

Diversity in Sinitic Languages

OUP UNCORRECTED PROOF – REVISES, 27/10/2015, SP1

Diversity in Sinitic Languages

Edited by
HILARY M. CHAPPELL

OXFORD
UNIVERSITY PRESS



Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

© editorial matter and organization Hilary M. Chappell 2015
© the chapters their several authors

The moral rights of the author have been asserted

First Edition published in 2015

Impression: 1

All rights reserved. No part of this publication may be reproduced, stored in
a retrieval system, or transmitted, in any form or by any means, without the
prior permission in writing of Oxford University Press, or as expressly permitted
by law, by licence or under terms agreed with the appropriate reprographics
rights organization. Enquiries concerning reproduction outside the scope of the
above should be sent to the Rights Department, Oxford University Press, at the
address above

You must not circulate this work in any other form
and you must impose this same condition on any acquirer

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data
Data available

Library of Congress Control Number: 2015938213

ISBN 978-0-19-872379-0

Printed in Great Britain by
Clay Ltd, St Ives plc

Links to third party websites are provided by Oxford in good faith and
for information only. Oxford disclaims any responsibility for the materials
contained in any third party website referenced in this work.

Contents

<i>Preface</i>	vii
<i>Acknowledgements</i>	ix
<i>List of abbreviations</i>	xiii
<i>Notes on contributors</i>	xvi

Part I. Approaches to Diversity in Sinitic Languages

1 Introduction: Ways of tackling diversity in Sinitic languages <i>Hilary M. Chappell</i>	3
2 Linguistic areas in China for differential object marking, passive, and comparative constructions <i>Hilary M. Chappell</i>	13
3 Grammatical change in Sinitic languages and its relation to typology <i>Alain Peyraube</i>	53

Part II. Typological Studies of Sinitic Languages

4 The semantic differentiation of demonstratives in Sinitic languages <i>Yujie Chen</i>	81
5 Bare classifier phrases in Sinitic languages: A typological perspective <i>Wang Jian</i>	110
6 The comparative construction in Sinitic languages: Synchronic and diachronic variation <i>Hilary M. Chappell and Alain Peyraube</i>	134

Part III. Individual Studies of Linguistic Micro-Areas in China

7 Language contact in Nanning: Nanning Pinghua and Nanning Cantonese <i>Hilário de Sousa</i>	157
8 On the origin of special numerals for ‘one’ in southeastern China: [kei ²¹³] in the northwestern Min dialect of Shaowu <i>Sing Sing Ngai</i>	190

9 Complex pronouns in Wu Chinese: Focalization and topicalization <i>XuPing Li</i>	226
10 Comparative constructions of inequality in the Southern Min dialect of Hui'an <i>Weirong Chen</i>	248
<i>References</i>	273
<i>Index</i>	303

Preface

This volume brings together ten chapters that in the main represent some of the major research outcomes during the four and half years of the European Research Council (ERC) Advanced Grant project, SINOTYPE (2009–2013): ‘The hybrid syntactic typology of Sinitic languages’.

The majority of the chapters in this volume were in fact first presented at the Lyons workshop on ‘Ecology, Population Movements, and Language Diversity’ held from 26–27 May 2011 at the Université Lumière-Lyon 2 during the 4th Biennial AFLICO conference (*Association française de la linguistique cognitive*), except for the chapters by Yujie Chen who was on fieldwork in China at the time, and by Weirong Chen who joined the project in its later stages. Alain Peyraube, as one of the invited speakers at this workshop, has also kindly contributed his article.¹

The workshop was co-organized by Professor Salikoko Mufwene of the University of Chicago who was, at the time, a Fellow at the Institute of Advanced Studies, Collégium de Lyon in 2011, and by Dr François Pellegrino, Senior Researcher at the CNRS, and Director of the Research Centre, *Dynamique du Langage*, associated with the Université Lumière-Lyon 2.

Some of the hotly debated current issues that were proposed as focal themes for the *Linguistic Diversity* workshop included the challenge of considering issues such as:²

- the relationship between genetic and typological classifications;
- the connection between linguistic typology and areal linguistics;
- the treatment of languages with a mixed typology;
- intra-linguistic variation and its significance for linguistic typology;
- explanations for the emergence of linguistic diversity;
- the connection between population movements, ensuing language contact, and language typology.

Although we do not claim to answer any of these quite monumental questions exhaustively, all of them find, in fact, their echo in this volume of ten chapters which grew out of the SINOTYPE project. Each chapter has been substantially reworked into a complete analysis, as SINOTYPE project members continued to intensify their research

¹ There were two other papers presented at the Diversity workshop in Lyons. These were by William Croft (*University of New Mexico*), ‘The continuity between crosslinguistic variation and language-internal variation’, and Johanna Nichols (*UC Berkeley*) ‘Growing diversity and complexity: The contributions of geography, sociolinguistics, demography, and time.’

² These topics were judiciously suggested by Professor Salikoko Mufwene and Professor Alain Peyraube for the *Ecology, Population Movements, and Linguistic Diversity* workshop.

efforts and linguistic analysis based on both many months of fieldwork as well as several hundreds of hours of seminar discussion on the project topics.

On the SINOTYPE project:

The SINOTYPE project (Advanced Grant N° 230388) was fully funded by the FP7 Programme of the European Research Council (ERC) from 2009 to 2013. The seven team members each chose a Sinitic language to work on that had either never been described, or had only received the basic minimum of a phonological sketch, in order to undertake a typological description of the entire grammar. This task included a second general aim which was to try to uncover the degree of variation possible for a large number of morphological and syntactic parameters, according to established crosslinguistic continua and simultaneously to test out established implicational universals in these domains.

This brand new orientation in Chinese linguistics is thus diametrically opposed to the tradition that has established itself over the past decades of taking Standard Mandarin as the main, if not the only, point of reference for Sinitic languages in typological studies in the West, sometimes with inclusion of data from the better-known ‘Chinese dialects’ such as Hong Kong Cantonese, Hokkien, or Shanghainese. Until recently, this trend has even persisted in the proper domain of Chinese linguistics. Consequently, these detailed grammatical descriptions of little-known Sinitic languages, including the analyses in this volume, represent one of the first steps in revealing the nature and extent of diversity for this language taxon.

Over the four-and-a-half-year period of the project, the team members have each undertaken regular field trips, amounting to over one year’s intensive fieldwork on location in China, staying with the families of their native speaker informants, if not ‘next door’ in village or county inns. The final outcome is to publish descriptive grammars of the following seven languages:

A grammar of Xianghua, a Sinitic language of Northwestern Hunan

Hilary Chappell

A grammar of Central Plains Mandarin: the Shangshui dialect

Yujie Chen

A grammar of the Hui'an dialect of Southern Min

Weirong Chen

A grammar of Nanning Southern Pinghua

Hilário de Sousa

A grammar of Gan Chinese: the Yichun dialect

XuPing Li

A grammar of Shaowu, a Sinitic language of Northwestern Fujian

Sing Sing Ngai

A grammar of the Jixi dialect of Hui

Wang Jian

Map o.1 indicates the location of these Sinitic languages in China.

Acknowledgements

I would like to thank first and foremost the European Research Council (ERC) under the European Community's Seventh Framework Programme (FP7/2007–2013) for the funding which has supported the research leading to the findings presented in this anthology and, significantly, which enabled the creation of the SINOTYPE team in 2009. At its peak, SINOTYPE was composed of seven full-time researchers including one professor (the director), five postdoctoral fellows, and a doctoral student. Thus, all but one of the chapters in this volume form part of the research carried out for this ERC Advanced Grant agreement, entitled 'The hybrid syntactic typology of Sinitic languages (No. 230388: Sinotype)' (2009–2013).

Chapter 3 has been the recipient of separate funding from the 'blue sky' *programme blanc* of the French Agence Nationale de la Recherche from which Chapters 1, 2, and 6 have also benefited. This was for the projects 'Diachronic change in Southern Min, a Sinitic language / Changement diachronique en min méridional, langue sinitique' (n° ANR-08-BLAN-0174 CSD 9: Diamin) (2009–2011) and 'Typologization of synchronic and diachronic processes in Southern Min (Sinitic language) / Typologie des processus synchronique et diachronique en Min-Sud (langue sinitique)' (n° 11-ISH2-001-01:Tysomin) (2012–2015), both international bilateral projects with the National Science Council of Taiwan, since renamed the Ministry of Science and Technology.

For their stimulating questions and comments, the SINOTYPE team and Alain Peyraube join me in thanking Professors William Croft, Salikoko Mufwene, Johanna Nichols, François Pellegrino, William S-Y. Wang, and other members of the audience at the Lyons workshop on 'Ecology, Population Movements, and Language Diversity' held at the Université Lyon 2, 26–27 May 2011, where most of the chapters in this volume were presented for the first time. The workshop was organized by Professor Salikoko Mufwene (University of Chicago/I.A.S., Collégium de Lyon) and Dr François Pellegrino, Director of the Research Centre, *Dynamique du Langage*, CNRS/Université Lumière-Lyon 2 to whom we express our gratitude for this important opportunity to air our research.

Sing Sing Ngai and the volume editor would also like to thank the Australian Academy of Social Sciences for its kind permission to reproduce, in Chapter 8, Map B-12 of the *Min Supergroup of Dialects* from the 1987 edition of the *Language Atlas of China*, a groundbreaking atlas that was jointly compiled with linguists from the Chinese Academy of Social Sciences in Beijing under the direction of Professors Stephen Wurm and Li Rong.





MAP 0.1 Fieldwork locations in China of the SINOTYPE team

OUP UNCORRECTED PROOF – REVISES, 27/10/2015, SP1

List of abbreviations

The abbreviations for grammatical glosses follow the *Leipzig Glossing Rules* in the main, apart from categories which are common in Sinitic languages and for which the SINOTYPE team created new labels such as VCL for ‘verbal classifier’, RDP for ‘reduplication’, or TENT for ‘tentative aspect’.

1/2/3 SG/PL	first/second/third person singular/plural
ACC	accusative
ACOM	the comparee
ADV	adverb
ASP	aspect marker
ATTR	attributive
BA/OM	object marker
BSTA	the standard of comparison
CAUS	causative
CL	classifier
CM	comparative marker
COMP	completive aspect marker
COMPR	comparative
COP	copula
CRS	currently relevant state marker
DEGR	degree marker
DEM	demonstrative
DIM	diminutive
DIR	directional complement
DIST	distal
DMC	dependent-marker of comparison
DO	direct object
DOM	differential object marker
eBr	elder brother
EXP	experiential
FOC	focus
FocP	focus phrase

xiv *List of abbreviations*

GEN	genitive marker
HMC	head-marker of comparison,
HORT	hortative
IDEO	ideophone
IMP	imperative
IO	indirect object
IP	inflectional phrase
LOC	locative
LW	linking word
MED	medial
MOD	particle for linking modifiers with head noun
MW	measure word
N	head noun
NEG	negative
NEG:IMP	negative imperative modal verb
NEUT	neutral
NMZ	nominalizer
NOM	nominative
NP	noun phrase
NP _A	noun phrase referring to the comparee
NP _B	noun phrase referring to the person or thing acting as the standard or benchmark
NSG	non-singular
NUM	numeral
OM	object marker
ONOM	onomatopoeia
PASS	passive marker
PFV	perfective aspect marker
PL	plural
PN	place/proper name
POSS	possessive marker
PRED	comparative predicate
PREF	prefix
PREP	preposition
PRF	perfect aspect

List of abbreviations xv

PRO	pronoun
PROX	proximal
PRT	particle
Q	question particle
RDP	reduplication
REL	relative clause
RES	resultative complement
RVC	marker of resultative verb compound
SE	Standard British English
SFP	sentence-final particle
SG	singular
SUF	suffix
SUPL	superlative
SUR	surpass
SVC	serial verb construction
TENT	tentative aspect
TOP	topic
V	verb
VCL	verbal classifier
VCOMPL	verb complement
VOC	vocative
VP	verb phrase
XP	X Phrase in generative grammar, where X is a variable that can represent any category
yBr	younger brother

Notes on contributors

HILARY CHAPPELL (曹茜蕾) holds the Chair in Linguistic Typology of East Asia at the École des Hautes Études en Sciences Sociales (EHESS) in Paris, France. Her main research is on the typology of Sinitic languages (or Chinese dialects) with the goal of investigating the extent of their linguistic diversity. To this end, she has been carrying out fieldwork over the past seven years in Hunan province, China, on the Xianghua language. A second, related research area pertains to the diachronic syntax of Southern Min (Hokkien), making use of a corpus of late sixteenth-and early seventeenth-century materials. Her publications include an earlier volume on *Sinitic Grammar: Synchronic and Diachronic Perspectives* (OUP, 2001) and *A Grammar and Lexicon of Hakka*, co-written with Christine Lamarre (EHESS, 2005).

YUJIE CHEN (陈玉洁) is Associate Professor at Zhejiang University in Hangzhou, China. After being awarded her doctoral degree in 2007 by the Chinese Academy of Social Sciences on the typology of demonstrative paradigms in Sinitic languages, she spent four years from 2009–2012 as a postdoctoral fellow with the SINOTYPE project in Paris. Her research interests focus on the syntax, semantics, and pragmatics of Sinitic languages from a typological perspective with a special interest in the unusual Central Plains Mandarin dialects. Dr Chen’s major publication is the monograph *Typological Research on Chinese Demonstratives* (in Chinese) (Beijing, CASS, 2010).

WEIRONG CHEN (陈伟蓉) is a lecturer in the School of Chinese Language and Literature, at the University of International Business and Economics in Beijing, China. She has also been a postdoctoral fellow with the ERC SINOTYPE project at the École des Hautes Études en Sciences Sociales (EHESS) in Paris from 2011–2013, after obtaining her doctoral degree at the University of Hong Kong. Her primary areas of research are morphosyntax, grammaticalization, and Chinese linguistics. She is a specialist in Southern Min dialects.

HILÁRIO DE SOUSA (苏沙) obtained his PhD from the University of Sydney on the Menggwa Dla language of New Guinea, and has held postdoctoral positions at the Max Planck Institute for Psycholinguistics (Nijmegen), studying the language of perception, on the SINOTYPE project at the École des Hautes Études en Sciences Sociales (EHESS), to write a grammar of the Southern Pinghua language in Nanning, China, and at Radboud University’s Center for Language Studies (Nijmegen) to carry out research on the language of olfaction in Cantonese. His research interests range across the domains of semantic typology (sense and emotion), phonology, and morphosyntax of languages in the Asia-Pacific region, as well as areal and historical linguistics.

XUPING LI (李旭平) completed his postdoctoral studies in Paris on the ERC SINOTYPE project at the École des Hautes Études en Sciences Sociales (EHESS), and a PhD at Bar Ilan University before taking up the post of Associate Professor at Zhejiang University in Hangzhou, China. His research falls into two main areas: formal semantics and Chinese dialectal syntax, with a

book published on *Numerical Classifiers in Chinese: The Syntax-Semantics Interface* (Mouton de Gruyter, 2013) and several articles in journals, including *Lingua* and *Language and Linguistics*.

SING SING NGAI (倪星星) is currently a doctoral student at the École des Hautes Études en Sciences Sociales (EHESS) in Paris, France, working under the supervision of Prof. Hilary Chappell. The title of her thesis is ‘A grammar of Shaowu, a Sinitic language of Northwestern Fujian’, which also constitutes a part of the ERC SINOTYPE project of which she was a member. Her present research focuses on describing and analysing major syntactic features of the Shaowu dialect and comparing those with its neighbouring dialects and dialect groups. She is interested in language typology, language evolution, syntax, and semantics in general.

