neng2*	grammatical function
语料库 *yu3liao4ku4*	corpus
语码 *yu3ma3*	code
语态 *yu3tai4*	mood; voice
语缀 *yu3zhui4*	affix
预设 *yu4she4*	presupposition
元音 *yuan2yin1*	vowel
元语言否定 *yuan2yu3yan2fou3ding4*	metalinguistic negation
韵律 *yun4lü4*	prosody
在 *zai4*	ZAI
照应链 *zhao4ying4lian4*	anaphoric chain
照应语 *zhao4ying4yu3*	anaphor
着呢 *zhe0ne0*	ZHE_NE
真值条件 *zhen1zhi2tiao2jian4*	truth conditions
真值 *zhen1zhi2*	truth value
正在 *zheng4zai4*	ZAI
之 *zhi1*	DE
直接宾语 *zhi2jie1bin1yu3*	direct object
直接受词 *zhi2jie1shou4ci2*	direct object
直示 *zhi2shi4*	deictic
指称 *zhi3cheng1*	reference
指代形式 *zhi3dai4xing2shi4*	pro form
指示代词 *zhi3shi4dai4ci2*	demonstrative
终结 *zhong1jie2*	telic
主位 *zhu3wei4*	theme

助词 *zhu4ci2*	particle
助动词 *zhu4dong4ci2*	auxiliary (verb)
专有名词 *zhuan1you3ming2ci2*	proper noun
转喻 *zhuan3yu4*	metonymy
状态(动词) *zhuang4tai4*(dong4ci2)	state (verb)
状语 *zhuang4yu3*	adverbial
着 *zhe0*	ZHE

1

Preliminaries

Chu-Ren Huang and Dingxu Shi

A grammar is the system of knowledge of the relation between what people do and what people know when they use a particular language. Since what people know in the context of their language use is often implicit, linguistic theories are proposed as a foundational hypothesis to enable the explicit explanation of a grammar of any particular language. This presents an underlying dilemma in the writing of any grammar. On the one hand, descriptive work is the foundation of any scientific study and is crucial to the language sciences. Modern linguistics emerged as a result of a conscientious effort to move from prescriptive to descriptive studies of language. On the other hand, once any theoretical framework or account is adopted, a grammar becomes prescriptive in the sense that it imposes a set of conceptual primitives and structure of rules prescribed by a sub-set of linguists. How to capture the system of implicit knowledge without prescribing an *a priori* theoretical framework remains the biggest challenge to any descriptive grammar.

A Reference Grammar of Chinese meets this challenge with an empirical approach focused on describing what people do when they use Mandarin Chinese, while allowing generalizations to emerge from our descriptions as well as from the readers' observation of the data. We believe that a keenly observed description of the generalizations and tendencies based on the observation of the extensive data of language use will lead to capturing the implicit knowledge people share without prescribing an explicit rule. To achieve this goal, corpora and Web-extracted examples are used extensively, with an occasional supplementation of made-up sentences. These data have been carefully examined by our authors for their distributional patterns and tendencies. None of the examples cited in this grammar are single instances of language use; rather, they were chosen as an illustrative representation based on a set of similar examples selected by the authors. In other words, this reference grammar is intended to be read like a guide to the Chinese language, mediated by an extensive set of extracted examples for each grammatical point we make. Readers can consult the example database when they read the grammar to strengthen both their understanding of the generalizations and the complexity of language in use.

This grammar assumes a minimal set of theoretical concepts, which include grammatical categories, basic grammatical functions, and some intuitive semantic concepts, such as the thematic roles of agent, theme, goal, etc. A more detailed discussion of the grammatical categories is presented in Chapter 2.

1.1. The Chinese language

Chinese, or Mandarin Chinese, has the most native speakers in the world as well as one of the longest cultural heritages. Mandarin Chinese also has become one of the most learned foreign languages in the world. The 2005 version of *The Language Situation in China* claimed that more than 30 million people in the world were learning Chinese as a foreign language. The need for a linguistically felicitous and accessible reference grammar of the Chinese language is clear and urgent. Authored by leading Chinese linguists in each topic area, this volume serves as a comprehensive and accessible reference grammar of Chinese in that it aims to cover all the important linguistic facts of the language; moreover, these facts are presented in a way that does not presuppose knowledge of a particular linguistic theory or grammar of Chinese.

This grammar provides a synchronic, descriptive grammar of present-day Standard Mandarin Chinese. It shares some of the major design philosophy with that of *The Cambridge Grammar of the English Language* (Huddleston and Pullum 2002).

1.1.1. Standard Mandarin Chinese

Standard Mandarin Chinese has the phonological system of the Beijing variant of Northern Mandarin as its norm of pronunciation. Historically, Mandarin (官话 *guan1hua4*) has been the common language adopted by officials, yet it developed into different variants among the areas where it was primarily spoken, such as Northern Mandarin, Southern Mandarin, and Sichuan Mandarin. In common English usage today, however, Mandarin and Chinese are used interchangeably and loosely to refer to Standard Chinese. The Chinese described in this reference grammar refers to the Standard Mandarin Chinese that is generally accepted in a wide range of public discourse, such as government, education, broadcasting, and publishing. This standard language is referred to as Putonghua (普通话 *pu3tong1hua4* 'common language') in Mainland China, Singapore, and Macau, and as GuoYu (国语 *guo2yu3* 'national language') in Taiwan, while both terms are used in Hong Kong. Broadly speaking, Putonghua follows Mainland China conventions and GuoYu follows Taiwan conventions, and they do differ from each other occasionally, not unlike the contrast between UK and US English. While focusing on the widely accepted usages as reflected in standard written and spoken Chinese, and on the common usages shared by Putonghua and GuoYu, we will point out significant distinctions when necessary.

1.1.2. Synchronic description of present-day Chinese

The earliest record of a well-developed system of Chinese writing dates back to more than three thousand years ago. Although linguists do not agree on all the details, the history of the Chinese language can be divided into four stages: 上古汉语 *shang4gu3han4yu3* 'Old Chinese' (Shang Dynasty to Han Dynasty, sixteenth century BCE–220 CE), 中古汉语 *zhong1gu3han4yu3* 'Middle Chinese' (Southern and Northern Dynasties to Five Dynasties and Ten Kingdoms, CE 220–960), 近代汉语 *jin4dai4han4yu3* 'Early Modern Chinese' (Song Dynasty to May Fourth Movement, 960–1919), and 现代汉语 *xian4dai4han4yu3* 'Modern Chinese' (1919–present). Throughout these periods, substantial changes took place in all linguistic aspects of Chinese. In terms of grammar, Old Chinese can be identified by the lack of more complex constructions, which developed later. Middle Chinese is the period when many new constructions and forms emerged, including the 把/将 *ba3*/*jiang1* constructions, the 被 *bei4* passive, and a pronoun system that is very similar to that of present-day Chinese. During this period, Chinese prepositional phrases also moved from the predominantly post-verbal position to the pre-verbal position. In Early Modern Chinese, aspectual markers such as 了 *le0* and 着 *zhe0* and phrasal suffixes such as 的 *de0* and 地 *de0* were widely used. Writing based on vernacular Chinese (白话文 *bai2hua4wen2*) emerged in the Tang Dynasty, but Classical Chinese (文言文 *wen2yan2wen2*) was still used in formal writing until Modern Chinese, particularly after the May Fourth Movement in 1919, when most publications in China started to use the vernacular language.

The historical change of Chinese is of great linguistic importance and interest, but as a synchronic grammar, this volume limits the description to present-day Modern Chinese, especially the language since 1991, because all the generalizations of this grammar are based on corpora with natural Chinese data collected from that time. It is important to bear in mind, however, that conventionalized historical forms are still used in Modern Mandarin Chinese, especially in formal registers. As such, they are part of present-day Chinese and will be covered in this grammar.

1.1.3. Varieties of Chinese

The term World Chineses (全球华语 *quan2qiu2hua2yu3*), though not as common as World Englishes, is becoming more and more widely used with the increasing popularity of Chinese as a second language and with the Chinese diaspora spreading and growing. Despite the same linguistic heritage, Mandarin Chinese in different regions has evolved in different ways as a result of the political, economic, cultural, and social development of each region. Variations can be found in pronunciation, lexicon, and syntax. While it is important to investigate these

differences, this reference grammar aims to present the shared core of grammar. In the rare cases where the variations render the description of a shared core difficult or if the variations present a special challenge to learners and speakers, observations and descriptions will be provided.

The varieties of Chinese also differ orthographically in adopting traditional characters (Taiwan, Hong Kong, and Macau) or simplified characters (Mainland and Singapore). Orthography projects language identity through the formation of a community of practice and may introduce language variations, although strictly speaking it is not part of the grammar. In this grammar, we adopted simplified characters as the most commonly learned system. It is, however, important to note that although a meaning-preserving mapping from traditional characters to simplified characters can be performed without ambiguity, the same cannot be said for mapping simplified characters to traditional characters.

