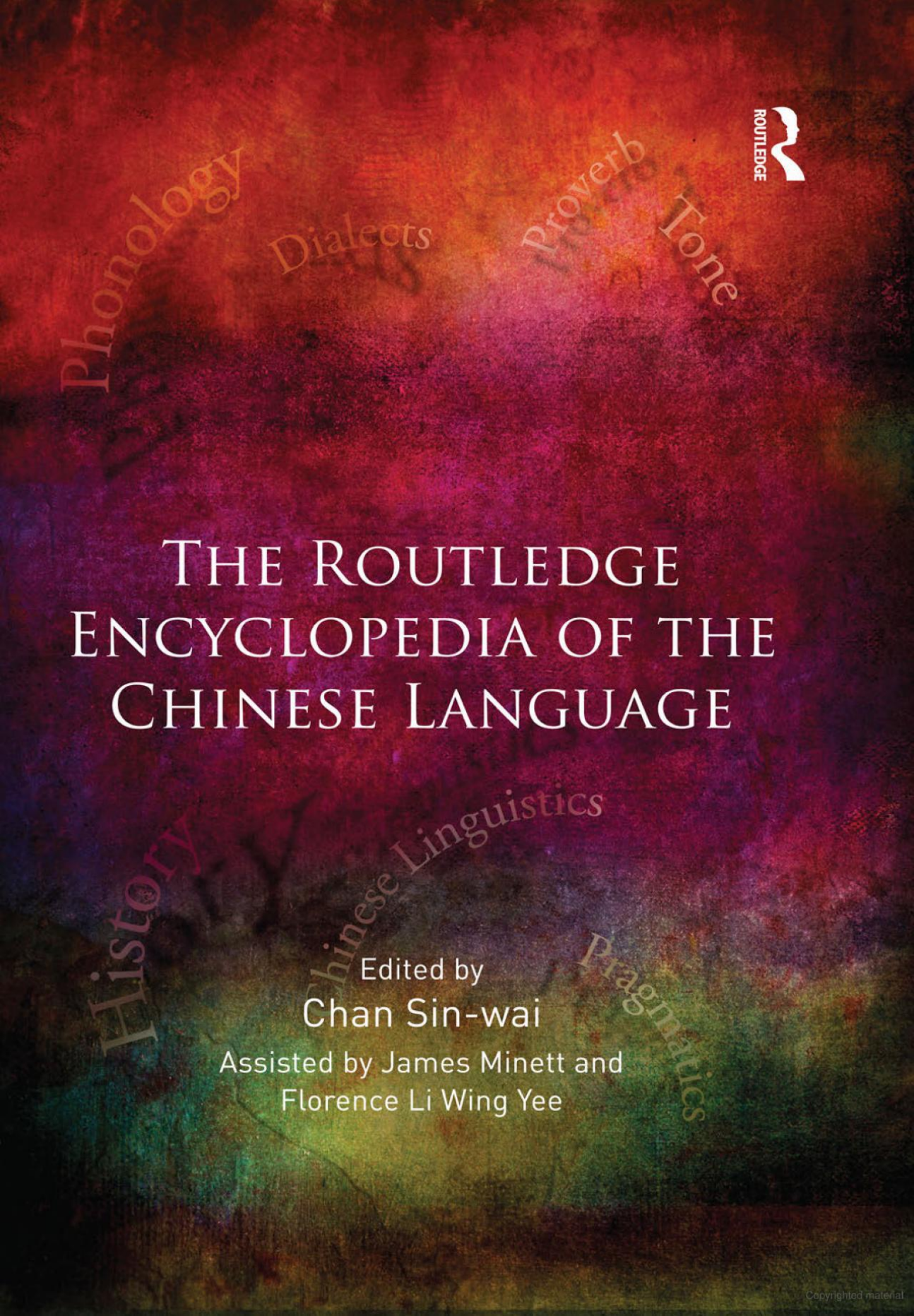

The Routledge Encyclopedia of the Chinese Language

Edited by
Chan Sin-wai
Assisted by James Minett and
Florence Li Wing Yee

THE ROUTLEDGE ENCYCLOPEDIA OF THE CHINESE LANGUAGE

Edited by Chan Sin-wai

ASSISTED BY JAMES MINETT AND FLORENCE LI WING YEE

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published 2016
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

and by Routledge
711 Third Avenue, New York, NY 10017

Routledge is an imprint of the Taylor & Francis Group, an informa business

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data
The Routledge encyclopedia of the Chinese language / edited by Chan Sin-wai; assisted by Florence Li Wing Yee and James Minett.
pages cm
Includes bibliographical references and index.
1. Chinese language–Encyclopedias. 2. Chinese philology–Encyclopedias. 3. Chinese characters–Encyclopedias. 4. Chinese language–Acquisition. I. Chan, Sin-wai, editor. II. Li Wing Yee, Florence, editor. III. Minett, James W., editor.
PL1031.R68 2015
495.1′3–dc23

2015017474

ISBN: 978-0-415-53970-8 (hbk)
ISBN: 978-1-315-67554-1 (ebk)

Typeset in Times New Roman
by Graphicraft Limited, Hong Kong

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF CONSULTANT EDITORS

William S.-Y. Wang
Department of Chinese and Bilingual Studies
Hong Kong Polytechnic University, Hong Kong, China

Huang Chu-Ren
Department of Chinese and Bilingual Studies
Hong Kong Polytechnic University, Hong Kong, China

Feng Shengli
Department of Chinese
The Chinese University of Hong Kong, Hong Kong, China

Cao Guangshun
Institute of Linguistics
Chinese Academy of Social Sciences, China

LIST OF CONTRIBUTORS

Bisang, Walter
Walter Bisang Dr. phil. I (Zürich) has been Professor of General and Comparative Linguistics at the University of Mainz (Germany) since 1992. He was the Director of a Collaborative Research Center on 'Cultural and Linguistic Contact' from 1999 to 2008 in Mainz. His research interests focus on linguistic typology, grammaticalization, language contact/areal typology, and the comparison of different theoretical approaches to language. His languages of interest are East and mainland Southeast Asian languages, Caucasian languages (Georgian and others), Austronesian languages (Bahasa Indonesia, Tagalog, Yabêm, Paiwan), and Yoruba (together with Remi Sonaiya).

Ceccagno, Antonella
Antonella Ceccagno is associate professor of Chinese Linguistics at the University of Bologna, Italy, where she also teaches Sociology of East Asia. In the field of Chinese Linguistics her interests revolve around Chinese morphology. More recently, she is also focusing analytical attention on the interactions between language and society. Her publications on Chinese linguistics and sociology of language include: 'Farewell to *tuhao*, welcome to *tuhao*: Language and Society in China as they Emerge from the Buzzwords of the Last Decade' 2014, in Magda Abbiati and Federico Greselin (eds) *Il Liuto e i libri*, Venezia, Cafoscari, 193–204; 'Sino-Tibetan: Mandarin Chinese' in Rochelle Lieber and Pal Stekauer (eds) *The Oxford Book of Compounding*, Oxford: Oxford University Press, 2009, 478–90, (with Bianca Basciano); 'Metacompounds in Chinese', *Lingue e Linguaggio* 8(2), 2009: 195–212; 'The Chinese Language and Some Notions from Western Linguistics', *Lingue e Linguaggio*, 8(1), 2009: 105–35 (with Bianca Basciano); *Shuobuchulai, la formazione delle parole in cinese* ('*Shuobuchulai* (*Word Formation in Chinese*)'), Bologna: Serendipità, 2009 (with Bianca Basciano); 'Compound Headedness in Chinese: An Analysis of Neologisms', *Morphology*, n 17, 2007: 207–31 (with Bianca Basciano); 'Classification, Structure and Headedness of Chinese Compounds' in *Lingue e linguaggio* 2, 2006: 233–60 (with Sergio Scalise).

Chan, Sin-wai
Chan Sin-wai is Professor at the School of Humanities and Social Science, The Chinese University of Hong Kong, Shenzhen. His teaching and research interests lie mainly in the areas of translation studies, translation technology, and bilingual lexicography. He is the

Chief Editor of *Journal of Translation Technology*, published by The Chinese University of Hong Kong.

He has published more than 44 books in 54 volumes, mainly dictionaries and scholarly monographs, and translated works in different fields. He edited *An Encyclopaedia of Translation* and *The Routledge Encyclopedia of Translation Technology* (2015), revised *Longman Dictionary of English Language and Culture* (bilingual edition), and authored *A Dictionary of Translation Technology* and *A Chinese–English Dictionary of the Human Body*. His book translations from Chinese into English include *An Exposition of Benevolence*, *Palaces of the Forbidden City*, *Letters of Prominent Figures in Modern China*, *Paintings and Calligraphy of Jao Tsung-I*, *Stories by Gao Yang*, *An Illustrated History of Printing in Ancient China*, *Famous Chinese Sayings Quoted by Wen Jiabao*, and *Selected Works of Cheng Siwei: Economic Reforms and Development in China*, Vol. 2. He also translated *My Son Yo Yo* from English into Chinese. His most recent co-edited books include *Style, Wit and Word-Play* (2012) and *The Dancer and the Dance* (2013), both published by Cambridge Scholars Publishing.

Chappell, Hilary

Hilary Chappell is currently Chair Professor in Linguistic Typology of East Asian Languages at the *Ecole des Hautes Etudes en Sciences Sociales* in Paris, an appointment she took up in 2005 after teaching in the Linguistics Department at La Trobe University in Melbourne, Australia, for 18 years. She was originally awarded her doctoral degree in 1984 by the Australian National University in Canberra for her thesis entitled 'A Semantic Analysis of Passive, Causative and Dative Constructions in Standard Chinese' and has over 60 publications on Chinese linguistics and typology, including four edited books and a volume on Hakka, co-authored with Christine Lamarre.

Her main research is on the typology of Sinitic languages (Chinese dialects) and the extent of their linguistic diversity, for which she directed an ERC Advanced Grant project from 2009 to 2013. She has also held two projects with Professor Alain Peyraube (CNRS-EHESS) to investigate the diachronic syntax of Southern Min (Hokkien), using a corpus of late sixteenth and early seventeenth century materials. In 2010, she was elected member of the Academia Europaea.

Chen, Lilly Lee

Lilly Lee Chen received her doctoral and master's degrees in linguistics from the University of Illinois, Urbana-Champaign, and her bachelor's degree from the Department of Foreign Languages and Literatures at National Taiwan University. She taught for 30 years in the Department of Linguistics and the Center for the Study of Languages at Rice University. In the course of her teaching career at Rice she established a variety of language oriented new programs and taught linguistics and Chinese language and culture, including synchronic and diachronic syntax, structure and symbolism of the classical novel, Chinese cinema, and Taiwanese language, culture, and linguistics from a cognitive perspective. She has served as President of the Linguistic Association of Canada and the United States (2013), President of the Chinese American Educational Research and Development Association (2000), and Founding President of the Chinese Language Teacher Association of the South (2006). She was also the Founding Director of the Institute for Chinese Language Teacher Training (2007–2010).

Cheng, Siu-Pong

Cheng Siu-Pong is Research Assistant at T. T. Ng Chinese Language Research Centre of the Institute of Chinese Studies, The Chinese University of Hong Kong. His research interests focus on Cantonese discontinuous constructions, the cartographic approach, and empty categories.

Cheung, Chi-Hang Candice
Candice Chi-Hang Cheung is Research Assistant Professor in the Department of Chinese and Bilingual Studies at the Hong Kong Polytechnic University. Her research interests lie in formal syntax, the syntax-semantics interface, the syntax-information structure interface, and parametric syntax with special focus on Cantonese, Mandarin and other varieties of Chinese. She has published articles in a number of linguistic journals, such as the *Journal of East Asian Linguistics* and *Natural Language and Linguistic Theory*.

Cheung, Kam-Siu
Cheung Kam-Siu received his PhD in Chinese Language and Literature from The Chinese University of Hong Kong in 2007. He is Associate Professor in the Department of Chinese Language and Literature and Honorary Research Associate of D. C. Lau Research Centre for Chinese Ancient Texts at The Chinese University of Hong Kong. His research interests focus on Chinese philology, ancient Chinese texts studies, and lexicography. He has recently published two books on the research of the prominent Qing semanticist and textual critic Wang Niansun (1744–1832), including *On Wang Niansun's Collated Edition of Ancient Texts* (Shanghai Ancient Books Publishing House, 2014), and *Collected Annotations on Wang Niansun's Collated Edition of the Hanfeizi* (The Chinese University Press, 2014). He is also an editorial board member of *Newsletter of Chinese Language*.

Chung, Karen Steffen
Karen Steffen Chung (史嘉琳 Shǐ Jiālín), originally from St. Paul, Minnesota, USA, has taught English, linguistics, and phonetics in the Department of Foreign Languages and Literatures of National Taiwan University since 1990. She is currently an Associate Professor at National Taiwan University. She received her BA in East Asian Languages from the University of Minnesota in 1976, her MA in East Asian Studies from Princeton University in 1981, and obtained her PhD in Linguistics from Leiden University in 2004. The title of her doctoral dissertation is: 'Mandarin Compound Verbs'.

Dong, Yan
Yan Dong received her MA in Linguistics and Applied Linguistics from Renmin University of China in 2008 and joined the PhD program in Linguistics at the University of Michigan in the same year. In July 2015 she recently defended her dissertation, which is entitled 'The prosody and morphology of elastic words in Chinese: annotations and analyses'.

Duanmu, San
San Duanmu is Professor of Linguistics, University of Michigan. He received his PhD in Linguistics from MIT in 1990 and has held teaching posts at Fudan University, Shanghai (1981–1986) and the University of Michigan, Ann Arbor (1991–present). He is the author of *The Phonology of Standard Chinese* (2nd ed., Oxford, 2007) and *Syllable Structure: The Limits of Variation* (Oxford, 2008).

Feng, Shengli
Feng Shengli began his academic career as an undergraduate student in the History Department of Beijing Normal University in the class of 1977. He was accepted by the Graduate Program in the Department of Language and Literature at the Beijing Normal University in 1979, majoring in Ancient Chinese Exegesis under the supervision of Professor Lu Zongda. After completing his MA degree, he taught Classical Chinese and Language and Literature in the

Department of Language and Literature at Beijing Normal University. He received his PhD in Linguistics from the University of Pennsylvania in 1995. He taught Chinese language, Chinese linguistics, and Chinese culture in the Department of East Asian Languages and Cultures at the University of Kansas from 1994, and in 1999 he was promoted to Associate Professor. He was appointed in 2003 as Professor of the Practice and Director of the Chinese Language Program at Harvard University and Director of Harvard-Beijing Academy. In 2005 he was appointed as Yangtze Scholar Adjunct Professor (长江学者讲座教授) at Beijing Language and Culture University. He is the author of *The Prosody, Morphology, and Syntax of Chinese* 漢語的韻律、詞法與句法 (1997), *Prosodically Constrained Syntactic Changes in Early Archaic Chinese* (1996), *The Prosodic Syntax of Chinese* (2002), and *A Preliminary Theory of Chinese Poetic Prosody* 漢語的韻律詩體學論稿 (2015). He has edited *Written Chinese: the Present and the Past* (2013), *New Exploration on Chinese Prosodic Grammar* (2015). He currently serves as associate editor for *Language and Linguistics*.

He, Yun Amy

Amy He Yun is a visiting research fellow in the Centre for Intercultural Politeness Research at the University of Huddersfield, the United Kingdom, where she has previously worked as a postdoctoral research associate. She obtained her PhD degree from Loughborough University for studying politeness in Mandarin Chinese, co-funded by an Overseas Research Studentship Award (Higher Education Funding Council for England), University Research Studentship, and Great Britain–China Educational Trust: Chinese Student Award. Her recent interests are in (im)politeness and face, but her research cuts across pragmatics, discourse analysis, sociolinguistics, intercultural communication, and translation studies. She has completed five research projects and has published widely in the aforementioned areas of research, including 'Different generations, different face?' (*Journal of Politeness Research* 8(1), 2012) and a chapter on identity in *Pragmatic explorations in identity and communication* (ed. Chen, X. Beijing: Higher Education Press, 2013). She is currently working on a book with the provisional title *Generational Variation of Politeness in Contemporary Chinese*.

Healy, Alice F.

Alice F. Healy is College Professor of Distinction and Director of the Center for Research on Training at the University of Colorado Boulder. She received her doctorate from the Rockefeller University in 1973, and she was on the faculty of Yale University from 1973 to 1981, when she joined the faculty of the University of Colorado. Dr. Healy has served as Editor of the journal *Memory and Cognition*, Chair of the Psychology Section of the American Association for the Advancement of Science, President of the Rocky Mountain Psychological Association, President of the Division of Experimental Psychology (Division 3) of the American Psychological Association (APA), and Chair of the Society of Experimental Psychologists. She has published over 250 journal articles and book chapters, is co-author of *Cognitive Processes* (Prentice Hall, 1986) and *Train Your Mind for Peak Performance* (APA, 2014), is editor of the volume *Experimental Cognitive Psychology and its Applications* (APA, 2005), and is senior editor of the two-volume *Essays in Honor of William K. Estes* (Erlbaum, 1992), *Learning and Memory of Knowledge and Skills* (Sage, 1995), *Foreign Language Learning* (Erlbaum, 1998), the *Experimental Psychology* volume of the *Handbook of Psychology* (Wiley, 2003 and 2013), and *Training Cognition* (Psychology Press, 2012).

Herzberg, Larry

Larry Herzberg has been a professor of Chinese at Calvin College in Grand Rapids, Michigan for the past 30 years, where he is also Director of Asian Studies. After studying Chinese for

five years at Vanderbilt University, he did his PhD work in Chinese language and literature at Indiana University. Together with his Chinese wife, Xue Qin, he has published three books, including *The Chinese Survival Guide*, *Basic Patterns of Chinese Grammar*, and *Chinese Proverbs and Popular Sayings*. He and his wife have also produced three films on China, including 'China Today: Issues that Trouble Americans at the Start of the 21st Century', 'The China Threat: Perception versus Reality', and 'Chinese Dialects and Their Influence on the Pronunciation of Mandarin'. In 2011 he received the Presidential Award for Exemplary Teaching, the highest honor bestowed on faculty members by Calvin College.

Hsieh, Shu-Kai

Hsieh Shu-Kai is an associate professor in the Graduate Institute of Linguistics at National Taiwan University, Taiwan. He received his PhD in Computational Linguistics from the University of Tübingen in Germany. He has chaired and served on several program committees for linguistics at international conferences, and has given a number of invited talks on computational lexical semantics, corpus, and ontology at different universities. His primary research interests include computational lexical semantics and ontology. His publications include two books on computational treatment of Chinese characters and translation of the Generative Lexicon, and more than 30 refereed research papers published in journals, books, conferences, and workshops. He is a member of the editorial board and the linguistic section chair of IJCLCLP, an international journal of computational linguistics and Chinese language processing. He is a member of the council of the Linguistics Society of Taiwan and Global WordNet Association.

Huang, Chu-Ren

Huang Chu-Ren is a Chair Professor of Applied Chinese Language Studies at the Hong Kong Polytechnic University. He is Fellow and President of the Hong Kong Academy of the Humanities, a permanent member of the International Committee on Computational Linguistics, and a former President of the Asian Association of Lexicography. His main research areas are ontology, corpus and computational linguistics, lexical semantics, Chinese grammar, language resources, and digital humanities. The conference series in which he plays advisory and/or organizing roles include ALR, ASIALEX, CLSW, CogALex, COLING, IsCLL, LAW, OntoLex, PACLIC, ROCLING, and SIGHAN. His research output includes 12 licensable language resources, ten searchable online language databases, 20 books or edited volumes, over 130 journal articles or book chapters, and over 380 refereed conference papers. He is currently Chief Editor of the Journal *Lingua Sinica*, and the book series: *Studies in Natural Language Processing* and *The Humanities in Asia*. He received a Doctor *honoris causa* from Aix-Marseille University in 2013.

Jiang, Yan

Yan Jiang received his BA in English Language and Literature and his MA in modern English from Fudan University, Shanghai. He obtained his PhD in linguistics from London University. He is now a lecturer in linguistics and the languages of China at the Department of Linguistics, School of Oriental and African Studies, University of London. He has published studies on the Chinese adverb of scope 'dou', abduction and pragmatic inference, dexis and anaphora, Chinese counterfactuals, and quantification in Chinese. His book-length publications are listed below: Jiang Yan and Pan Haihua (1998, 2005) *Introduction to Formal Semantics*, China Social Science Press; Jiang Yan (2008) *Translation of Dan Sperber and Deirdre Wilson: Relevance: Communication and Cognition*, China Social Science

Press; and Jiang Yan (2011) (ed.) *Approaching Formal Pragmatics*, Shanghai Educational Publishing.

Jiao, Liwei
Jiao Liwei has a PhD in Chinese Linguistics from Nankai University, China. Since 1998 he has taught at Renmin University of China, the University of Durham, Princeton in Beijing, and the Duke Study in China Program. He is currently a Lecturer in Foreign Languages at the University of Pennsylvania. Among his many publications are *500 Common Chinese Idioms* (co-authored, 2010), *500 Common Chinese Proverbs and Colloquial Expressions* (co-authored, 2014), and *A Cultural Dictionary of the Chinese Language: 500 Proverbs, Idioms and Maxims* (forthcoming, 2017), all with Routledge, as well as *Chinese Phonetics in the 20th Century* (co-authored, 2004) with Shuhai Press, China. He is the Chinese translator of Professor Yuen Ren Chao's oral history: *Chinese Linguist, Phonologist, Composer and Author*. His research fields include lexicology, relationships between language and culture, and Chinese phonetics.

Kirkpatrick, Andy
Andy Kirkpatrick is Professor in the Department of Languages and Linguistics, at Griffith University in Brisbane. He has worked in tertiary institutions across Asia including six years at the Hong Kong Institute of Education as Chair Professor. He has been a Visiting Scholar at Peking University and Bei Hang University in Beijing. His translation and commentary of Chen Kui's *Wenze* won the International Society for the History of Rhetoric's inaugural prize for best article. He is the editor-in-chief of the *Journal of Multilingual Education* and of a book series (with Springer) of the same name. His most recent book is *Chinese Rhetoric and Writing: An Introduction for Teachers of Writing*, co-authored with Xu Zhichang.