ALAIN PEYRAUBE (贝罗贝) is currently Emeritus Senior Researcher at the Centre National de la Recherche Scientifique (CNRS, France) and Chair Professor of Chinese Linguistics at the École des Hautes Études en Sciences Sociales (EHESS) in Paris. His main research interests are on Chinese diachronic syntax and semantics, and on the typology of East Asian languages. His latest research has been carried out within a broadly functional and cognitive framework, using a crosslinguistic perspective and includes projects on grammatical change in Early Southern Min and language contact in Northwest China.

WANG JIAN (王健) was originally awarded his doctoral degree in 2005 by Peking University. Following his doctoral work, he conducted his research at the Chinese Academy of Social Sciences (CASS) from 2005–2007, Fudan University from 2007–2009, and at the École des Hautes Études en Sciences Sociales (EHESS) in Paris from 2009–2013 as a postdoctoral fellow on the ERC SINOTYPE project. He is currently Professor of Linguistics in the School of Humanities at Shanghai Jiaotong University. His research areas include language typology, Chinese dialectology, computational linguistics, and applied linguistics. Wang Jian’s major work is on the *Comparative Dialectal Grammar of the Jiangsu and Anhui Linguistic Area* (Commercial Press, 2014) (in Chinese).

OUP UNCORRECTED PROOF – REVISES, 27/10/2015, SP1

Part I

Approaches to Diversity in Sinitic Languages

OUP UNCORRECTED PROOF – REVISES, 27/10/2015, SPL

1

Introduction: Ways of tackling diversity in Sinitic languages

HILARY M. CHAPPELL

This volume represents one of the first major outcomes of the ERC SINOTYPE project (2009–2013) which brought together seven researchers from China, France, and Australia to carry out empirically-based typological studies on the vast, yet little explored domain of Sinitic languages.¹

Challenging the tradition in Chinese linguistics, the SINOTYPE project has consistently treated Chinese dialects as a separate, though unquestionably affiliated, group of languages belonging to the Sino-Tibetan language family and having the technical name of ‘Sinitic’. They are generally considered to be a first-order, sister branch to the Tibeto-Burman languages, the latter lying to the west and southwest of the core area in China where Sinitic languages are found. Details of their classification and a linguistic map are presented in Chapter 2 for the ten recognized second-order branches of Sinitic: Mandarin, Jin, Wu, Hui, Gan, Xiang, Yue, Hakka, Min, and Pinghua.

The issue of linguistic diversity, genetic relationship, and areal linguistics has been of fundamental importance to the SINOTYPE project, since any typological classification that involves a set of parameters applied to a number of languages located in contiguous zones, related or not, will have a tendency to transcend genetic relationships and, in the case of Sinitic, to be positively oriented towards the setting up of linguistic areas within China, grouping together languages and dialects from different branches of Sinitic. Admittedly, in this instance, we are treating related languages and dialects and, thus, cases of apparent ‘interdialectal’ language contact and diffusion, yet the same principles relevant for areal linguistics in general (Clark 1989; Dahl 2001; Enfield 2005) and grammaticalization zones (Heine and Kuteva 2005, 2006) can be seamlessly applied in such cases, as Chapter 2 will argue. The relevant connections between genes, family tree relationships, languages, and diffusibility

¹ A description of the SINOTYPE project is given in the preface to this volume.

are treated in both Matisoff (2001) and Peyraube (2007) for Southeast Asia and East Asia respectively.

In reality, the Sinitic languages are, *grosso modo*, as distinct from one another as European languages, a point argued in Chappell (2001a). Even within the second-order groups of Sinitic, such as the large Mandarin dialect area with its estimated 800 million L₁ speakers today, a further division into eight main subgroups is necessary to account for its diversity—and this merely in the form of differences in their phonological inventories, tone systems, and lexicon, to which the *Language Atlas of China* amply bears testimony (Wurm et al. 1987, 2nd edition 2012). The work on syntactic differences is only beginning.

Moreover, in comparing any two representative dialects for these subgroups of Mandarin, we could quite appropriately discuss issues related to ‘interdialectal’ or ‘intervarietal’ differences and the outcomes of contact, just as we might for the varieties of Irish and Scottish English, or regional dialects such as Yorkshire and Geordie, in relation to Standard British English (SE). Similarly to the case for SE, only a small proportion of the Mandarin population are actually native speakers of the variety closest to Standard Mandarin or *pǔtōnghuà* 普通话, based on the Beijing dialect. In contrast to this, if we compared features of the Puxian isolate of the Min dialect group, located in the southeastern coastal province of Fujian, with the Xi'an dialect of Central Plains Mandarin in Shaanxi, northern China, some 1,800 km away, this would be more akin to a comparison between English and Swedish, or Portuguese and Romanian, and thus truly interlinguistic, despite the evident genetic relationship in all cases.

The closely intertwined connection between typology, areal linguistics, and linguistic diversity has become an indisputable reality today, since the findings of typological classification clearly feed into the identification of linguistic areas and these, in turn, identify the regions where the greatest—or the least—diversity is to be found. Dryer's detailed studies (2003, 2008) on the typological features of Sino-Tibetan and mainland Southeast Asian languages are a case in point. By way of contrast, this diversity—which is only starting to be measured and appreciated for Sinitic languages—has to be ultimately linked with migrations and consecutive episodes of language contact, particularly in the North-to-South direction in China over the past two millennia (Chappell 2001a and You R. 1992 on migration patterns in China and the formation of Sinitic; LaPolla 2001 on migration patterns and the formation of Sino-Tibetan).

In other words, the explanations for the emergence of linguistic diversity in China have to be precisely formulated in terms of sociolinguistic parameters, including population movements, social behaviour, and social attitudes towards language, that repeatedly arise in language contact situations and condition external borrowing, if not the adoption of another language in part or in whole (Thomason and Kaufman 1988; Mufwene 2001). These considerations of language-external causes leading to diversity undeniably include features otherwise not to be predicted for Sinitic and are principally discussed in Chapter 7 with respect to the area of Far Southern China and

Introduction: Ways of tackling diversity in Sinitic languages

5

the interaction between Pinghua, Cantonese Yue, and Northern Zhuang (Tai-Kadai). The case is also debated for non-Sinitic languages as the putative source of word order change for ditransitives and adverbial placement in certain dialect groups in Chapter 3, while the possibility of an external model for Sinitic *Surpass* comparatives is weighed up in Chapter 6. Language contact in Sinitic languages and some of its consequences in the form of stratification, convergence, and hybridization are treated in an earlier article of mine (Chappell 2001b).

The problem of the mixed typology of Sinitic languages is a general one and has been specifically addressed by the SINOTYPE project as a crucial area to target.

Hence, one of the objectives has been to find explanations for the contradictory mixture of word order correlations in Sinitic, some of which tally with VO while others tally with OV. The project has thus set out to examine and seek explanations for this apparent hybrid typology found in most Sinitic languages, all of which reveal a perplexing mixture of head-final and head-initial features for languages whose basic word order is SVO.

Consequently, Sinitic languages pose several striking counterexamples to classic Greenbergian word order correlations, in particular, the word order for the main type of comparative construction which is *Marker-Standard-Adjective/VP*. As is well known, this is the order which correlates with SOV languages (Dryer 1992) and is discussed in Chapter 2, Chapter 6, and Chapter 10. The same disharmonic situation applies for prenominal relative clauses which merely reflect the overall head-final requirement for the noun phrase in Sinitic, yet once again, does not tally with its SVO clause-level word order. Preliminary findings show that, in contrast to Standard Chinese, many Sinitic languages show the use of postnominal relative clauses and head-marked comparatives which evidently harmonize ‘correctly’ with SVO basic word order. Many of these theoretical issues have been separately treated in Chappell, Li, and Peyraube (2007).

Diachronic grammar provides another important methodology for explaining how grammatical change creates linguistic diversity in Sinitic languages. This source, including the language-internal mechanisms of reanalysis, analogy, and exaptation (the recycling of a form into a new unrelated function), but also external borrowing, is championed in Chapter 3.

Facets of diversity are also reflected, albeit in different manners and on a wide variety of topics, in the detailed descriptive studies handled in a typological framework in this volume. Some of the new research findings revealed in this volume are the following:

- (i) Based on a corpus of over 300 Sinitic languages, the description and classification of a wide variety of demonstrative systems in Sinitic whose paradigms vary from neutral one-term to highly differentiated types with four or five terms (Y. Chen, Chapter 4).

- (ii) The diversity of behaviour with respect to definiteness and word order for bare classifier noun phrases in a sample of over 100 Sinitic languages. Bare classifier phrases are the product of omission of either a demonstrative or the numeral ‘one’ from the position immediately preceding the classifier and its head noun. While this has been described for individual Chinese dialects, it has never been explored on such a scale before (Wang Jian, Chapter 5).
- (iii) The highly unusual case in Sinitic of triple sets of pronominal paradigms that are found in certain Northern Wu dialects and include two special series of complex fused forms whose use is determined by specific discourse functions (X.P. Li, Chapter 9).
- (iv) The analysis of comparatives of inequality in Southern Min given in Chapter 10 (W. Chen) captures an impressive array of diversity for just this one structure, treated mainly from an intra-linguistic but firmly typological viewpoint for the Min dialects: the Hui'an dialect possesses no less than six different comparative constructions including head-marking, dependent-marking, double-marking, and zero-marking structures. Certain of these can be shown to be the product of innovation or of interdialectal borrowing and also evince diachronic change in progress.
- (v) In an enclave formed by a group of different languages and dialects in central and northern Fujian and adjacent areas, an unusual source for the basic numeral ‘one’ is found in the form of GÈ 个, which appears to be related to the general classifier in Standard Mandarin. Additional forms for ‘one’ are found as well in this micro-area, all of whose sources derive from words for ‘lone’ or ‘unique’, including a typical Min form, SHÙ 蜀. In the Shaowu dialect of Northwestern Min, the numeral GÈ 个 operates in complementary distribution with the form, SHÙ 蜀 ‘one’, while it also serves as a general classifier, leading to some interesting conjectures on the diachronic development of GÈ 个 (S. Ngai, Chapter 8).
- (vi) Intra-linguistic variation is considered as an integral part of the description of linguistic diversity for a substantial number of phonological, lexical, and morphosyntactic features in the Southern Pinghua dialects of Guangxi, in addition to language contact issues for the linguistic micro-area they form and share with the Cantonese and Nanning Yue dialects as well as Northern Zhuang (Tai-Kadai/Kra-Dai) languages (H. de Sousa, Chapter 7).

In the chapters which follow, as mentioned, some of the analyses take a macroscopic, typological perspective to analyse large-scale samples of Sinitic languages in terms of their crosslinguistic (or cross-dialectal) variation, while others concentrate on one particular Sinitic language or group of dialects, if not on one particular linguistic area.

For the thematic reasons outlined above, following the two scene-setting chapters in Part I, the large-scale typological studies of Sinitic languages have been united to

Introduction: Ways of tackling diversity in Sinitic languages

7

form Part II, while the studies on individual languages, language areas, or dialect groups are presented in Part III.

The unifying thread in this volume, however, regards the findings of all these chapters which attest to the previously rather meagrely perceived, if not entirely unheeded, linguistic diversity of Sinitic languages which is slowly but methodically being uncovered through the work of many scholars, be they in China or in the ‘West’.

The detailed chapter summaries follow:

The first main chapter by **Hilary Chappell** has as its main objective to show the great diversity inherent in Sinitic languages from both angles of morphosyntax and grammaticalization, a diversity which has been little described in the past for Chinese dialects, apart from the important exceptions of phonology and the lexicon. She describes the widespread variation in differential object-marking, passive and comparative constructions in terms of the lexical source of the marker in the case of the first two construction types and in terms of both lexical source and structural type for the comparative of inequality.

Her aim is to see if it is possible to discern linguistic areas within China on this basis. She tentatively proposes that, in terms of grammatical behaviour, at least five principal areas exist in China, rather than just a simple North–South division: in addition to the Northern area, the South is split into the three linguistic areas of Southwestern, Far Southern, and Southeastern. Furthermore, in between North and South China lies an intermediate zone, aptly named the Central Transitional area which demonstrates considerable turbulence in its typological features. Within the Northern area a small enclave of grammatically conservative Jilu and Jiaoliao dialects of Shandong peninsula is also to be found. H. Chappell observes that this refinement of earlier classifications of linguistic areas in China is but the first step, nonetheless, an undoubtedly necessary first step that can be used as a basis for further research to verify, re-adjust, or re-align the boundaries.

Alain Peyraube links diachronic grammar and linguistic typology as two important domains of research on Chinese. He observes that the three basic mechanisms involved in grammatical change for Sinitic languages are the processes of reanalysis (including grammaticalization and exaptation), analogy (including the phenomena of lexicalization and/or degrammaticalization), and external borrowing through language contact. A. Peyraube points out, however, that at first glance these do not appear to involve linguistic typology, *stricto sensu*, or only minimally. The same applies to the motivations for grammatical change such as semantic-pragmatic change and structural requirements.

This leads A. Peyraube to consider precisely what connects the two domains. Beginning in the 1980s, diachronic syntax has not in fact ceased to impact on typological research in China, providing it with a new *élan* in the form of a framework

upon which hypotheses can be built concerning the common properties which Sinitic languages share, or the basic differences which separate them.

To illustrate these points, examples are given from Sinitic languages for the relation between causatives and passives, postverbal and preverbal adverbs, double object constructions with verbs of giving, and finally, the reanalysis of verbs of saying into complementizers. All these topics are discussed from a diachronic point of view, successfully relating this in synchronic terms to typology.

Yujie Chen has undertaken the first large-scale study of demonstrative paradigms in 303 Sinitic languages, classifying them into neutral one-term, two-term, three-term, and multiple systems, along the lines of Diessel (1999). She shows that while two-term systems are widespread in Sinitic, three-term paradigms are not uncommon. In the Gan dialect group whose heartland is in Jiangxi province in Central China, three-term, four-term, and five-term systems can be found which may use tone sandhi and reduplication to distinguish the paradigm members for distance: *close-proximal-distal-yonder-further than yonder*. Strikingly, one-term neutral systems may be found to interact with another more complex demonstrative paradigm in the same language.

In general, these paradigms are sensitive to the parameter of distance from a given reference point while a small minority are sensitive to visibility as well. Y. Chen also examines the relation of ‘yonder’ and ‘close’ demonstratives to distal and proximal categories in terms of implicational hierarchies. She argues that, in the case of Sinitic languages, ‘yonder’ demonstratives do not necessarily imply the presence of distal ones, as in certain languages they may be used independently and non-contrastively, whereas ‘close’ demonstratives consistently imply the presence of a proximal category.

She concludes by proposing two universals to explain (i) the dominance of the feature of distance in Sinitic demonstrative systems and (ii) the fact that distal demonstratives can be differentiated for a greater number of semantic attributes than are proximal, this being the general case for the other main language families in Southeast Asia.