1.1.4. Chinese dialects vs. Sinitic languages

"Chinese" as the language spoken by ethnic Han people is traditionally divided into seven major groups: Mandarin (or Northern Chinese), Wu, Xiang, Gan, Kejia (Hakka), Yue (Cantonese), and Min. Speakers of different groups are mutually unintelligible in terms of speaking, although they share the same written language and the grammar of written Chinese for each group does not differ substantially from that of Standard Chinese. Such facts bring about the question of whether they should be referred to as dialects or as languages (i.e. Sinitic languages), a linguistic issue with strong cultural, political, and societal implications. Since this grammar concentrates on present-day Standard Chinese, when references to other varieties is necessary, only the language/dialect name will be used, without explicit reference to its language/dialect status.

1.1.5. Descriptive account

The descriptive account of this grammar is succinct and theory-neutral, based on corpus observation and reflecting how the language is actually used. For non-standard or ungrammatical usages, the grammar reports that the usages are rarely or not found in the corpora, rather than providing created examples in contrast to the standard and grammatical examples.

1.1.6. Grammar

We agree with Pullum and Huddleston (2002: 4) that a grammar is divisible into syntax and morphology; the former is concerned with the way words combine to form phrases, clauses, and sentences, while the latter is concerned with the formation of words. Because a word plays a prominent role as a basic unit at

different levels, this grammar pays attention to the definition and identification of words, in addition to other important linguistic facts of the language.

This grammar includes a chapter on classifiers (Chapter 7), a grammatical category less familiar in grammars of English and other Western languages. Classifiers are essential components of noun phrases in Chinese and they represent important selectional and semantic information. In addition, following the conventions established in Chao (1968), this grammar assigns the semantic correspondents of adjectives in English and many other languages, such as 'small' (小 *xiao3*) and 'quite' (安静 *an1jing4*), as well as the category of (state) verbs, and reserves the category of adjectives for the non-predicative types, such as 超级 *chao1ji2* 'super' (see Chapter 10).

1.2. A data-driven and corpus-based reference grammar

While more and more linguistic research and reference grammars have adopted corpus-based empirical approaches, this grammar takes a further step by being both data-driven and corpus-based. The developments in the past thirty years have made it possible for people to access and extract generalizations from corpora. The large-scale corpora accessible for this grammar include the POS-tagged Chinese Gigaword Corpus (1,400 million characters from Mainland China, Taiwan, and Singapore; data collected during 1991 to 2004, Huang 2009) and the manually tagged Sinica Corpus (10 million words collected primarily since 1996, Chen et al. 1996). The availability of large-scale corpora also enables the use of computational tools, such as Word Sketch Engine, to extract grammatical information directly from the corpora. For this grammar, the authors' expertise and judgments have been greatly enhanced by their access to both the Sinica Corpus and the 2nd edition of the Chinese Gigaword Corpus, through the corpus interface of Chinese Word Sketch (Huang et al. 2005). Most of the examples in this grammar, with very few exceptions, have been carefully selected from the corpora with minimal modification. In this sense, this grammar is the first Chinese grammar written based on corpus data. It is also among the first such reference grammars in the world.

As a reference grammar, our emphasis is to get the facts and generalizations right, especially those facts or generalizations missed or mischaracterized by previous grammars. This goal was achieved through a two-pronged empirical approach. First, the authors had access to the largest available Chinese corpora as well as the most powerful corpus interface. In addition, they were encouraged to consult the Web through Google when in doubt. This ensured that the widest range of language data was accessed. Second, each chapter was drafted by a designated author(s) who has done extensive work on the topic area. After the initial draft, the chapter was presented and discussed at authors' workshops, and each chapter

also underwent extensive comments and review by at least two other experts. Lastly, each chapter was reviewed and revised by the two chief editors. In sum, each and every chapter reflects the collective research knowledge of four or more leading Chinese linguists in the field, each contributing to the consistency and comprehensive coverage of the facts.

This reference grammar is anchored by illustrative example sentences. All of these corpus-extracted realistic examples are presented in the standard four-line format: the first line consists of the text in Chinese characters; the second line consists of a word-for-word Pinyin transcription; the third line is aligned with the second line to provide a gloss; and the last line provides faithful free translation. The example [1] is the sentence [2b] from Chapter 4.

[1] 有空的时候, 到公园里去走一走, 呼吸呼吸新鲜空气。

you3kong4	de0	shi2hou0	dao4	gong1yuan2	li3	qu4
have_time	DE	when	go_to	park	in	go

zou3yi1zou3	hu1xi1hu1xi1	xin1xian1	kong1qi4
walk_a_walk	breathe_breathe	fresh	air

'When you have some time, have a walk in the park and breathe some fresh air.'

Following linguistic conventions, the example sentences are discussed and grammatical information is explicated in the text immediately before or after the example given. The above example shows that we have ordered the examples in each chapter according to their order of appearance, and use a, b, c, etc. to differentiate a group of similar sentences given under the same example number.

This grammar will be accompanied by a periodically updated online example database in order to supplement the examples of the grammar and to add value to the restriction of the finite number of printed pages. The original database was constructed together with the grammar, when at least twenty example sentences were selected for each linguistic topic described in the grammar. Each sentence is not only annotated with the topic for which the sentence is selected, but also annotated and indexed with all other relevant linguistic topics covered in the grammar.

This reference grammar aims to make the underlying set of linguistic facts from which we built our generalizations sharable with others who may construct a parallel reference grammar with different design criteria. Therefore, in addition to the example database, a citation database was constructed based on the topics of the grammar. The periodically updated database consists of all the bibliography used for this grammar and all the articles from the major journals of Chinese linguistics, including 中国语文 *zhong1guo2yu3wen2* '*Chinese Language and Writing*'

and *Journal of Chinese Linguistics*. With such a database, this reference grammar will remain relevant with respect to the most updated research topics, even if the printed version is not updated as frequently.

1.3. Chinese writing system

The Chinese writing system is the longest continuously used system in the world. The system is composed of characters (汉字 *han4zi4*, kanji in Japanese as well as in common English translation), which are logographic symbols encoding both phonetic and semantic information. Unlike phonological writing systems, each Chinese character is grounded with some conceptual knowledge information that was conventionalized at the time the character was created. Furthermore, the writing system is considered a cultural symbol that unifies the Chinese people speaking mutually unintelligible varieties of Chinese. By this design, the Chinese writing system is not as arbitrary as phonological writing systems, which are common among other languages in the world. This non-arbitrariness, in turn, has allowed the Chinese writing system to reflect more about the grammar of Chinese; hence, some discussion of the writing system in this reference grammar is required.

It is important to debunk the myth that the Chinese writing system consists of Chinese characters only. This may have been the case as recently as fifty years ago; however, most contemporary Chinese dictionaries nowadays include a few hundred so-called alphabetic words (字母词 *zi4mu3ci2*). These alphabetic words are *bona fide* entries in the lexicons of modern Mandarin Chinese, with full grammatical functions in the category to which they belong, as discussed in Chapter 3. These words can be composed of all alphabets (Q *kyu* 'to have sustainable good texture when chewed on'; HSK *eich-es-kei* Hanyu Shuiping Kaoshi/'Mandarin Standard Test'; IBM *ai-bi-emu* 'International Business Machine'), or start with one or more alphabets but end with a character (AA 制 *ei-ei-zhi4* 'to go Dutch'; K 书 *kei-shu1* 'to hit the books hard'), or start with a character but end with an alphabet (阿 Q *a1-kyu* 'a fatalist, a fictional protagonist in Lu Xun's novel'; 卡拉 OK *ka3la1-ou-kei* 'karaoke'). Unlike Japanese katakana, these alphabetic words are not restricted to loanwords, although many have loanword origins (especially those referring to new technology or brand names). However, the blended translation alternative remains a desirable alternative, so many loanwords are actually represented by Chinese characters (such as 可乐 *ke3le4* 'able-enjoy cola'; 爱疯 *ai4feng1* 'love-crazy iPhone'). It is worth noting that many alphabetic words originate from either Pinyin-based abbreviations or vivid imitations of/associations with the sounds/shapes of the alphabets. Alphabetic words also have a unique linguistic feature in that they do not conform to the phonological integrity of Chinese. The alphabetic parts of the alphabetic words are typically pronounced without an assigned tone, and many

of them represent syllables which are not part of the Chinese phonological repertoire, such as *emu* (from IBM), *kei* (from K 书), and *kyu* (from Q, 阿 Q). They simply represent a conventionalized way to pronounce these alphabets. These alphabetic words can be a noun, a verb, or an adjective, and the alphabet-plus-character template seems to be the most productive pattern of neologism.

1.3.1. A brief history of Chinese script

The Chinese character script has been continuously used for more than 3,000 years. The oracle bone script (甲骨文 *jia3gu3wen2*) of the Shang Dynasty (1600–1046 BCE) is the earliest surviving evidence of a well-developed writing system of Chinese and is directly related to the subsequent Chinese scripts. Oracle bone inscriptions are found mainly on turtle shells (甲 *jia3*) or ox and other large animal bones (骨 *gu3*), hence 甲骨文 (*jia3gu3wen2* 'shell-bone-writing'). These inscriptions are the records of the answers from the divinatory practice of the royal family communicating with their ancestral spirits. Evolving from the oracle bone script, bronze inscription script (金文 *jin1wen2*) is found on ritual bronze vessels of the late Shang Dynasty to Zhou Dynasty (1100–403 BCE). After the Zhou Dynasty, the writing systems of different parts of China diverged until the Qin Dynasty (221–207 BCE) unified China and set the Qin variant of seal script (篆书 *zhuan4shu1*) as the national standard script. For easier and faster writing in government bureaucracy, clerical script (隶书 *li4shu1*) was adopted and it gradually replaced seal script in the Han Dynasty (206 BCE–220 CE). Clerical script is structurally and rectilinearly very similar to the modern Chinese scripts and is thus considered the ancestor of modern scripts. Clerical script then evolved into standard script (楷书 *kai3shu1*) and was replaced by standard script during the Southern and Northern Song Dynasties (420–589 CE). Since then, standard script has been used as the standard form of orthography in China.