Lau, Chun Fat
Lau Chun Fat is Professor at the Department of Chinese Language and Literature at Xiamen University, China. He was born in the city but was brought up in a small Hakka village in the middle of the New Territories of Hong Kong. He studied Biology at The Chinese University of Hong Kong and received his Dr. rer. nat. degree at the Free University of Berlin. After spending more than a decade in the laboratory, he changed his career in the mid-1990s, when he found that practically no children were speaking Hakka around him. He received his master's and PhD in Linguistics from the Polytechnic University of Hong Kong, and went to teach at Xiamen University in 2002. He has published more than 80 papers and authored eight books. His main interests are sociolinguistics, historical linguistics, Chinese dialectology, and documentation of endangered Chinese dialects around Hong Kong, as the whole area has shifted to Cantonese in the past three decades. He is also chairman of the Association of Conservation of Hong Kong Indigenous Languages.

Li, Lan (Chapter 31)
Li Lan is a Fellow of the Chartered Institute of Linguists, UK, with MPhil and PhD degrees in Applied Linguistics from the University of Exeter. She works as an Associate Professor at the Department of English, Hong Kong Polytechnic University. She is also a committee member of ASIALEX. She has been teaching and conducting research in semantics, lexicology, and bilingual lexicography. Her research interests and publications cover lexicology, lexicography, metaphor studies, computer-mediated professional communication, sociolinguistics, and corpus linguistics.

Li, Lan (Chapter 34)
After gaining his PhD in 1995 from the Chinese Academy of Social Sciences (CASS) in Beijing on the subject of 'Variation in the Tones and Initials of Southwestern Mandarin Dialects', Li Lan was offered a position at the CASS, where he continues to work as a Senior Research Fellow in the Institute of Linguistics and head of the Dialect Research Section.

Apart from many articles on dialectology and typology of Sinitic languages, which have been published in the top-ranking linguistics journals in China such as *Zhongguo Yuwen* 《中国语文》(*Studies of the Chinese Language*) and *Fangyan* 《方言》(*Dialects*), Li Lan has also published two books on *The Guiyang Dialect* (1997) and on *The Miao (Hmong) Language of Chengbu, Hunan* (2004). He won the Wang Li prize at Peking University in 2005 for this latter monograph and has been the recipient of a large number of projects on linguistics funded by the Social Sciences Foundation of China.

Li Lan has also undertaken research visits to the Hong Kong University of Science and Technology and to the Centre de Recherches Linguistiques sur l'Asie Orientale in Paris.

Li, Chris Wen-chao
Chris Wen-Chao Li is Professor of Linguistics at San Francisco State University, where he teaches linguistics, news writing, and translation-interpretation. He received his doctoral degree in General Linguistics and Comparative Philology from Oxford University, where his research focused on Mandarin sound change and Chinese phonology. He is the author of *A Diachronically-Motivated Segmental Phonology of Mandarin Chinese* (New York: Peter Lang, 1999) and *Media Chinese* (Taipei: Lucky Books, 2005). His translations and scholarship have appeared in *Renditions*, *The Chinese Pen*, *Target*, *Language and Communication*, and the *Journal of the American Oriental Society*. His current research interests include sound change, language contact, diglossia, standardization, phonological translation, Chinese Romanization systems, and the phonological description of Mandarin Chinese varieties.

Li, Yafei
Li Yafei is Professor of Linguistics at the University of Wisconsin-Madison, USA, and Siyuan Chair Professor (part-time) at Nanjing University, China. Born in Baoding, China, he received a BA in 1982 and an MA in 1985 from Shandong University and a PhD in Linguistics from MIT in 1990. Specializing in syntax and morpho-syntax, he has taught at Brandeis University, Cornell University, University of Wisconsin-Madison, and Nanjing University. His publications include articles (in *Language*, *Linguistic Inquiry*, *Natural Language and Linguistic Theory*, *The Linguistic Review*, *Journal of East-Asian Linguistics*, 《中国语文》(*Studies of the Chinese Language*) and 《当代语言学》(*Contemporary Linguistics*)) and two books, *X^0: A Theory of the Morphology-Syntax Interface* from MIT Press and the co-authored *The Syntax of Chinese* from Cambridge University Press. Currently, he is working on his third book, *The Fact-first Principle*, which evaluates the empirical foundation of the Chomskyan model of syntax and articulates novel analyses of facts from multiple languages that have not been accurately or adequately dealt with in the mainstream literature.

Liu, Jin
Liu Jin is Associate Professor of Chinese at Georgia Institute of Technology. She received her PhD in East Asian Literature and Culture from Cornell University and her MA and BA in Chinese Linguistics from Peking University. Her interdisciplinary research studies contemporary Chinese popular culture and media culture from the perspective of language, sound, voice, and music. She is the author of the book *Signifying the Local: Media Productions*

Rendered in Local Languages in Mainland China in the New Millennium (Brill, 2013). She co-edited and contributed to the volume *Chinese Under Globalization: Emerging Trends in Language Use in China* (World Scientific, 2012). She has published articles in journals including *Positions: Asia Critique*, *Modern Chinese Literature and Culture*, *Journal of Modern Literature in Chinese*, *Chinese Language and Discourse*, and *Harvard Asia Pacific Review*.

McEnery, Tony

Tony McEnery is Distinguished Professor of English Language and Linguistics at Lancaster University. He has published a number of books on corpus linguistics including *Corpus Linguistics* (with Andrew Wilson, 1996), *Corpus-Based Language Studies* (with Richard Xiao and Yukio Tono, 2006), and *Corpus Linguistics, Method Theory and Practice* (with Andrew Hardie, 2011). A key feature of his research is that he has used the corpus approach to investigate a wide range of languages including, in collaboration with Richard Xiao, Mandarin Chinese. Tony is a Fellow of the Academy of Social Sciences and is currently Director of the ESRC Centre for Corpus Approaches to Social Science at Lancaster University. Previously he was Director of Research at the UK's Arts and Humanities Research Council.

Miao, Ruiqin

Miao Ruiqin is currently Associate Professor of Linguistics in the School of Foreign Languages at Shanghai Jiao Tong University, China. She received her BA in English Language and Literature from Hunan University (1993), MA in English Linguistics and Applied Linguistics from the same university (1996), and PhD in Linguistics from the State University of New York at Stony Brook (2005). Her research focuses on second language acquisition, phonetics/phonology, language contact, sociolinguistics, and language pedagogy. She has published in refereed journals and edited volumes including *Journal of Asian Pacific Communication*, *ITL-International Journal of Applied Linguistics*, and *Developments in Applied Linguistics/Series on Western Research in the Humanities and Social Sciences* (Jianguo Ji and Nan Jiang, eds. Beijing: China Renmin University Press, 2007). She has written *Foreign Language Listening Comprehension: Theory and Practice* (co-authored with Liu Longgen, Beijing: Foreign Language Teaching and Research Press, 2011). She has also edited several English textbooks used in China and presented papers at various international conferences.

Moore, Robert L.

Robert L. Moore is a professor of anthropology at Rollins College and Director of International Affairs for the college's Hamilton Holt School. He has a BA in anthropology from Tulane University and a PhD from the University of California at Riverside. In 1993–1994 he was a visiting professor on the faculty of Qingdao University in China. He has undertaken research in Hong Kong, Qingdao, and Beijing and published a number of academic articles on linguistics, Chinese culture, and romantic love in non-Western societies. His current research focuses on contemporary youth cultures and family relationships in China.

Orton, Jane

Jane Orton, PhD, is Director of the Chinese Teacher Training Centre (CTTC) at the University of Melbourne, Australia, where she coordinated modern languages education for 20 years. Jane has an honours degree in Chinese language and literature from the University of Melbourne. She taught at Capital Normal University in Beijing from 1981 to 1983 and has been a frequent guest lecturer, conference speaker, and researcher in China over the ensuing 30 years. The CTTC undertakes research and professional development into Chinese language programs, teacher education, and innovative, research-grounded resource development. Jane's

particular research interests are the learning demands of Chinese as a Second Language, especially in oral skill development. Jane is a board member of the association Chinese as a Second Language Research (CASLAR) and a member of the editorial board of the CASLAR Journal. Her recent publications include 'Comparing teachers' judgments of learners' speech in Chinese as a foreign language', *Foreign Language Annals*, 47(3), 2014: 507–26.〈澳大利亚中小学汉语教学的复杂性〉(The Complexities of Teaching Chinese Language in Australian Primary and Secondary Schools)《云南师范大学学报(对外汉语教学与研究版)》(*Journal of Yunnan Normal University* (*Teaching and Research of Chinese as a Foreign Language*)) 4, 2014: 1–5; 'Developing Chinese Oral Skills – A Research Base for Practice', in Istvan Kesckes (ed.) *Research in Chinese as a Second Language*, Berlin: Mouton de Gruyter, 2013: 3–26.

Packard, Jerome L.
Jerome L. Packard is Professor of Chinese, Linguistics and Educational Psychology, and a faculty affiliate at the Beckman Institute for Advanced Science and Technology at the University of Illinois, USA. He received his BA in Chinese Language and Literature from the University of Massachusetts in 1976 and his PhD in Linguistics from Cornell University in 1984. He specializes in Chinese word structure, Chinese psycholinguistics, and Chinese language teaching and learning. He has taught Chinese at the University of Massachusetts, Cornell University, the University of Pennsylvania, and the University of Illinois, and was a Fulbright dissertation scholar in China. He has performed research on Chinese aphasia and dyslexia, the first language acquisition of Chinese, and the acquisition of Chinese as a second language. His current research interests include Chinese sentence processing, sentiment analysis of Chinese texts, and how children in China learn to read and write. He has received several grants and is the author of numerous publications relevant to Chinese language, including the books *New Approaches to Chinese Word Formation*, *The Morphology of Chinese*, *Chinese Children's Reading Acquisition* (with W. Li and J. Gaffney), and *Processing and Producing Head-final Structures* (with H. Yamashita and Y. Hirose).

Peyraube, Alain
Alain Peyraube is currently emeritus Director of Research at the Centre National de la Recherche Scientifique (CNRS, Paris, France), Chair Professor of Chinese Linguistics at the Ecole des Hautes Etudes en Sciences Sociales (EHESS, Paris), and Director of the Institute of Advanced Studies 'Collegium de Lyon'. He was Adjunct Professor at the University of Science and Technology of Hong Kong (2005–11) and Founding Member of the Scientific Council of the European Research Council (ERC) (2005–13), and he has been Honorary Professor at the Peking University since 2007, and Distinguished Member of the Chinese Academy of Social Sciences since 2009. He became a member of the European Academy (Academia Europaea) in 2006 and Chair of its Linguistics Section in 2013. As a specialist in Chinese historical syntax and linguistic typology of Sinitic languages, he has authored five books and more than a hundred articles on Chinese studies, mainly with respect to Chinese linguistics. His latest research has been done within a broadly functional and cognitive framework from a cross-linguistic perspective. He was conferred the title of Chevalier de la Légion d'honneur in 2008.

Shi, Dingxu
Shi Dingxu works in the Department of Chinese and Bilingual Studies of the Hong Kong Polytechnic University. He received his MA in applied linguistics from the University of

Pittsburgh, and MA and PhD in linguistics from the University of Southern California. His research interests are on syntax, typology, language change induced by contact, interface between syntax and semantics, and pedagogical grammar of Chinese. His publications cover a wide range of topics on Chinese in these fields.

Simpson, Andrew
Andrew Simpson is Professor of Linguistics and East Asian Languages and Cultures at the University of Southern California. His research focuses on the comparative grammar of East and Southeast Asian languages, and issues of language planning. He is the editor of *Language and National Identity in Asia* and *Language and National Identity in Africa* (Oxford University Press), and is the joint general editor of the *Journal of East Asia Linguistics*.

Su, Xinchun
Su Xinchun is the Head of the College of Humanities and Communications College of Jiageng College of Xiamen University, and also the Head of Language Education and Teaching Material Branch of National Language Resources Monitoring and Research Center (CNLR). His major teaching and research interests cover Chinese lexicology, ancient and modern Chinese vocabulary, cultural lexicology, metrology of lexicology, dictionary and corpus linguistics, terminology, and so on. He has published over 20 books and more than 100 articles on Chinese lexicology studies.

Tang, Gladys
Gladys Tang is Professor of Linguistics at the Department of Linguistics and Modern Languages, The Chinese University of Hong Kong. She is also Director of The Centre for Sign Linguistics and Deaf Studies. She has been researching sign language and how deaf children acquire language, signed and spoken. In recent years, she has also embarked on developing a model of education that benefits both deaf and hearing children. The sign bilingualism and co-enrollment approach promotes the early learning of signed language and spoken language to support children's linguistic, cognitive and socio-emotional development.

Tang, Sze-Wing
Tang Sze-Wing received his BA and MPhil from The Chinese University of Hong Kong and completed his doctoral dissertation at the University of California, Irvine. He is Associate Professor and Vice-Chairman of the Department of Chinese Language and Literature, The Chinese University of Hong Kong, and concurrently serving as Director of T. T. Ng Chinese Language Research Centre of the Institute of Chinese Studies, The Chinese University of Hong Kong, and Chief Editor of *Studies in Chinese Linguistics* and *Newsletter of Chinese Language*. His research interests lie primarily in Chinese syntax, theoretical approaches to the study of Chinese dialects, and comparative grammar.

Tao, Hongyin
Tao Hongyin is Professor of Chinese Linguistics and Applied Linguistics at the University of California, Los Angeles. He has published a dozen books/edited volumes/special journal issues and over 90 articles in journals and edited volumes on Chinese discourse and grammar, applied linguistics, corpus linguistics, sociolinguistics, and English discourse grammar. Some of his books include *Units in Mandarin Conversation: Prosody, Discourse, and Grammar* (1996) and *Current Trends in Sociolinguistics* (1997/2004, co-edited). His most recent books are *Working with Spoken Chinese* (2011) and *Chinese under Globalization: Emerging Trends*

in Language Use in China (2012, co-edited). In addition to being the president of the Chinese Language Teachers Association, USA (2014), he is on over a dozen editorial boards and is the Executive Editor of the journal *Chinese Language and Discourse* and its companion book series, *Studies in Chinese Language and Discourse*, both with John Benjamins.

Tao, Liang

Liang Tao is associate professor of linguistics and coordinator of the Chinese Language Program in the Department of Linguistics at Ohio University. She received her doctorate from the University of Colorado Boulder in 1993, and was a post-doctoral research associate in the Healy-Bourne Research Laboratory from 1993 to 1996, when she joined the faculty of Ohio University. Dr. Tao has broad research interests in the areas of theoretical linguistics, psycholinguistics, and applied linguistics, focusing on discourse-functional approaches in her research. She pioneered the study of repair in Chinese conversations, proposed the hierarchical structure of language processes against a popular iconic proposal of topic continuity, and proposed the interface of phonology and syntax (phono-syntactic conspiracy) in Mandarin grammaticalization. With Dr. Alice Healy she examined basic native and foreign language processes, using findings from experimental studies to propose, for instance, transfer of cognitive strategies in processing native and non-native languages, the developmental progression of the unitization effect, and the varied role of words in reading processes. She also conducted research on language development and cognitive approaches to Chinese pedagogy. Dr. Tao has published extensively in journal articles and book chapters, and has served as a co-editor of the volume *Current Issues in Chinese Linguistics*.

Wang, William S.-Y.

William S.-Y. Wang is Chair Professor at the Hong Kong Polytechnic University, based in its Department of Chinese and Bilingual Studies. He is also Professor Emeritus of the University of California at Berkeley (where he was Professor of Linguistics for 30 years), and Honorary Professor at Peking University and, Beijing Language and Culture University, and the Chinese University of Hong Kong. He is Editor of the *Journal of Chinese Linguistics*, which he founded in 1973. In 1992, he was elected President of the International Association of Chinese Linguistics at its formation. In the same year he was elected to the Academia Sinica in Taiwan.

His central interest is in language within an evolutionary perspective. He has published some 200 papers and ten books in diverse areas of theoretical and applied linguistics. These have appeared in general magazines, such as *American Scientist*, *Nature*, *Proceedings of the National Academy of Sciences (USA)*, *Scientific American*, etc., in specialized journals, such as *Brain and Language*, *Diachronica*, *Language*, *Lingua*, *Language and Cognitive Processes*, *Neuropsychologia*, *Journal of Phonetics*, etc., and in various encyclopedias. His writings have been translated into many languages. He has lectured widely in America, Asia, and Europe.

Wu, Weiping

Wu Weiping is Director of the Yale-China Chinese Language Center at The Chinese University of Hong Kong (CUHK) and a guest professor in four universities in China. He obtained his PhD in Linguistics from Georgetown University, and worked as a Research Associate and Director of the Chinese Language Testing Program at the Center for Applied Linguistics in Washington DC before moving to CUHK. His academic profile includes teaching, research, and publication in linguistics, teaching Chinese as a Second Language (CSL), and assessment. Among his contributions are articles in referred journals such as *Forensic Linguistics*,

Language Teaching and Research, *TCSOL Studies*, book and book chapters in CSL and sociolinguistics, the Computerized Oral Proficiency Assessment (COPA) in Putonghua and Cantonese for learners with background in English, Japanese, and Korean, and volumes of textbooks under the Pragmatic Framework. He is founding President of the CSL Teaching and Research Society (CSLTARS), a member of the editorial committee of several core journals in CSL, and series editor of *Linguistics and Teaching Chinese as a Second Language*.

Xiao, Richard

Richard Xiao is Reader in Corpus Linguistics and Chinese Linguistics in the Department of Linguistics and English Language at Lancaster University. His major research interests cover corpus linguistics, Chinese linguistics, contrastive linguistics, and translation studies. He has published numerous books in these research areas, including *Aspect in Mandarin Chinese: A Corpus-based Study* (2004), *Corpus-based Language Studies: An Advanced Resource Book* (2006), *A Frequency Dictionary of Mandarin Chinese: Core Vocabulary for Learners* (2009), *Corpus-Based Contrastive Studies of English and Chinese* (2010), *Using Corpora in Contrastive and Translation Studies* (2010), *Corpus-Based Studies of Translational Chinese in English-Chinese Translation* (2012), *Translation and Contrastive Linguistic Studies at the Interface of English and Chinese* (2014), and *Corpus Pragmatics in Translation and Contrastive Studies* (2015). He is a member of editorial boards of a range of international journals, such as *Chinese Language and Discourse*, *Corpora*, *Corpus Linguistics*, *Foreign Language Learning: Theory and Practice*, *Glossa*, *Global Chinese*, *International Journal of Corpus Linguistics*, and *Languages in Contrast*.

Xing, Zhiqun Janet

Janet Zhiqun Xing graduated from the University of Michigan in 1993, and is a Professor of Chinese and Linguistics at Western Washington University. Her research interests are in grammaticalization, semantic change, historical linguistics, discourse pragmatics, and Chinese language pedagogy. Her representative publications include two edited volumes – *Newest Trends in the Study of Grammaticalization and Lexicalization* (Mouton de Gruyter, 2012) and *Studies of Chinese Linguistics – Functional Approaches* (Hong Kong University Press, 2008); two monographs – *A Contrastive Model of Teaching Chinese to Native English Speakers* (Peking University Press, 2011) and *Teaching and Learning Chinese as a Foreign Language*: *A Pedagogical Grammar* (Hong Kong University Press, 2006); and several dozens of refereed journal articles related to grammaticalization, semantic change, and Chinese language pedagogy.

Yin, Jing-hua John

John Jing-hua Yin graduated from Beijing Capital Normal University in 1982, and he received his MEd and PhD in Foreign Language Education from the State University of New York at Buffalo in 1984 and 1995 respectively. He then began teaching Chinese at the Chinese Summer School at Middlebury College, the University of Oregon, and the College of William and Mary. Since 1997, he has been teaching Chinese language and literature at the University of Vermont. He is Professor of Chinese and Chair of the Department of Asian Languages and Literatures. His book *Fundamentals of Chinese Characters* was published by Yale University Press in 2006, and his book *Practical Rhythmic Chinese* was published by Foreign Language Teaching and Research Press in 2008. He has also published articles on his studies and the teaching of Chinese characters to college students in the USA.

Zhang, Grace

Grace Zhang is an Associate Professor at Curtin University in Australia, where she teaches and supervises postgraduate students in Chinese linguistics. She has 30 years of tertiary teaching experience in Chinese language at various universities in Australia, China, New Zealand, and the United Kingdom. Her research interests include Chinese linguistics and pragmatics. She was awarded a PhD in linguistics by the University of Edinburgh in 1996, and has published extensively on linguistics. Her recently published books include: *Using Chinese Synonyms* (2010) by Cambridge University Press; *Request Strategies*: *A Comparative Study in Mandarin Chinese and Korean* (2008, co-author) by John Benjamins; *Critical Perspectives on Language Education*: *Australia and Asia Pacific* (2014, co-editor) by Springer; *Elastic Languages: How and Why We Stretch Our Words* (2015) by Cambridge University Press; *Communicating through Vague Language: A Comparative Study of L1 and L2 Speakers* (2015, co-author) by Palgrave Macmillan. Of specific relevance to this article is her co-authored journal article on the translation of Chinese *xiehouyu*: Liu, Chiung-wen and Grace Zhang (2006) 'Translation of Chinese *xiehouyu* (sayings) and Relevance Theory', *Across Languages and Cultures* 7(1): 49–76.

Zhang, Xiaoheng

Dr. Xiaoheng Zhang has a PhD degree in computer science from Oxford Brookes University, UK, an MSc degree in computer science from Hunan University, China, and a BA degree in English, Hunan University. His research interests include Modern Chinese writing systems, computational linguistics, computer-assisted language learning, and computer-assisted translation. He has published three books and over 60 research papers in these areas. Dr. Zhang is an assistant professor in the Department of Chinese and Bilingual Studies, Hong Kong Polytechnic University. The subjects he teaches at the undergraduate and post-graduate levels include Interface between Chinese and Information Technology, Computer-Assisted Chinese Language Teaching, Computer Tools for the Language Professionals, and Modern Chinese Characters and Information Technology. He is a member of the Editorial Committee of the *Journal of Modernization of Chinese Language Education* and a visiting professor of Hunan University.