Wang Jian presents the results of a typological analysis of bare classifier phrases based on a sample of 120 Sinitic languages and dialects in order to formalize what has already been described for isolated cases of Chinese dialects about this phenomenon. Bare classifier phrases equate to cases where the numeral ‘one’ or a demonstrative has been omitted, engendering either an indefinite or a definite reading, depending on certain other factors, which can be explained as follows:

On the basis of the interplay of the two crucial parameters of (i) syntactic position with respect to the main verb—preverbal or postverbal—and (ii) the interpretation of definiteness for the bare classifier noun phrase in question, seven main types of Sinitic languages are identified with respect to the behaviour of this type of noun

phrase. At one end of the spectrum, we find Min dialects which do not allow bare classifier phrases at all in any position, while at the other end, languages mainly located in the Central Transitional zone of China allow all four possibilities for bare classifier phrases. These include Wu, Hui, Gan, Xiang, Hakka, and Jianghuai Mandarin. The unmarked and most common type includes many of the Northern Sinitic languages from both Mandarin and Jin subgroups, but also adjacent languages of Central China. These allow only indefinite, postverbal bare classifier NPs.

Wang Jian concludes his discussion by setting up three important implicational universals connected with these bare classifier phrases for the properties of definiteness versus indefiniteness and pre- or postverbal position. Universal 1 states that if a language possesses preverbal bare classifier noun phrases, it will also possess postverbal ones. Two further universals are related to the tendency in Sinitic languages for preverbal constituents to be definite and postverbal, indefinite, this configuration being the unmarked case: Universal 2 accounts for the fact that if a language allows preverbal bare classifier phrases to have an indefinite interpretation, it also allows for a definite one in this position. Finally, the third universal, Universal 3, accounts for the fact that if a language allows a definite interpretation for postverbal bare classifier phrases, it also allows an indefinite one.

Hilary Chappell and Alain Peyraube set out to explore the interrelationship between synchronic and diachronic aspects of the two main comparative construction types in Sinitic languages, viz, the Type I *Compare* comparative with 比 *bǐ* and the Type II *Surpass* comparative with 過 *guò* which, broadly speaking, align China on a North–South basis, opposing a dependent-marking strategy to a head-marking strategy. First, they argue for the establishment of a new cognitive schema of a *Compare* comparative, a schema which has not been attested nor consequently included in any of the main typological studies on this topic, yet represents a crucial kind for Sinitic languages. An overview of the distribution of these two types of comparative structures in China and their markers is next provided with a brief but pertinent digression on the *Surpass* comparative in the non-Sinitic languages of southern China as well as in those of other Southeast Asian language families.

In the diachronic treatment of this topic, they observe that the earliest written records reveal that Archaic Chinese began as a head-marking Type II comparative language. They show that the *Compare* comparative is essentially a late Northern development which became dominant in the early Yuan period of the thirteenth century. In the present era, it is steadily encroaching on all major dialect areas within China, and even on the non-Sinitic languages of southern China, competing with the native types, principally the *Surpass* comparative. They hypothesize that during evolution of the *Compare* comparative into the major comparative structure of inequality in the north of China, the *Surpass* comparative was able to independently develop and flourish in central and southern China without any initial competitors.

That it must have had a broader distribution at some earlier time can be seen in the fact that it is still found in peripheral areas, such as in the Jiaoliao and Jilu Mandarin spoken on the Shandong peninsula or in isolated dialects in Shaanxi province, both within the Northern area of China.

Hilário de Sousa presents a finely detailed investigation of the interaction between three main languages in the city of Nanning: Nanning Southern Pinghua, Nanning Cantonese—which are both Sinitic languages—and the unrelated Kra-Dai (or Tai-Kadai) language of Northern Zhuang, the major language spoken in the Guangxi Autonomous Region. The prolonged language contact amounts to many centuries for Zhuang and Pinghua speakers, the latter arriving in this area of Southern China as early as the eleventh century. In contrast, there is much less time depth for the Cantonese speakers who arrived relatively recently in the nineteenth century. Nonetheless, the results of this investigation are unexpected in terms of the lexicon versus morphology and syntax. While borrowing freely from Northern Zhuang vocabulary, Pinghua shows much less influence from Zhuang in terms of grammatical features than do the two varieties of Cantonese, Nanning and Standard. Furthermore, Nanning Cantonese proves to be even more Zhuang-like than Standard Cantonese, for which claims of a Tai substratum have been well-substantiated. It is particularly for the ordering of gender affixes with their head nouns, direct and indirect objects in ditransitive (or double object) constructions, and the extent of polysemy in the functions of classifiers, where these differences emerge. De Sousa proceeds to explain these outcomes in terms of a specific set of sociolinguistic parameters conditioning the language contact outcomes.

Sing Sing Ngai considers the possible pathways of evolution classifier *kei*²¹³ 个 and its use as the numeral ‘one’ in the Shaowu dialect of Western Min, a dialect spoken in an inland region of Fujian province on the frontier with the Gan dialect group. She treats, in particular, this unusual etymon for ‘one’, which is quite distinct from the corresponding term in most other Sinitic languages, where *yī* — ‘one’ is widespread. A third synonym, constituted by SHŪ 蜀 ‘one’ is also discussed. SHŪ 蜀 is in fact a common form for the numeral ‘one’ in the Min dialects. All three forms for ‘one’ being found serendipitously to co-exist in the Shaowu dialect, they can be usefully compared for their syntactic distribution, with the added significance of *kei*²¹³ 个 having a classifier usage as well with both nouns and verbs. Several further unusual forms for ‘one’ are discussed for the Central and Western Min region, showing an intriguing diversity for this small area.

Since the frequency and distribution of the numeral use is higher than that of the classifier use, S. Ngai claims that [kei²¹] 个 may actually be the indigenous form for ‘one’ in Shaowu which lies in the northwestern corner of Fujian province, while observing that this form, used with the same meaning of ‘one’, may be found in adjacent Wu, Hakka, and Gan-speaking areas.

Introduction: Ways of tackling diversity in Sinitic languages 11

Ngai argues that the classifier and numeral senses of the morpheme *kei*²¹³ 个 represent a bifurcated pathway of development, and that the numeral use has most likely evolved from an earlier meaning of ‘lone’ or ‘unique’, evidenced in unrelated language families as a common source for the basic numeral ‘one’.

XuPing Li’s contribution concentrates on the Fuyang dialect of Northern Wu to tease apart the argument roles and semantic values of the stressed and the non-stressed uses of complex emphatic personal pronouns. In the Taihu subgroup of Northern Wu dialects, paradigms of complex fused pronouns exist side by side with a simple pronominal series, and include several dialects of the greater Shanghai area as well as Fuyang, Shaoxing, and others. Emphatic pronouns of the complex series in Fuyang Wu are described as having a special discourse prominence, while stress is defined as a strictly acoustically-defined feature.

X.P. Li demonstrates that stressed or emphatic complex pronominal forms are freer in their syntactic distribution than the corresponding unstressed forms. They may occur in a larger number of syntactic positions and construction types, including contrastive topic and focus. Contrastive focus, for example, licenses their use in the basic postverbal object position, but also in the preverbal object position of the KE object-marking construction, as well as in cleft and pseudo-cleft constructions. The unstressed counterparts may not occur in these positions but rather only in preverbal primary and secondary topic positions, although this too effectively crosscuts subject and object argument roles.

X.P. Li also briefly treats diachronic aspects and upholds the view that these complex pronouns are historically the result of fusion of the simple series of pronouns with a preceding copular verb /z/ in bare cleft constructions. He proposes a cycle of stress placement and focus, whereby the originally focus-sensitive complex pronouns are de-stressed and de-focalized, thence undergoing reanalysis to be subsequently used in plain topic positions. In the next stage in the cycle, these complex pronouns are able to receive stress anew, and so may occur once more not only in contrastive topic positions, but also in contrastive focus positions, and thus in a wider range of syntactic positions than that of simply topic—primary or secondary—as is the case for the unstressed emphatic pronouns.

Southern Min languages are well known for their special and often, unique, linguistic characteristics which set them apart from other Sinitic languages. Weirong Chen’s chapter presents no exception in considering the large number of structurally different types of comparative construction found in the Hui’an dialect of Southern Min, on the basis of data gathered from discourse recordings. In all, six different comparatives are analysed in order of frequency of use, being described from the essential angles of syntactic configuration and specific semantic properties but importantly also in terms of typological parameters such as word order harmony and the position of relators. These include both head- and

dependent-marking types, the hybridized or double-marked comparative and the zero-marked type.

The adopted Northern-style *比* *Compare* comparative proves to be the least frequent in her database while the hybridized structure using both dependent-marking with preposition *比* [*pi*³] ‘compared to’ and the adverbial head-marking strategy with *kha?*⁷ 怪 is the most common. In fact, three of the Hui'an comparative constructions involve the use of this adverbial marker *kha?*⁷ 怪, which W. Chen hypothesizes may be a case of syntactic reanalysis from a degree adverb ‘a bit, fairly’ to a formal marker of the comparative.

W. Chen is able to reveal some new types that have not been fully described or identified in the typological literature on comparatives and accounts for some of the less frequent structures in the Hui'an dialect in terms of borrowing from Mandarin, and generational change.

2

Linguistic areas in China for differential object marking, passive, and comparative constructions

HILARY M. CHAPPELL

2.1 Introduction

This chapter considers the relation between diversity and typology in the context of areal linguistics of the Sinitic languages of China, also called ‘Chinese dialects’. It will specifically examine the feasibility of a classification into five areas as a refinement of the traditional North-South division originally proposed by Hashimoto (1976a, 1986).¹ The aim is to reveal the great diversity of forms in Sinitic languages that are used to express one and the same constructional meaning.

Traditional classifications of Chinese dialects have mainly been based on phonology—and the lexicon to a lesser extent. In this analysis, the illustrations will be given in terms of morphosyntax and grammaticalization pathways in order to see if these can yield a new approach to dialect classification in China. To this end, differential object marking, passive and comparative constructions are examined for the ten main Sinitic languages.

2.1.1 *Classifications of the Sinitic branch of Sino-Tibetan and the issue of linguistic areas*

The general consensus is that Sinitic or Chinese languages form a sister branch to Tibeto-Burman languages in the Sino-Tibetan language family. Broadly speaking, these

¹ In Chappell (forthcoming c), I also discuss smaller linguistic areas such as the Qinghai-Gansu border area of Northwestern China where Mandarin Chinese dialects have been subject to radical morphosyntactic change due to contact with Turkic and Mongolian languages. These more geographically limited ‘micro-areas’ will not fall under general discussion, however, in this introductory chapter (see also Peyraube forthcoming).

languages are tonal and analytic in make-up. While SVO is a common word order in Sinitic, (S)OV is equally common in the Wu and Min dialect groups (Liu D. 2003a). Sinitic languages exhibit both head-final and head-initial features: they tend to be almost entirely head-final for nominal structure while a mixture of both types is present in predicate syntax. All modifiers precede the head noun including relative clauses whereas at the predicate level, adverbs and most types of prepositional phrase precede the main verb, yet certain verb complements and locative adjuncts follow it.

Derivational morphology and compounding are common processes as opposed to the use of inflectional marking for categories such as tense/aspect, person, gender, and number. Less well known is the fact that fusional, portmanteau morphology is more widespread than thought, not to mention tone sandhi for coding grammatical features such as aspect, the plural form of pronouns, to distinguish types of demonstrative pronouns (Y. Chen, this volume) or even to confer a demonstrative function on a classifier (Wang Jian, this volume). As the background to the main discussion, the composition of the Sinitic branch of Sino-Tibetan is next briefly described.

According to the *Language Atlas of China* (Wurm et al. 1987, revised 2nd edition 2012, edited by Zhang Zhenxing), the ten main Sinitic languages (or Chinese dialect groups) currently recognized are Northern Chinese (Mandarin), Xiang, Gan, Wu, Min, Kejia (or Hakka), Yue, Jin, Pinghua, and Hui dialects. Note that these dialect groups have largely been established on phonological, lexical, and a few morphological features, as mentioned earlier. Furthermore, a certain amount of debate exists about this classification with respect to the independent status of the last three groups, Jin, Pinghua, and Hui, particularly for Pinghua, which some scholars still classify under Yue (for more details on this debate, see de Sousa, this volume). Table 2.1 presents the relevant information.

These ten main groups of Sinitic languages are represented in Map 2.1, which includes, importantly for any study of areal linguistics, the eight main subdivisions of the large Mandarin group. These are: Beijing (or Northern) 北京, Northwestern (Lanyin) 兰银, Jilu 冀鲁, Jiaoliao 胶辽, Northeastern 东北, Central Plains (Zhongyuan) 中原, Jianghuai (or Southern) 江淮, and Southwestern Mandarin 西南官话. There is also a part of southern Hunan province and adjacent northern Guangdong where varieties of patois or *tūhuà* 土话 are spoken and whose affiliation is as yet unknown, so too for Xianghua 乡话 (also known as Waxiang 瓦乡) of Northwestern Hunan.

2.1.2 Linguistic areas within China

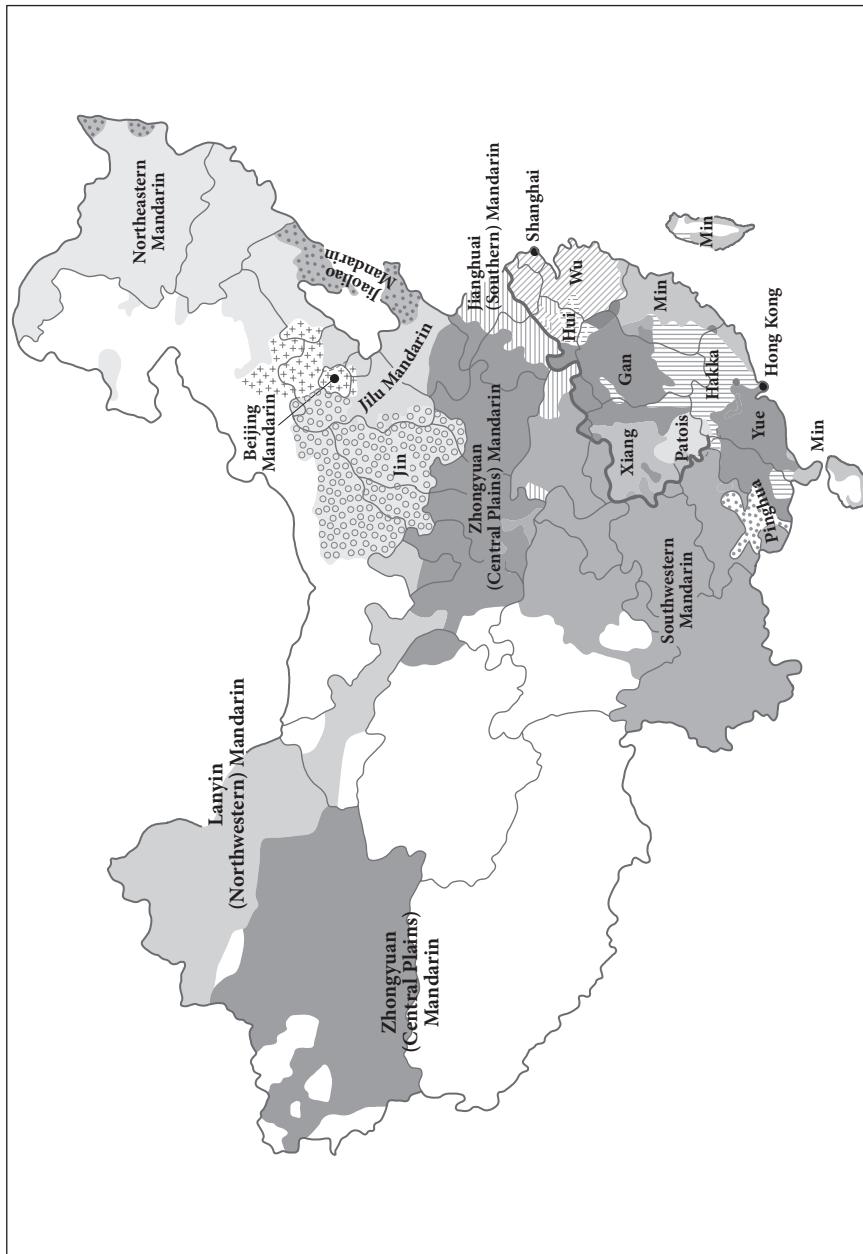
Making use of important typogeographical features in phonology, lexicon, and syntax, Mantaro Hashimoto (1976a, 1986) argued for a major North–South linguistic division in China. In his view, this was the result of many centuries of contact between Altaic and Sinitic languages in the north, and Tai languages in the south. His early works signalled the beginnings of areal typology in Chinese linguistics, subsequently further advanced by Norman (1982, 1988). Some of the tendencies he identified are listed in Table 2.2 for the Chinese languages found in the two zones.