Two crucial observations can be made based on the four example characters shown in Table 1.1 below. First, many characters are decomposed into components (such as 明 *ming2*, which consists of the two components 日 *ri4* 'sun' and 月 *yue4* 'moon'), and these components may or may not be characters themselves. This point will be addressed in more detail in the next section. Second, all of the different historical scripts are variants, and the internal structure of each character is largely preserved between script changes. For example, 明 *ming2* 'bright' is composed of the two components 日 *ri4* 'sun' to the left and 月 *yue4* 'moon' to the right. Regardless of how much the graph representing the character changes, both the composition and the left–right structure of these two components remain the same. The consistency of this component composition relation can also be observed to have occurred over time for 渔 *yu2* 'to fish,' as in Table 1.1. Furthermore, even for non-decomposable characters, such as 水 *shui3* 'water' and

Table 1.1 Evolution of Chinese scripts[1]

	'water'	'to fish'	'up'	'bright'
Oracle-bone script				
Bronze script				
Small seal script				
Clerical script	水	漁	上	明
Standard script	水	漁	上	明

[1] The Chinese scripts were extracted from Academia Sinica's Database of Chinese Characters Composition (漢字構形資料庫) at http://cdp.sinica.edu.tw/cdphanzi

上 *shang4* 'up,' it is possible to see that the internal structure of the critical components of the graph stays the same. It is with this consistency of the internal component composition relation that we can show that the Chinese writing system is a single, continuously used system. Any historical text, regardless of the style of script it is in, can be directly mapped to any other script to be read.

1.3.2. Structure of Chinese characters

A character is the smallest meaningful unit of the writing system in Chinese, compared with a morpheme, which is the smallest meaningful unit, and a word, which is the smallest unit with independent syntactic functions. In contrast to phonographic languages, such as English, that are mainly composed of symbols that encode phonetic values only, the character-based writing system of Chinese is featured as logographic in that it is mainly composed of logographic symbols that encode both phonetic and semantic values. Phonetically, each Chinese character represents a syllable, compared to English where a letter or a group of letters represents a phoneme. Semantically, a Chinese character usually encodes a lexical concept, which allows it to stand for the same (or similar) meaning regardless of language changes and variations.

The misconception that Chinese characters cannot be learned without rote memory covering the stroke-by-stroke order of all the strokes of a character has both added to the notoriety of Chinese (and to the myth of the complexity of Japanese, since it uses kanji as one of its complex writing systems) and lent

support to the proposal to convert Chinese to an alphabetic writing system. However, studies have shown that Chinese characters are composed of components (部件 *bu4jian4*). Each component can in turn be composed of smaller components or, eventually, a fixed number and order of strokes. What this means is that recognizing and writing a character only requires knowledge of the components of a character as well as how these components are put together once the basic components are known. There is also a general rule of the order of the composition of left-first, before top-first, and outside-in when other rules do not apply. For instance, the character 明 *ming2* 'bright' is formed by the composition of 日 (on the left) first, and 月 second. The character 盟 *meng2* 'alliance' starts with the same 日+月 sequence, with the third component 皿 *min3* 'basin' last and at the bottom. The character 萌 *meng2* 'to sprout' is formed with the grass radical ⺾ *cao3* on top, followed by the same 日 +月 sequence. These component sequences are largely preserved through the evolution of different scripts (including most cases of simplified characters) and even apply to some regional glyph variants. For instance, 峰 and 峯 *feng1* 'peak' are variants of the same character and they can be described by the same component composition rule of 山+夆 *shan1* + *feng1*, except that one variant follows the left–right order while the other follows the top-down order.

A Chinese character is not only formally composed of components, but its formal composition also follows rules of internal composition. 说文解字 (*shuo1wen2jie3zi4*, 121 BCE, literally, *Explanations of simple graphs and analyses of composite graphs*) compiled by the Eastern Han scholar Xu Shen was the first comprehensive dictionary to analyze the structure of Chinese characters. Xu Shen proposed six principles of Chinese character composition, of which four are firmly established in modern philology: pictographic (象形 *xiang4xing2*), ideographic (指事 *zhi3shi4*), semantic–semantic composition (会意 *hui4yi4*), and semantic–phonetic composition (形声 *xing2sheng1*). Pictographic characters such as 日 *ri4* 'sun' and 月 *yue4* 'moon' resemble the objects in the physical world. Ideographic characters such as 上 *shang4* 'up' and 下 *xia4* 'down' represent abstract ideas. A semantic–semantic compound is typically composed of two or three pictographic or ideographic characters and encodes a combination of the meanings of the characters. For example, 明 *ming2* 'bright' is a combination of the pictographic 日 *ri4* 'sun' and 月 *yue4* 'moon,' while 森 *sen1* 'forest' is composed of three 木 *mu4* 'tree,' which by itself is a pictographic character. Semantic–phonetic compounds typically consist of a phonetic unit and a semantic unit. For instance, 妈 *ma1* 'mother' is composed of the radical 女 *nü3* 'woman' and the phonetic 马 *ma3* 'horse,' representing the phonetic part, suggesting the sound of 妈 when the character was created. The semantic–phonetic composition is considered to be the most frequently used principle, estimated to represent more than 90 percent of the characters; moreover, the radical–phonetic

composition is sometimes taught as the general principle underlying Chinese character formation.

The 540 radicals (部首 *bu4shou3* 'class-head') proposed by Xu Shen, as well as some later modified variations, have become the canonical way to classify characters adopted by traditional Chinese dictionaries. However, whether a group of characters sharing the same radical necessarily has the same basic semantic concept or not has been hotly debated. This hypothesis seems intuitive and works reasonably well with radicals such as 言 *yan2* 'word,' which heads a group of characters, including 语 *yu3* 'language,' 论 *lun4* 'argue,' 讲 *jiang3* 'talk,' and 请 *qing3* 'invite,' which are all types of speech acts. However, this generalization is not as obvious for radicals such as 艸 *cao3*, instantiated as ⺾ *cao3* on top of a character. This radical is supposed to encode the concept '(leafy or grass-like) plant,' such as 兰 *lan2* 'orchid,' 葱 *cong1* 'green onion,' or 芥 *jie4* 'mustard green.' However, among the most frequently used characters with the radical ⺾, there are also characters such as 花 *hua1* 'flower,' 芳 *fang1* 'fragrance, fragrant,' 落 *luo4* 'fallen, to fall,' 菜 *cai4* 'vegetable,' and 药 *yao4* 'medicine' which do not represent any kind of plant. Nevertheless, these characters are semantically dependent on the meaning of the radicals in different ways: 花 *hua1* 'flower' and 萌 *meng2* 'sprout, to sprout' are parts of a plant; 菜 *cai4* 'vegetable' and 药 *yao4* 'medicine' are what plants are used for; and 芳 *fang1* 'fragrance, fragrant' and 落 *luo4* 'fallen, to fall' are salient states crucial for describing plants. Taking this telic (i.e. function and purpose-driven) view, it can be shown that all the characters in a radical group, at least in their original meaning, are semantically dependent on the basic meaning of the radical.

Another generalization underlining the semantic relevance of the Chinese character writing system is the fact that all disyllabic morphemes are written with two characters with identical radicals. There are more than sixty of these disyllabic morphemes, which are traditionally called 联绵词 *lian2mian2ci2* 'interlinked words.' Unlike more prevalent patterns of disyllabic words composed of two morphemes, these morphemes contain two non-morpheme syllables which cannot be used alone because they do not have any independent meaning of grammatical function. 蝴蝶 *hu2die2* 'butterfly' and 葡萄 *pu2tao2* 'grape' are two of the most common examples of a single syllable not being able to occur independently except when it serves as an abbreviation of the disyllabic morpheme. Such disyllabic morphemes can also be found in the verbal category, such as 怂恿 *song3yong3* 'to egg on,' or in the adjectival category, such as 狡猾 *jiao3hua2* 'cunning.' Maintaining the semantic dependency between a morpheme and its written form, the Chinese writing system assigns the two component syllables of a disyllabic morpheme the same radicals without exception, as exemplified by all four examples above. An interesting pair of disyllabic morphemes serves to underline the importance of

semantic dependency in the Chinese character system. It is known that 琵琶 *pi2pa2* 'Pipa, a Chinese stringed lute' is etymologically related to 枇杷 *pi2pa2* 'loquat,' not unlike the recent borrowing of the fruit blackberry by the device BlackBerry. Even though these two morphemes remain homophones, they are conventionalized to be represented by different written forms, the instrument with the 琴 *qin2* 'string instrument' radical (i.e. the top part of both 琵 and 琶) and the fruit with the 木 *mu4* 'wood' radical (i.e. the left part of both 枇 and 杷). Furthermore, the homophonic syllables are not substitutable, as in using one for the other (e.g. 枇 with 琵 for *pi2* or 琶 with 杷 for *pa2*). The component composition nature of characters as graphs and the semantic primitives as the motivation for the grouping of characters by radicals are the two most foundational facts for the linguistic description of the Chinese writing system.