Zhuo, Jing-Schmidt

Zhuo Jing-Schmidt received her BA and MA in German language and literature from Peking University, MA in Germanic Linguistics from the University of California, Los Angeles, and PhD in General Linguistics from the University of Cologne, Germany. She is associate professor of Chinese Linguistics at the University of Oregon. Her research is concerned with how language structure and its use are shaped by culture, society, and human psychology. The inquiry into the role played by human emotion in linguistic pragmatics constitutes an abiding focus of her work. Her corpus-based research methods have contributed to increased rigor in Chinese pragmatics and historical linguistics research. With scholarly experiences in three continents, she researches and publishes in English, Chinese, and German. She is the author of *Dramatized Discourse: The Mandarin Chinese Ba-Construction* (John Benjamins, 2005), and editor of *Increased Empiricism: Recent Advances in Chinese Linguistics* (John Benjamins, 2013), and has published numerous articles in leading linguistics journals including *Cognitive Linguistics* and *Journal of Pragmatics*.

THE CHINESE LANGUAGE: THE GLOBAL, HISTORICAL, AND LINGUISTIC ASPECTS

Chan Sin-wai

THE CHINESE UNIVERSITY OF HONG KONG, HONG KONG, CHINA

Introduction

This is an encyclopedia that will, we believe, meet the academic, linguistic, and pedagogical needs of millions of people in the world who are interested in the Chinese language in different capacities and for different reasons, such as Sinologists, Chinese linguists, and teachers and learners of Chinese as a second language. In the following, we will explain the contents of this encyclopedia of the Chinese language from a global, historical, and linguistic perspective.

The global aspects of the Chinese language

Chinese, as we all know, is a language spoken by about one fifth of the world's population (over a billion people) as China is a vast country with a population currently approaching 1.4 billion. According to statistics, the number of Chinese language speakers exceeds 1.4 billion, including 50 million people who have Chinese as their second language. As of 2014, there were 480 Confucius Institutes spread across the six continents, the number of Chinese language learners was over 40 million worldwide, and more than 3,000 higher education institutions in 109 countries offered Chinese language programs. All these figures show that the target users of this encyclopedia could be the countless language learners and linguists of Chinese worldwide, amateurs of Chinese literature and culture, and scholars in Chinese studies. It is not surprising then that the twenty-first century has sometimes been hailed as the century of the Chinese language.

This encyclopedia contains a chapter on 'Chinese Language in a Global Context', co-authored by Liu Jin of Georgia Tech and Tao Hongyin of the University of California, Los Angeles. This chapter explores the relationship between globalization and the Chinese language, and discusses some major issues such as the contact of Chinese with other languages in historical times, the ongoing impact of English on Chinese and the perceived crisis involving Chinese, the rapidly ascending status of the Chinese language worldwide, and the impact of nationalization and modernization on regional varieties. Finally, a series of major changes in contemporary Chinese is analyzed, as is the emergence of notable new

linguistic forms and their theoretical implications. Also related is the issue of the 'Chinese Language and National Identity', a chapter by Andrew Simpson of the University of Southern California. This chapter provides an overview of how Chinese has affected the development of national identity, and how the terms 'nation', 'nationalism', and 'national identity' apply in the context of China. It is shown that the relation of 'Chinese' to 'national identity' varies depending on the perspective of the Chinese nation that is adopted (ethnic vs. political), and that different emphases have been placed on the role of Chinese in the development of national identity at different times in modern history.

The globalization of the Chinese language is closely related to the teaching of it as a second language and the training of Chinese teachers. Wu Weiping of The Chinese University of Hong Kong contributed a chapter on 'Chinese Language Pedagogy' in which he discusses the topic both diachronically and synchronically. While the former approach explores issues related to the major components of the language, namely phonology, vocabulary, grammar, and the writing system with characters; the latter focuses on the shift from language structure to language use in teaching activities. A systematic approach to implement the Pragmatic Framework in CSL teaching, covering assessment, curriculum design, teaching materials preparation, and teacher training, is also provided as part of this chapter. Jane Orton of the University of Melbourne focuses on 'Chinese Language Education: Teacher Training'. She claims that certified schoolteachers of Chinese in the West generally graduate from generic courses in education studies and language teaching. They overwhelmingly comprise L1 users who need special training to manage the educational culture and learning styles of Western classrooms. The few dedicated Chinese teacher training programs available concentrate mainly on linguistic features. Analysis of Chinese as an object of learning, and the demands its very particular features make on speakers of European languages, are still to be incorporated in Chinese teacher training.

The historical aspects of the Chinese language

The Chinese language is usually divided into Ancient (or Classical) Chinese and Modern Chinese, which are drastically different in several aspects. This encyclopedia includes a chapter on 'Ancient Chinese' written by Alain Peyraube of the Ecole des Hautes Etudes en Sciences Sociales (EHESS), Paris. Ancient Chinese, according to Peyraube, is simply defined by Wang Li as 'the language of the documents of the past'. It covers a very long period, from the first Chinese inscriptions known to us, dated from the fourteenth century BCE, until the nineteenth century CE. Three basic stages are usually distinguished for Ancient Chinese: (i) the Archaic period; (ii) the Medieval period; and (iii) the Modern period. It was during the Archaic period, between the fifth and second centuries BCE, that what is known today as Classical Chinese *par excellence* was established, and it is this language, also called Late Archaic Chinese, that is discussed in this chapter. Covering the same period, but with a different approach, is the chapter on 'Classical Chinese' by Chris Li Wen-Chao of San Francisco State University. This chapter surveys the development of Classical Chinese from the language of early philosophical discourses to later imitations that have given rise to a written register distinct from the vernacular. Salient linguistic features of the classical language are explained, including the monosyllabicity of lexis, the fluidity of form class, and pronominal case grammar. The author also describes the shift in written language from Classical Chinese to the modern vernacular in the early twentieth century, viewed through the lens of diglossia.

The language that we use today is known as 'Standard Chinese', about which Shi Dingxu of the Hong Kong Polytechnic University has written a chapter. According to Shi, Standard

Spoken Chinese used to mean *Guanhua* 'official language' which was typically the dialect of the capital city. *Guoyu* 'national language' was created in the 1920s to take the role of the Standard Spoken Chinese and was succeeded by *Putonghua* 'common language' in the 1950s. The two versions were accompanied by *Zhuyin Fuhao* 'phonetic symbols' and *Hanyu Pinyin* 'Chinese phonetic alphabets' respectively. Another chapter on modern Chinese is written by Feng Shengli of The Chinese University of Hong Kong. This chapter, entitled 'Modern Chinese: Written Chinese', shows how written Chinese is different from spoken Chinese in modern times, how the formal register system of written Chinese has newly developed after the May Fourth Movement, and finally, what principles the formal register grammar must observe. The answer lies first on the argument that a formal style after the destruction of literary Chinese by the May Fourth Movement is established by creating a sense of expressive distance from everyday speech. That is, when linguistic expressions are used, the more distant the expressions are kept from everyday speech, the more formal sense they can create in face-to-face conversation. The demand for vernacular Chinese to function formally after the May Fourth Movement motivated and still motivates speakers and writers to use some classical forms to satisfy their urgent need for stylistic effect. This is the essential reason why the separation of writing (formal) from speaking (informal) has been resurrected in modern times. It is then argued that the formal features of written Chinese mainly consist of (i) Monosyllabic Words Used in Disyllabic Templates, (ii) Disyllabic Words Used in Disyllabic Copulates, and (iii) Formal Phrasal Patterns. Although these must be used following strict principles, such as Stylistic Coherence, Auditory Intelligibility (PAI), and Shaping by Prosody (PSP), they must also be mixed with some colloquial features in order to make the language natural. As a result, an amalgamation principle which modulates literary dictions with colloquial expressions is proposed here. During the analyses of the issues outlined above, the syllabic system of Prosodic Grammar and the Tripartite System of Stylistic-Register Grammar are introduced and their empirical applications are also discussed, such as the statistical measurement methods recently developed for degree of formality measurement, composition testing, readability scaling, style gradation, textbook compilation, L2 learning, literacy acquisition, etc. Finally, a discussion is held on what principles the formal register grammar must observe.

The linguistic aspects of the Chinese language

The subjects under this section are numerous. They include characters, dialects, areas of linguistics, grammar, and Romanization systems. However, before diving into the chapters on these individual subjects, an overall review of the study of Chinese language, entitled 'Chinese Linguistics', has been written by William S. Y. Wang of The Chinese University of Hong Kong. His chapter gives a synoptic view on the languages and peoples of China from a multidisciplinary perspective. It describes the geographical setting of the country, as well as the prehistoric and historic background of the current multi-ethnic inhabitants. It draws upon knowledge gained in anthropology and genetics in addition to linguistics. While the main focus is on varieties of Mandarin, the dominant speech in China for over 1,000 years, the other major dialects are also discussed. This chapter serves as an introduction to other aspects of the Chinese language.

Chinese characters

This encyclopedia has a chapter on 'Chinese Characters' by John Yin Jinghua of the University of Vermont, who introduces the history and development of Chinese characters from

ancient times to the present. It also explains the formation and structure of Chinese characters and presents the features that are unique to Chinese characters.

Dialects

The Chinese language, which, as mentioned before, is spoken by the largest number of people on earth as their mother tongue has hundreds of different dialects. It is estimated that other than Mandarin, dialects in China are spoken by 1.2 billion people. The ten most popular dialects are: Mandarin, Min, Wu, Jin, Yue, Gan, Kejia (Hakka), Xiang, Pinghua, and Hui. This encyclopedia comprises a number of chapters that deal with dialects. 'Mandarin and Other Sinitic Languages', by Hilary Chappell of the Ecole des Hautes Etudes en Sciences Sociales in Paris and Li Lan of the Chinese Academy of Social Sciences in Beijing, provides a description of the ten main dialect groups which belong to the Sinitic or Chinese branch of the Sino-Tibetan language family. Beginning with the large Mandarin group which enjoys a wide geographical distribution in China, each branch of Sinitic is briefly characterized in turn in terms of its historical relation to Middle Chinese with respect to phonology (syllable structure), its tone system, and its personal pronoun paradigm, in addition to several interesting grammatical features with respect to a chosen representative dialect. Mandarin, as the most widely spoken dialect, is also discussed in detail in 'Mandarin' by Shi Dingxu of the Hong Kong Polytechnic University, who holds that it is the dialect spoken in Beijing in a narrow sense and the Northern dialect spoken in North China in a broader sense. The phonetic and phonologic system of *Putonghua* ('common language') is based on that of Mandarin. The description of Mandarin presented here covers its sound inventory, tonal system, words and morphology, phrases, clauses, sentences, aspect system, negation, comparison, and information package constructions. Two other dialects are also discussed in depth. First, 'Cantonese', a chapter written by Cheng Siu-Pong and Tang Sze-Wing of The Chinese University of Hong Kong; second, 'The Hakka Dialect' by Lau Chun Fat of Xiamen University.

Dictionaries

Dictionaries are a necessary evil for language learners, translators, and language professionals, and lexicography has a particularly long history in China. According to Cheung Kam-Siu of The Chinese University of Hong Kong, who has contributed the chapter on 'Lexicography', this activity has evolved and developed in China for more than two millennia. This chapter gives an overview of various aspects of diachronic studies of Chinese lexicography, including dictionary compilation and theorization in China from the ancient time to the modern and contemporary periods, and the development of Chinese monolingual and bilingual dictionaries. With detailed illustrations, some of the classics in the history of Chinese lexicography are specifically introduced and analyzed.

Lexicography has changed vastly from the ancient period to the present age. Huang Chu-Ren and Li Lan of the Hong Kong Polytechnic University, and Su Xinchun of Xiamen University, in 'Lexicography in the Contemporary Period', express the belief that the Chinese language enjoys a long history of dictionaries. The presentation of word meaning, pronunciation, grammar, and usage has developed in tandem with modern lexicography. The application of language mega-corpora has contributed to the compilation of Chinese dictionaries, either in paper or electronic form, for both native and non-native language learners and users. The dichotomy, however, is the definition of lexical units; there are no clear rules to segment words in data. Due to this reason, both character dictionaries and word dictionaries exist in the Chinese language.

Grammar

Whether Chinese has the same parts of speech as the Indo-European languages has been the subject of much debate. This encyclopedia has a chapter on 'Chinese: Parts of Speech' by Candice Cheung Chi-Hang of the City University of Hong Kong that offers a comprehensive survey of the major parts of speech in Chinese, aiming to establish the parts of speech that are found both in Chinese and in the Indo-European languages, and those found only in Chinese. A significant implication of this study is that whereas some parts of speech may be viewed as universal, others are language specific.

Idioms

Idiomatic expressions in Chinese include idioms, end-clippers, proverbs, slangs, and taboos. All these topics are covered adequately in this encyclopedia.

Jiao Liwei of the University of Pennsylvania focuses on 'Idioms'. According to Jiao, most Chinese idioms (*chengyu*) consist of four characters, and typically their meaning is more than the sum of the parts. Based on idiomatic salience, Chinese idioms can be divided into five categories and 15 types. After Written Vernacular Chinese replaced Classical Chinese in the 1910s, idioms have become an elite part of the Chinese language and received tremendous attention from people from all walks of life ever since. The chapter on 'Chinese Proverbs and Popular Sayings' by Lawrence Herzberg of Calvin College in Michigan takes a different approach and provides an overview of the rich treasury of Chinese proverbs and popular sayings. Over 400 sayings are presented in the chapter, including both the ancient sayings of Confucius and Lao Zi as well as the homespun truths of everyday life. All of these pithy proverbs offer insights into the Chinese language and culture.

'Chinese Slang' is discussed by Robert Moore of Rollins College. Slang is a category of informal speech, characterized by attributes related to playfulness or identification with a specific social group. In Mandarin, the term *liyu* is commonly translated as 'slang', though there are some differences between Chinese *liyu* and English slang. The main difference is that *liyu* was traditionally thought to be linked to local dialects. However, a new version of *liyu* has emerged on the Internet which is not dialect-based but is closely associated with youth culture.

Xiehouyu, or end-clippers, are covered in 'Chinese *Xiehouyu*' by Grace Zhang of Curtin University in Australia. This chapter provides a general review of Chinese *xiehouyu*, including its definition, origin, types, content, structure, features, functions, and translation. *Xiehouyu* refers to Chinese metaphorical folk sayings which follow a two-part pattern of a metaphor. These two parts are separated by a pause in speech or a comma or dash in writing, providing time for or drawing attention to the intended meaning and enhancing the effect of humour or satire.

Finally, Amy He Yun of the University of Huddersfield in United Kingdom, on the other hand, contributes a chapter on 'Chinese Taboo'. This chapter presents an introduction to Chinese taboo, focusing particularly on its distinctive characteristics such as the hierarchical nature of death taboos in ancient China. Starting by sketching the milestone stages of historical development, the chapter provides detailed analyses of taboo words/expressions and their euphemistic or dysphemistic substitutes in five categories. Apart from the motivations and functions of taboo and prosecution of transgressors, she also discusses the diachronic and synchronic variation as well as punishment avoidance strategies with illustrative examples.

Linguistic areas

This encyclopedia covers the main areas of linguistics, such as semantics and pragmatics; a number of sub-areas, such as corpus linguistics, psycholinguistics, computational linguistics, and sign language; as well as other related fields, such as rhetoric and syntax.

Semantics is discussed in 'Chinese Linguistics: Semantics' by Hsieh Shu-Kai from the National Taiwan University in Taiwan. According to Hsieh, meaning is the key to the understanding of communication, human cognition, and culture. As a subfield of linguistics, contemporary studies of meaning are characterized by various formal modelling, abstract representation, and empirical evidence that can be verified or disproved in a scientific manner. This chapter aims to give a brief and non-technical introduction to modern linguistic semantics, giving priority to the core issues surrounding lexical meanings in Chinese, ranging from morpho-semantics to lexical semantics. Janet Xing of Western Washington University in the United States writes on 'Semantic Change in Chinese'. Through analysis of diachronic data, she demonstrates how new meanings are developed at the lexical, sentential, and discourse levels. The result of her study suggests that three mechanisms, namely meta[phoricalization], meto[nymization], and semantic reanalysis, are commonly present in the process of semantic change in Chinese, rather than the mere two (meta and meto) recognized in Indo-European languages. Furthermore, she argues that Chinese lexemes tend to develop polysemies accretively due primarily to their isolating-analytical characteristics.

Meanings carried by several types of words are indeed noteworthy. Zhuo Jing-Schmidt writes on metaphors under the chapter heading of 'Metaphor in Chinese: Cognition, Culture, and Society', which focuses on the sociopolitical function of metaphor in a sociohistorical context by providing a close examination of two Chinese metaphors that have shaped China's collective unconsciousness.

Neologisms are discussed in 'Chinese Neologisms: Word-formation Strategies in Chinese' by Antonella Ceccagno of the University of Bologna. This chapter presents some central word-formation strategies in Chinese by focusing on the word-formation patterns that emerge as a result of the crucial role played by compounding in Chinese. The word-formation strategies that are discussed include: metacompounding, reanalysis of syllables as morphemes, neologisms formed by selecting a new meaning for the constituents of existing compounds, and the creation of new morphemes or new meanings for existing morphemes. Most of these word-forming patterns have been singled out only recently, and their implications are explored.

In the history of the Chinese language, words from various languages, including both languages of foreign countries and non-Chinese languages within China, have been borrowed. In 'Loanwords', Miao Ruiqin of Shanghai Jiao Tong University claims that loanwords are adapted into Chinese through one of the following four approaches: (i) phonemic transliteration, (ii) meaning translation, (iii) graphic borrowing, or (iv) the hybrid method. Linguistic processes to nativize loanwords into Chinese mainly involve phonetic, phonological, and orthographic transformations, whereas semantic, morphological, and grammatical changes are much less common. The use of loanwords in modern Chinese demonstrates certain sociolinguistic features in terms of stylistic and regional variations. In addition, existence of multiple adaptations present challenges for language standardization and language policy.

Elastic words, whose length can vary from monosyllabic to disyllabic, without a change in meaning, are discussed by Duanmu San and Dong Yan, both of Michigan University. Elastic words have been known since Karlgren (1918), but some theoretical and empirical questions still remain. A precise definition of elastic words is provided, along with an estimate

of the percentage of elastic words in modern Chinese. Finally, theories of why these words are created in Chinese are compared.

Pragmatics is about the use of language in different contexts. Jiang Yan of the Hong Kong Polytechnic University writes on Chinese pragmatics. This chapter starts with a survey of some general concepts and their exemplifications in Chinese. It then moves on to a short summary of relevance theory, followed by reports on findings from three case studies in Chinese pragmatics. The first concerns the Chinese adverb *bai*, which triggers a presupposition that resembles the Principle of Relevance. The second studies 'list gem' as a unique figure of speech, which is thought to have a special role to play in the current debate between contextualism and semantic minimalism in the sense that each side needs to supply extra accounts to accommodate such a case. The third provides a new analysis of Chinese counterfactuals, which takes most Chinese counterfactual conditionals as falsifying contingent counterfactuals.

Related to the core areas of linguistics there are some subfields, including corpus linguistics, psycholinguistics, computational linguistics, and sign language. 'Corpus-based Study of Chinese' by Tony McEnery and Richard Xiao, both of Lancaster University, explores the state of the art in using corpora in Chinese linguistic investigations by focusing on areas where corpus linguistics has made its greatest impact, namely lexical study, grammatical study, and Chinese interlanguage research. 'Chinese Psycholinguistics', on the other hand, is a chapter written by Jerome L. Packard of the University of Illinois. This chapter offers an overview of the field of Chinese psycholinguistics. The chapter begins by describing the psychology of Chinese script processing, including the role of strokes and radicals, followed by psycholinguistic aspects of Chinese text reading. The psycholinguistics of lexical access in Chinese is then discussed, including the issue of phonetic and semantic activation during word formation and retrieval. A summary of sentence processing is then presented, followed by a synopsis of Chinese speech perception and production. An additional chapter on psycholinguistics, entitled 'Psycholinguistics: Reading Chinese' is provided by Liang Tao of Ohio University and Alice F. Healy of the University of Colorado. This chapter presents findings from experimental studies on reading Chinese, including examinations of processes in English and Chinese, two typologically different languages in their respective orthography, syntactic structures, and discourse grammar. This cross-linguistic comparison illustrates reading processes that reflect (i) universal cognitive behavior and (ii) unique characteristics in Chinese reading. The theoretical issues supported in the former include the unitization hypothesis and the Stroop effect. In the latter, Chinese orthographic features invoke special cognitive processes in word identification and in discourse processes.

Computational linguistics is discussed in a chapter written by Zhang Xiaoheng of the Hong Kong Polytechnic University. This chapter focuses on Chinese, and especially on its differences with other languages, such as English. Chinese computational linguistics is often referred to as Chinese information processing in China. This chapter introduces the important areas of Chinese character information processing, word segmentation, natural language understanding and generation, corpus linguistics, and machine translation.

'Sign language' forms another chapter in this area, introduced by Gladys Tang of The Chinese University of Hong Kong. According to Tang and her fellow researchers, linguistic study of Hong Kong Sign Language (HKSL) was begun in the 1990s by James Woodward (1993). Since then, research on the linguistic properties of HKSL at different linguistic levels has been flourishing. In this chapter, she first provides an account of the historical development of HKSL, and then focuses on the descriptions of HKSL at the phonological, morphological, and syntactical levels.

Other areas related to linguistics covered in this encyclopedia are morphology, syntax, and rhetoric. 'Chinese Morphology' is discussed by Jerome L. Packard of the University of Illinois. This chapter describes the properties of words in Mandarin Chinese. It begins with Chinese typological characteristics, and defines word as a syntactically free form. The four Chinese morpheme types (free content morpheme, free function morpheme, bound root morpheme, and affix) are then described, including how the processes compounding, composition, and affixation compose the word types compound, bound root word, derived word, and grammatical word. The chapter concludes with a discussion of borrowing, reduplication, abbreviation, and neologism in Chinese.