TABLE 2.1. Sinitic languages of China

	Language branch	Core region of China	Population (millions)	Representative varieties
I.	Mandarin 官话	North, Northeast, Southwest of China	799m	Beijing 北京, Nanjing 南京, Chengdu 成都
II.	Jin 晋	Shanxi 山西, Inner Mongolia 内蒙古, Shaanxi 陕西	63m	Taiyuan 太原, Huojia 荷嘉
III.	Xiang 湘	Hunan 湖南	36m	Changsha 长沙, Shuangfeng 双峰
IV.	Gan 赣	Jiangxi 江西	48m	Nanchang 南昌, Yugan 余干
V.	Hui 徽	Anhui 安徽	3.3m	Jixi 绩溪, Xiuning 休宁
VI.	Wu 吴	Zhejiang 浙江, Southern Jiangsu 江苏	74m	Shanghai 上海, Wenzhou 温州
VII.	Min 闽	Fujian 福建, NE Guangdong 广东, Taiwan 台湾	75m	Xiamen 厦门, Fuzhou 福州, Chaozhou 潮州
VIII.	Kejia 客家 (Hakka)	SW Fujian 福建, NE Guangdong 广东	42m	Meizhou 梅州, Changting 长汀
IX.	Yue 粤 (Cantonese)	Guangdong 广东 Eastern Guangxi 广西	59m	Hong Kong 香港, Taishan 台山
X.	Pinghua 平话 and Tuhua 土话	Guangxi 广西, Guangdong 广东 Southern Hunan 湖南	7.8m	Nanning 南宁, Guilin 桂林, Lanshan 蓝山

TOTAL 1,207,100,000 (1.2 billion)²

² These figures are based on those given by Xiong Zhenghui 熊证辉 and Zhang Zhenxing 张振兴 (2008: 97) and Zhang Zhenxing 张振兴 (2013) for the new version of the *Language Atlas of China* and have been rounded up. Xiong and Zhang explain that they have used the 2004 China Administrative Regions Yearbook (中国行政区划简册Zhongguo Xingzhengqu Huajiance) for the population figures. This compares with the 17th edition of *Ethnologue* (Lewis et al., 2014) which is based on an extrapolation of the 2000 China census, the last census to pose questions on language and dialect use. Lewis et al. estimate 840 million of speakers of the different Mandarin dialects, rather than 799 million as above. The 2010 China Census counted 1,339,724,852 inhabitants in China but did not, unfortunately, provide any data on language use. Of these, 91.5% are of Han Chinese nationality. According to Xiong and Zhang (2008), within the Han nationality, 66% speak one of the dialects of Mandarin, whereas Lewis et al. (2014) claim 70% are speakers. This accounts for the differences in the estimated figures for the Mandarin group.



MAP 2.1. Sinitic languages and dialects of China

TABLE 2.2. Tendencies in Sinitic languages according to Hashimoto's North–South division

ALTAICIZATION (North)	TAICIZATION (South)
Stress-based and fewer tones	Larger number of tones
Higher proportion of polysyllabic words	Higher proportion of monosyllabic words
Simpler syllable structure	More complex syllable structure
Smaller inventory of classifiers	Larger inventory of classifiers
Preponderance of MODIFIER-MODIFIED, including gender affixes on animal terms	MODIFIED-MODIFIER order possible, including gender affixes on animal terms
Preverbal adverbs	Possibility of postverbal or clause-final adverbs
DOUBLE OBJECT: IO-DO word order unmarked for prepositionless ditransitives	DOUBLE OBJECT: DO-IO word order unmarked for prepositionless ditransitives
COMPARATIVE: Marker-Standard-Adjective	COMPARATIVE: Adjective-Marker-Standard
PASSIVE MARKERS: based on causative speech act verbs	PASSIVE MARKERS based on verb 'to give'

This basic North–South division was later refined by Norman (1988) who added a third transitional zone where types intermediate between the two could be found. This included the Wu, Gan, Hui, and Xiang dialects in central China.

While this set of properties shows certain important trends, it cannot be viewed as definitive. Counterexamples exist where the claimed Northern feature is possessed by a purportedly Southern Chinese language. A case in point is the basic ditransitive order in Southern Min which is IO-DO and not DO-IO as Table 2.2 would suppose. The same ordering of IO-DO is, moreover, found even further south than Southern Min in Fujian province with respect to Nanning Pinghua, spoken in the Guangxi Autonomous Region (see de Sousa, this volume, §7.4.8). Similarly contradicting this prediction, certain subgroups of Mandarin use the DO-IO order (see Chappell and Peyraube, this volume §6.3; Chappell and Peyraube 2007; Chappell forthcoming a; Zhang Min 2011).

Another problem with this set of properties is that instances of modified-modifier constructions are largely overstated, given that they concern but a small number of place names, and a subset of compound words. In addition to these refinements, it will also be shown in §2.3 that passive markers in Sinitic languages may derive from six different lexical sources, and not only causative or *give* verbs, and in §2.4 that there are at least six structural types for the comparative of inequality, and thus not just two main types as listed in Table 2.2.

2.1.3 Linguistic areas within mainland East and Southeast Asia

Matisoff (1991: 386) further refined Hashimoto's basic classification by dividing the larger Southeast Asian zone into two main areas: the Sinospheric and the non-

Sinospheric or Indospheric area. The Sinospheric area includes Southern Sinitic (basically Sinitic languages south of the Yangtze River) and the language families which have been in close cultural contact with China such as Hmong-Mien, Tai-Kadai, Vietnamese (Mon-Khmer branch of Austroasiatic), and certain branches of Tibeto-Burman such as Lolo-Burmese. The non-Sinospheric languages include Austronesian languages, many Mon-Khmer languages, and Tibeto-Burman languages, for example, those found in Northeastern India and Nepal.

Matisoff (1991) identified a large number of grammatical features which unify the Southeast Asian area into a linguistic zone, including those in Table 2.3.

TABLE 2.3. Some major Southeast Asian linguistic features described by Matisoff (1991)

- 1) Development of modal verbs > desiderative markers, 'be likely to'
- 2) Development of verbs meaning 'to come' > motion towards a deictic centre
- 3) Development of verbs meaning 'to place, put' > durative and perfective aspect markers
- 4) Development of verbs meaning 'to dwell' > progressive aspect markers
- 5) Development of verbs meaning 'to finish' > perfective aspect markers
- 6) Development of verbs meaning 'to get, obtain' > 'manage', 'able to', 'have to'
- 7) Development of verbs of giving > causative and benefactive markers
- 8) Development of verbs of saying > complementizers, topic and conditional markers
- 9) Formation of resultative and directional compound verbs through verb concatenation

With respect to Sinitic, all of these pathways of grammaticalization are characteristic of Northern Chinese as well. The only exception would be the limited use of 'give' as a causative verb in Standard Mandarin (see also Chappell 2001b). This leaves us with an inventory of features that is extremely useful for characterizing the ensemble of East and Southeast Asia as a mega-linguistic area, rather than Southeast Asia on its own, with the caveat that it does not at present allow for any finer differentiation on a smaller geographical scale.³

2.1.4 Definition of a linguistic area

Before embarking on any investigation of linguistic areas for the Sinitic languages of China, the definition for a linguistic area adopted needs first of all to be given. A linguistic area is constituted when languages are grouped together for reasons of both geographic contiguity and the sharing of a significant number of structural properties. These languages should ideally include some which are unrelated,

³ Other major work on Southeast Asia as a linguistic area or with regard to shared typological features has been carried out by Marybeth Clark (1989), Walter Bisang (1996), and N.J. Enfield (2005), to mention just a few of the scholars involved in this type of research. For Sinitic languages, see also Yue-Hashimoto (1993) and Yue (2003).

according to the accepted definitions currently in use. The key properties distinguishing them from neighbouring languages are typically considered to be the result of the diffusion of traits (Dahl 2001: 1457; Enfield 2005).

Such consensus notwithstanding, we could apply this definition quite felicitously to just the case of languages and dialects within the same language family, or within the same branch of a language family (that is, a taxon), such as Sinitic. In the case of related languages, as opposed to linguistic areas built on languages that are mainly unaffiliated, diffusion of linguistic features over time can be shown to interact with those which are directly genetically inherited. The outcome is similarly the formation of an area whose key features are distinct in some way from those of the surrounding areas, though be they linguistically related.

The case differs somewhat from the canonical situation for one and the same language family or taxon, given that there are likely to be cognates of the form in the replica language or dialect for the borrowed function that is adopted from the model language. Such forms simply continue to be used in a different function from the borrowed one, if they have not already been relegated to a special register or even to an obsolete stratum (see Lien 2001 for the outcome of contact between Southern Min and Mandarin). Sinitic languages are well-endowed with cases of this type of contact-induced grammaticalization that takes place between *related* model and replica languages (Heine and Kuteva 2005).

In the main body of this description of linguistic areas discernible for Sinitic languages, I will examine the distribution of differential object-marking in §2.2, passives in §2.3 and comparative constructions of inequality in §2.4, according to the morphological markers they use with the purpose of evaluating to what extent they enable us to define smaller linguistic areas of diffusion and shared traits. Areal classifications will first be set up according to each individual construction. Then, in §2.5, a synthesis for the three constructions under investigation will be made in the attempt to set up five main linguistic areas in China. This is carried out by overlaying areal maps of the variation found for these three important grammatical constructions, one upon the other. Other studies on the areal distribution of linguistic features in Sinitic are also referred to at this point.

In the section which follows, the results of a study on the diversity of object-marking constructions in Sinitic languages are presented.

2.2 Differential object marking constructions in Sinitic

By the term ‘object-marking construction’, I refer in as neutral terms as possible to non-canonical constructions where the direct object is explicitly morphologically marked and precedes the main verb:

S – *Differential object marker* – O – Verb

This is known as the *chūzhishi* 处置式 in Chinese linguistics, and is often translated as the ‘disposal construction’ in English, or the *bǎ* 把 construction, if referring to Standard Mandarin. It contrasts with the typical SVO order in many dialect groups of Sinitic. These disposal or object-marking constructions appear in fact to ‘prepose’ the direct object of an SVO clause, marking its new position by a preposition. Despite these synchronic facts, historically they have evolved from serial verb constructions (SVC) where the first verb, a verb meaning ‘hold’ or ‘take’, grammaticalizes into a preposition which introduces the direct object. The first example is from Standard Mandarin.⁴

(1) Standard Mandarin object marking construction with *bǎ* 把

(NP_{CAUSE/SUBJECT}) – [DIFFERENTIAL OBJECT MARKER [*bǎ* 把] + NP_{DIRECT OBJECT}] – VP
當然是我們下江幫**把**最高級的理髮店包了。

Dāngrán shi wǒmen Xià Jiāng bāng bǎ zuì
of:course be 1PL (name) gang DOM most
gāojí de lǐfàdiàn bāo le.
classy MOD hairdresser's occupy CRS

‘Of course, it was our Xia Jiang gang who took over the classiest hairdressing salon.’ (Zhang and Sang 1987: 340)

In the basic type of object-marking construction, the preverbal direct object has a tendency to be referential and to be affected by the event coded by a highly transitive predicate, conforming to the parameters given in Hopper and Thompson (1980). By ‘affect’, the causative element which codes a change of state is intended, including changes of location (Chappell 1992). The exact parameters will necessarily depend on the particular Sinitic language or dialect. Since the morphological marking is not obligatory, and concerns just a subset of object nouns, its use being determined by the discourse and semantic features associated with the object noun, it can be classified as differential in nature, as defined in Bossong (1985), Aissen (2003), and Lazard (1994). Lazard appears to be the first to have applied this term to the marker in the Mandarin *bǎ* construction.

⁴ For the transcriptions of examples, I use *pinyin* romanization 汉语拼音 for Standard Mandarin and the International Phonetic Alphabet for the data from the majority of Sinitic languages, unless there is a romanization system already in widespread currency such as the Yale system for Hong Kong Cantonese. When quoting examples from published sources, the transcriptions have been faithfully reproduced. When small capitals are used for the Standard Mandarin *pīnyīn* romanization, this generally signifies that the original article has not provided any kind of phonetic transcription but only Chinese characters. However, I also use *pīnyīn* romanization in small capitals for *in-text references* to the common lexical sources of morphological markers in Sinitic languages in order to represent these at a more abstract ‘archi-morpheme’ level. This practice has been adopted in order to circumvent the problem of which phonological form to choose from an array of dialect cognates; for example, GUO for the comparative marker 過 (过), regardless of the dialect. Chinese characters used in the examples are presented in either traditional or simplified form, according to the original source.

In Chappell (2006, 2007a, 2013), I examined and described the crosslinguistic variation found in these object-marking constructions in over 200 Sinitic languages for both sources of these prepositional markers and constructional types. In terms of grammaticalization pathways, verbs of holding and taking such as *bǎ* 把 and *ná* 拿, which are well described for Mandarin and for Northern Wu, do not, however, turn out to be the sole source of object markers in Chinese languages. In fact, there are four dominant lexical fields which act as sources.⁵ Apart from *hold/take*, the second source is found in comitative prepositions, while a further two new lexical sources are found in verbs of giving and helping.⁶

Crosslinguistically, TAKE verbs constitute a common source for object markers, as for example, in the West African Benue-Kwa languages of the Niger-Congo family (see Lord 1993; Heine and Kuteva 2002). However, GIVE and HELP verbs are not well-attested in typological literature as a source of such markers. For example, such a use of GIVE verbs is not identified in the comprehensive study of its polysemy by Newman (1996), nor listed in Heine and Kuteva's lexicon of grammaticalization (2002). Comitatives are similarly not generally attested as sources for object markers, but rather for instruments, allatives, manner markers, and even their syntactic antonyms, agent markers (Stolz 1999).

An example of each of these four main lexical sources for object markers in Sinitic languages is presented in the following examples.