1.3.3. Simplification of Chinese characters

As mentioned above, the seeming complexity of the Chinese character writing system has prompted attempts during the first half of the twentieth century to replace Chinese characters with different alphabetic systems. These efforts of alphabetizing Chinese did not succeed. A more moderate change of adopting a simplified script in 1965 to promote literacy was more successful. In total, 2,235 characters are simplified, and different strategies have been adopted for simplification. For example, 穀 *gu3* 'grain' is simplified to 谷 *gu3*, which is an existing character with the same pronunciation; 書 *shu1* 'book' is simplified to 书 *shu1*, a simpler form based on cursive and informal handwriting; 廠 *chang3* 'factory' became 厂 *chang3* after the half-enclosed component 敞 *chang3* as well as the dot stroke on top were omitted; and 衛 *wei4* 'guard' has been replaced by the newly created form 卫. Of the simplified characters, more than 1,700 are derived from the simplification of the radicals these characters share. For example, 言 *yan2* 'word' as a radical is simplified to 讠, so characters such as 語 *yu3* 'language,' 話 *hua4* 'words,' and 說 *shuo1* 'speak' are changed to 语, 话, and 说, respectively.

Currently, places using simplified characters include Mainland China and Singapore, whereas traditional characters are still used in Taiwan, Hong Kong, and Macau. In countries where Chinese is learned as a foreign language, textbooks with simplified Chinese are increasingly popular, partially because of China's rise as an economic power. The ongoing debates on the merits of the traditional or simplified character systems not only discuss the linguistic benefits and drawbacks of learning and using simplified characters, but also involve the practical, aesthetic, cultural, and political effects of the two concurrent systems. Nonetheless, it is foreseeable that the two systems will coexist for a long time.

1.3.4. Direction of text and use of punctuation

Traditional Chinese text lines run top down, and the vertical text columns are ordered from right to left. Hence, a page starts from the top-right corner and ends at the bottom-left corner. In addition, punctuation was not used in old Chinese writing and printing; thus, 点书 *dian3shu1* (literally, 'to punctuate a book') was deemed to be one of the required basic skills in Chinese scholarship. After the vernacular Chinese movement, however, a modified punctuation system combining the 点书 *dian3shu1* annotation tradition and Western punctuation was introduced and conventionalized and is commonly used (see Appendix I). However, the full stop (i.e. the Chinese period '。') is used parsimoniously and often does not occur until the end of a topic group or a paragraph, with commas used more freely even at the end of a sentence. This is probably due to the influence of the 点书 *dian3shu1* tradition, where full stops mark the natural end when reading a text.

Horizontal texts are widely used now with the influence of Western texts and books. The left-to-right text orientation has become the standard for horizontal text, especially for scientific, translation, or new media texts. However, the vertical text is still commonly used for classical texts. Regions using traditional characters tend to use more vertical texts, while regions adopting simplified characters use predominantly horizontal texts. It should be noted that right-to-left horizontal texts can still be found, especially in traditional signs or banners.

2

Syntactic overview

Dingxu Shi and Chu-Ren Huang

The description of Chinese grammar in this book is based on the following principles: sentences are composed of parts, which may themselves have parts; sentence parts belong to a limited range of types; and different parts of a sentence have specific roles or functions within the larger parts to which they belong. This description is divided into two major components: morphology and syntax. The former deals with how words are put together, while the latter deals with how words are combined to make phrases, clauses, and then sentences.

2.1. Morphemes, words, and word classes

A morpheme is the smallest unit with meaning. A word is the smallest unit that has independent functions in syntax, while a sentence is the largest syntactic unit. The main constituent of a sentence is a clause, or several coordinated clauses, which represents a proposition and usually appears in the form of a subject–predicate construction. The other constituents in a sentence are final particles, which are attached to the end of the main clause to form a sentence while providing additional and necessary information. A clause can function as part of another clause, namely, as a subordinate clause. The clause that is not contained in any other clause is the main clause. The subject, predicate, and their modifiers are phrases, while the object inside the predicate is also a phrase. A phrase has one or more words as its immediate parts and one of the parts is its head, which determines its syntactic status. For example, a verb phrase has a verb as its head and an adjective phrase has an adjective as its head.

2.1.1. Morphemes

Most native Chinese morphemes are monosyllabic, like those in [1a]. Native disyllabic morphemes, such as those in [1b], are few in number and a large portion of them have either a consonant alliteration, such as 蜘蛛 *zhi1zhu1* 'spider,' or a vowel rhyme, such as 迷离 *mi2li2* 'bewildered.' Some disyllabic morphemes are borrowed from other languages, such as those in [1c]. Multisyllabic morphemes are predominantly loan items, such as those in [1d].

[1] a. 猪 *zhu1* 'pig,' 山 *shan1* 'mountain,' 铁 *tie3* 'iron,' 花 *hua1* 'flower,' 跑 *pao3* 'run,' 哭 *ku1* 'cry,' 扔 *reng1* 'throw,' 杀 *sha1* 'kill,' 追 *zhui1* 'chase,' 好 *hao3* 'good,' 旧 *jiu4* 'old,' 脏 *zang1* 'dirty,' 臭 *chou4* 'stinking,' 大 *da4* 'big'

b. 蜘蛛 *zhi1zhu1* 'spider,' 秋千 *qiu1qian1* 'swing,' 玫瑰 *mei2gui0* 'rose,' 迷离 *mi2li2* 'bewildered,' 嘹亮 *liao2liang4* 'loud and clear,' 从容 *cong2rong2* 'calm, sufficient'

c. 骆驼 *luo4tuo0* 'camel,' 克隆 *ke4long2* 'clone,' 扑克 *pu1ke4* 'playing cards'

d. 爱因斯坦 *ai4yin1si1tan3* 'Einstein,' 洛杉矶 *luo4shan1ji1* 'Los Angeles,' 伏特加 *fu2te4jia1* 'vodka'

2.1.1.1. Inflectional morphemes

Inflection does not play a major role in Chinese syntax. The few inflectional morphemes which do play a role include the plural marker -们 *men0* that is a suffix to personal pronouns and certain human nouns, aspectual markers -着 *zhe0*, -了 *le0*, and -过 *guo4* that are suffixes to verbs, and, to a certain extent, infixes like -得- *de0* '-able' or its negative form 不 *bu4* 'NEG' that appear inside a V(erb)-R(esultative) complex word.

2.1.1.2. Derivational morphemes and word formation

Most Chinese morphemes, no matter whether free or bound, or whether content or functional, are active in word formation. The derivation of Chinese words in terms of the status of their component morphemes is discussed in Chapter 3.

2.1.2. Words

A few thousand Chinese lexical items are simple words containing a single morpheme, although many of them are not frequently used any more. Many more Chinese words are compounds consisting of two or more morphemes. One common way to form a compound is by simply juxtaposing the morphemes in a predetermined order, like similar processes in many languages, as discussed in detail in Chapter 3.

Another way is reduplication, a productive lexical process in Chinese that duplicates the entire word or part of it. Reduplication does not typically change the inherent meaning of the root. It mostly targets monosyllabic or disyllabic roots and the output is characteristically in the form of Aa, AaBb, ABab, or AaB, where A represents the first syllable, B the second syllable, and a or b the duplicated syllable.

Monosyllabic roots that undergo reduplication can be verbs, as in [2a], or adjectives, as in [2b]. If a verb represents an action lasting for some time, its reduplicated

form typically stands for a shortened duration that bears some casual or tentative flavor. The 听听 *ting1ting0* in [2a] is thus more casual than the root 听 *ting1* 'listen.' If a verb represents an action that ends instantly, its reduplicated form typically means repeating the action, and it bears some casual flavor as well. The 跳跳 *tiao4tiao0* in [2a] thus means to repeat the action of 跳 *tiao4* 'jump' in a lighthearted way. If a gradable adjective undergoes reduplication, the process will alter the gradation of its meaning and the effect of the change will depend on the new form's syntactic function. The reduplicated version of 大 *da4* 'big' means less big but quite cute when it modifies a noun or functions as a predicate, as in 大大的眼睛 *da4da4 de0 yan3jing1* 'lovely biggish eyes' and 眼睛大大的 *yan3jing1 da4da4 de0* 'eyes are lovely big,' but it means bigger than normal when it modifies a verb, as in 大大地喝了一口 *da4da4 de0he1 le0 yi1 kou3* 'take a very big sip.'

[2] a. 听听 *ting1ting0* 'listen casually,' 跳跳 *tiao4tiao0* 'jump repeatedly'
b. 大大 *da4da4* 'rather big,' 红红 *hong2hong2* 'reddish'

Not many nouns can undergo reduplication, and many duplicated forms have become fixed expressions with derived meanings. The expression 山山水水 *shan1shan1shui3shui3* 'mountains and rivers' is predominantly used in its derived meaning of 'scenery,' even though it is indeed the duplicated version of 山水 *shan1shui3* 'mountain and river.'