'Syntax' is covered in a chapter by Li Yafei of the University of Wisconsin-Madison. Chinese syntax during the past 30 years has uncovered a wide range of new facts, explored various novel analyses, and come up with a non-trivial number of previously unknown questions. It has benefited from the general theory of linguistics as well as contributed to it. Both the similarities and differences between Chinese and other well-studied languages are becoming increasingly better understood. A question is how to improve the ways in which Chinese syntax helps shape the theory of the human languages. Another chapter on syntax, entitled 'Chinese Syntax', is written by Walter Bisang. This chapter starts out from a critical discussion of the extent to which there is syntax in Chinese and the question of the interaction between syntax and pragmatics. For that purpose, it briefly highlights some claims concerning subject–object asymmetry in coordinate-clause constructions and in Chinese relative clauses. Later on, it addresses the argument structure, nominal expressions, verbal structures, anaphora, information structure (focus and topic), the formation of questions, and issues of word-order typology (SVO vs. SOV). The presentation of these phenomena includes well-known syntactic constructions of Chinese like the *ba*-construction, the *bei*-construction, and numeral classifiers.

'Chinese Rhetoric', on the other hand, is a chapter written by Andy Kirkpatrick of Griffith University in Brisbane, Australia. His chapter provides an overview of the development of Chinese rhetoric from the earliest times to the present. It argues that China needs to rediscover its rich rhetorical tradition so that it may be adapted for public discourse in contemporary China. It includes some comparison and contrast with developments in Western rhetoric. Related to this area is a chapter on 'Poetic Prosody' by Feng Shengli of The Chinese University of Hong Kong, which deals with poetic prosody in Classical Chinese. It is shown that Archaic Chinese prosody had undergone a change from moraic foot structure (a foot consisted with two moras) to syllabic foot structure (a foot consisted with two syllables). More specifically, disyllabic Foot Formation developed during the Spring and Autumn period (722–481 BC). Only later did trisyllabic super-foot structure develop, requiring the maturity of disyllabic Prosodic Word Formation. It took even longer for Prosodic Word Compounding in the form of four-syllable structures to fully develop. Observing the development of the prosodic system from Old Chinese up to Middle Chinese, it would be expected that the development of pentasyllabic lines would have to wait for the maturity of super-foot formation and that heptasyllabic lines would have to wait for the maturity of PrWd Compounding. The former is seen in the trisyllabic compound formation of the Eastern Han dynasty (25–220), while the latter appears in tetrasyllabic verb formation during the Southern and Northern dynasties (420–589). According to the theory and facts presented here, the prosodically conditioned poetic effects might provide a basis for motivating a theory of prosodic stylistics in future studies.

Romanization systems

Romanization consists of the spelling of Mandarin in Roman letters. It is related to syllable initial and final in traditional Chinese phonology. There are over 20 Romanization systems for the Chinese language.

This encyclopedia contains chapters on the Romanization of Cantonese, the Hanyu Pinyin system, and the Wade–Giles Romanization system. In the chapter on 'Cantonese Romanization', Cheng Siu-Pong and Tang Sze-Wing of The Chinese University of Hong Kong hold that the diversity of Cantonese Romanization is a reflection of different backgrounds and purposes in rendering Cantonese into Roman letters. Some forms of Romanization are marked for their historical significance and influence on later proposals; whereas others are noteworthy for their current popularity, as they are often found in books and other materials. Individual frameworks are discussed in detail, and the concluding tables summarize these discussions. It follows that such diversity compensates for the lack of a predominant form of Cantonese Romanization.

Hanyu Pinyin, which is the most popular Romanization system in China and possibly in the world, is discussed in detail by Lilly Chen of Rice University. This chapter begins with an introduction to the official system for spelling Mandarin in Roman letters, and how it emerged from among many competing spelling systems, the earliest of which were devised by foreign missionaries. The structure of Pinyin is described in detail, including also the concepts of syllable initial and final in traditional Chinese phonology. The historical development from a morpho-syllabic to a phonemic system of sound representation is discussed, along with a functional evaluation of Pinyin.

Another system that is popular in Taiwan is the 'Wade–Giles Romanization System', explained by Karen Steffen Chung of National Taiwan University. The earliest efforts to Latinize Chinese began in the sixteenth century with the work of Italian Jesuit priests Michele Ruggieri and Matteo Ricci. Their system was modified over the centuries by other Western missionaries, most notably Robert Morrison (1782–1834). Sir Thomas Francis Wade (1818–1895) further refined Morrison's scheme, and Herbert Allen Giles (1845–1935) later popularized Wade's system mainly through his 1,415-page *A Chinese–English Dictionary*. The Wade–Giles system was soon widely adopted, until its eclipse in the 1970s by Hanyu Pinyin, which is now the undisputed worldwide standard.

Conclusion

It is hoped that these chapters on the global, historical, and linguistic aspects of the Chinese language, written by 61 scholars and experts from different parts of the world, will provide a comprehensive background and an authoritative reference to learners and researchers of the Chinese language.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my former colleague Professor David Pollard, who served as Chair Professor of Translation at The Chinese University of Hong Kong for some years, for offering me a chance to work on this very meaningful project, and to Andrea Hartill, Senior Publisher of Routledge, for initiating this encyclopedia. Without her support, the present volume would never have been completed and published. My thanks also go to Camille Burns, Editorial Assistant, and Geraldine Martin, production editor, both of Routledge, for their assistance in the editing and production of this volume.

I am grateful to all the Consultant Editors for their advice and support, to all contributors for their time and effort in writing articles for this volume, to Florence and James, assistant editors of this encyclopedia, for their help and hard work, and to Florence and Wei Yuxiang for compiling the Index.

Chan Sin-wai

The editor and publishers would like to thank the following copyright holders for permission to reproduce the following material:

Tables and material excerpted from 'Fundamentals of Chinese Characters'. © 2006 John Jing-hua Yin. Reproduced with kind permission of Yale University Press.

'ACTFL/CAEP 2012 Standards'. Reproduced with kind permission of American Council on the Teaching of Foreign Languages.

Figure 'Austronesian Languages in Taiwan', courtesy of Professor Paul Li.

Figure '235' from 'The Archaeology of Ancient China' by Kwang-chih Chang. © 1987 Yale University Press. Reproduced with kind permission.

Figures, Tables, and Excerpts from 'Love and War in Ancient China: Voices from the shijing' pp. xxii–xxxiii and p. 38 Figure 8b, © 2013 City University of Hong Kong. Reproduced with kind permission.

Figures, Tables, and Excerpts from 'On Modern Written Chinese' 37: 145–61, © 2009 Journal of Chinese Linguistics. Reproduced with kind permission.

Table (1) from 'Shoubuchulai, La formazione delle parole in cinese' p. 128, © 2009 Serendipita Editrice. Reproduced with kind permission.

Figures, Tables, and Excerpts from 'A Prosodic Explanation for Chinese Poetic Evolution' 2: 223–57, © 2011 Tsing Hua Journal of Chinese Studies. Reproduced with kind permission.

Material excerpted from 'Chinese Proverbs and Popular Sayings: With Observations on Culture and Language'. © 2012 Qin Xue Herzberg and Larry Herzberg. Reproduced with kind permission of Stone Bridge Press, Inc.

While the publishers have made every effort to contact copyright holders of material used in this volume, they would be grateful to hear from any they were unable to contact.

1
ANCIENT CHINESE

Alain Peyraube

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE, ECOLE DES HAUTES ETUDES EN SCIENCES SOCIALES, PARIS, FRANCE

1. Introduction

'Chinese is only one of a very few languages whose history is documented in an unbroken tradition extending back to the second millennium BC' (Norman 1988: ix). Chinese is usually divided into Ancient Chinese (*gǔdài hànyǔ* 古代汉语) and Contemporary Chinese (*xiàndài hànyǔ* 现代汉语). Ancient Chinese is simply defined as 'the language of the writings of the past' (Wang 1979: 1). It covers a very long period, from the oracle bone inscriptions (OBI, the first Chinese inscriptions known to us, dated from the fourteenth century BCE) until the nineteenth century. Three basic stages are generally distinguished for Ancient Chinese: (i) the Archaic period (*shànggǔ* 上古), until the second century BCE; (ii) the Middle or Medieval period (*zhōnggǔ* 中古), from the first century BCE to the middle of the thirteenth century CE; (iii) the Modern period (*jìndài* 近代), from the middle of the thirteenth century to the middle of the nineteenth century.

It was during the Archaic period that what is known today as Classical Chinese (*wényán* 文言) was standardized. This language, playing a role like Latin in Europe, remained the main written language used in literary texts until the beginning of the twentieth century. The period for Classical Chinese *par excellence* refers more precisely to the language used by the philosophers and scholars of the Warring States period (475–221 BCE) and it was probably not very different from the educated speech of the period. The gap between the written and the spoken language began to develop in Han times or Pre-Medieval Chinese (206 BCE – CE 220) and increased considerably with time. It is essentially Classical Chinese, also called Late Archaic Chinese (*hòu shànggǔ hànyǔ* 后上古汉语), that will be discussed in this chapter. I will nevertheless make several digressions regarding the period prior to the fifth century BCE, i.e. what is known as pre-Classical Chinese or Early Archaic Chinese (*qián shànggǔ hànyǔ* 前上古汉语), and about the Medieval period, especially in the phonological section.

2. Phonology

The history of Chinese phonology is usually divided into four periods based on the two important rhyme dictionaries, the *Qiēyùn*《切韵》of 601 and the *Zhōngyuán yīnyùn*《中原音韵》(*Rhymes according to the pronunciation of the Central Plains*) of 1324: Old Chinese, Middle Chinese, Old Mandarin, and Modern Chinese. Old Chinese refers to the period before

the *Qiēyùn* and is the earliest period; the *Qiēyùn* represents the beginning of Middle Chinese; the *Zhōngyuán yīnyùn* represents the beginning of Old Mandarin; and Modern Mandarin is the period leading up to present time, with no fixed dates. See Peyraube and Shen (forthcoming).

2.1. Old Chinese phonology

The source materials for Old Chinese phonological reconstruction, mainly based on rhyming materials (largely drawn from the *Shī jīng* 《诗经》 (*The Book of odes*), eleventh–sixth centuries BCE), do not allow us to clearly identify the reconstructed phonological forms in space and time.

The syllable in Old Chinese is analyzed as being composed of an *initial* and a *final*. Following Baxter (1992: 7), whose inventory of phonetic segments is given below, (i) the initial contains a *pre-initial* (treated as a prefix in Baxter and Sagart 1998) and an *initial* (Table1.1), and (ii) the final contains a *medial*, a *main vowel* (Table 1.2), a *coda*, and a *post-coda* (Table 1.3).

Table 1.1 37 Old Chinese initials (Baxter 1992: 177)

p	p^h	b	m	hm	w	hw
t	t^h	d	n	hn	l	hl
					r	hr
					j	hj
					z	s
ts	ts^h	dz				
k	k^h	g	ŋ	hŋ		
k^w	k^{wh}	g^w	$ŋ^w$	$hŋ^w$		
ʔ	x	ɦ				
$ʔ^w$						

Table 1.2 The six main vowels (Baxter 1992: 180)

i	ɨ	u
e		o
	ɑ	

Table 1.3 Old Chinese codas (Baxter 1992: 181)

zero	k	ŋ
j	t	n
w	wk	
	p	m

Three medial elements have been reconstructed: **-r-*, **-j-* (though the reconstruction of the medial **-j-* has now been replaced by a contrast of vowel length), and, marginally, **-l-*. The two post-codas are **-ʔ* and **-s*, which are the respective sources of the rising tone (*shǎngshēng* 上声) and of the departing tone (*qùshēng* 去声) in Middle Chinese.

This reconstructed Old Chinese phonemic system is far from being universally accepted. In fact, several specialists continue to consider that we are unable to do more than to approach a reconstruction of Old Chinese.

2.2. Middle Chinese phonology

The reconstruction of Middle Chinese phonology, in a more accurate sense, means the reconstruction of phonetic values for the existing categories, which are usually assumed to be phonemic.

Table 1.4 Middle Chinese initials (Baxter 1992: 45)

Labials	p	p^h	b	m			
Dentals/ Alveolars	t	t^h	d	n			
Lateral						l	
Retroflex stops	tr	tr^h	dr	nr			
Dental sibilants	ts	ts^h	dz		s	z	
Retroflex sibilants	tsr	tsr^h	dzr		sr	zr	
Palatals	tsy	tsy^h	dzy	ny	sy	zy	y
Velars	k	k^h	g	ŋ			
Laryngeals	ʔ				x	h	

The basic medials are the glides *-j-* and *-w-*.

The reconstruction of main vowels has changed several times. Efforts have been made to reduce the number of main vowels and to make the vowel system more natural in relation to linguistic universals.

Table 1.5 Middle Chinese main vowels (Baxter 1992: 61)

i	u
e	o
ɛ	
æ	

These main vowels may be followed by commonly reconstructed codas, which are either stop consonants or approximants (Table 1.6).

Table 1.6 Middle Chinese codas (Peyraube and Shen, forthcoming)

zero	w		j
ŋ		m	n
k		p	t

Middle Chinese has also a system of four tones which were first identified and named in the fifth century. These tones are called *píng* 平 'level', *shǎng* 上 'rising', *qù* 去 'departing or falling', and *rù* 入 'entering and short'. Each Chinese syllable belongs to one of these four tonal categories.

3. Morphology

It is often said that Ancient Chinese, being a monosyllabic and monomorphemic language, is a language with an impoverished morphology, a language in which the grammatical processes are almost totally syntactic, thus a language of that type called analytic or isolating. However, Ancient Chinese did indeed possess morphological processes, although none of them was fully productive. They actually may represent a vestige of older stages in which such a process was considerably more productive. These word-formation processes are of the same type as those found in Contemporary Chinese: compounding, reduplication (total reduplication such as *wēiwēi* 巍巍 'tall and grand', or partial reduplication such as *tángláng* 螳螂 'praying mantis'), and even affixation, which was not at all an unproductive process in Ancient Chinese. Having derivational and even inflectional affixes, Ancient Chinese could be thought of as more synthetic than later stages of the language.

Several prefixes, suffixes, and infixes have now been reconstructed for Ancient Chinese. These are derivational morphemes changing the meaning or part of the speech of the words to which they are attached (the ensuing discussion closely follows the treatment of Baxter and Sagart 1998).

A good example of a prefix that has been reconstructed is the prefix **N-* (causing a following voiceless obstruent to become voiced in Medieval/Middle Chinese) which, when attached to a verb (or even a noun in some cases), seems to produce an intransitive verb or adjective: thus, *kens* 见 'to see' > **N-kens* 'to appear'. Another prefix **k-*, added to verbs, could refer to more concrete actions taking place in a limited time frame: **ljuk* 鞠 'to nourish' > **k-ljuk* 'to breast feed'. See Sagart (1999). Another prefix **s-* derives causative verbs from non-causative verbs or even from nouns (denominative suffix): **m-lun-s* 顺 'obedient' > **s-lun* 'to make (a horse) obedient'.

Concerning the suffixes, there is in Old Chinese a reconstructed suffix **-s*, which is the source of the departing tone of Middle Chinese, which, when added to adjectives or verbs produces derived nouns: for example, **tsrek* 责 'to demand payment' > **tsrek-s* 'a debt'. See Mei (1989). Some gradable adjectives also have corresponding noun forms in which the suffix **-s* functions like English *-th*, occurring in pairs such as 'deep/depth', 'wide/width' (see Downer 1959; Mei 1980).

Two infixes have also been reconstructed in Old Chinese by Sagart (1993). The exact function of the first one, **-j-*, is difficult to establish, but forms with and without **-j-* do appear to be semantically related. The second infix, **-r-*, is said to produce forms that are plural or collective in the case of nouns, and iterative, durative, or indicating effort in the case of verbs.

There are also some bi-morphemic monosyllabic words in Classical Chinese, which result from the fusion of two morphemes. One of the most cited examples of this fusion phenomenon is the negative *fú* 弗, which is analyzed as a combination of the negative *bù* 不 'not' and the third-person pronoun *zhī* 之 'him, her, it'.

Among other contractions of this kind in Classical Chinese are to be found: *zhī* 之 'third personal pronoun' + *yú* 于 'to, at' > *zhū* 诸; *wú* 毋 'not' + *zhī* 之 'third personal pronoun' > *wù* 勿; *yú* 于 'at, to, from' + *zhī* 之 'third personal pronoun' > *yān* 焉; *hú* 胡 'why' + 不 *bù* 'not' > *hé* 盍. Examples:

(1) 乞诸其邻而与之(《论语》)

Qi	zhu	qi	lin	er	yu	zhi
ask-for	3SG+to	his	neighbour	and	give	3SG

'(He) asked his neighbour for it and gave (it to) him' (*Confucian Analects*, fifth century BCE)

(2) 得之则生, 弗得则死(《孟子》)
de zhi ze sheng fu de ze si
get 3SG then live NEG+3SG get then die
'(If he) gets it (he will) live, (if he) does not get it, (he will) die' (*Mencius*, fourth century BCE)

(3) 百亩之田, 勿夺其时(《孟子》)
bai mu zhi tian wu duo qi shi
hundred mu LIG field NEG+3SG deprive their time
'The hundred *mu* of fields, do not deprive them of their time (of cultivation)' (*Mencius*)

(4) 学焉而后臣之(《孟子》)
xue yan er hou chen zhi
learn from+him and later make-subject 3SG
'(He) learned from him and later made him (his subject)' (*Mencius*)

In fact, starting in the Pre-Medieval (Han) period, ca. first century BCE, all the above examples of fusion disappeared and a fission process came into being, which indicates undoubtedly a move from syntheticity to analyticity: *zhū* 诸 > *zhī* 之 + *yú* 于; *fú* 弗 > *bù* 不 + *zhī* 之; *wù* 勿 > *wú* 毋 + *zhī* 之; *yān* 焉 > *yú* 于 + *zhī* 之; *hé* 盍 > *hú* 胡 + *bù* 不.

4. Syntax

In Ancient Chinese, the subject (when expressed, as many sentences are without any subject) precedes the predicate, the verb precedes its object, and the modifiers precede the words they modify. Thus, Ancient Chinese is an SVO language (where S = subject, V = verb, and O = object).

4.1. Sentence types

The predicate might be complex, i.e. composed of more than one verb. Such cases involve serial verb constructions of the type V1 . . . V2 . . . (V3). . . . The semantic relationship between verbs in series is varied. In the following example, it involves an implication of purpose, where *yǐ* 'in order to' links the two verb phrases:

(5) 楚人伐宋以救郑(《左传》)
Chu ren fa Song yi jiu Zheng
Chu people raid Song in-order-to save Zheng
'The people of Chu raided Song in order to save Zheng' (*The Tradition of Zuo*, fifth century BCE)

Complex predicates may also involve a 'pivotal construction', in which the noun phrase object of the first verb is the subject of the second verb:

(6) 请君讨之(《左传》)
qing jun tao zhi
ask Prince attack 3PL
'[I] ask [you] the Prince to attack them' (*The Tradition of Zuo*)

An interesting characteristic of the subject–predicate construction in Classical Chinese is that the whole construction can be nominalized by inserting the particle (ligature) *zhī* 之 between the two constituents of the construction:

(7) 人之爱人求利之也(《左传》)
ren zhi ai ren qiu li zhi ye
person LIG love person pursue profit 3SG FIN-PART
'One person loving another person [would] pursue profit [for] him'
(*The Tradition of Zuo*)

A transitive verb can take one object (usually a noun phrase) or two, an indirect object (IO) and a direct object (DO) in that order (V+IO+DO). The double-object construction is restricted to those verbs with the semantic feature [+ give], [+ say] or [+ teach]: Apart from the pattern V+IO+DO, example (8), two other common orders, involving the prepositions *yǐ* (DO marker) or *yú* ('to', introducing the IO), are possible for the dative construction: (i) *yǐ* + DO+V+IO (9), and V+DO+*yú*+IO (10). See Peyraube (1987).

(8) 公赐之食(《左传》)
gong ci zhi shi
prince offer 3SG food
'The prince offered him food' (*The Tradition of Zuo*)

(9) 尧以天下与舜(《孟子》)
Yao yi tianxia yu Shun
Yao OM Empire give Shun
'Yao gave the Empire to Shun' (*Mencius*)

(10) 尧让天下于许由(《庄子》)
Yao rang tianxia yu Xu You
Yao leave Empire to Xu You
'Yao left the Empire to Xu You' (*Zhuangzi*, fourth century BCE)

Verbs may be followed by complements (*bǔyǔ* 补语), a term used for adjuncts when they follow the verb. When adjuncts precede the verb, they are denoted as adverbials (*zhuàngyǔ* 状语). Complements are divided into two types, depending upon whether or not they are introduced by a prepositional marker which forms a prepositional phrase (PP). Notable examples of PPs are locative complements, which are usually introduced by the preposition *yú* 于/於 'at, to', as in:

(11) 北学于中国(《孟子》)
bei xue yu zhong guo
north learn at central state
'He went to the north to learn [it] in the Central States' (*Mencius*)

The PPs can also be adverbials, especially when they are formed with the preposition *yǐ* 以, as in:

(12) 以羊易牛(《孟子》)
yi yang yi niu
with sheep change ox
'Change the ox for a sheep' (*Mencius*)

Sentences are also customarily divided into simple and complex types and can be further classified as declarative, interrogative, imperative, and exclamatory sentences. There are three basic types of interrogative sentences: (i) yes/no questions; (ii) WH-questions; and (iii) rhetorical questions. The first type is formed with final question particles (*hū* 乎, *yú* 欤, *yé* 耶), which essentially transform statements into questions. The second type (WH) contains a question word (one of the interrogative pronouns listed below), generally without any final particle. The third type is more complex. Some rhetorical questions are formed with a final particle (*hū*, *yú*, or *yé*), but also with an adverb of negation placed before the verb, implying an affirmative answer. Others are formed with the modal particle *qǐ* 岂 'how could' requiring an affirmative answer. The final particles *hū* and *zāi* 哉 are also generally used, though they may be omitted:

(13) 予岂好辩哉?(《孟子》)
yu qi hao bian zai
1SG how-could like debate FIN-PART
'How could I [be one who] loves debating?' (*Mencius*)

Imperatives are not syntactically marked as such in Classical Chinese. The subject is usually deleted, but this in itself is not a sufficient diagnostic of the imperative sentence. However, when the imperative is intended to be understood as a request, and not as an order or a prohibition, the verbs *yuàn* 愿 'wish' or *qǐng* 请 'beg' are used:

(14) 王请度之(《孟子》)
wang qing du zhi
prince beg measure it
'[My] Prince, please measure it' (*Mencius*)

The final particle *zāi* 哉 or sometimes *yǐ* 矣 is the usual marker of the exclamatory sentence. It can be added either to a declarative or to an interrogative. The subject–predicate order is usually inverted in exclamatory sentences. Example:

(15) 死矣盆成括!(《孟子》)
si yi Pencheng Kuo!
dead FIN-PART Pencheng Kuo
'He is dead, Pencheng Kuo!' (*Mencius*)

A special kind of sentence in Classical Chinese is the copular sentence. If one defines the copula as an overt word which, when used in equational sentences, links the subject to a nominal predicate, and expresses (i) an equivalence meaning or (ii) a property or classificatory meaning, then one can identify the presence of copulas in Classical Chinese, even if they are not strictly necessary.