(2) DIFFERENTIAL OBJECT MARKER < VERB OF TAKING AND HOLDING

Shanghainese 上海 (Wu 吴语): DOM *nɔ*⁵³ < 'take'

儂拿鈔票還拔伊。

non⁴² nɔ⁵³ tsʰɔ³⁴ pʰ iɔ³⁴ fiue²³ pəʔ⁵ fi²³.
2SG DOM money return give 3SG

'You give the money back to him.' (Xu and Tao 1999)

(3) DIFFERENTIAL OBJECT MARKER < VERB OF HELPING

Chenxi 辰溪 (Xiang 湘语): DOM *pau*⁴⁴ < 'help'

我帮月毛毛放哒床上,好吗?

ŋo³³ pau⁴⁴ nye²¹³mau²¹³mau²¹³ fau²¹⁴ ta³¹ dzau²¹³sa³¹ xau³¹ ma⁴⁴?
2SG DOM baby put.at be.on good Q

'I'll put the baby to bed, OK?' (Wu 2005: 204)

⁵ In a more recent study, based on a much larger corpus of over 650 Chinese dialects, Li and Cao (= Chappell) (2013a, 2013b) found at least eight main sources for these markers, including the four discussed above, the latter accounting nonetheless for the majority of Sinitic languages in this corpus. In addition to HOLD/TAKE, GIVE, HELP and comitatives, even smaller 'micro-areas' are to be found which use speech act verbs, allative directionals, CONNECT, and GET verbs.

⁶ Note that in Chappell (2005, 2013), verbs of giving and helping have been 'lumped' together, though they are evidently not identical semantically. This is due to their similar pathway of grammaticalization.

22 Hilary M. Chappell

(4) DIFFERENTIAL OBJECT MARKER < VERB OF GIVING

Qimen 祁门 (Hui 徽语): DOM fǎ¹¹ < 'give'

尔分钱摆好,不要跌失。

n¹¹ fǎ¹¹ tsʰí:⁵⁵ pa⁴²-xé⁴², pa⁵iú:²² ti:³⁵-ci⁰.

2SG DOM money place-proper NEG:IMP drop-lose

'Put your money away safely, and don't lose it.' (Hirata 1998: 306)

(5) DIFFERENTIAL OBJECT MARKER < COMITATIVE PREPOSITION

Xianghua 乡话, Hunan (unaffiliated Sinitic): DOM kai⁵⁵ < COMITATIVE

我跟橘子剥下皮。

u²⁵ kai⁵⁵ kon⁵⁵tsa pau⁴¹ ka³³ fa²⁵.

1SG DOM orange peel COMPL skin

'I peeled the orange (of its skin).' (Chappell et al. 2011)

There is one main exception to the use of object-marking constructions: in what we will identify below as the Far Southern area,⁷ serial verb constructions (SVC) are employed with a TAKE verb as V₁ that has not yet fully grammaticalized into an object marker. This is clearly the case, since these TAKE verbs retain all their verbal characteristics including aspect marking. This part of China largely aligns itself with the contiguous regions of mainland Southeast Asia where, similarly, the use of TAKE verbs in object marking constructions in Khmer, Vietnamese, Thai, and Hmong is not highly grammaticalized either, according to Bisang (1992). Hence, the predominance of serial verb constructions in the Far Southern Sinitic area clearly links it with the rest of the Southeast Asian *Sprachbund*. The next example shows the use of one such serial verb construction in a Pinghua dialect of the Far Southern area, for which nai⁵¹ 捅 'take' can be viewed as a potential or incipient object marker. Its use is evidently still verbal, since it is followed by a directional verb complex tʃ'y⁵¹-fu²² 'out-go (away from speaker)'.

(6) SERIAL VERB CONSTRUCTION WITH A VERB OF TAKING

'Take' serial verb construction in Xing'an Gaoshang Northern Pinghua 兴安高
尚桂北平话:

搊衣裳出去晾起。

nai⁵¹ i³⁵ʃion²¹ tʃ'y⁵¹-fu²² lon²²-khi⁵⁵.

OM clothes out-go_{DIR} dry-ASP

'Put the clothes out in the sun to dry.' (Literally: 'Take the clothes and put them out in the sun to dry.') (Lin Yi 2005: 223, 232)

⁷ In my earlier articles on object marking constructions (Chappell 2005, 2007a, 2013), I used the term 'Southern area' for the Yue, Hakka, and Pinghua dialects of Guangdong and Guangxi. For reasons of distinctiveness and appropriacy, I have modified this to 'Far Southern' in alignment with de Sousa (this volume) who coined this label.

According to the dominant patterns of use for grammaticalized object markers (OM), the following four areas can be tentatively discerned as the outcome, including one where the process of reanalysis for verbs of taking is in a young stage:⁸

1. NORTHERN AREA: DIFFERENTIAL OBJECT MARKER < HOLD and TAKE verbs Jin dialect group; Mandarin dialect group: Northeastern 东北, Beijing 北京, Northwestern or Lanyin 兰银, Jilu 冀鲁, Jiaoliao 胶辽, and some subgroups of Southwestern Mandarin 西南官话; also Northern Wu dialects 北吴.
2. CENTRAL TRANSITIONAL AREA: DIFFERENTIAL OBJECT MARKER < GIVE and HELP verbs Xiang 湘语, Gan, 赣 Hui 徽, Southern Wu 南吴, many non-Northern Mandarin dialects: Central Plains Zhongyuan 中原, Southern Jianghuai 江淮 and subgroups of Southwestern Mandarin 西南官话.
3. SOUTHEASTERN AREA: DIFFERENTIAL OBJECT MARKER < COMITATIVE prepositions Essentially a feature of Min dialects 闽语 but also of certain Hakka 客家话 and Eastern Wu dialects 东吴; Xianghua 瓦乡; scattered dialects of Jianghuai Mandarin 江淮 and Southwestern Mandarin 西南官话.
4. FAR SOUTHERN AREA: Serial verb constructions with TAKE

Use of serial verb constructions with ungrammaticalized TAKE verbs as V_i: Yue 貢语 and Hakka dialects 客家话 in Guangdong province, Pinghua 平话 dialects in the Guangxi Autonomous region.

Note that for the Southwestern Mandarin dialects of Hubei, Sichuan, and Guizhou provinces, classified as part of the Central Transitional zone, there is a high incidence of the use of BĀ 把. The data available in the main do not allow us to judge at this point of time whether this morpheme BĀ 把 is related to the lexeme ‘give’ or the lexeme ‘take’ as such ambiguity for this verb is characteristic of the Central zone (see §2.3.6 on this issue; also §5.1 in Chappell 2013a). In Yunnan province, a variety of sources has been identified in dialect surveys, including BĀ 把 < ‘hold’/‘give’, BĀNG 帮 ‘help’ and ÁI 挨 ‘be connected to’. For the present, we shall tentatively consider these areas as being part of the Central Transitional zone.

In the Far Southern China zone, apart from TAKE serial verb constructions, the literary Chinese and Standard Mandarin markers JIĀNG 將 < ‘hold’, and sometimes BĀ 把 < ‘hold’, can be found in use. These two forms have indeed been borrowed into the formal register of these languages and serve as ‘true’ object markers in this case. They are not, however, a native feature of the local vernacular languages (see Chappell 2013a for details).

⁸ Endo (2004) has also carried out extensive research on the sources and types of object marking constructions in Sinitic languages, and is one of the few scholars to have remarked upon the comitative source.

In terms of the areal nature of these features, these four linguistic areas for DOMs cross-cut dialect groupings within Sinitic. For example, Hakka dialects in close contact with Yue (Cantonese) in Guangdong province tend to behave linguistically like the Yue dialects, the dominant linguistic group, that is, they similarly make use of TAKE serial verb constructions, typical of Far Southern Sinitic. They do not behave at all like Hakka dialects further north in Hunan and Jiangxi provinces in the Central Transitional area which, in their turn, ‘copycat’ the local Xiang and Gan dialects, having developed object markers based on GIVE and HELP verbs. Further evidence reinforcing this areal trend is that Hakka dialects close to Min-speaking areas in Fujian and in Taiwan use comitatives.⁹

These four areas are indicated on Map 2.2, determined by the lexical source of the object markers in the basic S-Marker-O-V construction, found in all dialect groups of Sinitic, apart from the Far Southern.

2.3 Sinitic passives

Passives in many Southeast and East Asian languages have been described as ‘non-canonical’ in the work of Siewierska and Bakker (2013), not to mention in earlier work by Siewierska (1984). Taking Sinitic languages as an example, this is due to the fact that, two typical constructional features which go against the ‘canon’ are the following:

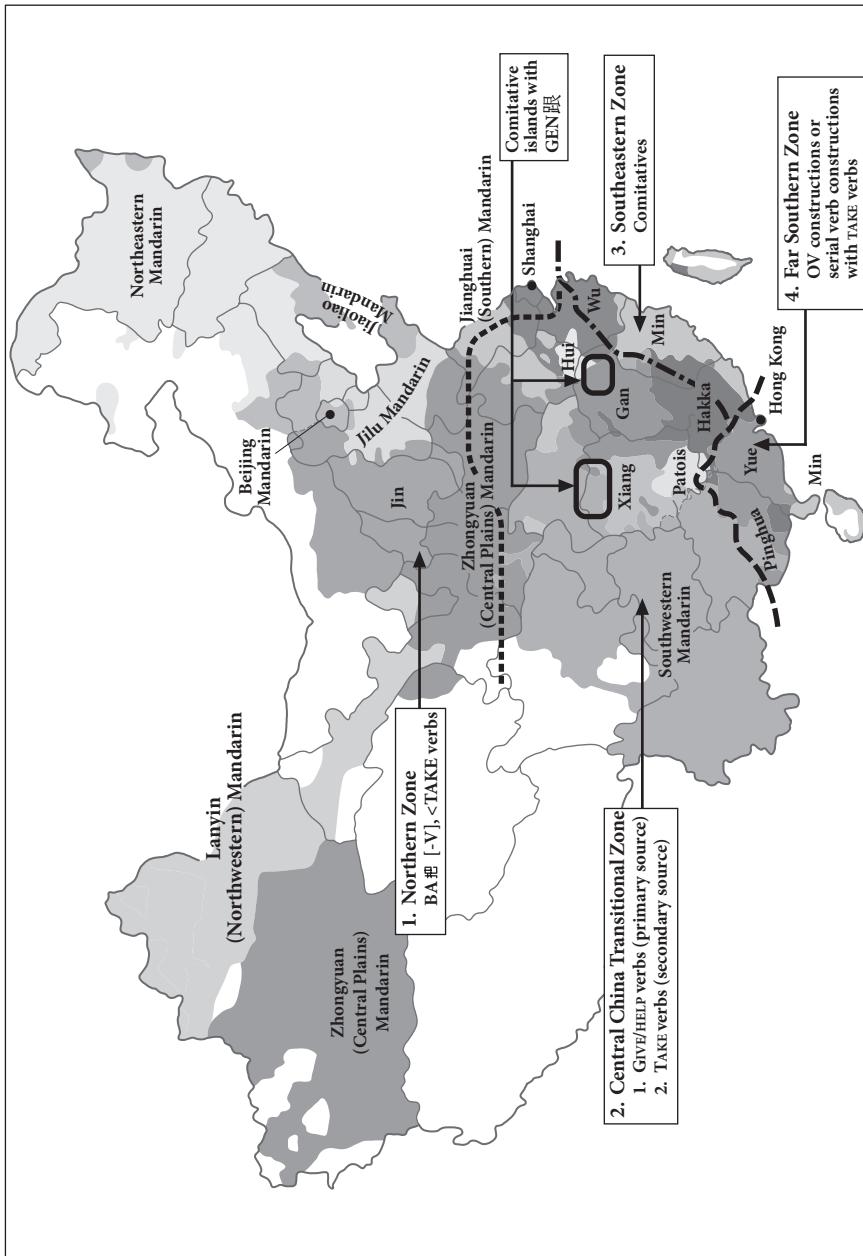
- (i) Agentiveness—the presence of an obligatory agent noun phrase
- (ii) Adversity—a semantic feature coded as part of the overall constructional meaning, to the effect that the event is detrimental in some way for the patient NP.

In canonical passive constructions or in basic passives, as in the analysis of Keenan and Dryer (2007: 338–9), the agent NP should be one that is optional while the event construal should be neutral without any affective interpretations such as adversity. Clark (1974, 1989) has depicted this second parameter as a feature of Southeast Asian passives, which could again be extended to East Asia since it is found not only in Burmese, Thai, Lao, Khmer, and Vietnamese, but also in Japanese, and in all ten branches of the Sinitic languages, not just Mandarin. It is possibly due to these non-canonical features that passives turn out to be textually infrequent (see Xiao et al. 2006 for Standard Mandarin statistics).

The basic form of the Sinitic passive for the most common agentive type has the syntactic configuration as below, followed by an example from Standard Mandarin.

NP_{1(PATIENT)} – [Preposition_{PASSIVE} – NP_{2(AGENT)}] – VP_{COMPLETIVE}

⁹ See Chappell (2006, 2007a, and 2013) for more detailed studies of these object-marking constructions, and particularly Chappell et al. (2011) for a study of comitatives as a source of DOMs.



MAP 2.2. Sources of object markers in Sinitic languages and their distribution

Source: The maps in this analysis display the dominant features for each linguistic area and thus necessarily represent generalizations of the data we have to hand. The boundaries for the areas need to be consequently viewed as approximations.

(7) 曹伟 被 她 吓 了一跳。

Cáo Wěi bì tā xià-le yī-tiào.
(name) PASS 3SG frighten-PFV one-VCL
'Cao Wei was startled by her.'

As the syntactic configuration given in (7) depicts, the agent noun phrase in Sinitic passives is overtly marked morphologically, being introduced by a prepositional marker that has generally grammaticalized from a verbal source. These will be referred to as ‘agent markers of the passive’ in the following discussion.

2.3.1 Sources of passive markers: A crosslinguistic view

With reference to the crosslinguistic studies of passive morphology presented in Haspelmath (1990), Heine and Kuteva (2002), Keenan and Dryer (2007), and Wiemer (2011), the following list of verbs and grammatical morphemes can be compiled as frequent diachronic sources of passive markers in the languages of the world. Of eleven main sources, five are relevant for Sinitic and correspond to the boldface forms in this list.

(8) DIACHRONIC SOURCES OF PASSIVE MARKERS – CROSSLINGUISTIC SYNTHESIS

- (i) stative and inchoative auxiliaries: be, become, have, receive, stay, sit ...
- (ii) come, go
- (iii) fall (down)
- (iv) see
- (v) **suffer, undergo**
- (vi) get
- (vii) pronominal forms: particularly 3PL, and reflexive pronouns
- (viii) comitatives
- (ix) eat
- (x) touch/hit
- (xi) **causative verbs/causative morphology including give**

2.3.2 Passive constructions in Sinitic languages and the sources for their agent markers

Despite the comprehensive nature of these studies, the list presented in (8) does not prove to be exhaustive of all the principal sources of agent markers found in Sinitic passives. As earlier mentioned, Hashimoto (1976a, 1986) also used the parameter of passive markers to argue for the North–South linguistic division for China: verbs of giving in the South, versus causative verbs in the North. However, since his path-breaking work on typology was first published, data newly available on the grammatical description of Sinitic languages allows us to identify at least seven sources and

thus to refine his areal description, noting that these include the highlighted in (8).¹⁰

- (9) Diachronic sources of passive markers for Sinitic Languages
- (i) suffer
 - (ii) get/obtain, take
 - (iii) eat
 - (iv) touch/hit, be in contact with, be close to
 - (v) give
 - (vi) speech act verbs ‘tell’, ‘call’, ‘ask’
 - (vii) wait

First, SUFFER verbs represent a common source in Sinitic, particularly in the Mandarin group of dialects. Second, in a very secondary fashion, GET verbs are also found in Sinitic (including here broadly verbs belonging to the semantic field of TAKE and OBTAIN), while third, there are a few rarer cases of EAT, as in one or two Xiang dialects of Hunan.¹¹ The fourth source of contact verbs, including TOUCH/HIT is widespread in Southwestern Mandarin while the fifth, GIVE, is above all the major source of passive morphology in the Wu, Min, Gan, Xiang, Hui, Hakka, and Yue branches of Sinitic, displaying a wide range of different forms. Speech act verbs which take on a causative function are also a common sixth source in Mandarin dialects. The seventh source of WAIT verbs is found in the Central Transitional zone above all.