When a disyllabic adjective undergoes reduplication, it usually takes the AaBb or ABab form, while it occasionally takes the AaB form. An adjective consisting of two coordinated morphemes predominantly takes the AaBb form in reduplication, as in [3a]. If the first morpheme is the modifier of the second one, then reduplication predominantly takes the ABab form, as in [3b], and sometimes takes the AaB form, as in [3c].

[3] a. 漂漂亮亮 *piao4piao4liang1liang1* 'beautiful,' 整整齐齐 *zheng3zheng3qi2qi2* 'orderly'
b. 雪白雪白 *xue3bai2xue3bai2* 'snow white,' 滚烫滚烫 *gun3tang4gun3tang4* 'boiling hot'
c. 通通红 *tong1tong1hong2* 'fire red,' 冰冰凉 *bing1bing1liang2* 'icy cold'

When it undergoes reduplication, a disyllabic verb generally takes the ABab form, as in [4a]. Like its Aa counterpart derived from a monosyllabic root, an ABab verb typically represents a shortened duration of action or a repeated instant action, and thus bears some casual or tentative flavor. Unlike its root or its Aa counterpart, an ABab verb generally takes a definite object; thus, the verbs in [5a] and [5b] cannot change places. Some V–O sequences are built with one or two

bound morphemes and thus behave like ordinary verbs but not verb phrases. Such a verb typically takes the AaB form in reduplication, as in [4b]. There are a small number of AaBb compound verbs, which form a closed class. A few of them have corresponding AB compounds as origin, as in [4c], but many of them seem to be formed through conjunction of reduplicated monosyllabic verbs Aa and Bb, as in [4d].

[4] a. 休息休息 *xiu1xi0xiu1xi0* 'rest for a while,' 讨论讨论 *tao3lun4tao3lun4* 'discuss informally,' 照顾照顾 *zhao4gu0zhao4gu0* 'give some special treatment, favor'

b. 打打拳 *da3da3quan2* 'play boxing casually,' 养养神 *yang3yang3shen2* 'cultivate spirit, rest,' 游游泳 *you2you2yong3* 'swim casually'

c. 蹦蹦跳跳 *beng4beng4tiao4tiao4* 'jump repeatedly,' 摇摇摆摆 *yao2yao2bai3bai3* 'swing, vacillate'

d. 跌跌撞撞 *die1die1zhuang4zhuang4* 'stumble,' 走走停停 *zou3zou3ting2ting2* 'walk and stop'

[5] a. 咱们先讨论一个问题。

zan2men0	xian1	tao3lun4	yi1	ge4	wen4ti2
we	first	discuss	one	CL	problem

'Let's discuss the problem first.'

b. 咱们先讨论讨论这个问题。

zan2men0	xian1	tao3lun4tao3lun4	zhe4	ge4	wen4ti2
we	first	discuss	this	CL	problem

'Let's discuss this problem first.'

A third way of constructing a compound is to arrange morphemes according to syntactic relations. All the structural relations found in clauses are also found inside compounds and the overall meaning of such a compound is derived on the basis of that relation. The verbal morpheme in [6a] takes the nominal morpheme as the object and such a V–O compound predominantly has the reading of an intransitive verb. Some V–O compounds such as 读书 *du2shu1* 'attend school' have obtained a metaphoric interpretation while some others such as 怀疑 *huai2yi2* 'suspect' have become transitive. The two morphemes in the compounds of [6b] also have a verb–object relationship and could produce a verb reading, but they also have a derived meaning of the person who does the action. The two morphemes in [6c] have a subject–predicate relationship. Most of these compounds represent events or natural phenomena, either as a noun or as a verb, but some

of them have undergone changes to represent something else, such as pampering represented by 心疼 *xin1teng2*.

[6] a. 怀疑 *huai2yi2* 'hold-doubt, suspect,' 用功 *yong4gong1* 'use-effort, work hard,' 读书 *du2shu1* 'read-book, attend school'

b. 司机 *si1ji1* 'control-machine, driver,' 司令 *si1ling4* 'control-order, commander,' 掌勺 *zhang3shao2* 'hold-spoon, chef'

c. 海啸 *hai3xiao4* 'sea-shout, tsunami,' 眼红 *yan3hong2* 'eye-red, jealous,' 心疼 *xin1teng2* 'heart-ache, feel-sorry-for, pamper'

The two morphemes in [7] are in a conjunctive pattern: two nominal morphemes conjoined to form a nominal compound like 人民 *ren2min2* 'people,' two verbal morphemes conjoined to form a verb like 算计 *suan4ji0* 'to plot against,' and two adjectival morphemes conjoined to form an adjective like 漂亮 *piao4liang0* 'beautiful.' The meaning of the compound can be a simple sum of its morpheme components, like that of 攻击 *gong1ji1* 'attack,' but it can also be metaphorically derived, like that of 江山 *jiang1shan1* 'territory, realm.'

[7] 人民 *ren2min2* 'person-person, people,' 江山 *jiang1shan1* 'river-mountain, territory, realm,' 攻击 *gong1ji1* 'attack-hit, attack,' 漂亮 *piao4liang0* 'beautiful-bright, beautiful,' 辛苦 *xin1ku3* 'spicy-bitter, hard'

The first verb morpheme in [8] represents an action and the second one represents the resulting status brought about by the action, just like in resultative constructions. Such a compound characteristically has the meaning and function of a verb.

[8] 推翻 *tui1fan1* 'push-overturn, overthrow,' 冻僵 *dong4jiang1* 'freeze-stiff, frozen stiff,' 改良 *gai3liang2* 'change-good, improve,' 说明 *shuo1ming2* 'speak-clear, explain'

The second nominal morpheme in [9] is modified by the first morpheme, as a nominal like 钢梁 *gang1liang2* 'steel-beam,' an adjectival like 红旗 *hong2qi2* 'red-flag,' or a verbal like 行人 *xing2ren2* 'pedestrian.' Such a compound is predominantly a noun. The second verbal morpheme in [10] is modified by the first one, as a verbal like 回顾 *hui2gu4* 'back-look, reflect' or as a nominal like 席卷 *xi2juan3* 'mat-roll, conquer totally.' Such compounds are predominantly verbs. The meaning of such compounds can be a sum of its components, like that of 刀劈 *dao1pi1* 'cut with a knife,' but it can also be derived metaphorically, like that of 席卷 *xi2juan3*, which means to take control over a large area as if rolling the whole place up like a mat.

[9] 钢梁 *gang1liang2* 'steel-beam,' 电影 *dian4ying3* 'electricity-shadow, movie,' 红旗 *hong2qi2* 'red-flag,' 黑手 *hei1shou3* 'black-hand, evil backstage manipulator,' 行人 *xing2ren2* 'walking-person, pedestrian,' 跑车 *pao3che1* 'running-car, sports car'

[10] 回顾 *hui2gu4* 'back-look, reflect,' 赡养 *shan4yang3* 'back-up-raise, support parent,' 刀劈 *dao1pi1* 'knife-cut, cut with a knife,' 席卷 *xi2juan3* 'mat-roll, conquer totally'

There is no overt marking within a compound to indicate the syntactic relationship between its roots and this may lead to ambiguity as in the case of 烤肉 *kao3rou4*, which consists of a verb 烤 *kao3* 'roast' and a noun 肉 *rou4* 'meat.' It could be understood as a V–O compound to mean the action of roasting meat, but it could also be interpreted as a noun modified by a verb to mean barbequed meat.

The interpretation of the relation between roots of a compound is subject to constraints such as semantic compatibility, cultural tradition, world knowledge, and idiomatic usage. For instance, both 买方 *mai3fang1* 'buy-side, buyer' and 买凶 *mai3xiong1* 'buy-killer, hire assassin' have a V–N structure with the same verb but the interpretations of the role of N differ. The 方 *fang1* 'side' in 买方 *mai3fang1* 'buy-side, buyer' is a bound root that refers to a participant in an activity, and it is predominantly used in nominal compounds like 卖方 *mai4fang1* 'sell-side, seller,' 军方 *jun1fang1* 'army-side, the military' and 警方 *jing3fang1* 'police-side, the police' to represent such entities. The 凶 *xiong1* 'killer' in 买凶 *mai3xiong1* 'hire assassin' is also a bound nominal root standing for killer or killers. It could be used either in nominal compounds like 真凶 *zhen1xiong1* 'real killer' to represent entities or in V–O compounds like 追凶 *zhui1xiong1* 'chase-killer, assassin-hunt' to represent actions.

2.1.3. Word classes

The role of a word is determined by its meaning and the syntactic properties derived from its meaning. Words are therefore categorized according to their meaning: verbs, nouns, numerals, classifiers, adjectives, adverbs, prepositions, coordinators, interjections, sentence-final particles, and onomatopoeia.

2.1.3.1. Verbs

Prototypic verbs represent actions. A verb will appear in its root form or bear an aspect marker if it is the head of a predicate. The root form is the only choice if the verb directly modifies a noun, a verb, or an adjective.