The most common way of creating copular sentences is to add the final particle *yě* 也 at the end of a sentence (see Peyraube and Wiebusch 1995 for a detailed account of the history of copulas in Ancient Chinese, and especially for a discussion of the status of *yě* as a copula):

(16) 彼丈夫也,我丈夫也(《孟子》)
bi zhangfu ye wo zhangfu ye
that reliable-man FIN-PART 1SG reliable-man FIN-PART
'They were reliable men, I am a reliable man [too]' (*Mencius*)

In addition to *yě*, other copulas are attested in Classical Chinese. Thus, the negative copula *fēi* 非 'to not be' is required in all negative nominal predicate sentences. In affirmative copular sentences, the verb *wéi* 为, which also means 'to do, to regulate, to act, to consider as', and so forth also acts regularly as a copula, and finally the copular verb *shì* 是 'to be', still used today, and which comes from the demonstrative pronoun *shì* 'this' through a process of grammaticalization, is already attested no later than the Qin dynasty (second century BCE).

Concerning the passive sentences, there are semantic passives in Classical Chinese, i.e. passives expressing passivity without any overt morphological marker: a transitive verb can be made passive by placing its object (the patient) in subject position. However, there are also passive structures marked with some formant, such as a preposition, or an auxiliary verb. The most common way to form a passive construction in Classical Chinese is to use the preposition *yú* 于/于 'by' to introduce the agent (V + *yú* + Agent), as in:

(17) 治于人者食人治人者食于人(《孟子》)
zhi yu ren zhe si ren zhi ren zhe si yu ren
rule by other the-one-who feed other rule other the-one-who feed by other
'Those who are ruled by others feed others, those who rule are fed by others' (*Mencius*)

Four other passive structures are commonly attested: *wéi* 为 + V, *jiàn* 见 + V (where *wéi* and *jiàn* are best considered to be auxiliary verbs, and where no agent is expressed), and *wéi* + Agent + V and *wéi* + Agent + *suǒ* 所 + V, which became the most common passive forms from the second century BCE on. Examples:

(18) 盆成括见杀(《孟子》)
Pencheng Kuo jian sha
Pencheng Kuo PASSIVE kill
'Pencheng Kuo was killed' (*Mencius*)

(19) 后则为人所治(《史记》)
hou ze wei ren suo zhi
late then PASSIVE other PASSIVE control
'[If I react] late, [I] will then be controlled by others' (*Records of the Historian*, first century BCE)

Yet another passive form appears at the end of the Classical period: *bèi* 被 + V, where *bèi* is a verb meaning 'to suffer; to be affected'. It will first become an auxiliary verb expressing passivity, like *wéi* and *jiàn* without any agent, and several centuries later, around the Early Medieval period, a preposition introducing a noun phrase agent, after being grammaticalized. For a more detailed analysis of the passive forms in Ancient Chinese, see Peyraube (1989). Example of *bèi* being a passive preposition:

(20) 亮子被苏峻害(《世说新语》)
Liangzi bei Su Jun hai
Liangzi PASSIVE Su Jun kill
'Liangzi was killed by Su Jun' (*Shi shuo xin yu*, fifth century CE)

Complex sentences are composed of two or more clauses joined through coordination or subordination. The joining of clauses can be accomplished without any overt marking, as in example (21) or with a connective marker *ér* 而 'and, but' or *yì* 亦 'also', as in (22):

(21) 老者安之, 朋友信之, 少者怀之(《论语》)
lao zhe an zhi pengyou xin zhi shao zhe huai zhi
old the-one-who soothe 3PL friend trust 3PL young 3PL care 3PL
'As for the old, soothe them; as for friends, trust them; as for the young, care for them' (*Confucian Analects*)

(22) 人民少而禽兽众(《韩非子》)
renmin shao er qin shou zhong
people few but bird beast numerous
'People are few but [wild] animals are numerous' (*Han Feizi*, third century BCE)

Subordination may be indicated by subordinating conjunctions or particles, which can occur in the first clause, in the second, or in both. In the case of conditional sentences, the conjunctions *rú* 如, *ruò* 若 or *gǒu* 苟 'if' may appear in the first clause (*if*-clause), and the markers *zé* 则 or *sī* 斯 'then' in the main clause (for an exhaustive analysis of the conditionals in Classical Chinese, see Harbsmeier 1981: 229–87). In concessive sentences, the most commonly used conjunction of concession is *suī* 虽 'although, even if'. In the main clause one often finds *ér* 而, which then has its adversative meaning 'yet'.

In sentences expressing cause, the 'because' clause may be introduced by the preposition *yǐ* 以, and the main clause may contain the connective *gù* 故 'so, therefore':

(23) 以其不争故天下莫能与之争(《老子》)
yi qi bu zheng gu tian xia mo neng yu zhi zheng
because 3SG NEG compete therefore Heaven under nobody can with 3SG compete
'Because he does not compete, nobody can compete with him under Heaven' (*Laozi*, fourth century BCE)

4.2. Word order

Saying that Classical Chinese is an SVO language (see above) is not uncontroversial. The problem of word order and word order change in Chinese has been much debated since the 1970s. The discussions have essentially surrounded the hypothesis of Li and Thompson (1974), according to which Archaic Chinese was a SOV language, which might have changed to SVO, before shifting back again to SOV, the last stage still being underway: a) SOV > SVO; b) SVO > SOV.

The hypothesis b) SVO > SOV has been criticized, and it has been shown that, synchronically, Chinese still is and remains a SVO language and that the OV order is a marked word order coding contrastiveness. The first change a) SOV > SVO has been challenged less because the SOV order for Archaic Chinese seems plausible, insofar as this order is found in Classical Chinese under special conditions. Consequently, it has been claimed that Proto-Chinese must have been SOV and, therefore, Proto-Sino-Tibetan too, since almost all the Tibeto-Burman languages have a verb-final order (see LaPolla 1993; Mei 1997).

Contrary to this viewpoint, the first attested documents of Pre-Archaic Chinese, i.e. the Oracle Bone inscriptions (OBI) (fourteenth–eleventh centuries BCE), show a regular order of

SVO, and this stage of the language is indeed more SVO than later stages of Chinese, such as Early or Late Archaic. To claim, then, that in a more ancient stage, before the oracle bone inscriptions, the basic order could have been SOV, is a pure surmise, and cannot be empirically proved (see Shen 1992; Peyraube 1997a, 1997b; and Djamouri 2001), who gives the following figures: 93.8% of the sentences in the OBI are (S)VO vs. 6.2% (S)OV).

There are also several clear indications for SVO order being the more basic and dominant word order than SOV in Early and Late Archaic, i.e. in Classical Chinese. When the object is a full lexical noun phrase (NP), the basic order is almost exceptionlessly VO, as in:

(24) 君必失国(《左传》)
Jun bi shi guo
Prince certainly lose state
'The prince (will) certainly lose the State' (*The Tradition of Zuo*)

There are some cases in which the NP-object is in a preverbal position, but these cases are marginal and the OV order is then marked [+ contrastive]. The same situation applies when the noun object is followed by a preverbal marker, usually *shì* 是 or *zhī* 之, as in (6):

(25) 今吴是惧(《左传》)
jin Wu shi ju
now Wu OM afraid
'Now [they] are afraid of [the state of] Wu' (*The Tradition of Zuo*)

However, there are also cases of OV order in Archaic Chinese, not found in Contemporary Chinese, in which the object is a pronoun, being either (i) an interrogative pronoun (26), (ii) the demonstrative pronoun *shì* 是 'this' or (iii) a pronoun in a negative sentence (27):

(26) 吾谁欺?欺天乎?(《论语》)
wu shei qi? qi tian hu?
I who deceive deceive Heaven INT-PART
'Whom should I deceive? Should I deceive Heaven?' (*Confucian Analects*)

(27) 不吾知也(《论语》)
Bu wu zhi ye
NEG 1SG understand FIN-PART
'(You) don't understand me' (*Confucian Analects*)

What can be said about these various OV orders involving pronouns? There are first some statistics that show that the OV order has always been non-dominant, since it is only well-attested in the case of pronoun objects. In a corpus of 2,767 VO or OV sentences drawn from the bronze inscriptions, 88.56% of objects (O) are nouns, only 3.3% are pronouns (see Guan 1981: 88). The ratio of pronoun objects is certainly higher in other later documents of Classical Chinese, but it never exceeded 15% of the entire body of VO and OV constructions. It is also known that in many languages, the position of pronouns is different from that of noun phrases, and that 'unstressed constituents, such as clitic pronouns, are often, cross-linguistically, subject to special positioning rules only loosely, if at all, relating to their grammatical relation, so sentences with pronouns can be discounted in favor of those with full noun phrases' (Comrie 1989: 89).

4.3. Full words and empty words

Classical Chinese words are traditionally divided into two categories: *shízì* 实字 'full words' and *xūzì* 虚字 'empty words'. The former are content words (which carry semantic content) and form an open class; included in this category are nouns, verbs, and adjectives. The latter are function words or grammatical words, used to express grammatical relationships. They include pronouns, adverbs, prepositions, conjunctions, and particles (see Peyraube 2004). Words can be used in functions customarily reserved for other words. This does not imply, however, as some scholars have assumed, that there are no parts of speech in Classical Chinese, and that words can be used indifferently in any grammatical category.

4.3.1. Full words

Chinese nouns typically function as subjects or objects. However, under certain conditions, they may function like verbs, such as the predicates in (28) and (29), or like adverbs, such as the adverbials in (30):

(28) 君君臣臣父父子子(《论语》)
jun jun chen chen fu fu zi zi
ruler ruler minister minister father father son son
'The ruler acts as a ruler, the minister as a minister, the father acts as a father, and the son as a son' (*Confucian Analects*)

(29) 物物而不物于物(《庄子》)
wu wu er bu wu yu wu
thing thing and NEG thing PASSIVE thing
'Treat things as things and do not be treated as things by things' (*Zhuangzi*, fourth century BCE)

(30) 豕人立(《左传》)
shi ren li
pig man stand-up
'The pig, like a man, stood up' (*The Tradition of Zuo*)

Chinese verbs are fundamentally predicative in nature. Unlike nouns, which are negated by the adverb of negation *fēi* 非 'to not be', verbs are negated by the simple adverb *bù* 不 'not'. One particular use of intransitive verbs in Classical Chinese is in a causative function: thus, *huó* 活 'to live': 'to make (people) live'; *xíng* 行 'to go': 'to put into motion'; *yǐn* 饮 'to give to drink': 'to give something to drink'; *dòu* 斗(鬬) 'to fight': 'to make (people) fight'. Example:

(31) 不如先斗秦赵(《史记》)
buru xian dou Qin Zhao
better first (make-)fight Qin Zhao
'It is better to make the Qin and Zhao fight first of all' (*Records of the Historian*)

Another verbal subclass is composed of modal auxiliary verbs (*qíngtài zhùdòngcí* 情态助动词), also sometimes called 'can-wish verbs' (*néngyuàn dòngcí* 能愿动词). They are verbs that take other verbs as their objects and express the modality of the following

verb phrase. This modality (ability, possibility, probability, certainty, obligation, volition, etc.) can be characterized as epistemic, deontic, or dynamic. Auxiliary verbs form a closed list and can be classified into the following four semantic groups: (i) verbs expressing mainly possibility and permission, including *kě* 可, *néng* 能, *zú* 足, *dé* 得, *huò* 获, *kěyǐ* 可以, and *zúyǐ* 足以 (32); (ii) the four verbs of volition, *gǎn* 敢, *kěn* 肯, *yù* 欲, and *yuàn* 愿 (33); (iii) the two auxiliaries of necessity, obligation and requirement (certainty and obligation), *yí* 宜 and *dāng* 当; and (iv) the passive auxiliaries *jiàn* 见, *wéi* 为, and *bèi* 被. For a detailed analysis of mood and modality in Chinese, see Chappell and Peyraube (forthcoming).

(32) 天子不能以天下与人(《孟子》)
Tian zi bu neng yi tianxia yu ren
Heaven son NEG can OM Empire give other
'The Emperor cannot give the Empire (to) others' (*Mencius*)

(33) 子欲居九夷(《论语》)
Zi yu ju Jiu Yi
Master intend-to live Jiu Yi
'The Master intends to live in Jiu Yi' (*Confucian Analects*)

Adjectives can also be considered as a subcategory of verbs. Indeed, they are intransitive verbs of quality, being negated by the adverb *bù* 不:

(34) 名不正则言不顺(《论语》)
ming bu zheng ze yan bu shun
name NEG correct then word not justified
'If names are not correct, then words cannot be justified' (*Confucian Analects*)

Like intransitive verbs, adjectives can also have a causative use:

(35) 王请大之(《孟子》)
Wang qing da zhi
King beg great it
'Your Majesty, [I] beg [you] to make it great' (*Mencius*)

In addition, adjectives are also typically found as noun phrase modifiers, as in *bái mǎ* 白马 'white horse', or as verb phrase modifiers, for instance *jí zǒu* 急走 (lit. rapid-run) 'run rapidly'.

Finally, one can consider that numerals constitute a subclass of the category of adjectives. They do indeed behave syntactically like adjectives; thus, they can form predicates and are negated by the adverb *bù*. Most commonly, however, they function as modifiers of nouns:

(36) 年已七十矣(《孟子》)
nian yi qi shi yi
age already seven ten FIN-PART
'[He] is already 70 years old' (*Mencius*)

(37) 吾何爱一牛?(《孟子》)
wu he ai yi niu
I why begrudge one ox
'Why (should) I begrudge one ox?' (*Mencius*)

4.3.2. Empty words

Pronouns are usually analysed as empty words. Several types of pronouns can be identified: personal, demonstrative, interrogative, and indefinite. Personal pronouns characteristically occur in different forms. The most common ones, with no distinction being made between singular and plural are: (i) for the first person, *wú* 吾, *wǒ* 我, and *yú* 余(予); (ii) for the second person, *rú* 汝, *ruò* 若, and *nǎi* 乃; (iii) for the third person, *zhī* 之, *jué* 厥 and *qí* 其. Several scholars have tried to characterize their different usages according to case (nominative, accusative, or genitive), but dialectal variation also is a factor: in most instances, the different usages of the pronouns depend on the different texts in which they occur.

The most common demonstrative pronouns are: (i) *shì* 是, *cǐ* 此, *sī* 斯, *zhī* 之, and *zī* 兹 'this, these, here'; and (ii) *bǐ* 彼, *fū* 夫, and *qí* 其 'that, those, there'. Here too, it is difficult to explain formal differences without considering dialectal variation. All of these demonstratives can be used as adjectivals (modifying the following nouns or noun phrases), or as subjects or objects.

The interrogative pronouns are divided into two categories: (i) those that replace subjects or objects (which are usually nouns), *shéi* 谁 'who', *shú* 孰 'which, who', *hé* 何 'what'; and (ii) those that replace predicative verbs or adverbs, *hú* 胡 'why, how', *xī* 奚 'why', *hé* 曷 'how, why', *ān* 安 'where, how', *yān* 焉 'how, where'. See Peyraube and Wu (2005).

The class of indefinite pronouns includes *huò* 或 'some, someone, something', *mò* 莫 'none, no one, nothing', and *mǒu* 某 'some, a certain one'.

Adverbs are also considered as empty words. Usually positioned in preverbal position, they typically modify the predicate of the sentence. One can distinguish several types: (i) adverbs of degree (*jí* 极 'extremely', *zuì* 最 'most', *shǎo* 少 'little', *shén* 甚 'very', etc.); (ii) adverbs of quantification and restriction (*jiē* 皆 'all', *gě* 各 'each', *dú* 独 'only', etc.); (iii) adverbs of time or aspect (*yǐ* 已 'already', *cháng* 尝 'once', *jiāng* 将 'be going to', *nǎi* 乃 'then', *fāng* 方 'just', etc.); (iv) adverbs of negation (*bù*, *fú*, *fēi*, *wú*, *wèi* 未, etc.). See above for the processes of fusion for some of these adverbs of negation.

Chinese prepositions are all verbal in origin (i.e. they arise from verbs through a process of grammaticalization). There are two commonly occurring prepositions in Classical Chinese: (i) *yú* 于/於 'at, to, in, from, towards, than, by', etc.; and (ii) *yǐ* 以 'with, by means of, in order to, because', etc. The first of these, *yú*, can be used to form locative, ablative, dative, comparative, or passive constructions; the second, *yǐ*, primarily instrumental, also expresses purpose and several other grammatical relationships. An important characteristic of *yǐ* is that it can also introduce the direct object of a double-object construction (32). Additional prepositions are *yòng* 用 'with', *wéi* 为 'for, on behalf of, for the sake of, because', *yǔ* 与 'with', *zì* 自 'from', among still others. Generally, simple juxtaposition is sufficient to coordinate nouns or noun phrases, as in *fù mǔ* 父母 (lit. father mother) 'father and mother'; or verbs and verb phrases in serial verb constructions. However, some coordinative conjunctions also occur, such as *jí* 及 and *yǔ* 与 'and' for coordinating noun phrases, or *ér* 而 'and' and *qiě* 且 'and, moreover' for coordinating verb phrases or clauses. Liu and Peyraube (1994) have argued that the conjunctions *jí* and *yǔ* did not directly develop from verbs, but from prepositions, which were themselves derived from verbs. In other words, two processes of grammaticalization have occurred sequentially: verb > preposition > conjunction (Chinese conjunctions are thus more grammaticalized than prepositions).

Finally, the category of particles (*zhùcí* 助词) is also obviously included in empty words. This category is usually divided into structural particles (*zhī* 之, *suǒ* 所, *zhě* 者) and modal particles that can occupy the initial, the medial, and, in most cases, the final position of

a sentence. These final particles can be divided according to the sentence types in which they occur – declarative, interrogative, exclamatory, and so on. In declarative sentences, one often finds *yǐ* 矣 (a particle of the perfect aspect), *yě* 也 (transforming a statement into an assertion, a judgment), *ěr* 耳, and *yān* 焉. *Hū* 乎, *yú* 欤, *yě* 也, and sometimes *zhě* 者, are more typically used in interrogative sentences. *Zāi* 哉 occurs in exclamatory sentences.

5. Lexicon

The overall lexicon of Ancient Chinese is quite different from that of Contemporary Chinese. The former is composed of: (i) words that are still attested in the contemporary language, like *shān* 山 'mountain' or *shuǐ* 水 'water'; (ii) words that only exist in the ancient language and have disappeared from the modern language, such as *yuē* 曰 'say'; (iii) words that are still used today, but with different meanings, like *zǒu* 走 'run' (Ancient Chinese) > 'walk' (Contemporary Chinese). Of the three types, the second are rare and the last are numerous (see He and Jiang 1980: 3).

5.1. Historical development of the lexicon

The Ancient Chinese lexicon has changed considerably since the Pre-Archaic period. From the vocabulary of everyday life (lexemes for food, clothing, housing), we find only 15 words in the oracle bone inscriptions (fourteenth–eleventh centuries BCE) that are still used today, 71 in the bronze inscriptions (tenth–sixth centuries BCE), and 297 in *Shuō wén jiě zì* 说文解字 (Explain the figures [single characters] and interpret the characters [compound characters], second century CE). On the varying richness of Chinese vocabulary in different periods, see He and Jiang (1980: 9).

According to these two authors (1980: 136–7), Classical Chinese has an identifiable basic vocabulary of about 2,000 full words, of which 1,100 occur quite commonly. From four major works of the Late Archaic period, including *Lún yǔ* 《论语》 (*Confucian Analects*) and *Mèngzǐ* 《孟子》 (*Mencius*), they have isolated 4,466 distinct words, estimating that about half are proper nouns such as personal names or place names. There is no implication that the vocabulary of Classical Chinese is impoverished compared with that of Contemporary Chinese – simply different. For example, there is only a single verb meaning 'to wash' in Contemporary Chinese (*xǐ* 洗), whereas there are five in Classical Chinese: *mù* 沐 'to wash (the hair)', *yù* 浴 'to wash (the body)'; *hui* 沫 'to wash (the face)'; *zǎo* 澡 'to wash (the hands)'; *xǐ* 洗 'to wash (the feet)'.

During the long history of Ancient Chinese, several different processes have led to changes in the lexicon. The major processes of internal development include: (i) compounding, a highly productive process beginning in the first century CE (Han period); (ii) semantic extension (e.g. *zú* 卒 'foot soldier' > *zú* 卒 'all sorts of soldiers'); and (iii) semantic reduction (e.g. *zǐ* 子 'child' (boy or girl) > *zǐ* 子 'son'). In addition, the Ancient Chinese lexicon was enlarged by borrowing words from other languages.

5.2. Inherited elements and loanwords

There has been a strong tendency in the past to view the Ancient Chinese lexicon as a monolithic linguistic entity, resistant to influences from all surrounding foreign languages. This is certainly a fallacy. Without going as far as Norman (1988: 17), who states: 'the fact

that only a relatively few Chinese words have been shown to be Sino-Tibetan may indicate that a considerable proportion of the Chinese lexicon is of foreign origin', we can undoubtledly rightly assert that the Ancient Chinese lexicon contains numerous loanwords. Nevertheless, the identification of such words and their sources is often uncertain. Below we mention a few uncontroversial examples of loanwords.