Despite this large array of source semantic fields, these can in reality be reduced to three main types, as a function of the semantic change involved in their converging grammaticalization pathways as they develop into agent markers: these are (i) GIVE; (ii) SUFFER; and (iii) causative verbs. The reasons are as follows:

- (a) the CONTACT class verbs develop into SUFFER-type adversative passives due to the inferred feature of affectedness. This can be construed as the result of the actions associated with TOUCH, HIT (A TARGET), MEET (A SURFACE) etc.
- (b) the GET/OBTAIN class of verbs develop the meaning of GIVE as part of a widespread pattern of polysemy in the Central Transitional area and so may be merged under the GIVE source (see also §2.3.2)
- (c) the WAIT class and speech act verbs undergo the same pathway of grammaticalization as for the causative speech act verbs listed in (iv) and may

¹⁰ This analysis is based on the descriptions and data from Li Lan (2006), Li Rulong and Zhang Shuangqing (1997), Wu Yunji (2005), and Cao Zhiyun et al. (2008), in addition to fieldwork data of the ERC SINOTYPE team, including my own.

¹¹ See Y. Wu (2005: 192) on passive markers in Xiang dialects, Chapter 6. The Xiang dialect of Longhui uses ‘eat’ as a passive marker. In older vernacular texts, such as the thirteenth century *Wu Deng Hui Yuan* «五灯会元», (a Song dynasty history of Zen Buddhism in five books), this marker is equally in evidence (Li Lan 2006). Note that in Archaic Chinese, the verb *jiàn* 見 ‘to see’, 為 wéi ‘to do’, ‘to be’ and even *yú* 於 ‘to go’ in Early Archaic Chinese also served as passive markers (see Peyraube 1989a) but we have not so far found these sources in the contemporary dialects.

consequently be combined with them, and ultimately with GIVE as well which passes through a causative stage before becoming a passive marker.

The distribution and characteristics for each of the main lexical sources are next discussed in turn.

2.3.3 GIVE verbs as a lexical source

A large number of Sinitic languages in Central and Southeastern China use verbs of giving as a source of markers used to introduce the agent NP in the passive. Such is the case in the Hunan Xiang dialect of Hengshan which uses *tsæ²⁴* < 'give'.

- (10) Hengshan dialect of Xiang (湘语衡山话)

他今日会得他爷骂。
t^ha³³ t̪iəŋ³³ŋi²⁴ fu²⁴ tsæ²⁴ t^ha³³ iã¹¹ ma²⁴.
3SG today will PASS_{<GIVE} 3SG father rebuke
'He will be rebuked by his father today.' (Wu 2005: 197)

Table 2.4 provides a sample of some of the high frequency verbs belonging to the semantic field of verbs of giving. In each case, focal locations of use are indicated.

Map 2.3 clearly shows how GIVE verbs preponderate as a source for passive markers in Southeastern China and parts of Central China, where Mandarin is not spoken as L₁, yet their distribution also significantly extends to some of the Jianghuai Mandarin dialects spoken in Jiangsu province. This source for an agent marker is unusual typologically. Chappell and Peyraube (2006c) argue that the grammaticalization pathway for GIVE-passives has to pass through a causative verb stage, once the GIVE verb in question has developed a permissive LET causative sense.

TABLE 2.4. GIVE verbs as the source of agent markers in Sinitic languages

GIVE > PREPOSITION INTRODUCING THE AGENT IN THE PASSIVE

Marker	Dialect group
<i>pei²</i> 界	Majority of Yue dialects, many Hui dialects
<i>khit⁴</i> 乞	Most Min dialects in northern and eastern Fujian; Southern Min dialects in Guangdong, the Leizhou peninsula and Hainan, certain Southern Wu dialects
<i>hou⁷</i> 與	Southern Min dialects in the Xiamen (Amoy) area; Taiwanese Southern Min
<i>pun</i> 分	Hakka in Guangdong province and Guangxi Autonomous Region
<i>pø?</i> 拨	Shanghainese and many Northern Wu dialects
<i>ts⁴</i> 得, <i>pa</i> 把,	Xiang dialects, many Gan dialects ¹
<i>pa.ts⁴</i> 把得	
<i>gei³</i> 紿	Hui dialects in Southern Anhui, Jianghuai Mandarin dialects in Jiangsu

¹ This set of verbs may also have the meaning of 'get' or 'obtain', depending on the dialect.

To our knowledge, outside of Sinitic, GIVE has only been attested for the broader Asian region in colloquial Malay, as spoken in northern and western parts of peninsular Malaysia (*bagi* < ‘give’) (Yap and Iwasaki 2003, 2007) and in Manchu (Nedjalkov 1993) and a few other Tungusic languages, where the causative and passive verb morphology is identical. Despite this, the reanalysis of GIVE > CAUSATIVE is on the other hand, extremely common in Southeast Asia, for example, in Thai (Tai-Kadai), Jinghpao (Tibeto-Burman), Khmer (Austro-asiatic), Vietnamese (Austro-asiatic), and Yao (Hmong-Mien) (Jenny, 2015) but appears to stop at this stage.

Oddly enough, GIVE-constructions have developed into impersonal or agentless passives in certain non-standard German dialects, including Luxembourgish and Moselle Franconian (Moselfränkisch) dialects (Glaser 2006; Lenz 2008).¹² Note however that both the construction type and the grammaticalization pathway are quite distinct from that found in Sinitic and can be explained in terms of an initial development into a copular auxiliary use (Lenz 2008). Sinitic GIVE-passives are agentful due to their origin in causative constructions and generally do not allow the agent noun to be omitted as described above in §2.3.:

$$\text{NP}_1(\text{patient}) - \text{Preposition}_{\text{passive}} (<\!\!\text{V}_1(\text{give})\!>) - \text{NP}_2(\text{agent}) - \text{V}_2 - \text{X}$$

Finally, Wiemer (2011) observes that GIVE may form a reflexive-permissive passive in West Slavic languages, including Polish and Czech, but that these have not yet proceeded to his two final stages of modal and ‘real’ passives. This resembles more closely the Sinitic pathway for the GIVE-passive, which, has by way of contrast, reached maturity.

2.3.4 SUFFER and CONTACT verbs as a lexical source

The second main source of agent markers in Sinitic passives belongs to the field of SUFFER verbs, including verbs of contact. In fact, SUFFER-passives in Sinitic appear to be intimately related to verbs that mean ‘be close to’, ‘meet’, or ‘be in contact with’. As observed above, CONTACT class verbs extend their meaning to the adversative sense of ‘suffer’ during the course of their semantic evolution. The main marker of the passive in Standard Mandarin is *bèi* 被, one of whose earliest meanings was ‘to put on the body’, ‘to cover’ (Peyraube 1989a) and hence, similarly, has its source in a CONTACT verb, which came to mean ‘to suffer’ through the process of semantic extension.

Table 2.5 illustrates some of the main lexical verb sources for agent markers of the passive in Sinitic languages.

¹² I thank Elvira Glaser and Volker Dellwo for kindly bringing this important information to my attention during a research talk which I presented at the University of Zurich on 9 May 2014 and for their data on Luxembourgish and the Trier dialect respectively.

TABLE 2.5. SUFFER and CONTACT verbs as the source of agent markers in Sinitic languages

'SUFFER' > PREPOSITION INTRODUCING THE AGENT IN THE PASSIVE

MARKER		DIALECT GROUP
被 BEI	'suffer' < 'cover', 'put on body'	Standard Mandarin, formal registers of many dialects; and dialects along northern borders of Yangtze River in Jiangxi, Anhui, and Jiangsu
着 ZHUO	'to hit the target', 'to touch'	Southwestern Mandarin dialects (e.g. in Hubei, Hunan, Guizhou, Sichuan, and Yunnan); a handful of Jiaoliao and Jilu Mandarin dialects in Shandong
挨 AI	'be close to'	Southwestern Mandarin and Pinghua dialects in Guangxi; Jianghuai Mandarin dialects in Jiangsu province
遭 ZAO	'to suffer', 'to meet'	Ningxia Mandarin (Central Plains)

What should be observed from comparing Table 2.5 with Map 2.3 is that even though the *bēi* passive, a prototypical example of a SUFFER or adversative passive, appears to be largely a Mandarin trademark, it is not widespread in the Northern area as the native or 'local' colloquial marker of the passive (see §2.3.5 on causative sources). It should be pointed out, however, that the marker *bēi* is closely associated with the history of passive constructions in Chinese, and is attested in this use from the beginning of the Medieval period (third–thirteenth century, Peyraube 1989a).

The *bēi* passive also tends to have a relatively higher frequency in corpora of contemporary Standard Mandarin or *pǔtonghuà*, based mainly on written genres and particularly in its agentless form, when compared with other passive constructions (for statistics, see Xiao et al. 2006; on syntactic and semantic properties, see Chappell forthcoming b).

(11) Agentless form of the *bēi*-passive in Standard Mandarin

NP_{Subject/Patient} – Passive Marker – [Ø]_{Agent} – VP_{Transitive}
那個鬼就**被**趕出來了。

Nāge guǐ jiù **bēi** gǎn.chū.lai le.
that.CL ghost then PASS chase.out.come CRS
'The ghost was chased out.'

By way of contrast, it is the Southwestern area where SUFFER verbs come to the fore as colloquial passive markers, particularly in Yunnan, Guizhou, Sichuan, Hubei, and western Hunan, but also in the western parts of Guangxi Autonomous Region. This incorporates a large area where Southwestern Mandarin is spoken and ZHUO 着 is used as the marker (Li Lan 2006), although it is not the only SUFFER-verb, since in Guangxi, many Pinghua dialects make use of a marker from the same semantic field, AI 挨 'to be in contact with'. An example follows for each of these CONTACT verbs.

- (12) Yiyang dialect of Xiang (湘语 益阳话)

他在路上着狗咬了一口。

xa³³ tai¹¹ ləu¹¹-lɔ¹¹ ts^ho⁴⁵ kau⁴¹ ɳa⁴⁵ ka i⁴⁵ k^hau⁴¹.
3SG at road-on PASS dog bite ASP one CL

‘He was bitten by the dog along the way.’ (Cui Zhenhua 1998: 140)

- (13) Yining Pinghua (义宁平话)

个蟆挨(只蛇)吃去呃。

ko³³ maʒ⁴ ɳai³¹ (tci⁵ ɳiə³¹) hiə?⁵ həu³³ e³³.
CL frog PASS CL snake eat away PFV

‘The frog was eaten up (by the snake).’ (Zheng and Lin 2005: 252–3)

It can be noted again that this southern area of China is contiguous with the linguistic area of Southeast Asia in which passive markers have evolved from the similar source of a CONTACT verb: these include Khmer (Austroasiatic), Thai and Lao (both Southwestern Tai), Hmong (Hmong-Mien), and also Tibeto-Burman languages such as Jinghpao (Kachin). Vietnamese uses the marker *bị*, part of its extensive Sino-Vietnamese vocabulary.¹³

It is significant, however, that the use of 被 BÉI is found scattered across certain Mandarin-speaking regions of Central China along the north side of the Yangtze River in Hubei, Anhui, and Jiangsu provinces (Map 95, Cao et al. 2008) rather than causative speech act verbs as the preferred vernacular form in Northern China. This suggests a peripheral area for its use, combined with the fact that BÉI is intermingled with GIVE-class verbs as markers in the adjacent transitional zone, as displayed in Map 2.3 (cf. Hashimoto 1987, 1988 on passives). In contrast to this, in the case of non-Mandarin dialects, the use of 被 BÉI is typically a Mandarin borrowing, for example, in the Hunan Xiang dialects (cf. Wu 2005) or in formal registers of Hong Kong Cantonese (Matthews and Yip 2001).

2.3.5 Speech act causative verbs as a lexical source

The most typical source of agent markers in Northern Chinese passives is not *bèi*, however, but the coercive causative verb *jiào 叫* ‘to make’ that has its source in a speech act verb ‘to tell’. This distribution can clearly be observed from Map 2.3. Secondly, the use of *ràng 让* ‘to let’ has developed a permissive causative use from its earlier meanings of ‘to yield’, ‘to request/ask’. Its use is intermingled with that of *jiào 叫* in the Northeast (Manchuria) and in Shanxi and Shandong provinces, but is spread over a slightly less extensive area than for *jiào 叫* (see Map 95 in Cao Zhiyun et al. (2008) for the precise areas and localities). In the northern parts of Jiangxi province, many Gan dialects use WAIT verbs as their source, which can be shown to pass through the stage of a permissive LET causative (Chappell and Liu, to appear).

¹³ This marker quite likely represents a borrowing of the morpheme BÉI from the Medieval period, when Vietnam was under Chinese domination.

TABLE 2.6. Causative verbs as the source of agent markers in Sinitic languages

CAUSATIVE VERB > MARKER INTRODUCING THE AGENT IN THE PASSIVE

MARKER		DIALECT GROUP
叫 JIAO	'to tell' > 'make'	Mandarin – all northern groups: Northeastern, Beijing, Jiaoliao, Jilu, Central Plains, and Northwestern; Jin dialect group
让 RANG	'to request/ask' > 'let'	Mandarin – Northeastern, Shandong, Shanxi provinces
等 DENG	'to wait' > 'let'	Gan, Xiang, Wu, and Hakka dialects
听任 TÍNGRÈN	'to let it be'	Xiang
盡 JIN	'let'	Xiang, Mandarin

A representative list of causative verbs which develop into agent markers in the passive is provided in Table 2.6.

From original lexical meanings of 'tell' (*jìao* 叫), 'request', or 'allow' (*ràng* 让), these lexical verbs evolve from their use in pivot constructions of the form 'tell/allow someone to do something' to the constructional meaning of a causative: 'make/have or let someone do something'. (14) is an example from a Jin dialect for the use of *jìao* 叫 as an agent marker.

NP_{Subject/Patient} – Passive Marker_{<Causative} – NP_{Agent} – VP_{Transitive}

- (14) Jinyuan dialect of Jin 晋源方言

碗儿叫我打咧。

vang⁴²æ¹¹ tçiau³⁵ γy⁴² ta⁴² lie¹¹.
 bowl PASS<_{TELL} 1SG strike PFV

'The bowl was broken by me.' (Wenqing Wang 2002: 194)

The less well-known causative source of WAIT verbs may, even so, be traced back to the Southern Song dynasty (1127–1279) for DAI 待 and to the Yuan dynasty for DENG 等 (1206–1368) (Peyraube and Liu 2013). According to textual material from these periods, the two verbs could already express the permissive 'let' sense, hence providing the ideal conditions for a further development into an agent marker in the passive, just as for the speech act and GIVE class verbs. Here is a contemporary example from Yichun Gan (data from XuPing Li).