Verbs can be classified according to the number of arguments they take or the situation type they describe. Certain words are classified as verbs even though they do not actually represent actions. One type is copulas such as 是 *shi4* 'be' and linking verbs such as 属于 *shu3yu2* 'belong to.' Another type is modal auxiliaries, which can be classified into epistemic modals such as 会 *hui4* 'will' and 可能 *ke3neng2* 'can'; deontic modals such as 必 *bi4* 'must' and 可以 *ke3yi3* 'may'; and dynamic modals such as 肯 *ken3* 'be willing' and 愿意 *yuan4yi4* 'be willing.'

2.1.3.2. Nouns

Nouns represent entities that can be concrete, abstract, or imaginary. Pronouns represent entities indirectly and are treated as a subcategory of nouns.

A noun can function as the head, the complement, or the modifier in a compound. It can also function as the ultimate head of a nominal phrase. The prototypical function of a nominal phrase is to be an argument in a clause. The ultimate head of a nominal phrase can sometimes be a deverbal element like the 检修 *jian3xiu1* 'maintaining' in [11], which represents an action but has the syntactic properties of a noun. Although 检修 usually takes an argument with a patient role, the patient phrase cannot appear after 检修 as an object in [11] but instead has to appear before it as a modifier.

[11] 工程处没有进行设备的检修。
gong1cheng2 chu4 mei2you3 jin4xing2
engineering department NEG perform
she4bei4 de0 jian3xiu1
equipment DE maintaining
'The Engineering Department did not carry out the equipment maintenance.'

Another function of deverbal nouns is to modify nouns in compounds. Arguments of such a deverbal word can no longer function as the subject or the object. Instead, they become the modified like the 工人 *gong1ren2* 'worker' in [12] or the modifier like the 汽车 *qi4che1* 'car' in [12].

[12] 汽车修理工人
qi4che1 xiu1li3 gong1ren2
car repairing worker
car-repairing worker

2.1.3.3. Numerals and classifiers

Numerals represent the result of counting, while classifiers are the measurement units or the natural division units with which the counting is done.

Numerals and classifiers predominantly appear as components of a nominal phrase with the order of determinative–numeral–classifier–noun, like the one in [13a]. A numeral–classifier sequence occasionally functions as a predicate like the 三十八个 *san1shi2ba1 ge4* 'thirty-eight pieces' in [13b].

[13] a. 我要那两张桌子。
wo3 yao4 na4 liang3 zhang1 zhuo1zi0
I want that two CL desk
'I want those two tables.'

b. 三筐南瓜一共三十八个。
san1 kuang1 nan2gua1 yi1gong4 san1shi2ba1 ge4
three CL pumpkin altogether thirty_eight CL
'There are altogether thirty-eight pumpkins in these three baskets.'

2.1.3.4. Adjectives

Adjectives predominantly represent properties of entities, including size, shape, color, age, state, hardness, heaviness, dimension, speed, kindness, cruelty, simplicity, complication, and so on.

A major function of most adjectives is to be the head of a predicative phrase, which characteristically consists of an adjective and a degree adverbial, like the 聪明 *cong1ming0* 'smart' and 很 *hen3* 'very' in [14a]. An adjective phrase can also be a postverbal descriptive expression, as in [14b].

[14] a. 这小女孩很聪明。
zhe4 xiao3 nü3hai2 hen3 cong1ming0
this little girl very smart
'This girl was very smart.'

b. 那位老师讲得非常清楚。
na4 wei4 lao3shi1 jiang3 de0 fei1chang2
that CL teacher explain DE very
qing1chu3
clear
'The teacher provided an explanation that made it very clear.'

Most adjective phrases can function as subjects or objects, as in [15a] and [15b]. Certain adjective phrases can function as adverbials, as in [15c].

[15] a. 过分小心就不值得了。
guo4fen4 xiao3xin1 jiu4 bu4 zhi2de2 le0
over cautious thus NEG worth LE
'It is not worth it to be overcautious.'

b. 他一直在追求完美。
ta1 yi1zhi2 zai4 zhui1qiu2 wan2mei3
he all_the_time be_at strive_for perfect
'He has been striving for perfection all the time.'

c. 火车慢慢停了下来。
huo3che1 man4man4 ting2 le0 xia4lai2
train slowly stop LE down
'The train slowly came to a halt.'

An adjective can directly modify a noun as in 小河 *xiao3he2* 'small river, stream.' An adjective phrase cannot modify a nominal phrase directly but can do so as part of a relative clause.

2.1.3.5. Adverbs

Adverbs describe particular aspects of an action or a state, depicting the degree, scope, frequency, and contextual relations of a predicate or indicating the speaker's stance about a proposition.

An adverb can modify a verb or an adjective in a compound, as in 常备 *chang2bei4* 'constantly prepare' and 绝妙 *jue2miao4* 'absolutely wonderful.' An adverb phrase can modify a predicate, as in [16a], or modify a clause, as in [16b]. Although the 居然 *ju1ran2* 'unbelievably' in [16c] modifies the predicate, it actually represents the speaker's stance on the issue.

[16] a. 我们从来没有深入交谈过。
wo3men0 cong2lai2 mei2you3 shen1ru4 jiao1tan2
we ever NEG deeply communicate
guo4
GUO
'We have never talked to each other in depth.'

b. 总之这件事已经结束了。
zong3zhi1 zhe4 jian4 shi4 yi3jing1 jie2shu4 le0
In_a_nutshell this CL issue already finish LE
'In a nutshell, the incident has come to an end.'

c. 这么一大笔钱居然就不见了。
zhe4me0 yi1 da4 bi3 qian2 ju1ran2 jiu4
so one big CL money unexpectedly thus
bu4jian4 le0
disappear LE
'Unbelievably, such a large sum of money vanished.'

2.1.3.6. Prepositions

A preposition establishes the relationship between its object and the action represented by a verb phrase. The object of a preposition is typically a nominal phrase, as in [17a], but it can also be an adjective phrase or a verb phrase, as shown in [17b]. Some prepositions can take a clause as a complement, and the 跟 *gen1* 'with' in 跟...没有关系 *gen1...mei2you3 guan1xi0* 'have no relation with' in [17c] is a typical case.

[17] a. 他又朝杯子里倒了一些红酒，灌了下去。

ta1	you4	chao2	bei1zi0	li3	dao4	le0	yi1xie1
he	again	PREP	cup	inside	pour	LE	some

hong2jiu3,	guan4	le0	xia4qu4.
red_wine	gulp	PERF	down

'He added some wine to the cup, and gulped it down.'

b. 我们只好从进攻转入防守。

wo3men0	zhi3hao3	cong2	jin4gong1	zhuan3ru4
we	can_only	PREP	offense	switch_to

fang2shou3
defense
'We had no choice but to turn from offense to defense.'

c. 我跟公司亏损没有任何关系。

wo3	gen1	gong1si1	kui1sun3	mei2you3	ren4he2
I	PREP	company	loss	NEG_YOU	any

guan1xi0
relation
'I had nothing to do with the company's losing money.'

2.1.3.7. Coordinators and connective adjuncts

Coordinators mark the conjunction of two or more elements but contribute little to the overall meaning of the coordination. Connective adjuncts are mostly adverbs and adjectives that signify the logical relation between clauses, while maintaining certain modifying or predicative functions.

Coordinators that signify conjunctive relations include 和 *he2* 'and,' 跟 *gen1* 'and,' 同 *tong2* 'and,' 与 *yu3* 'and,' 及 *ji2* 'and,' and 以及 *yi3ji2* 'and' that conjoin nominal elements. 和 *he2* 'and' is used in most contexts, 跟 *gen1* 'and' and 同 *tong2* 'and' are mainly used in colloquial speech, and 与 *yu3* 'and' typically appears in formal writing. When two nominal elements are conjoined by 及 *ji2* 'and' or 以及 *yi3ji2* 'and,' the one appearing before the coordinator is typically given more

importance than the one after it, as in the case of [18b]. 并 *bing4* 'and,' 并且 *bing4qie3* 'and,' and 而且 *er2qie3* 'and' predominantly mark the conjunction of verbal, adjectival, or clausal elements, as in [19].

[18] a. 老师和学生 *lao3shi1he2xue2sheng1* 'teachers and students,' 生姜跟大蒜 *sheng1jiang1gen1da4suan4* 'ginger and garlic,' 司机同乘客 *si1ji1tong2cheng2ke4* 'driver and passengers,' 名词与动词 *ming2ci2yu3dong4ci2* 'nouns and verbs'

b. 美国总统奥巴马及随行记者
mei3guo2 zong3tong3 ao4ba1ma3 ji2 sui2xing2
American president Obama and accompany
ji4zhe3
reporter
'American President Obama and accompanying reporters'

[19] a. 会议讨论并通过了2010财政预算。
hui4yi1 tao3lun4 bing4 tong1guo4 le0 2010
meeting discuss and pass LE 2010
cai2zheng4 yu4suan4
finance budget
'The meeting discussed and passed the budget for the 2010 fiscal year.'

b. 消防车早已进入火场, 而且突击队已经找到了火源。
xiao1fang2che1 zao3 yi3 jin4ru4 huo3chang3
fire_engine early already enter scene_of_fire
er2qie3 tu1ji1dui4 yi3jing1 zhao3dao4 le0
and commando already find LE
huo3yuan2
source_of_fire
'The fire engines have entered the scene of the fire and the commandos have found the source of the fire.'