There are two common words for 'dog' in Ancient Chinese: *quǎn* 犬, which is probably the native Chinese word, and *gǒu* 狗, which appears at the end of the Warring States period. *Gǒu* 狗 is a loanword from a language ancestral to the Modern Miao-Yao languages (Hmong-Mien). The word 虎 *hǔ* 'tiger' might have been borrowed from an Austronesian language in prehistoric times (see Norman 1988: 17–20). Other words of non-Chinese origin are *xiàng* 象 'elephant' (borrowed from a Tai language?), *pútáo* 葡萄 'grape' (from Old-Iranian?), *mòlì* 茉莉 'jasmine' (from Sanskrit), *luòtuó* 骆驼 'camel' (possibly from an Altaic language).

6. Conclusion

The above chapter on syntax and lexicon is concerned chiefly with the Classical Language, as represented by the Warring States period, i.e. the Late Archaic period (fifth–second centuries BCE). From the time of the Early Medieval period (second–sixth centuries CE), one can consider that the vernacular language is actually distinct from the literary, deserving a separate description of its own. For example, certain important grammatical structures which did not exist in Classical Chinese are found to have developed in this later vernacular form, prior to the sixth century CE. For a detailed review of the developments in the language between these two stages, on the birth of the disposal construction (NP1-Agent + OM [object marker: *qǔ* 取, *chí* 持, *zhuō* 捉, *jiāng* 将, *bǎ* 把] + NP2-patient + VP), on the locative PPs, on the resultative construction, or on the birth and development of the classifier system, etc., see Peyraube (1996).

Acknowledgements

This research is part of a project entitled 'Typologie des processus synchronique et diachronique en Min-Sud (langue sinitique)' funded by the French National Research Agency's (ANR) 'blue sky' *programme blanc*: TYSOMIN n° 11-ISH2-001-01, 2012–14, projet bilatéral France (ANR) et Taiwan (National Research Council), 2013–15. I would also like to thank Hilary Chappell for her comments and questions.

List of abbreviations

1SG = 1st singular personal pronoun
3SG = 3rd singular personal pronoun
3PL = 3rd plural personal pronoun
FIN-PART = final particle
INT-PART = interrogative particle
LIG = ligature
NEG = negation
OM = object marker
PASSIVE = passive marker

References

Baxter, William H. (1992) *A Handbook of Old Chinese Phonology*, Berlin: Mouton de Gruyter.

Baxter, William H. and Laurent Sagart (1998) 'Word Formation in Old Chinese', in Jerome Packard (ed) *New Approaches to Chinese Word Formation: Morphology, Phonology and the Lexicon in Modern and Ancient Chinese*, Berlin: Mouton de Gruyter, 35–76.

Chappell, Hilary and Alain Peyraube (forthcoming) 'Mood and Modality in Sinitic Languages', in J. van der Auwera and Jan Nuyts (eds) *The Oxford Handbook of Mood and Modality*, Oxford: Oxford University Press.

Comrie, Bernard (1989) *Language Universals and Linguistic Typology* (2nd edition), Oxford: Basic Blackwell.

Djamouri, Redouane (2001) 'Markers of Predication in Shàng Bone Inscriptions', in Hilary Chappell (ed) *Sinitic Grammar – Synchronic and Diachronic Perspectives*, Oxford: Oxford University Press, 143–71.

Downer, Gordon (1959) 'Derivation by Tone-change in Classical Chinese', *Bulletin of the School of Oriental and African Studies* 22(2): 258–90.

Guan, Xiechu (1981) *Xi-Zhou jinwen yufa yanjiu*《西周金文语法研究》(*Research on the grammar of the bronzes of the Western Zhou*), Beijing: Shangwu yinshuguan.

Harbsmeier, Christoph (1981) *Aspect of Classical Chinese Syntax*, London: Curzon Press.

He, Jiuying and Jiang Shaoyu (1980) *Gu hanyu cihui jianghua*《古汉语词汇讲话》(*Talks on the lexicon of Ancient Chinese*), Beijing: Beijing chubanshe.

LaPolla, Randy (1993) 'On the Change to Verb-Medial Word Order in Proto-Chinese: Evidence from Tibeto-Burman', Paper presented at the 26th International Conference on Sino-Tibetan Languages and Linguistics, Osaka.

Li, Charles N. and Sandra A. Thompson (1974) 'An Explanation of Word Order Change: SVO > SOV', *Foundations of Language* 12(2): 201–14.

Liu, Jian and Alain Peyraube (1994) 'History of Some Coordinative Conjunctions in Chinese', *Journal of Chinese Linguistics* 22(2): 179–201.

Mei, Tsu-lin (1980) 'Sisheng bieyi zhong de shijian cengci'〈四声别义中的时间层次〉('Chronological Strata in Derivation by Tone-change'), *Zhongguo yuwen*《中国语文》(*Studies of the Chinese Language*) 6: 427–43.

Mei, Tsu-lin (1989) 'The Causative and Denominative Functions of the **s-* Prefix in Old Chinese', in *Proceedings of the Second International Conference on Sinology*, Taipei: Academia Sinica, 31–51.

Mei, Tsu-lin (1997) 'Typological Changes in the History of the Chinese language', *Chinese Languages and Linguistics* 4: 85–104.

Norman, Jerry (1988) *Chinese*, Cambridge: Cambridge University Press.

Peyraube, Alain (1987) 'The Double-object Construction in *Lunyu* and *Mengzi*', in *Wang Li Memorial Volumes,* Hong Kong: Joint Publishing, 331–58.

Peyraube, Alain (1989) 'History of the Passive Construction in Chinese until the 10th Century', *Journal of Chinese Linguistics* 17(2): 335–72.

Peyraube, Alain (1996) 'Recent Issues in Chinese Historical Syntax', in James C. T. Huang and Audrey Y. Li (eds) *New Horizons in Chinese Linguistics*, Dordrecht: Kluwer Academic Press, 161–213.

Peyraube Alain (1997a) 'On Word Order and Word Order Change in Pre-Archaic Chinese', *Chinese Languages and Linguistics* IV: 105–24.

Peyraube Alain (1997b) 'On Word Order in Archaic Chinese', *Cahiers de Linguistique – Asie Orientale* 26(1): 3–20.

Peyraube, Alain (2004) 'Ancient Chinese', in Roger D. Woodard (ed) *The Cambridge Encyclopedia of the World's Ancient Languages*, Cambridge: Cambridge University Press, Chapter 41, 988–1,014.

Peyraube, Alain and Shen Zhongwei (forthcoming) *History of the Chinese Language*, Cambridge: Cambridge University Press.

Peyraube, Alain and Thekla Wiebusch (1995) 'Problems Relating to the History of Different Copulas in Ancient Chinese', in Matthew Y. Chen and Ovid J. L. Tzeng (eds) *Linguistic Essays in Honor of William S.-Y. Wang*, Taipei: Pyramid Press, 383–404.

Peyraube, Alain and Wu Fuxiang (2005) 'Origin and Evolution of Question-words in Archaic Chinese: A Cognitive Approach', *Cahiers de linguistique – Asie Orientale* 34(1): 3–24.

Sagart, Laurent (1993) 'L'infixe *–r–* en chinois archaique', *Bulletin de la Société Linguistique de Paris* 88(1): 261–93.

Sagart, Laurent (1999) *The Roots of Old Chinese*, Amsterdam and Philadelphia: John Benjamins Pubishing Company.

Shen, Pei (1992) *Jiagu buci yuxu yanjiu*《甲骨卜辞语序研究》(*Word order in the Oracle Bone Inscriptions*), Taipei: Wenjin chubanshe.

Wang, Li (1979) *Gudai hanyu changshi*《古代汉语常识》(*Elements of Ancient Chinese*), Beijing: Renmin jiaoyu chubanshe.

2
CANTONESE

Cheng Siu-Pong and Tang Sze-Wing

THE CHINESE UNIVERSITY OF HONG KONG, HONG KONG, CHINA

The old language–dialect debate is of particular relevance to Cantonese, or Yue, given the ambiguous status it holds in places like Hong Kong. This leads to an inquiry into its distinctive features. They might be connected to its origin, which relates to its divergence from other Chinese varieties. Whereas the subclassification of Yue reflects the intra-group variation, the comparison between Cantonese and Mandarin reveals not only their differences in sounds and words, but also in the lesser-known aspects of grammar.

1. Cantonese: dialect or language?

In English, the term 'Cantonese' either refers to the speech of Canton, modern-day Guangzhou city of Guangdong province, or the entire Yue group of similar tongues, or just any variety within the group. It is often used as a convenient term for the mainstream Cantonese spoken in Hong Kong. The native speakers of Cantonese (in the broad and narrow sense) are overwhelmingly of Han Chinese ethnicity.[1] Depending on the communities, they refer to what they speak as *gwongzauwaa* / *gwongfuwaa* 'Guangzhou speech', *gwongdungwaa* 'Guangdong speech', *jyutjyu* 'Yue language', *baakwaa* 'colloquial speech', *tongwaa* 'Tang (dynasty) speech', or simply as *zungman* 'Chinese language'.[2] Like any other varieties of Chinese, the definition of Cantonese is tied to age-old disputes over whether it should be classified as a language or a dialect. Linguistically, it is not at all difficult to define 'dialect': a dialect is a subdivision of language. At issue is how the relationship between such terms as 'Chinese' and 'Cantonese' is incorporated into these notions. If 'language' is defined as a set of *mutually intelligible* varieties, many Chinese varieties, Cantonese included, can be conceived as separate languages. There is a very low degree of mutual intelligibility between Mandarin and Cantonese. In fact, the major 'dialect groups' of Chinese are all unintelligible to one another. This interpretation of 'language' implies that the term 'Chinese', when encompassing Cantonese (as well as Mandarin, etc.), should be treated as a language family rather than a single language. Thus, Cantonese is said to be a member of the Chinese or Sinitic languages. In most cases where Chinese is so defined, Cantonese is also deemed a language. Nonetheless, there is a need for this Cantonese language to exclude a subset of in-group varieties. This is because the group of Yue varieties contains a number of subdivisions whose mutual intelligibility is at best marginal. In this regard, the kind of Cantonese spoken by the great

majority of Hong Kong people is a single language. The mainstream Cantonese as used in Guangzhou, Hong Kong, and Macao can also be grouped as one language, given their close similarities, but outside this subgroup which they form (to be discussed below), the inclusion of other varieties depends on one's judgment on how similar they should be in order to be considered 'intelligible'.

The status of Cantonese also relates to the standard remark that 'a language is a dialect with an army and a navy', a quote attributed to Max Weinreich. Spoken in China, which now includes the two subnational entities of Hong Kong and Macao, Cantonese is also widely spoken by the Chinese diaspora in various parts of southeast Asia (Malaysia in particular) and the Western world (notably in Vancouver, Canada). Cantonese is not named as a national language in any of these countries, and no national or subnational governments explicate in their constitutional document that 'Cantonese' (the exact wording) serves as an official language. While the constitution of the People's Republic of China does not touch on the issue of national or official languages, the *Law on the Standard Spoken and Written Chinese Language*, effective 2001, states that 'the standard spoken and written Chinese language means Putonghua (a common speech with pronunciation based on the Beijing dialect) [. . .]'. Although Hong Kong is part of China, this law, together with the vast majority of other national laws, does not apply in the territory. Instead, the constitutional *Basic Law* states that '[i]n addition to the Chinese language, English may also be used as an official language by the executive authorities, legislature and judiciary of the Hong Kong Special Administrative Region'. A similar article is found in Macao's *Basic Law*, with 'Portuguese' stated in place of 'English'. The ambiguity of using 'Chinese' in the *Basic Law*s as well as other local laws (e.g. Hong Kong's *Official Languages Ordinance*) raises questions over whether Cantonese *per se* is official. There is little doubt that in Hong Kong and Macao, Cantonese is the *de facto* official spoken variety of Chinese. Cantonese is spoken in courts and legislative council meetings almost whenever 'Chinese' is used there. Even so, the answer to the question of whether Cantonese by itself attains *de jure* status is not clear. There is one complicating factor. In formal contexts, government and business documents published in Chinese are given in standard written Chinese, which is based on Mandarin and comprehensible to all literate Chinese speakers worldwide. Written Cantonese, i.e. the written form of Cantonese expressed in Chinese characters, is mostly used in informal register. At first glance, this helps prove the subordinate status of Cantonese. However, since the Chinese script is a logographic writing system, the characters for standard written Chinese can be read with their Cantonese pronunciations. It is easier to come up with the view that it is rather the literary counterpart of the colloquial speech (known as *shumianyu* 'literary language'), and its use does not undermine the status of Cantonese. In sociolinguistic terms, Cantonese is obviously an *abstand language* with respect to another Chinese variety, e.g. Mandarin; the variation between Cantonese and Mandarin in sounds and vocabulary is so great that they are on par with such European languages as French and Spanish. At issue is to what degree Cantonese is 'autonomous' enough to be an *ausbau language*. In mainland China, the answer is simple: Cantonese is considered by the Chinese government as a 'local dialect'. In the best scenario, the official policies tolerate (but never promote) the use of Cantonese in broadcasting and, to a lesser degree, in schools. In Hong Kong and Macao, Cantonese is used in a variety of domains, ranging from educational to official settings. Nonetheless, given the unresolved issues surrounding the meaning of 'Chinese', the Cantonese variety spoken there should be viewed as more an abstand than an ausbau language. The Weinreichian notion of 'language' does not accord well with the reality of Cantonese regardless of where it is spoken. A final remark is on the term 'dialect' itself. While most Chinese-language publications indicate that

Cantonese is a dialect, the original term they use in Chinese is *fangyan*, which literally means 'regional speech'. It is construed as a notion relative to (and often subordinate to) *gongtongyu* 'common language'. The possible semantic discrepancies between 'dialect' and *fangyan* had led DeFrancis (1984) to coin and adopt the term *regionlect* to capture the sense of the original word.[3] This is another way to get over the dialect–language dichotomy. In the following sections, this dichotomy is set aside, and Yue and Cantonese are variably called 'dialect' or 'language' in a non-committal manner.

2. Origin and distribution

Cantonese, or Yue, belongs to the Sino-Tibetan language family. While all natural languages are descended from some *proto-language*, the discussions of the origin here focus on the emergence of Yue dialects in present-day southeast China where they are chiefly spoken. A vast area of southern China was known to be occupied by the *Bai Yue* 'One Hundred Yue' tribes. There is still no consensus as to which languages these people spoke in the absence of written records. They were speakers of Austroasiatic-, Kam-Tai-, Hmong-Mien-, or possibly Austronesian-related languages (Norman and Mei 1976; Blust 1984–1985; J. Li 1990; LaPolla 2001), or there was an assortment of speakers from some or all of these linguistic groups. Historical records show that Han Chinese migrated from their homeland in the north to this area starting from the third century BCE. While most scholars agree that Modern Cantonese is descended from an earlier form of Chinese, it is still a matter of great debate as to what extent the Chinese variety first spoken in this area bears a relation to Modern Cantonese. The divergence of the earliest Yue variety from those of other parts of China is argued to have taken place from as early as the third century BCE when substantial numbers of Han Chinese arrived (J. Li 1990) to as late as the thirteenth century CE (Lau 2000). In those accounts that posit a later divergence time, it is argued that another wave of Han Chinese migration was vital to its formation. Historical studies are also concerned about the influence of non-Han languages on the early Yue. Evidence based on some colloquial words has been provided to show that they form Tai or Hmong substrata (see Yue-Hashimoto 1976, 1991; Bauer 1987; J. Li 1990, etc.). The substratum influence of non-Han languages can be expressed in more neutral terms: Yue has experienced a prolonged period of language contact vis-à-vis non-Han languages.

In modern times, the main area for the Yue dialect group is the central and western parts of Guangdong province and the southeastern part of Guangxi Zhuang Autonomous Region. This area also includes Hong Kong and Macao, which are now the Special Administrative Regions of the People's Republic of China. The Yue dialects are spoken by around 4 to 5% of China's 1.3 billion people. *Language Atlas of China* (2012) gives a figure of approximately 59.58 million speakers. In southeast Asia, there are considerable populations of Yue speakers in Malaysia, Vietnam, Singapore, and Indonesia. In particular, around 5% of Malaysia's population are from the Cantonese speech group (Tan 2005), and an even larger proportion of Malaysians can speak Cantonese with varying degrees of fluency. Outside Asia, the Yue speakers amount to around 1% or more in Canada, the United States, Australia, and New Zealand. According to *Language Atlas* (1987), there were approximately 11 million overseas speakers in total during the time surveyed. *Language Atlas* (2012) states that there are over 68 million Yue speakers worldwide. Major works on Yue or Chinese dialects in general, such as Zhan (1991, 2002), Yuan et al. (1960/2001), X. Li (1994), and *Language Atlas* (1987, 2012), differ on the subclassification of Yue dialects. The following classification includes the four major Yue branches of: (i) Yuehai; (ii) Siyi; (iii) Gao-Yang; and (iv) Guinan; this

classification is 'eclectic' in nature given that these four branches are not subordinate to any others in the above works, while other branches, e.g. Goulou Branch, are subsumed elsewhere into one of them.[4]

2.1. Yuehai Branch

Also called the Guangfu Branch, it is mainly spoken in the Pearl River (Zhujiang) Delta. This area includes Guangzhou, the capital of Guangdong province, together with Hong Kong and Macao. For this reason, this branch is named after Guangzhou, otherwise known as *Yuehai* or *Guangfu*. Yuehai Branch is the most well-known and widely spoken branch of Yue dialects. *Language Atlas* (2012) gives a figure of about 28.34 million speakers (excluding the speakers of the Goulou Branch; to be discussed below). The same source indicates that approximately 20.72 million of them live in Guangdong. In addition, around 7 million people from Hong Kong and Macao can speak this variety. Depending on the extension of this subgroup, its area also covers the central and northern parts of Guangdong, and scattered parts of Guangxi such as downtown Hezhou and Wuzhou. As a result of migration from this core area, the Yuehai Branch is the predominant form of Yue spoken in overseas Chinese communities. In Malaysia, it is a major Chinese variety spoken in Kuala Lumpur, Ipoh, Kuantan, Seremban, and Sandakan. In most Chinatowns in North America, the Yuehai Branch has been the major Chinese variety (amid a growing number of Mandarin-speaking immigrants), often achieving its prominent status at the expense of the Siyi Branch. Particularly in Metro Vancouver, Canada, its use has been scaled up to the city level. The city of Richmond is an example *par excellence*.

Accentual difference is found between the varieties spoken in downtown Guangzhou and the outskirt areas of the city, especially those districts newly (re-)attached to Guangzhou around this millennium, e.g. Panyu. A variety closely similar to the one used in downtown Guangzhou is also predominantly spoken in Hong Kong. Guangzhou and Hong Kong Cantonese are mutually intelligible to a very high degree, despite some lexical differences. However, it should be noted that the emergence and rise of Hong Kong Cantonese is resulted from the combined factors of immigration and language shift. Right before the beginning of the colonial time in the nineteenth century, a different Yue dialect, i.e. the Guan-Bao variety, was chiefly spoken. A gradual shift to Guangzhou-style Cantonese also took place in Macao. The rise of Cantonese to greater prominence in the latter half of the last century was mainly due to the social and economic development of Hong Kong. The extended influence of Hong Kong's trade and entertainment industries has promoted the use of Cantonese in a greater area of Guangdong and overseas Chinese communities. As a result, many people in traditionally non-Cantonese-speaking regions, such as the eastern part of Guangdong, came to understand it and often speak it as a second language.

Guan-Bao Branch

Zhan (1991) classifies it as a separate branch of Yue dialects. It is traditionally spoken in the eastern part of the Pearl River Delta. Precisely, it is the indigenous Yue variety of Hong Kong and Shenzhen, formerly parts of Xin'an (later named Bao'an), and of the neighboring Dongguan, hence the name 'Guan-Bao'. In Hong Kong, it is known as *Punti* or *waitauwaa*. The number of its speakers is diminishing to the point that it has become endangered, and its use is now restricted to the older generation in the more rural parts of Hong Kong. A decline in use is also observed in the rapidly developing Shenzhen and Dongguan. The

mutual intelligibility between Guan-Bao and Guangzhou Cantonese is less than ideal. Although downtown Guangzhou and Hong Kong are about 130 kilometers apart, the mainstream Cantonese in Hong Kong is much closer to Guangzhou Cantonese than it is to Guan-Bao.

Xiangshan Branch

Another independent branch of Yue given in Zhan (1991), it is spoken in the western part of the Pearl River Delta, including Zhuhai and Zhongshan, which had previously been called Xiangshan before it was named after Sun Yat-sen (Sun Zhongshan). Formerly, the Xiangshan Branch was also spoken in Macao.

Goulou Branch

First proposed around the time when the first edition of *Language Atlas* (1987) was published, it is spoken in the peripheral areas outside the Pearl River Delta, to the north and west of the heartland of Yuehai Branch. This area extends into Guangxi, and occupies the eastern portion of the region. It is named after the mountain range of Goulou in southeastern Guangxi bordering Guangdong. According to *Language Atlas* (2012), the number of Goulou speakers amounts to approximately 10.30 million.

2.2. Siyi Branch

Siyi 'four counties' refers to the region southwest of Guangzhou that lies to the west of the Pearl River Delta. The 'four counties' include Taishan, Kaiping, Enping, and Xinhui. In modern times, they are all under the jurisdiction of the city of Jiangmen, which also includes Heshan, so together they are called *Wuyi* 'five counties'. The population of Siyi speakers in Guangdong is approximately 3.88 million, according to *Language Atlas* (2012). The main Siyi-speaking area includes modern-day Jiangmen and the western part of Zhuhai. Migration of Siyi speakers out of this region contributed to its prominence elsewhere. In Hong Kong after the Second World War and the subsequent Chinese Civil War, the influx of Siyi immigrants led to a time when the Siyi Branch was more spoken before it was largely displaced by Hong Kong Cantonese a few decades later. Besides, the Siyi Branch was once the most influential Chinese variety in the American Chinese communities. Up until half a century ago, it was spoken by the great majority of Chinese immigrants to the United States (Bauer and Benedict 1997). Within the Siyi Branch, the Taishan subdialect was the most prominent variety spoken overseas. Although its status was largely replaced by the Yuehai Branch and Mandarin, the use of Siyi remains significant in some American Chinatowns, especially among the earlier migrants.