- (15) Yichun dialect of Gan 益春方言 (赣)

车子等人家骑走哩。

tçia³⁴tsi tñ⁴²⁻³³ nin⁴⁴ka tç^{hi}⁴.tsεu li.
 bike PASS<_{WAIT} other ride.away PFV

'The bike was ridden away by someone.' (data from XuPing Li)

The passive use is found in the Gan and Wu groups above all, but also in a few Hakka and Xiang dialects (see Hu and Ge 2003; Zheng Wei 2007; Chappell and Liu to appear).

2.3.6 TAKE~GIVE verbs as a lexical source

A residual class of TAKE verbs represents a semantic field less well-exploited in Sinitic for passive markers than its antonym GIVE. In fact, dialectal data need to be handled very carefully in this case, since the forms do not necessarily have the same meaning as their cognates in Standard Mandarin: for example, *ná* 拿 ‘take’, ‘hold’, *bǎ* 把 ‘hold’, and *dé* 得 ‘get’, ‘obtain’ may have ‘give’ as their primary meaning in the relevant dialects of Gan and Xiang or Hakka.

This class of ambiguous verbs which develop into agent markers of the passive is represented in Table 2.7.

TABLE 2.7. TAKE ~ GIVE verbs as the source of agent markers in Sinitic languages

‘TAKE’ > PREPOSITION INTRODUCING THE AGENT IN THE PASSIVE

Marker		Dialect group
拿	NA	‘take’, ‘hold’
得	DE	‘to obtain’ ~ ‘to give’
捞	<i>lau</i> ³³	‘to dredge up’
搁	<i>laq</i> ⁷	‘give’ or ‘take’

As shown on Map 2.3, the ambiguity between TAKE- and GIVE-class verbs is strikingly evident in Jiangxi and in the adjacent eastern parts of Hunan province (see example (10) with *tsæ*²⁴ 得 ‘give’). Furthermore, the identical forms may be used as object markers in the structurally isomorphic ‘disposal’ construction in one and the same dialect (see Chappell 2007b; Wu 2013).

In Changning, which is a Gan/Hakka dialect spoken in Hunan, 得 *te*³³ ‘give’ acts as a marker of both the agent in the passive and the patient in the object marking construction, whence the ambiguity of the following example.

- (16) Changning dialect of Gan/Hakka (常宁话)

- (a) Object marker reading:

NP ₁	AGENT	-	[Preposition	-	NP ₂	DIRECT OBJECT] -	VP _{TELIC}
爷爷	得	佬佬	打	哒	一	餐	饱	个。
ia ¹¹	ia ¹¹	te ³³	lo ⁴⁴	lo ⁴⁴	ta ⁴⁴	ta	i ³³ ts ^h å ⁴⁵	po ⁴⁴
father	OM	brother	hit	ASP	one-CL	enough	PRT	ke.

‘(My) father gave my younger brother a big slap.’

(b) *Passive marker reading:*

NP₁ _{PATIENT} – [Preposition – NP₂ _{AGENT}] – VP_{TELIC}

爷爷得佬佬打哒一餐饱个。

ia¹¹ia¹¹ te³³ lo⁴⁴ lo⁴⁴ ta⁴⁴ ta i³³ ts^hā⁴⁵ po⁴⁴ ke.
father PASS brother hit ASP one-CL enough PRT

‘My father was given a big slap by my younger brother.’

(data from Wu Yunji 2005: 201–2)

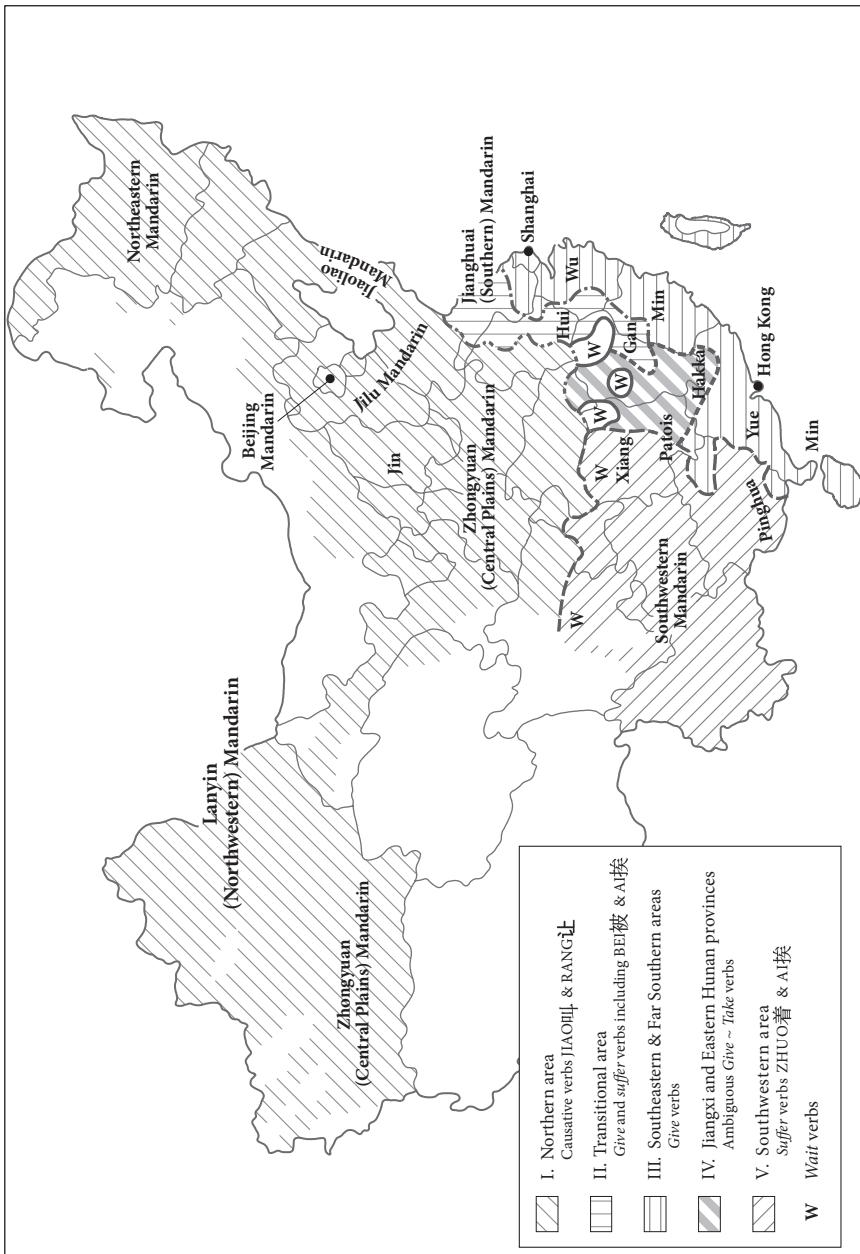
As an object-marking construction, the subject NP, *ia¹¹ia¹¹* ‘father’, is understood as the agent of the action who slaps his son, whereas under the interpretation of a passive construction, it is instead the father who is in the role of the patient and receives the slap from his son.

To sum up this discussion on agent markers found in Sinitic passive constructions, the five main linguistic areas are represented in Map 2.3. Note that zones II, III, and IV could be merged together as one larger GIVE area with the internal divergences, as indicated on the map, the outcome of overlapping with SUFFER areas or causative verb areas (including WAIT) or due to the presence of a zone of GIVE~TAKE ambiguity.

In conclusion, we have set up five main areas for the lexical sources of agent markers in the Sinitic passive constructions. The Northern area of causative verbs, and the SUFFER verb, BĒI, in the standard language, is the largest geographically, encompassing all the provinces and autonomous regions from the northwest of China across to Manchuria in the northeast, and thence south to the Yangtze River. This area includes the following subgroups of Mandarin: Lanyin (Northwestern), Central Plains, Beijing, Northeastern, Jiaoliao, and Jilu as well as dialects belonging to Southwestern Mandarin. Notably, the Jin dialect group falls into this area as well.

The Southwestern area of CONTACT class verbs forming SUFFER-type passives covers Yunnan, Guizhou, the southern region of Sichuan, parts of Hubei, and the western regions in both Hunan and the Guangxi Autonomous Region. It includes a large proportion of the Southwestern Mandarin subgroup, some Pinghua dialects in Guangxi and a few Xiang dialects. We pointed out that it is precisely this area which is contiguous with a large number of unrelated Southeast Asian languages that also use CONTACT verbs as a source for agent markers.

The Southeastern and Far Southern areas of GIVE verbs encompass the coastal provinces of Jiangsu, Zhejiang, Fujian, and Guangdong, displaying a variety of forms from this lexical field which first develop a causative use then grammaticalize into agent markers of the passive. This is a typologically unusual feature found, to date, almost exclusively in Sinitic, as far as the East and Southeast Asian linguistic area is concerned, with the exception of colloquial Malay to the south and certain Tungusic languages to the north.



MAP 2.3. Sources of agent markers in Sinitic passives

Source: This map is based on an analysis of the data given in Cao Zhiyun et al. (2008), *Linguistic Atlas of Chinese Dialects*—Map 95, in addition to the cited reference materials.

The area for GIVE verbs is adjacent to two other smaller, highly mixed areas that also implicate the use of GIVE verbs. The first one is the transitional area in southern Jiangsu and adjacent Anhui provinces where both GIVE verbs and the SUFFER verbs, BĒI and AI are found intermingled. Jianghuai Mandarin and Hui dialects are well represented in this area. The second mixed area is in eastern Hunan and Jiangxi provinces where Xiang, Gan, Hakka, and Wu dialects are all to be found. In this area, ambiguous GIVE and TAKE verbs act as the source for agent markers, as well as causative verbs derived from WAIT, found principally in the Gan dialects of northern Jiangxi, but also in some of the contiguous Wu dialects and Xiang and Hakka as well. Both these two highly mixed areas might have been subsumed under a larger, more generalized GIVE area. Renouncing the potential of such an economy, we have instead chosen to distinguish them as being turbulent transitional areas, sandwiched between the Northern area of causative verbs, on the one side, and the two combined areas of Southwestern SUFFER and Southeastern GIVE on the other side.

In the final section which follows, we discuss the different structural and cognitive types of comparative constructions to be found in Sinitic languages.

2.4 Sinitic comparatives

The comparative constructions of inequality fall structurally into seven main types in Sinitic languages and these seven types involve four cognitive schemas, following Heine (1997) in the main. The ones relevant for Sinitic are given in boldface script in Table 2.8.¹⁴

TABLE 2.8. Cognitive schemas for the comparative of inequality (based on Heine 1997)

Type	Cognitive schema	Example
1:	Source	'from'
2:	Goal	'to'
3:	Location	'at'
4:	Compare	'compared to'
5:	Action	'surpass, defeat'
6:	Polarity	'X is A, Y is not A'
7:	Sequence	e.g. Germanic 'than'
8:	Similarity	'as, like'
9:	Topic	'X and Y, Y is A'

¹⁴ Heine (1997) presents a set of eight cognitive schema to which Chappell and Peyraube (this volume) have argued for the addition of a dependent-marked *Compare* comparative.

I use the following notation to describe the different structural types for the comparative, as do Chappell and Peyraube (this volume):

CM	=	comparative marker
NP _A	=	noun phrase referring to the comparee
NP _B	=	noun phrase referring to the person or thing acting as the standard or benchmark
VP	=	verb phrase filled principally by an adjective or a verb as well as other predicative elements
DEGR	=	degree marker

Table 2.9 presents the seven different structural types of comparatives found in Sinitic and their corresponding cognitive schema.

TABLE 2.9. Structural types of comparatives in Sinitic

Structural type	Structural configuration	Cognitive schema
I. Prepositional	NP _A [CM NP _B]VP	Compare
II. Transitive	NP _A VP CM NP _B	Action _(i) (Surpass)
III. Zero-marked	NP _A VP NP _B (Q-CL)	Action _(ii)
IV. Adverbial	NP _A CM _{MORE} Verb NP _B	Action _(iii)
V. Hybridized	NP _A [CM NP _B]CM _{MORE} V	Compare + Action _(iii)
VI. Topic-comment	NP _B // <i>copula</i> NP _A VP	Topic
VII. Contrastive conjoined clauses	NP _A VP _x CL _{PL} , NP _B VP _{~x} CL _{PL}	Polarity

Next, I briefly discuss and exemplify each of these types in turn.

2.4.1 Type I: Prepositional comparative: Compare schema

Similar to the case for the object-marking construction (§2.2), the best described comparative construction for Sinitic is the Mandarin type using the marker *bì* 比 ‘compared to’, in the *Compare* schema.

Type I: Preverbal Prepositional Phrase: *Compare* schema

- (17) NP_A – CM – NP_B – VP (Standard Mandarin)

你比我大。

Nǐ_A bǐ_{CM} wǒ_B dà_{ADJ.}

2SG compared:to 1SG old<^{big}

‘You’re older than me.’ (*Literally: you-compared:to-me-old*)

The comparative marker is part of a prepositional phrase formed with the standard NP, NP_B, and as such represents a dependent-marking strategy which is at odds

with the typological profile of Sinitic languages. As Li Lan (2003), Ansaldi (2010), and Chappell and Peyraube (this volume) point out, Type I is widespread in Northern China, while it is gradually being adopted elsewhere in other Sinitic language groups, in particular, for the Wu, Hui, Xiang, and Gan dialect groups in the Central Transitional area where the native and the Mandarin replica form exist side by side. In local varieties of Taiwanese Southern Min spoken by the younger generation, the *Compare* comparative is on the point of accomplishing the task of wholly replacing the native comparative (see Li and Lien 1995).

2.4.2 Type II: Transitive structure: Action_(i) schema with Surpass- class verbs

This comparative type structurally equates to a transitive clause, in which the comparative marker acts as a complement to the verb, in original V₂ position. It has a transparent synchronic relation to a fully lexical verb in most dialects with a meaning such as ‘surpass’, ‘defeat’, or ‘win’. Type II exemplifies the Action cognitive schema. It has widespread usage in southern and southwestern China in particular in the Yue, Hakka, and Southwestern Mandarin groups, not to mention in the Central Transitional area for many Xiang, Gan, and Wu dialects. As Li Lan (2003) points out, this comparative may also be found in the Jiaoliao and Jilu Mandarin dialects of Shandong province which suggests a peripheral area of retention.

The comparative marker GUO [k^wɔ³³] 过 in Hong Kong Cantonese, exemplified in (18), has the lexical verb meaning of ‘cross (over)’ and ‘surpass’ as well as serving as a directional complement verb ‘over’ in the original V₂ position.

- (18) NP_A– VERB – CM –NP_B (Hong Kong Cantonese)

我	老	过	你。
ŋɔ ¹³	lou ¹³	k ^w ɔ ³³	lei ¹³ .
1SG	old	CM	2SG

‘I’m older than you.’ (*Literally: I-old-surpass-you*)

The fact that Type II *Surpass* comparatives are predominant in the Far Southern area of China also serves to recall the contiguity of this area with unrelated Southeast Asian languages, where this type is equally common (see §6.4.3 in Chappell and Peyraube, this volume). Furthermore, in terms of areality, this feature of *Surpass* comparatives overlaps with the one for absence of grammaticalized object-marking constructions in these two adjacent areas (discussed in §2.2).