但 *dan4* 'but' and 但是 *dan4shi4* 'but' predominantly conjoin two verbal, adjectival, or clausal elements and indicate an adversative relation, as in [20a]. Although 可 *ke3* 'but' can be a modal and 可是 'but' can be an adverb, they can also be used as a coordinator to indicate an adversative relation, as in [20b]. 而 *er2* 'and' is mainly used in formal written genres to represent a distinctive or adversative relation, as in [21].

[20] a. 我的工作不错, 但工资低了一点。
wo3de0 gong1zuo4 bu4cuo4 dan4 gong1zi1 di1
my job NEG_bad but salary low
le0 yi1dian3
LE a_bit
'My job is not too bad, but the pay is a little low.'

b. 这套三居室的房子不大, 可是却非常实用。
zhe4 tao4 san1ju1shi4 de0 fang2zi0 bu4 da4
this CL three_bedroom DE apartment NEG big
ke3shi4 que4 fei1chang2 shi2yong4
but yet very practical
'This three-bedroom apartment is not big, but it's very practical.'

[21] 这种杂交西瓜大而不甜。
zhe4 zhong3 za2jiao1 xi1gua1 da4 er2 bu4
this CL hybrid water_melon big but NEG
tian2
sweet
'The hybrid water melon is big but not sweet.'

Coordinators that indicate disjunctive relations include 或 *huo4* 'or' and 或者 *huo4zhe3* 'or.' They mark the disjunction of two nominal elements, as in [22a], or two verbal or clausal elements, as in [22b].

[22] a. 这款新车可以用汽油或天然气推动。
zhe4 kuan3 xin1 che1 ke3yi3 yong4 qi4you2 huo4
this CL new car can use gasoline or
tian1ran2qi4 tui1dong4
natural_gas propel
'The new car is powered by gasoline or natural gas.'

b. 医药费由保险公司先行支付, 或者由肇事者垫付。
yi1yao4fei4 you2 bao3xian3 gong1si1 xian1xing2
medical_bill PREP insurance company in_advance
zhi1fu4 huo4zhe3 you2 zhao4shi4zhe3 dian4fu4
pay or PREP perpetrator pay_in_advance
'The medical bill should be settled by the insurance company first or be settled by the person who caused the accident.'

Although it is common to use coordinators in conjunctions, it is also quite normal not to use any marker in conjunctive constructions. The three nominal phrases 红烧肉 *hong2shao1rou4* 'red cooked pork,' 炒鸡蛋 *chao3ji1dan3* 'scrambled eggs,' and 家酿米酒 *jia1niang4mi3jiu3* 'home-made rice liquor' in [23] are conjoined together without any markers, as are the two clauses in this compound sentence.

[23] 桌上摆着红烧肉、炒鸡蛋、家酿米酒, 散发着诱人的香味。
zhuo1shang4 bai3 zhe0 hong2shao1rou4
table_on place ZHE red_cooked_pork
chao3ji1dan4 jia1niang4mi3jiu3 san4fa1 zhe0
scrambled_eggs home_made_rice_liquor emit ZHE
you4ren2 de0 xiang1wei4
tempting DE smell
'On the table lay red cooked pork, scrambled eggs, and home-made rice liquor, and they smelled wonderful.

Connective adjuncts mainly signify the logical relation between clauses in compound sentences and certain types of complex sentences, while they also function as adverbials, modals, or predicates. They typically work in pairs but sometimes half of a pair can represent the same logical relation. There are also some connective adjuncts that usually work alone.

Connective adjuncts are categorized according to the logical relationship they represent. The relationship between two clauses in a compound sentence can be an equal one represented by 既 . . . 又 . . . *ji4* . . . *you4* . . . 'while . . . while . . . ,' as in [24a], a temporal sequence represented by 一 . . . 就 . . . *yi1* . . . *jiu4* . . . 'as soon as . . . then . . . ,' a progressive sequence signified by 不仅 . . . 而且 . . . *bu4jin3* . . . *er2qie* . . . 'not only . . . but also . . . ,' or an adversative one indicated by 不仅不 . . . 反而 . . . *bu4jinbu4* . . . *fan3er2* . . . 'not only not . . . but on the contrary . . . ,' as in [24b].

[24] a. 处理事故时, 你们既要有原则性, 又要有灵活性。
chu3li3 shi4gu4 shi2 ni3men0 ji4 yao4 you3
handle accident when you also need have
yuan2ze2xing4 you4 yao4 you3
quality_of_sticking_to_principles but_also need have
ling2huo2xing4
flexibility
'When handling the accident, you should uphold the principles while exercising enough flexibility.'

b. 小偷不仅没有逃走，反而将失主打伤。
xiao3tou1 bu4jin3 mei2you3 tao2zou3 fan3er2
thief not_only NEG run_away instead
jiang1 shi1zhu3 da3shang1
BA loss_owner beat_hurt
'The thief did not run away but instead beat the victim.'

Other connective adjunct pairs of equal status include 边...边...*bian1*...*bian1*...一面...一面...*yi1mian4*...*yi1mian4*...and 一方面...一方面...*yi1fang1mian4*...*yi1fang1mian4*...all of which are based on the space metaphor 'on the one side...on the other side,' 起先...后来...*qi3xian1*...*hou4lai4*...'at the beginning,' 首先...然后...*shou3xian1*...*ran2hou4*...'at first...then...,' 接着...然后...*jie1zhe*...*ran2hou4*...'right after...then...,' and 然后...终于...*ran2hou4*...*zhong1yu2*...'then...at last...,' which indicate a temporal sequence; 不但 *bu4dan4*/不单 *bu4dan1*/不独 *bu4du2*/不光 *bu4guang1*/不只 *bu4zhi3*...甚至...*shen4zhi4*/并且 *bing4qie3*...'not only...but also...,' which represents a progressive sequence; and 不但...反而...*bu2dan4*...*fan3er2*...'not only...on the contrary...' and 不但不...反而...*bu2dan4bu4*...*fan3er2*...'not only not...on the contrary...,' which represent an adversative relation.

Connective adjuncts that typically appear alone include 同时 *tong2shi2* 'at the same time,' 此外 *ci3wai4* 'in addition,' and 另外 *ling4wai4* 'in addition' for equal status; 接着 *jie2zhe* 'following that,' 甚至 *shen4zhi4* 'even,' and 然后 *ran2hou4* 'afterwards' for a progressive sequence; and 反之 *fan3zhi1* 'on the contrary,' 反而 *fan3er2* 'on the contrary,' 反倒 *fan3dao4* 'on the contrary,' 然而 *ran2er2* 'however,' 然则 *ran2ze2* 'however,' and 则 *ze2* 'on the other hand' for adversity.

In disjunctive compound sentences, the logical relation between the clauses can be a pure choice that is marked by pairs like the 是...还是...*shi4*...*hai2shi4* 'either...or...' in [25a] or 要么...要么...*yao4me0*...*yao4me0*...'either...or...' Sometimes the choice is a random one, such as...也好...也好...*ye3hao3*...*ye3hao3* '...as well...as well' occurring after the compared items. The choice can also be a preference introduced by pairs, such as 宁愿...也不...*ning4yuan4*...*ye3bu4*...'would rather...but not...' or pairs like the 与其...还不如...*yu3qi2*...*hai2bu4ru2*...'rather than...would be better...' in [25b].

[25] a. 咱们是从东京走还是从首尔走?
zan3men shi4 cong2 dong1jing1 zou3 hai2shi4 cong2
we be via Tokyo go or via
shou3er3 zou3
Seoul go
'Should we go via Tokyo or via Seoul?'

b. 与其在街上闲逛, 还不如去打篮球。
yu3qi2 zai4 jie1shang4 xian2guang4 hai2bu4ru2
rather_than be_at street_on wander rather
qu4 da3 lan2qiu2
go play basketball
'It is better to play basketball than to wander aimlessly on the street.'

Other disjunctive marker pairs for choice include the copular 不是...就是...*bu2shi4*...*jiu4shi4*...'either...or...', or the predicative...也罢...也罢...*ye3ba4*...*ye3ba4* '...as well...as well' following the choice items. Additional disjunctive modal pairs for preference include 宁肯...也要...*ning4ken3*...*ye3yao4*...'would rather...but would...', 宁可...也不...*ning4ke3*...*ye3bu4*...'would rather...but would not...' 与其...倒不如...*yu3qi2*...*dao4bu4ru2*...'rather than...but would be better...,' and 与其...宁可...*yu3qi2*...ning4ke3...'rather than...but would be better...'

The disjunctive relation between clauses can be represented by a single marker as well. In addition to half-pair markers like 还是 *hai2shi4* 'or be,' 还不如 *hai2bu4ru2* 'or would be better,' 宁肯 *ning4ken3* 'would rather,' and 宁可 *ning4ke3* 'would rather,' disjunctive markers like 抑 *yi4* 'or,' 抑或 *yi4huo4* 'or,' and 再不 *zai4bu4* 'otherwise' can also work alone.