2.3. Gao-Yang Branch

Also called the Gao-Lei Branch, it is the predominant Chinese variety spoken in the southwestern part of Guangdong, which includes Gaozhou and Yangjiang. The core areas of Gao-Yang, however, exclude the Leizhou Peninsula, together with some linguistic exclaves nearby, where the Gao-Yang Branch might be spoken but Hakka or Min dialects are the main speech. *Language Atlas* (2012) states that the Gao-Yang Branch is spoken by approximately 6.43 million people. The Wu-Hua Branch, considered a separate branch in *Language Atlas*, has about 1.28 million speakers. It is classified as part of the Gao-Yang-speaking region in the other accounts.

2.4. Guinan Branch

Guinan literally means 'southern Guangxi'. The first edition of *The Encyclopedia of China* uses it as an umbrella term to cover all subdivisions of Yue dialects spoken in Guangxi. According to *Language Atlas* (2012), approximately 16.86 million Guangxi people speak the Yue dialects. The Yue varieties in Guangxi are further divided into the branches of: (i) Yuehai; (ii) Goulou; (iii) Yong-Xun; and (iv) Qin-Lian. The Goulou Branch is spoken in a continuous area across Guangdong and Guangxi, and the predominant use of Yuehai Branch in Guangxi is found in the linguistic enclaves, or 'dialect islands'. Only the Yong-Xun and Qin-Lian Branches are specific to Guangxi.

Yong-Xun Branch

It is used as the majority variety in scattered places along the rivers of Yongjiang and Xunjiang. Most of these places have served as inland river ports. *Language Atlas* (2012) gives a figure of approximately 4.69 million speakers.

Qin-Lian Branch

It is spoken in the traditional areas of Qinzhou and Lianzhou (modern-day Hepu) along the southern (and the only) coast of Guangxi. Situated to the west of Guangdong coastal towns, this region was part of Guangdong until after the founding of the People's Republic of China. According to *Language Atlas* (2012), the number of speakers of this branch amounts to approximately 3.90 million.

3. Differences between Cantonese and Mandarin

It has been well documented that Cantonese differs from Mandarin in terms of sounds and vocabulary. In general, the sound system of Yue dialects is comparatively more complicated by having more rhymes and tones. Bauer and Benedict (1997) identify 58 rhymes in Cantonese including two syllabic nasals. By contrast, a standard Mandarin rhyme table includes 35 rhymes. A major reason for the difference lies in the fact that in Cantonese /p, t, k, m/ are found in the syllable-final position, yet these codas are absent in Mandarin. Regardless of tones, a regular Mandarin dictionary contains a number of 410 syllables, which is less than the 584 Cantonese syllables given in Bauer and Benedict (1997) for the standard Chinese characters, not to mention others solely used for colloquial speech. The number of tonal categories in Yue dialects varies from six to ten (Zhan 1991). As for Hong Kong Cantonese (as well as other Yue varieties), the category of Entering tones (tones for syllables with a coda from /p, t, k/) might be subsumed into the other categories, depending on the categorization criteria, and the Entering tones are viewed as the short or extra-short variants of the other tones (Zee 1999). Even so, it contains a total of six tones. This contrasts with Mandarin, which has only four. Historically, tonal split and merger have given rise to varying numbers of tones across different Chinese varieties. As in other modern dialects, Yue experienced the split of tonal categories primarily as a result of the loss of voicing distinction of initials. Subsequent tone mergers, however, were more restrained in the case of Yue dialects. In modern-day Hong Kong Cantonese, there are 19 consonants, including plosives /p, p^h, t, t^h, k, k^h, k^w, k^{wh}/, nasals /m, n, ŋ/, affricates /ts, ts^h/, fricatives /f, s, h/, and 'semi-vowels' as well as other approximants /j, w, l/. Vowels include both front and back,

and high and low monothongs: /iː, e, yː, uː, o, ɛː, ɔː, œː, ɵ, ɐ, aː/. In addition to these monothongs, there are also ten gliding vowels (or diphthongs): /iw, ɵɥ, uɥ, ej, ɔj, ow, ɐj, ɐw, aːj, aːw/. The maximal syllable takes the form of 'initial consonant + vowel + final consonant + tone', or C_iVC_f/T (Bauer and Benedict 1997). In other words, a *consonant cluster* is not found within an indigenous syllable. Both the initial and final consonants might be absent, and a final consonant (from /m, n, ŋ, p, t, k/) is disallowed when the vowel is a gliding one. A syllable always contains a vowel, with the exception of two *syllabic* nasals: /m, ŋ/. Assuming that the three Entering tones, namely Upper Dark Entering /5/, Lower Dark Entering /3/, and Light Entering /2/, are respectively subsumed into /55/, /33/, and /22/, the modern-day Hong Kong variety of Cantonese contains a total of six tones, which are /55/, /25/, /33/, /21/, /23/, and /22/.[5]

In terms of vocabulary, Hou (2002) indicates that about two thirds of basic words in Yue dialects are common to both Yue and Standard Chinese, or, in his wording, the 'common national language'. Out of the other one third or so of words that are different, over 90% of them were arisen as a result of the innovation of Yue speakers, whereas some others are old Chinese cognates that are no longer in use in modern Standard Chinese. Many scholars also note that some of the colloquial words owe their origin to non-Han languages including Tai or Hmong. The Yue dialects of major ports, notably Hong Kong Cantonese, stand out from other Chinese varieties by having more words borrowed from a European language, English in particular. Apart from English, Japanese is another major loan language. These words entered the Cantonese vocabulary in the form of semantic translation, phonetic borrowing, or in the case of most Japanese loanwords, the direct use of Japanese-derived words with the Chinese characters (*kanji*) read in Cantonese pronunciation. Structurally, the relative order of morphemes in a disyllabic word in Mandarin is sometimes reversed in Cantonese. Examples include *ganjiu* 'important' and *siuje* 'night-time snack'. In Mandarin, the morpheme order is reversed and they are rendered respectively as *yaojin* and *yexiao*. Another well-quoted example is *gaigung* 'rooster'. Although *gunggai* 'male + chicken' is also used, *gaigung* 'chicken + male' is indigenous. The morphemic ordering of this indigenous word in Yue and also in other Southern dialects has drawn speculation that the influence of non-Han languages was at work (see Yuan et al. 1960; Ouyang 1991, etc.). Other morphological variation includes the higher tendency for Cantonese, as compared to Mandarin, to have monosyllabic words, and the lower tendency for it to use certain suffixes, e.g. *zi*.

It is often assumed that all Chinese varieties share the same grammar. Despite the general perception of close similarities in grammar among them, there are significant differences that also set them apart. Functional words are important contributors to the syntactic variation. The following parts are based on Tang and Cheng (2014) and focus on the distinctive grammatical features of Cantonese that are less salient or cannot be found at all in Mandarin. The variables involved include (i) structural particles; (ii) directional verbs; (iii) aspect markers; (iv) definiteness; (v) double-object construction; (vi) comparative sentences; (vii) passives; and (viii) post-verbal elements.

3.1. Structural particles

Homophonous particles in Mandarin might be audibly differentiated when translated into Cantonese. Zhu (1961) classifies three types of *de* in Mandarin: de_1 follows an adverb to form an adverbial unit; de_2 follows an adjective to form an adjectival phrase; de_3 follows different phrases to form a nominal unit.

(1) Ta zhengtian buting-de shuo. *de*$_1$-adverbial
he all-day non-stop-DE speak
'He speaks non-stop all day'

(2) Zhe duo hua honghong-de. *de*$_2$-adjectival
this Cl flower red-DE
'This flower is a bit red'

(3) Xinxian-de ye you. *de*$_3$-nominal
fresh-DE also have
'There is also something fresh'

Zhu (1980) further suggests that the three subcategories of *de* correspond to *gam*, *dei*, and *ge* in Cantonese. The three sentences can be translated into Cantonese as follows.

(4) Keoi sengjat mting-gam gong. *gam*-adverbial
he all-day non-stop-GAM speak
'He speaks non-stop all day'

(5) Ni do faa hunghung-dei. *dei*-adjectival
this Cl flower red-DEI
'This flower is a bit red'

(6) Sansin-ge dou jau. *ge*-nominal
fresh-GE also have
'There is also something fresh'

In Cantonese, *gam*, *dei*, and *ge* are clearly separate morphemes as they differ in their pronunciation. With this in view, Zhu argues for the three subsets of *de* in Mandarin. As far as these structural particles are concerned, Cantonese is more 'marked' in their division of labor.

3.2. Directional verbs

Yuan et al. (1960) note that Mandarin tends to use *shang* 'go up' or *dao* 'arrive' as the main verbs, and *qu* 'go' and *lai* 'come' usually follow the object, as they do in (7) and (8). In Cantonese, on the other hand, *heoi* 'go' and *lai* 'come' are mostly used directly as the main verb, as in (9) and (10).

(7) Wo shang Beijing qu. (Mandarin)
I go-up Beijing go
'I go to Beijing'

(8) Wo jiejie shi zuotian dao Guangzhou lai de.
I sister be yesterday arrive Guangzhou come DE
'My elder sister came to Guangzhou yesterday'

(9) Ngo heoi Bakging. (Cantonese)
I go Beijing
'I go to Beijing'

(10) Ngo gaaze hai camjat lei Gwongzau ge.
I sister be yesterday come Guangzhou GE
'My elder sister came to Guangzhou yesterday'

According to Yuan et al., the use of *qu* 'go' and *lai* 'come' as the main verbs is also found in Mandarin, but this is obviously under the influence of the Southern dialects.[6]

3.3. Aspect markers

C. N. Li and Thompson (1989) identify four Mandarin aspect markers, *le*, *zai*, *zhe*, and *guo*: the perfective aspect is marked by *le*, the progressive aspect by *zai*, the durative aspect by *zhe*, and the experiential aspect by *guo*. Cheung (1972/2007) identifies six Cantonese aspect markers, in addition to a zero marker. They are perfective *zo*, experiential *gwo*, progressive *gan*, habitual *hoi*, durative *zyu*, and inchoative *hei(soeng)lai*. Matthews and Yip (1994/2011) and X. Li et al. (1995) added two more: delimitative *haa* and successive *lokheoi*. A linkage can be drawn between Mandarin and Cantonese perfective, experiential, and progressive / durative markers. Nonetheless, it is worth noting the distinction between Mandarin progressive / durative *zai* / *zhe* and Cantonese progressive / durative *gan* / *zyu*. Previous studies indicate that activity verbs take *zai* and stative verbs take *zhe* (Paris 1981; C. N. Li and Thompson 1989). Smith (1991/1997) also notes that in some Mandarin varieties the use of *zhe* has gradually become more extended and can convey both the progressive and durative aspect. In contrast, neither *gan* nor *zyu* has taken over the function of one another in Cantonese. While *gan* is a typical progressive maker, *zyu* has a static meaning and is used as a marker for durative aspect, as in (11) and (12).

(11) Zoengsaam daa-gan bo.
Zoengsaam play-Prog ball
'Zoengsaam is playing ball'

(12) Zoengsaam daai-zyu ngaangeng.
Zoengsaam wear-Dur eyeglasses
'Zoengsaam is wearing eyeglasses'

While Mandarin *zhe* can be used in the equivalents of both sentences above, Cantonese progressive *gan* and durative *zyu* are unambiguously distinctive and cannot be used interchangeably.

3.4. Definiteness

Like Mandarin classifiers, Cantonese classifiers precede nouns. In both varieties, the [classifier + noun] sequence follows a numeral, which itself might be preceded by a demonstrative, as in the following examples.

(13) zhe san ben shu (Mandarin)
this three Cl book
'these three books'

(14) ni saam bun syu (Cantonese)
this three Cl book
'these three books'

The addition of demonstratives and numerals is not compulsory. The bare [classifier + noun] form is possible in both Mandarin and Cantonese. As indicated by Cheng and Sybesma (1999), the difference between Cantonese and Mandarin in this nominal phrase is a contrast in the interpretations. In Cantonese, a nominal phrase of this form can either have a definite or indefinite reading. In Mandarin, only the indefinite reading is available. Therefore, the [classifier + noun] form is found in the subject position of Cantonese, but not Mandarin, given that this position normally limits the interpretation of nominals to a definite one.

(15) Wo xiang mai ben shu. (Mandarin)
I want buy Cl book
'I want to buy a book'

(16) *Ben shu hen hou.
*Cl book very thick

(17) Ngo soeng maai bun syu. (Cantonese)
I want buy Cl book
'I want to buy a book'

(18) Bun syu hou hau.
Cl book very thick
'The book is thick'

On the other hand, Mandarin bare nouns can receive definite, indefinite, or generic readings. Bare nouns in Cantonese can also have indefinite or generic readings, but the definite interpretation is largely prohibitive. For the definite reading, it is better for a bare noun to go with a classifier.[7]

(19) Gou yao guo malu. (Mandarin)
dog want cross road
'The dog wants to cross the road'

(20) Zek gau jiu gwo maalou. (Cantonese)
Cl dog want cross road
'The dog wants to cross the road'

3.5. Double-object construction

It has been noted that Cantonese has a double-object construction different from the one found in Mandarin (Yuan et al. 1960; Gao 1980). An animate object precedes an inanimate one in Mandarin, but in Cantonese the order appears to be in reverse.

(21) Ta gei-le wo yi-ben shu. (Mandarin)
he give-Perf I one-Cl book
'He gave me a book'

(22) Keoi bei-zo jat-bun syu ngo. (Cantonese)
he give-Perf one-Cl book I
'He gave a book to me'

In most cases, the inanimate and animate objects are respectively analyzed as the direct object (DO) and indirect object (IO). According to Cheung (1972), Cantonese places the DO before the IO most prevalently when the verb is *bei* 'give'. Note that *bei* can be used in two different positions in the same sentence. The same is true for Mandarin *gei*.

(23) Ta gei-le yi-ben shu gei wo. (Mandarin)
he give-Perf one-Cl book to I
'He gave a book to me'

(24) Keoi bei-zo jat-bun syu bei ngo. (Cantonese)
he give-Perf one-Cl book to I
'He gave a book to me'

While the first *gei* / *bei* 'give' is a verb, the second *gei* / *bei* 'to' is regarded as prepositional, as it often co-occurs with many other verbs to indicate IO.[8] Traditionally, the form of double-object construction as illustrated in (22) is known to have been 'inverted', since the order found in Mandarin is reversed. On the other hand, more recent studies (Tang 1998a, 2003; Liu 2001) show that the 'DO + IO' order in Cantonese actually originated from the 'DO + *bei* + IO' construction through the deletion of *bei* 'to'. That is to say, the original form of (22) is (24). Therefore, what distinguishes Mandarin and Cantonese is not the word order of objects, but the deletability of the preposition.[9]

3.6. Comparative sentences

Generally speaking, Cantonese and Mandarin use different comparison words, which results in different structures for comparative sentences.

(25) Zhangsan bi Lisi gao. (Mandarin)
Zhangsan compare Lisi tall
'Zhangsan is taller than Lisi'

(26) Zoengsaam gou-gwo Leisei. (Cantonese)
Zoengsaam tall-pass Leisei
'Zoengsaam is taller than Leisei'

In the Cantonese example, the comparison word is *gwo* 'pass', which follows the adjectival predicate *gou* 'tall'. The Cantonese comparative sentences are in the form of 'X + predicate + comparison word + Y'. In the Mandarin example, the comparison word *bi* 'compare' precedes both Y and the adjective; the comparative sentences in Mandarin are in the form of 'X + comparison word + Y + predicate'. This construction, however, is not impossible in Cantonese. As stated in Matthews and Yip (1994), there is an alternative maker in Cantonese, i.e. *bei* 'compare'. It is based on the *bi* construction in Mandarin, and 'has the formal flavor of Mandarin-based syntax', as shown in their example (1994: 167).

(27) Aawan bei keoi muimui leng.
Aawan compare she sister pretty
'Aawan is prettier than her younger sister'

It shows that the order of 'X + comparison word + Y + predicate' is also found in Cantonese, although the comparative construction with *gwo* 'pass' is arguably more common and 'authentically' Cantonese.[10]

3.7. Passives

Following the distinction drawn by Ting (1995, 1998), passive sentences in Mandarin are classified into what are known as 'long passives' and 'short passives'. A passive morpheme, *bei*, is used in both categories. Their difference is that an identified agent follows *bei* in the long passives, but such an agent is absent in the short ones.

(28) Zhangsan bei Lisi piping-le. (Long Passive)
Zhangsan BEI Lisi criticize-Perf
'Zhangsan was criticized by Lisi'

(29) Zhangsan bei piping-le. (Short Passive)
Zhangsan BEI criticize-Perf
'Zhangsan was criticized'

Long and short passives are not equally acceptable in Cantonese. Unlike its counterpart in Mandarin, the indigenous passive marker *bei2* (in high rising tone /25/) must contain an agent. Otherwise, the sentence will become ungrammatical.

(30) Zoengsaam bei2 *(ngo) naau.
Zoengsaam BEI I criticize
'Zoengsaam was scolded (by me)'

Cantonese passives without an identified agent will be acceptable only if the passive morpheme is given in a different form, *bei6*, with the low level tone /22/.

(31) Zoengsaam bei6 zizaak.
Zoengsaam BEI censure
'Zoengsaam was censured'

Passives with *bei6* are normally used in formal speech and in literary contexts. While *bei2* is the native form, *bei6* is assumed to be borrowed from Mandarin. As a result, its usage is highly restricted. In everyday Cantonese, the use of long passives, rather than the short ones, is predominant. Even if the agent needs to be 'hidden' for some reason, it is more natural to use a long passive containing an unspecified agent, such as *jan* 'person', than to use a short passive with *bei6*.[11]

3.8. Post-verbal elements

Cantonese is rich in post-verbal elements in comparison to Mandarin. There are a large number of suffixes and sentence-final particles, and also a high degree of co-occurrence of post-verbal elements. Cheung (1972) identifies 15 predicative suffixes. Matthews and Yip (1994) include seven major aspect particles and 29 verbal particles. The number of sentence-final particles is even higher. As noted by Fung (2000), previous studies show that it may range from 30 to as many as 90 (Kwok 1984; Ouyang 1990; Leung 2005). Furthermore,

multiple particles often co-occur in the sentence-final position, as in the following example (Leung 2005: 87).

(32) Nei waa zeoido to do go leng jyut tim ge zaa aamaa?
you say at-most delay more Cl more month TIM GE ZAA AAMAA
'You said (I can) delay for one or more months at most, right?'

As for their usage, Kwok (1984) connects them with sentence functions. The particles are categorized as the signals of speaker's attitude. Luke (1990) discusses three particles (*laa*, *lo*, and *wo*), which manifest different interactional goals. The particle *laa* is used to establish common grounds in conversation; *lo* is associated with interactional achievements; and *wo* is used to signal expectation and noteworthiness. Fung (2000) groups the majority of sentence-final particles into three main categories, i.e. Z-, L-, and G- families. Particles sharing the same initial belong to the same family, the distinctive core features of which are based on both their semantic and contextual meanings. Z-particles, including *ze*, *zek*, *zaa*, etc., have the core semantic feature of 'restriction'; the core feature of L-particles, which include *lei*, *laa*, etc., is to signal temporal and aspectual meanings; the core function of G-particles, including *ge* and *gaa*, is to mark a situation given in a context.[12]

Cantonese is rich in both pre-verbal and post-verbal elements, and several pre-verbal and post-verbal elements share similar semantic properties. Take *caa'mdo* 'almost' and *gamzai* 'almost' as examples. They can co-occur in the same sentence.

(33) Ngo caa'mdo gong-jyun gamzai.
I almost speak-finish almost
'I almost finished talking'

Tang (2006a, 2006b, 2007) suggests that post-verbal elements like *gamzai* 'almost' and pre-verbal ones like *caa'mdo* 'almost' form a constituent known as the 'discontinuous construction'. He argues that in such a construction, the combination of pre-verbal and post-verbal elements 'appears to be redundant'. That means the apparent redundancy of the pre-verbal and post-verbal elements should be understood as their being close in meaning. So far, over 40 different pre- and post-elements have been identified to belong to this kind of construction. In comparison, the discontinuous constructions are rare in Mandarin.[13]

4. Implications

Many scholars contend that Chinese varieties in essence only differ in sounds and vocabulary (Yuan et al. 1960; Chao 1968; Lü 1982). Since it is easier to observe the phonetic and lexical differences, many previous studies on Chinese dialects focus on these two aspects. Comparatively speaking, the outcomes of syntactic variation studies are less fruitful. On the surface, there are more similarities than differences in the grammar of different Chinese varieties. However, some 'similarities' are in fact more apparent than real. A number of obvious grammatical variables have been described, but there is still plenty of room for further research. Given the data from Cantonese as well as other dialects, we can discover many linguistic phenomena that are absent in Mandarin. This allows us to have a better understanding of Chinese as a whole. As shown in Zhu (1980), and Cheng and Sybesma (1999), etc., the comparative studies between Cantonese and Mandarin can help identify the grammatical subtleties of one another.

Notes

1 Many of the non-Han speakers of Cantonese belong to the Zhuang ethnic group, which mostly lives in the Guangxi Zhuang Autonomous Region.
2 The Romanization of Cantonese in this chapter is based on Standard Cantonese, and follows the Linguistic Society of Hong Kong (LSHK) Cantonese Romanization Scheme, also known as *jyutping*.
3 See also Mair (1991) for his proposal of using *topolect* as 'an exact, neutral translation of *fangyan*'.
4 Those varieties whose status is contestable are excluded, e.g. Danzhouhua and Tuhua.
5 Tone numerals are used. The non-Entering tones of /55/, /25/, /33/, /21/, /23/, and /22/ are respectively known as Dark Flat, Dark Rising, Dark Departing, Light Flat, Light Rising, and Light Departing tones by their traditional names. /25/, /21/ and /23/ might be given as /35/, /11/ and /13/. For simplicity, the description above follows Bauer and Benedict (1997).
6 See Yiu (2005) for the spatial characteristics and semantic extension of Cantonese directional verbs.
7 Simpson (2010) observes that some native speakers of Cantonese seem to be able to interpret bare nouns as definite, and gives the following example to illustrate this.

Coeng	hai	mai	saan-zo?
window	be	not-be	closed-Perf

'Is the window closed?'

See Y.-H. A. Li (1997), Cheng and Sybesma (1999), Au Yeung (2005), Sio (2006), Wu and Bodomo (2009), and Tang (2011) for the syntactic analysis of definiteness in Cantonese and Mandarin.
8 See Huang and Ahrens (1999) for the contrasting view that the post-DO *gei* in Mandarin is verbal.
9 See Zhang (2011) for typological evidence to support the deletion approach along these lines.
10 See Mok (1998) for the syntactic analysis of *gwo* comparatives.
11 See Chin (2009) for the historical development of *bei* that connects its functions in passives and double-object constructions.
12 See S.-P. Law (1990), Tang (1998b), Y.-K. A. Law (2004), and B. Li (2006) for the proposals on the syntactic positioning of sentence-final particles.
13 See Tang (2009) for the syntax and hierarchical order of discontinuous constructions.

References

Australian Academy of the Humanities and Chinese Academy of Social Sciences (1987) *Language Atlas of China*, Hong Kong: Longman.

Au Yeung, Wai-Hoo (2005) 'An Interface Program for Parameterization of Classifiers in Chinese', Doctoral Dissertation, Hong Kong University of Science and Technology.

Bauer, Robert S. (1987) 'Kadai Loanwords in Southern Chinese Dialects', *Transactions of International Conference of Orientalists in Japan* 32: 95–111.

Bauer, Robert S. and Paul K. Benedict (1997) *Modern Cantonese Phonology*, Berlin: Mouton de Gruyter.

Blust, Robert (1984–1985) 'The Austronesian homeland: a linguistic perspective', *Asian Perspectives* 26(1): 45–67.

Chao, Yuen-Ren (1968) *A Grammar of Spoken Chinese*, Berkeley and Los Angeles: University of California Press.

Cheng, Lai-Shan Lisa and Rint Sybesma (1999) 'Bare and Not-so-bare Nouns and the Structure of NP', *Linguistic Inquiry* 30(4): 509–42.

Cheung, Hung-Nin Samuel (1972/2007) *Xianggang Yueyu Yufa de Yanjiu* (*A Grammar of Cantonese as Spoken in Hong Kong*), Hong Kong: The Chinese University Press.

Chin, Chi-On (2009) 'The Verb GIVE and the Double-object Construction in Cantonese in Synchronic, Diachronic and Typological Perspectives', Doctoral Dissertation, University of Washington.

DeFrancis, John (1984) *The Chinese Language: Fact and Fantasy*, Honolulu: University of Hawaii Press.

Fung, Suk-Yee Roxana (2000) 'Final Particles in Standard Cantonese: Semantic Extension and Pragmatic Inference', Doctoral Dissertation, Ohio State University.

Gao, Huanian (1980) *Guangzhou Fangyan Yanjiu* (*A Study of the Cantonese Dialect*). Hong Kong: The Commercial Press.

Hou, Jingyi (2002) *Xiandai Hanyu Fangyan Gailun* (*An Outline of Modern Chinese Dialects*), Shanghai: Shanghai Jiaoyu Chubanshe.

Huang, Chu-Ren and Kathleen Ahrens (1999) 'The Function and Category of *Gei* in Mandarin Ditransitive Constructions', *Journal of Chinese Linguistics* 27(2): 1–26.

Institute of Linguistics of the Chinese Academy of Social Sciences, Institute of Ethnology and Anthropology of the Chinese Academy of Social Sciences, and Language Information Sciences Research Centre of the City University of Hong Kong (2012) *Zhongguo Yuyan Dituji* (*Language Atlas of China*), 2nd ed., Beijing: The Commercial Press.

Kwok, Helen (1984) *Sentence Particles in Cantonese*, Hong Kong: Center of Asian Studies, University of Hong Kong.

LaPolla, Randy J. (2001) 'The Role of Migration and Language Contact in the Development of the Sino-Tibetan Language Family', in Alexandra Y. Aikhenvald and R. M. W. Dixon (eds) *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*, New York: Oxford University Press, 225–54.

Lau, Chun-Fat (2000) 'Xiandai Yueyu yuanyu Songmo yimin shuo' ('A Theory that Modern Cantonese Derived from Late Song Dynasty Immigrants'), in *Proceedings of 7th International Conference on Yue Dialects*, Beijing: The Commercial Press.

Law, Sam-Po (1990) 'The Syntax and Phonology of Cantonese Sentence-final Particles', Doctoral Dissertation, Boston University.

Law, Yan-Kei Ann (2004) 'Sentence-final Focus Particles in Cantonese', Doctoral Dissertation, University College London.

Leung, Chung-Sum (2005) *Dangdai Xianggang Yueyu Yuzhuci de Yanjiu* (*A Study of the Utterance Particles in Cantonese as Spoken in Hong Kong*), Hong Kong: Language Information Sciences Research Centre, City University of Hong Kong.

Li, Boya (2006) 'Chinese Final Particles and the Syntax of the Periphery', Doctoral Dissertation, Leiden University.

Li, Charles N. and Sandra A. Thompson (1989) *Mandarin Chinese: A Functional Reference Grammar*, Berkeley and Los Angeles: University of California Press.

Li, Jingzhong (1990) 'Yueyu shi Hanyu zuqun zhong de duli yuyan' ('Cantonese as an Independent Language in the Family of Chinese Languages'), in *Proceedings of 2nd Conference on Yue Dialects*, Guangzhou: Jinan University Publishers.

Li, Xinkui (1994) *Guangdong de Fangyan* (*Dialects of Guangdong*), Guangzhou: Guangdong Renmin Chubanshe.

Li, Xinkui, Huang Jiajiao, Shi Qisheng, Mai Yun, and Chen Dingfang (1995) *Guangzhou Fangyan Yanjiu* (*A Study of the Guangzhou Dialect*), Guangzhou: Guangdong Renmin Chubanshe.

Li, Yen-Hui Audrey (1997) 'Structures and Interpretations of Nominal Expressions', Ms., University of Southern California.

Liu, Danqing (2001) 'Hanyu geiyulei shuangjiwu jiegou de leixingxue kaocha' ('A Topological Study on the Ditransitive Construction of Chinese Give Verbs'), *Zhongguo Yuwen*《中国语文》(*Studies of the Chinese Language*) 5: 387–98.

Lü, Shuxiang (ed) (1982) *Yuwen Changtan* (*Elementary Knowledge of Language*), Hong Kong: Joint Publishing.

Luke, Kang-Kwong (1990) *Utterance Particles in Cantonese Conversation*, Amsterdam and Philadelphia: John Benjamins Publishing Company.

Mair, Victor H. (1991) 'What Is a Chinese 'Dialect/Topolect'? Reflections on Some Key Sino-English Linguistic Terms', in *Sino-Platonic Papers* 29: 1–31. Philadelphia: Department of East Asian Languages and Civilizations, University of Pennsylvania.

Matthews, Stephen and Virginia Yip (1994/2011) *Cantonese: A Comprehensive Grammar*, London and New York: Routledge.

Mok, Sui-Sang (1998) 'Cantonese Exceed Comparatives', Doctoral Dissertation, University of California, San Diego.

Norman, Jerry and Mei Tsu-Lin (1976) 'The Austroasiatics in Ancient South China: Some Lexical Evidence', *Monumenta Serica* 32: 274–301.

Ouyang, Jueya (1990) 'Guangzhouhua de yuqi zhuci' ('The Modal Particles in Cantonese'), in *Wang Li Xiansheng Jinian Lunwenji* (*Wang Li Memorial Volumes*), Beijing: The Commercial Press, 464–76.

Ouyang, Jueya (1991) 'Yunyong diceng lilun yanjiu shaoshu minzu yuyan yu Hanyu de guanxi' ('Study of the Relationship between Minority Languages and Han Chinese with the Substratum Theory), *Minzu Yuwen* (*Minority Languages of China*) 6: 23–9.

Paris, Marie-Claude (1981) *Problèmes de Syntaxe et de Sémantique en Linguistique Chinoise*, Paris: Collège de France.
Rao, Bingcai (1988) 'Yue fangyan' ('Yue Dialects'), in *Zhongguo Dabaike Quanshu – Yuyan Wenzi* (*Encyclopedia of China – Language Volume*), Beijing: Zhongguo Dabaike Quanshu Chubanshe.
Simpson, Andrew (2010) 'The Definite Bare Classifier Construction: A Cross-linguistic Study', Paper presented at Workshop on Definiteness in Chinese, The Chinese University of Hong Kong.
Sio, Ut-seong Joanna (2006) *Modification and Reference in the Chinese Nominal*, Utrecht: LOT.
Smith, Carlota S. (1991/1997) *The Parameter of Aspect*, Dordrecht: Kluwer Academic Publishers.
Tan, Chee-Beng (2005) 'Chinese in Malaysia', in Melvin Ember, Carol R. Ember, and Ian Skoggard (eds) *Encyclopedia of Diasporas: Immigrant and Refugee Cultures Around the World*, New York: Kluwer, 697–706.
Tang, Sze-Wing (1998a) 'On the "Inverted" Double Object Construction', in Stephen Matthews (ed) *Studies in Cantonese Linguistics*, Hong Kong: The Linguistic Society of Hong Kong, 35–52.
Tang, Sze-Wing (1998b) 'Parametrization of Features in Syntax', Doctoral Dissertation, University of California, Irvine.
Tang, Sze-Wing (2003) *Hanyu Fangyan Yufa de Canshu Lilun* (*A Parametric Theory of Chinese Dialectal Grammar*), Beijing: Peking University Press.
Tang, Sze-Wing (2006a) 'Yueyu *dakzai*, *matzai*, *gamzai* shifou shuyu tongyige jiazu?' ('Do *dakzai*, *matzai*, and *gamzai* in Cantonese Belong to the Same Family?'), *Studies in Chinese Linguistics* 1: 1–11.
Tang, Sze-Wing (2006b) 'Yueyu yiwenju *sin* de jufa tedian' ('Syntactic Properties of *Sin* in Cantonese Interrogatives'), *Zhongguo Yuwen*《中国语文》(*Studies of the Chinese Language*) 3: 225–32.
Tang, Sze-Wing (2007) 'Yueyu kuangshi xuci de jubuxing he duochongxing' ('Locality and Multiplicity of Function Words in the Discontinuous Construction in Cantonese'), in *Proceedings of the 10th International Conference on Yue Dialects*, Beijing: Zhongguo Shehui Kexue Chubanshe.
Tang, Sze-Wing (2009) 'The Syntax of Two Approximatives in Cantonese: Discontinuous Constructions Formed with *zai6*', *Journal of Chinese Linguistics* 37(2): 227–56.
Tang, Sze-Wing (2011) 'On Gerundive Nominalization in Mandarin and Cantonese', in Foong-Ha Yap, Karen Grunow-Harsta, and Janick Wrona (eds) *Nominalization in Asian Languages: Diachronic and Typological Perspectives*, Amsterdam and Philadelphia: John Benjamins Publishing Company, 147–60.
Tang, Sze-Wing and Cheng Siu-Pong (2014) 'Aspects of Cantonese Grammar', in C.-T. James Huang, Y.-H. Audrey Li, and Andrew Simpson (eds) *Handbook of Chinese Linguistics*, Hoboken, New Jersey: Wiley-Blackwell, 601–28.
Ting, Jen (1995) 'A Non-uniform Analysis of the Passive Construction in Mandarin Chinese', Doctoral Dissertation, University of Rochester.
Ting, Jen (1998) 'Deriving the Bei-construction in Mandarin Chinese', *Journal of East Asian Linguistics* 7(4): 319–54.
Wu, Yicheng and Adams Bodomo (2009) 'Classifiers ≠ Determiners', *Linguistic Inquiry* 40(3): 487–503.
Yiu, Yuk-Man Carine (2005) 'Spatial Extension: Directional Verbs in Cantonese', Doctoral Dissertation, The Hong Kong University of Science and Technology.
Yuan, Jiahua, et al. (eds) (1960) *Hanyu Fangyan Gaiyao* (*An Outline of the Chinese Dialects*), Beijing: Wenzi Gaige Chubanshe.
Yuan, Jiahua, et al. (eds) (2001) *Hanyu Fangyan Gaiyao* (*An Outline of the Chinese Dialects*), 2nd ed., Beijing: Yuwen Chubanshe.
Yue-Hashimoto, Anne (1976) 'South Chinese Dialects – the Tai Connection', *Computational Analyses of Asian and African Languages* 6: 1–9.
Yue-Hashimoto, Anne (1991) 'The Yue Dialects', in William Shi-Yuan Wang (ed) *Languages and Dialects of China* (*Journal of Chinese Linguistics* Monograph 3), Berkeley: Project on Linguistic Analysis, 237–93.
Zee, Yun-Yang Eric (1999) 'Chinese (Hong Kong Cantonese)', in International Phonetic Association (ed) *Handbook of the International Phonetic Association*, Cambridge: Cambridge University Press, 58–60.
Zhan, Bohui (ed) (1991) *Hanyu Fangyan ji Fangyan Diaocha* (*Chinese Dialects and the Dialect Investigation*), Wuhan: Hubei Jiaoyu Chubanshe.

Zhan, Bohui (ed) (2002) *Guangdong Yue Fangyan Gaiyao* (*An Outline of Yue Dialects in Guangdong*), Guangzhou: Jinan University Press.

Zhang, Min (2011) 'Hanyu fangyan shuangjiwu jiegou nan-bei chayi de chengyin: leixingxue yanjiu yinfa de xin wenti' ('Revisiting the Alignment Typology of Ditransitive Constructions in Chinese Dialects'), *Bulletin of Chinese Linguistics* 4(2): 87–270.

Zhu, Dexi (1961) 'Shuo *de*' ('On *De*'), *Zhongguo Yuwen*《中国语文》(*Studies of the Chinese Language*) 12: 1–15.

Zhu, Dexi (1980) 'Beijinghua, Guangzhouhua, Wenshuihua he Fuzhouhua li de *de* zi' ('*De* in the Beijing, Guangzhou, Wenshui, and Fuzhou Dialects'), *Fangyan* (*Dialects*) 3: 160–5.

because they change the meaning of a word when one consonant is replaced by another. Take the (un)aspirated bilabial stops as examples. Unlike English, aspirated /p^h/ (like *p* in *pie*) and unaspirated /p/ (like *p* in *spy*) are two phonemes, and a minimal pair between them can be found, e.g. /p^haː55/ 'crouch' vs. /paː55/ 'daddy'. Other plosives are the same and as a result there are two sets of plosives in Cantonese, one aspirated and another unaspirated. /ts, ts^h/ may be realized as [tʃ, $tʃ^h$], but they are either *allophones* (when the following vowel is rounded, especially when it is /yː/) or simply free variants of /ts, ts^h/, depending on the speaker's accent. All of these consonants are syllable-initial. Nasals are also syllable-final, and syllable-final plosives /p, t, k/ are unreleased, expressed as [$p^{\neg}$, $t^{\neg}$, $k^{\neg}$]. The 19 consonants reflect the *standard* pronunciation. Many native speakers do not articulate /k^w, k^{wh}, n, ŋ/ syllable-initially and /k, ŋ/ syllable-finally (except in rhymes /ek, ok, eŋ, oŋ/; see Bauer and Benedict 1997). This is of particular interest as far as Romanization is concerned, because many Cantonese dictionaries and textbooks consider the drop or replacement of these sounds to be *non-standard* or even 'sloppy' and retain the standard way of pronunciation.[1]

Vowels in Cantonese include both front and back, and high and low monothongs /iː, e, yː, uː, o, ɛː, ɔː, œː, ɵ, ɐ, aː/. Depending on which analysis is adopted, the use of the vowel length mark may be optional or conditional,[2] and even the total number of vowels can be reduced as some of them occur in *complementary* phonetic contexts. Whether or not they are deemed phonemes depends on how these sounds are perceived. For example, /e/ may lose its phonemic status and be argued as an allophone of /iː/, since they are not only similar but also complementary: /e/ before /k, ŋ/ and /iː/ elsewhere. This is the same for the sound pair of /o, uː/. In the standard, conservative accent, the sound pair of /œː, ɵ/ follows a similar distribution, although /œː/ can stand alone without a final consonant and monothong /ɵ/ is always followed by /t/ or /n/. This part being introductory, the academic debate associated is set aside. What concerns us is the fact that many Romanization schemes use a single letter for both sounds of the pair. In addition to these monothongs, there are also ten gliding vowels in Cantonese: /iw, ɵɥ, uɥ, ej, ɔj, ow, ɐj, ɐw, aːj, aːw/. In some analyses, the final glides are rendered as a vowel symbol. Notice that in the same syllable, a gliding vowel (or diphthong) is not followed by a consonant. Only a monothong precedes a consonant coda. For this reason, the final glide in a gliding vowel is often considered as a coda as well. The vowels above are all phonemic, with the apparent 'exceptions' of /ɵɥ, uɥ/. Minimal pairs can be found only when they are preceded by /k/, and also /k^h/ if tones are disregarded. For the sound pairs of /ɐj, aːj/ and /ɐw, aːw/, they are clearly distinctive, and it is easy to find a minimal pair to differentiate the two members within. But given that they are similar in sound quality, these distinctive sounds, together with /ɵɥ, uɥ/, are sometimes represented by a single *grapheme* when they are rendered in Romanization. This will be further addressed in the following sections. In Cantonese, a syllable always contains a vowel, with the exception of two *syllabic* nasals /m, ŋ/. They can stand alone without a vowel to follow, in which case they are rendered as [m̩, ŋ̍].

Cantonese is a tonal language. There are six or nine tones in Cantonese, depending on one's approach to categorize them. Traditionally, tones in Chinese are classified into Flat, Rising, Departing, and Entering. In Cantonese, the first three groups contain two tones (one Dark and one Light), and together with three Entering tones (Upper Dark, Lower Dark, Light), they amount to a total of nine tones. For the non-Entering tones, Cantonese has Dark Flat tone /55/, Light Flat tone /21/, Dark Rising tone /25/, Light Rising tone /23/, Dark Departing tone /33/, and Light Departing tone /22/.[3] For the Dark Flat tone, there is also a variant which is high-falling /53/. The two Dark Flat tones are either not distinctive (Cheung 1986), or in fact, the falling variant might have been fading out altogether (Matthews and


image
not
available


image
not
available

1958, it has been reprinted several times over a period of four decades. In its sequel *Book II*, Huang states that the Romanization system was developed by Kok. Based on the year of its first publication, it is reasonable to assert that the scheme came into being no later than the late fifties. Huang's later work *Cantonese Dictionary: Cantonese–English, English–Cantonese* also adopted this system. Cantonese is one of the languages Romanized under a system created at Yale University. Other Yale Romanizations involve East Asian languages such as Mandarin, Korean, and Japanese. Since its introduction, the Yale system has been extensively used in English-language textbooks, dictionaries, and glossaries made for Cantonese learners. These include some of the most popular titles currently in circulation. Comparatively speaking, the use of the Yale system is less prevalent in Chinese-language publications. While the native speakers sometimes do need to know how to pronounce a particular Chinese character, they are often referred to another homophonous character. Even when Romanization is consulted (not very often), the Yale system is not the one needed. So far, it is chiefly used to study Cantonese as a *foreign* language, rather than offer a pronunciation guide for the native speakers. Thus, Cantonese learners using this system might encounter a situation where they know this Romanization much better than the native speakers.[6]

Under the Yale system, the 19 consonants are expressed in a way that makes a distinction between aspirated and unaspirated consonants. The plosives /p, p^h, t, t^h, k, k^h, k^w, k^{wh}/ are respectively rendered as *b*, *p*, *d*, *t*, *g*, *k*, *gw*, and *kw*. This contrasts with the Government Romanization. For example, in the Government version, *p* represents both /p/ and /p^h/, but Yale Romanization uses two separate letters *b* and *p*. In this case, the Yale system has one-to-one mapping between letters and phonemes. To achieve this goal, the Yale system uses letters traditionally reserved for voiced consonants to represent the voiceless ones. Affricates /ts, ts^h/ are given as *j* and *ch*. Similarly this distinction is not observed in the Government Romanization, and they are both expressed as *ch* or *ts*. Nasals /m, n, ŋ/, fricatives /f, s, h/, and /j, w, l/ are straightforwardly rendered as *m*, *n*, *ng*, *f*, *s*, *h*, *y*, *w*, *l*. With /j/ Romanized as *y*, the Yale system resembles the English orthography.

Phoneme-to-grapheme consistency is also achieved in the vowel representation, although a single letter may in separate situations be used to represent distinct sounds. Letter *a*, for example, is used for both /aː/ and /ɐ/. Yet /aː/ is expressed in single letter *a* only when it stands alone in the rhyme. Gliding vowels with /aː/ are given in double letters *aa*, and *aa* is also used when /a/ precedes a consonant coda. Since /ɐ/ (always given as *a*) never occurs syllable-finally, a distinction can still be made between the two vowels. Letter *i* is used for /iː, e/, which is common to most Romanization schemes as they are in complementary distribution. The same applies to *u*, which represents both monothongs /uː/ and /o/. It should be noted, however, that *u* is also used in *yu* as a grapheme representing /yː/. The *y* in *yu* does not represent the semi-vowel (approximant), except in the case where the initial consonant is /j/. To illustrate, /syː/ is expressed as *syu*, whereas /jyː/ is given as *yu*, instead of the incorrect *yyu*. The 'shorthand' rule takes advantage of the fact that /juː/ does not exist as an indigenous syllable in Cantonese, but problems may arise when loanwords are involved. /juː/, for example, can mean 'university' in daily speech. Apparently, the Yale system does not assign a separate Romanized form for this syllable. Letter *u* is also used for other vowels when it is associated with *e*. Letter *e* as a single letter represents /ɛː/, but when combined with *u* to form *eu*, it refers to both /œː/ and /ɵ/: /lœk/ is given as *leuk* and /lɵt/ is given as *leut*. Letter *o* is used for monothong /ɔː/. The same principles are generally extended to the gliding vowels, so /iw, ej, ɔj, ɐj, ɐw, aːj, aːw/ are expressed as *iu*, *ei*, *oi*, *ai*, *au*, *aai*, *aau*, with *u* representing the rounded glide and *i* for the unrounded one. However, /ɵɥ, uɥ/

GENERAL INDEX