Certain complex sentences are marked with the conjunctive adjunct to indicate the logical relation between the main clause and the subordinate clause. The relation can be that of cause and effect signified by pairs like 因为...所以...*yin1wei2*...*suo2yi3*...'because...therefore...' in [26a] or 由于...于是...*you2yu2*...*yu2shi4*...'for the reason that...as a consequence....' It can also be that of an induction marked by 既然...也就...*ji4ran2*...*ye3jiu4*...'given that...therefore...' in [26b].

[26] a. 因为旅客人数众多, 所以铁路局加开了两趟列车。
yin1wei4 lü3ke4 ren2shu4 zhong4duo1
because passenger number_of_people many
suo3yi3 tie3lu4ju2 jia1kai1 le0 liang3 tang4
so Railroad_bureau add LE two CL
lie4che1
train
'Because there were too many passengers, the Railroad Bureau added two trains.'

b. 既然嘉宾们都来了, 宴会也就提前开始。
ji4ran2 jia1bin1men0 dou1 lai2 le0 yan4hui4
since honored_guests all come LE banquet

ye3jiu4 ti2qian2 kai1shi3
therefore in_advance start
'Since all the guests had come, the banquet started earlier.'

Other connective adjunct pairs marking a cause–effect relation are more formal in style and include 所以 . . . 在于 . . . *suo2yi3* . . . *zai4yu2* . . . '(the reason) why . . . is caused by . . . ' and 之所以 . . . 是因为 . . . *zhi1suo2yi3* . . . *shi4yin1wei4* . . . '(the reason) why . . . it is because . . . .' Other adjunct pairs marking an induction include 既然 . . . 于是/则/那就/那么 . . . *ji4ran2* . . . *yu2shi4/ze2/na4jiu4/na4me* . . . 'given that . . . therefore . . . .' The cause–effect relation can be marked by half of a pair, either the half indicating the cause, such as 由于 *you2yu2* 'because' and 因为 *yin1wei4* 'because,' or the half marking the effect, such as 于是 *yu2shi4* 'therefore,' 那么 *na4me0* 'therefore,' and 那就 *na4jiu4* 'therefore.' There are also some cause–effect adjuncts that work alone, such as 以致 *yi3zhi4* 'leading to,' 以至于 *yi3zhi4yu2* 'leading to,' and 从而 *cong2er2* 'therefore.'

Connective adjuncts are also used in conditional sentences. Such a sentence could include a simple consequence marked by 如果 . . . 就 . . . *ru2guo3* . . . *jiu4* . . . 'if . . . then . . . ' or 要不是 . . . 那么 . . . *yao4bu4shi4* . . . *na4me0* . . . 'if not . . . then (not) . . . ,' such as those in [27a] and [27b]. The condition set out in the subordinate clause can also be a counterfactual one, as in the case of [27c].

[27] a. 如果发生八级以上地震, 大桥就会倒塌。
ru2guo3 fa1sheng1 ba1ji2 yi3shang4
if happen eight_magnitude above

di4zhen4 da4qiao1 jiu4hui4 dao3ta1
earthquake big_bridge then collapse
'If there is an earthquake measured at magnitude 8 or above, the bridge will collapse.'

b. 要不是急着用钱, 那么谁也不会现在卖。
yao4bu2shi4 ji2 zhe0 yong4qian2 na4me0 shui2
if_not urgent ZHE use_money then who

ye3bu4hui4 xian4zai4 mai4
would_not now sell
'If not in urgent need of cash, no one would sell (it) now.'

c. 要是我早点知道真相, 就不会把钱投进去了。
yao4shi4 wo3 zao3dian3 zhi1dao4 zhen1xiang4 jiu4
if I earlier know truth then

bu4 hui4 ba3 qian2 tou2 jin4qu4 le0
NEG will BA money invest into LE
'Had I learned the truth earlier, I would not have invested any money in it.'

Other connective adjunct pairs indicating a condition and its consequence include 假如/假使/假若/如/若/倘若...那么/就/则/便...*jia3ru2/jia2shi3/jia3ruo4/ru2/ruo4/tang3ruo4*...*na4me/jiu4/ze2/bian4*...'if...then...,' 要不是...就/也...*yao4bu2shi4*...*jiu4/ye3*...'if not...then...,' and 若不是...那...*ruo4bu4shi4*...*na4*...'if not...then....' Some adjuncts indicating the condition can work alone, such as 假如/假使/假若/如/若/倘若 *jia3ru2/jia2shi3/jia3ruo4/ru2/ruo4/tang3ruo4* 'if' and 要不是/若不是 *yao4bu2shi4/ruo4bu2shi4* 'if not.'

If a conditional sentence is marked with the pair 只有...才...*zhi2you3*...*cai2*...'only if...then...,' as in [28a], the subordinate clause will represent a necessary condition. If the conditional sentence is marked with the pair 只要...就...*zhi3yao4*...*jiu4*...'as long as...then...,' as in [28b], the subordinate clause will represent a sufficient condition. The subordinate clause can also represent an exhaustive condition, and a common marker pair for such a sentence is 无论...都...*wu2lun4*...*dou1*...'no matter...all...,' as in [28c].

[28] a. 只有赶快跳车, 她才不会被烧死。
zhi3you3 gan3kuai4 tiao4che1 ta1 cai2
only_if hastily jump_from_the_car she then
bu4 hui4 bei4 shao1si3
NEG will BEI burn_to_death
'She will not be burned to death unless she jumps out of the car.'

b. 只要有姓名地址, 我们就可以找到她的电话号码。
zhi3yao4 you3 xing4ming2 di4zhi3 wo3men0 jiu4
as_long_as have name address we then
ke3yi3 zhao3dao4 ta3 de0 dian4hua4 hao4ma3
can find she DE telephone number
'As long as her address and name are known, we will be able to find her phone number.'

c. 无论是否下雨, 你们都必须准时到达集合地点。
wu2lun4 shi4fou3 xia4yu3 ni3men0 dou1
no_matter whether_or_not rain you all
bi4xu1 zhun3shi2 dao4da2 ji2he2 dian4dian3
must on_time arrive gather place
'You should arrive at the gathering point on time no matter whether it rains or not.'

Other connective adjunct pairs that mark a necessary conditional sentence include 惟有 . . . 才 . . . *wei2you3* . . . *cai2* . . . 'if only . . . then . . . ,' 除非 . . . 才 . . . *chu2fei1* . . . *cai2* . . . 'unless . . . then . . . ,' and 除非 . . . 不然/否则 . . . *chu2fei1* . . . *bu4ran2/fou3ze2* . . . 'unless . . . would not . . . .' Those marking a sufficient conditional sentence include 只要 . . . 都/便 . . . *zhi3yao4* . . . *dou1/bian4* . . . 'as long as . . . all/then . . . ,' 一旦/一经 . . . 就/便 . . . *yi1dan4/yi1jing1* . . . *jiu4/bian4* . . . 'once . . . then . . . ,' and 一 . . . 就/便 . . . *yi1* . . . *jiu4/bian4* . . . 'once . . . then . . . .' Other marker pairs marking exhaustive conditional sentences include 不论/不管 . . . 都/总/也 . . . *bu2lun4/bu4guan3* . . . *dou1/zong3/ye3* . . . 'no matter . . . all . . . .' The adjunct marker of the main clause can work alone in a necessary conditional sentence or a sufficient conditional sentence.

The subordinate clause can state the purpose of the action stated in the main clause. Such a complex sentence is commonly marked with 为了 . . . 就 . . . *wei4le0* . . . *jiu4* . . . 'in order to . . . then . . . ,' as in [29a]. In some cases, only the purpose clause is marked with an adjunct, such as 为的是 . . . *wei4de0shi4* . . . 'it is for the purpose of . . . ,' 以求 . . . *yi3qiu2* . . . 'in order to . . . ,' 藉以 . . . *jie4yi3* . . . 'so as to . . . ,' and the 以便 *yi3bian4* 'in order to' in [29b]. The marker 以免 *yi3mian3* 'in order to avoid' in [29c] marks a purpose clause describing a situation to be avoided. Markers with a similar function include 免得 *mian3de0* 'in order to avoid' and 省得 *sheng3de0* 'saving the trouble of.'

[29] a. 为了省钱, 小两口就决定自己安装厨柜。

wei4le0	sheng3qian2	xiao3liang3kou3	jiu4	
in_order_to	save_money	the_young_couple	then	
jue2ding4	zi4ji3	an1zhuang1	chu2gui4	
decide	self	install	cupboard	

'The young couple decided to install the cupboard themselves in order in order to save money.'

b. 她在墙上贴满了英语单词, 以便随时复习。

ta1	zai4	qiang2shang4	tie1man3	le0	ying1yu3
she	at	wall_on	cover_with	LE	English
dan1ci2	yibian4	sui2shi2	fu4xi2		
word	in_order_to	at_any_time	review		

'She covered the wall with English words, in order to review them at any time.'

c. 请勿大声交谈, 以免妨碍其他读者。

qing3	wu4	da4sheng1	jiao1tan2	yi3mian3	fang2ai4
please	do_not	loud	speak	lest	disturb