There are at least two types of comparative constructions in Sinitic which are typologically rare in the languages of the world, according to predictions made on the basis of a new sample of languages in Haspelmath et al. (2013). These correspond to Types III and IV, next discussed.

2.4.3 Type III: Zero-marked structure: Action (ii) schema

A majority of Sinitic languages and dialect groups may use this third comparative type as a *secondary* strategy (Li Lan 2003). In other words, the Zero-marked comparative is not areally defined within China. Its first striking feature is that it is devoid of any morphological marking whatsoever to explicitly code the comparative. There is neither a degree nor a standard marker.

A second important feature of the Zero-marked comparative is that it is mainly used with stative or quality verbs derived from adjectives such as ‘tall’, ‘old’, ‘fat’, and ‘rich’, that is, attributive adjectives used as predicative ones. Furthermore, a third, syntactic, feature concerns the unusual deployment of such stative verbs in a transitive clause; see (19). Evidently, transitive SVO syntax is not otherwise normally permitted with this class of stative verbs in Sinitic languages. It could be seen as a kind of semantic coercion of this verb class which places them in a transitive ‘straitjacket’, thus creating a new comparative structure. The Type IV *Adverbial* comparative, to be discussed next, shows similar syntax and constraints (§2.4.4).

Apart from Southern Min, the stative verbs found in Type III are typically quite semantically restricted to those which denote physical characteristics, size, and age. Furthermore, a quantifier phrase may be required in clause-final position after the standard NP, the case in most branches of Sinitic. For example, Standard Mandarin requires a quantifier phrase as in (20) whereas, once again, Southern Min stands apart from other Sinitic dialect groups in not requiring it, as shown in (19). The use of transitivized stative verbs points to the Action schema once again.

- (19) Hui'an 惠安 (Southern Min 闽南)

NP _A	Verb _{STATIVE}	NP _B	(Q-CL)
伊	大	我。	
i ³³	tua ⁴²	gua ⁵⁵ .	
3SG	old< ^{big}	1SG	

‘S/He is older than me.’ (*Literally: She-olds-me*)

- (20) Standard Mandarin 普通话

NP _A	Verb _{STATIVE}	NP _B	Q-CL
哥哥大我三岁。			
Gēgē	dà	wǒ	*(sān sui).
brother	old< ^{big}	1SG	three year

‘My brother is three years older than me.’

The Type III comparative appears to be a form that has reached a stage of maturity in Southern Min dialects not yet evident in many other branches of Sinitic. This is shown in its ability to take a wider range of stative predicates than is possible in most other dialect groups of Sinitic, and in the fact that it does not require a quantifier

phrase (see §10.5 in W. Chen, this volume).¹⁵ It is noteworthy that this structure has been in use since the period of Classical Chinese (fifth–third centuries BC) (Peyraube 1989b).

2.4.4 Type IV: *Adverbial comparative: Action (iii) schema with ‘more’*

This fourth type of comparative similarly belongs to the Action schema. Its first interesting feature, which distinguishes it from Type III, is that it makes use of an intensifying adverb with the general meaning of ‘more’ as its comparative marker of degree. A second distinctive feature is that it is only found in a small area of China, in Min and in certain Hakka dialects.

Otherwise, the *Adverbial* comparative resembles the Zero-marked one in being used with stative or quality verbs in transitive clauses, but in this case, in conjunction with the adverbial marker ‘more’, as shown in (21): Adverb_{CM} + Verb_{STATIVE}.

- (21) Hui'an 惠安 (Southern Min 闽南)

NP _A	CM _{<MORE}	Verb _{STATIVE}	NP _B
伊	較	富	我。
i ³³	k'a? ⁴	pu ⁵⁵	gua ⁵⁵ .

3SG more_{CM} rich 1SG

‘She is richer than me.’ (*Literally: she-more-riches-me*)

Type IV is highly typical of many Southern Min dialects. W. Chen (this volume, §10.4) suggests that Type IV might be native to the Southern Min area, given that this comparative structure is also attested in historical documents recording this language. Uncommon both in Sinitic and crosslinguistically, this type has been predicted to be rare by Haspelmath et al. (2013).¹⁶

2.4.5 Type V: *Hybridized construction: Compare + Action schemas*

Type V generally represents a blend of Types I and IV, that is, a combination of the *Compare* and the *Adverbial* comparatives which results in a new hybrid form. For languages using this hybridized comparative, Type V is in fact a borrowing of both the syntax of the Northern strategy and its morphological marker *bì* ‘compared to’, combined with the marker of the *Adverbial* comparative, relevant for the given Sinitic

¹⁵ Apparently, the Zero-marked Type III and the *Adverbial* Type IV are neither well attested nor well described in the typological literature. In a recent study by the Comparative Constructions Consortium, the prediction is made that this zero-marked type is not possible on the basis of a new sample of 230 languages, complementing Stassen’s sample of 167 (Haspelmath et al. 2013). The Comparative Consortium has proposed Universal C1: «No language lacks both a degree marker and a standard marker» e.g. ‘The dog is big the pig.’ (2013: 8). This structure corresponds exactly to that of our Type III.

¹⁶ Haspelmath et al. (2013) propose the following Universal C2: *Languages with only a degree marker are rare* (page 9), which can be upheld for Sinitic, particularly for Type IV in the Min dialects.

language.¹⁷ Since the markers for these two different types of comparative construction occur in different parts of the clause, in a preverbal prepositional phrase in Type I, and in the predicate in Type IV respectively, the process could be seen to take place harmoniously, without loss of any constituents. Interdialectal contact thus leads to hybridization through the creation of a new form using the Type I *bì* 比 structure, in which the native adverbial marker has been simply retained in its original place in the predicate, modifying the verb in preverbal position.

The example which follows comes from Meixian Hakka, the prestige variety spoken in northeastern Guangdong province. The two comparative markers are, as already foreshadowed, the Mandarin preposition *pi*³¹ ‘compared to’ in preverbal position and the Hakka adverbial *kuo*⁵³⁻⁵⁵ ‘more’ in the predicate,¹⁸ derived from the ‘surpass’ or ‘exceed’ verbal meaning.

(22) *Hybridized comparative construction in Meixian Hakka* (梅县客家话)

NP _A [CM _(i) NP _B] CM _(ii) more Verb
這隻比那隻過好。
e ³¹ tsak ¹ pi ³¹ e ⁵³⁻⁵⁵ tsak ¹ kuo ⁵³⁻⁵⁵ hau ³¹ .
this-CL CM that-CL ADV _{MORE<‘surpass’} good
‘This one is better than that one.’
(Literally: this-one-compared-to-that-one-more-good)

Another example for the outcome of interdialectal contact is the Southern Min *Hybridized comparative* which similarly uses the marker *pi*² 比 ‘compared to’ but a different adverb, *khah*⁴ 較 ‘more’ (= *k'a?*⁴ in (21)).

(23) *Hybridized comparative construction in Taiwanese Southern Min*

NP _A [CM _(i) NP _B] CM _(ii) (more) V
我比伊較高。
gua ² pi ² i ¹ khah ⁴ kuaiN ⁵ .
1SG CM _(i) 3SG CM _(ii) tall
1SG compare 3SG more tall
‘I am taller than him.’

The morpheme for the intensifying or degree adverb ‘more’ varies across dialects, for example, it may be *k'a?*⁴ 較 + ADJ, as in many Southern Min dialects, *kau*⁴ 較 + ADJ,

¹⁷ Some of the dialects concerned may only possess the short variant of the Type IV *Adverbial Comparative* with NP_A+ADVERB_{CM}+ADJECTIVE, that is, an absolute comparative such as found in Southern Min (see W. Chen, this volume). In these cases, the process of hybridization is even more straightforward, since no rearrangement of constituents is required, the standard NP not being overtly mentioned in this simpler structure.

¹⁸ Li Lan (2003: 217) lists another double-marked type of comparative for which only data from Guangdong Hakka is given. In this variant, the stative adjective is marked by both the preverbal adverb JIAO 較 ‘more’ and the postverbal GUO 过 ‘surpass’, that is, it is a hybridization of the Type II *Surpass* and Type IV *Adverbial* comparatives. Only one example is given without transcription from Fengshun Hakka.

or *ko*⁴ 過+ADJ, as in Hakka, and even *kien*⁵ 更 + ADJ in certain Gan dialects (see Li and Zhang 1992: 450 for Gan and Hakka dialects and W. Chen, this volume, §10.2, on the double-marked comparative in the Hui'an dialect of Southern Min).

2.4.6 Type VI: Topicalization

A small number of Wu and Hui dialects show a predilection for this unusual strategy in Sinitic whereby the comparee and the standard NP have their order reversed: the standard NP is placed first in topic position. Note that, apart from the pan-Sinitic zero-marked Type III, this is the only strategy in this group of dialects for coding the comparative of inequality.

- (24) Fuyang dialect of Northern Wu

NP_B copula NP_A VP

法国，还是中国大。

fǎi⁵¹kuo²⁵¹ fua²³ zì³¹³ tçion⁵³kuo²⁵¹ du³¹³
France still be China big

‘Compared with France, China is bigger.’ (Data from XuPing Li)

(Literally: France: still-be-China-big)

Li Lan (2003) explains this type in terms of the conflation of two conjoined clauses after the comparee, ‘China’, is omitted from the putative first clause: ‘Comparing China with France: it is still the case that China is bigger.’ > As for France, China is bigger.’

2.4.7 Type VII: Polarity schema

In Xianghua, an unclassified Chinese language spoken in Northwestern Hunan, the basic comparative strategy uses a polarity schema, with conjoined clauses containing the standard NP and the comparee NP respectively: ‘X is A, Y is B’. The stative verbs in the predicates of each clause are antonyms of each other, and the plural classifier *sa*⁵⁵ 些 ‘a.little’ is obligatorily present.

- (25) Polarity schema: NP_A VP_x CL_{PL}, NP_B VP_{~x} CL_{PL}

Xianghua

你肥些，我瘦些。

n̄i²⁵ fi²¹³ sa⁵⁵, wu²⁵ ua⁵⁵ sa⁵⁵
2SG fat a.little 1SG thin a.little

‘You’re fatter and I’m thinner,’ => ‘You’re fatter than me.’ (fieldwork data of the author)

This is the regular pattern for the comparative in Xianghua, representing a type that has not yet been well studied for Sinitic. Similar monoclausal patterns of the

short comparative variant (the absolutive) can be observed in other dialect groups but these usually do not constitute the main comparative strategy, in such cases.

- (26) Standard Mandarin absolutive construction

他快一点。

Tā kuài yídiǎn.

3SG fast a:little

‘She’s a little faster.’

Apart from the ‘native’ comparative, the Mandarin *Compare* comparative has also been borrowed into Xianghua, just as in Southern Min and Hakka dialects. In a similar fashion to the *Hybridized Type V* with double marking, the Mandarin comparative is combined with the only overt marker in the Xianghua *Polarity* comparative: the postverbal plural classifier, sa⁵⁵ 些. This could be viewed as a kind of hybridization, though on a much smaller scale than for Type V.

- (27) NP_A pi²⁵(CM₁) NP_B VP sa⁵⁵ (CM₂)

你比我 肥些。

nǐ²⁵ pi²⁵ wu²⁵ fì²¹³ sa⁵⁵.

2SG CM₁ 1SG fat a:little

‘You’re fatter than me.’

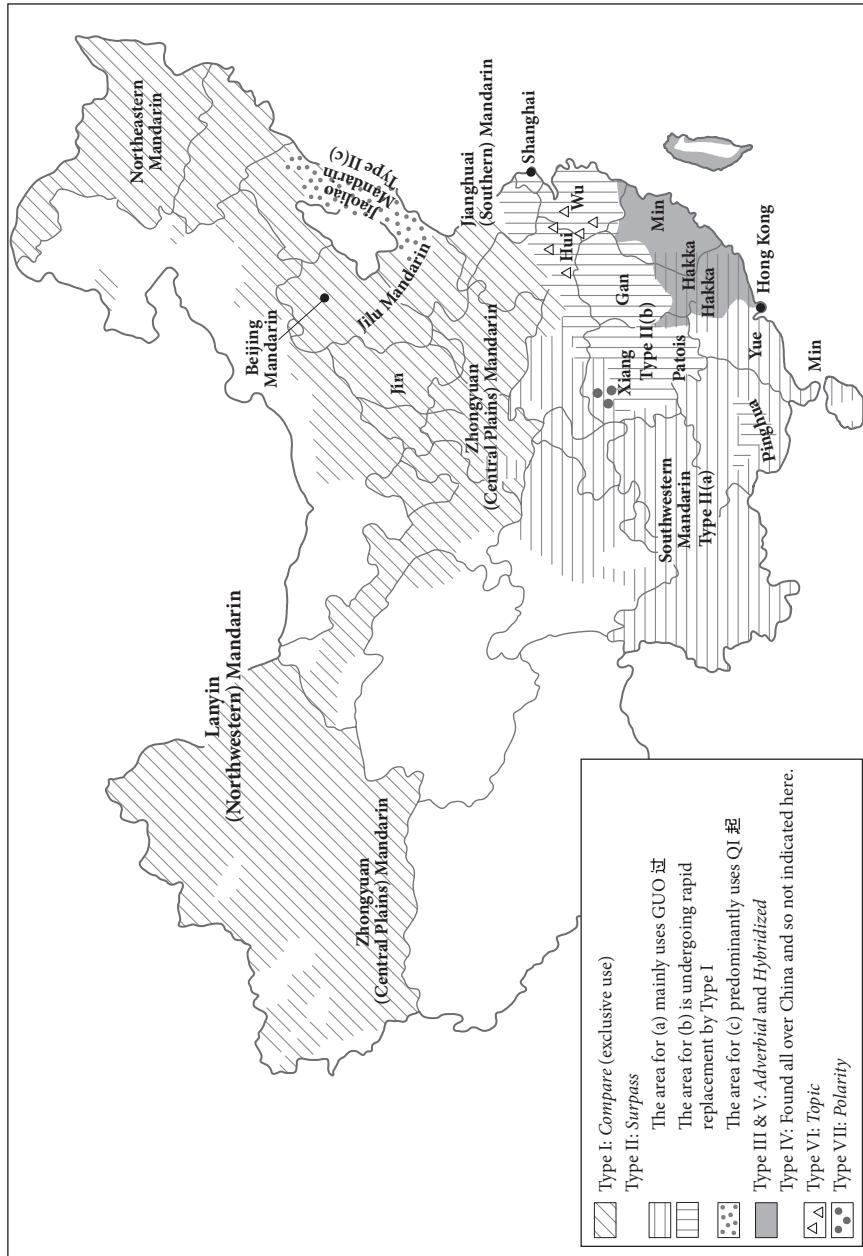
To sum up this discussion, the distribution of the main structural types of Sinitic comparatives is indicated on Map 2.4.

2.4.8 Summary of comparatives

Unlike the case for the object-marking construction, most dialects will have at least two common colloquial strategies for coding the comparative meaning. In the Northern area, which is largely Mandarin territory, we find Type I, the prepositional *Compare* comparative, *She compared to me is tall*, co-existing with Type III, the zero-marked comparative *She tall-s* me by one inch.

Type II, the transitive *Surpass* comparative, *I tall-surpass you*, is widely used across the Far Southern area and the Southwest where Yue, Hakka, Pinghua, and Southwestern Mandarin are spoken. It is similarly used alongside the Type III zero-marked comparative.

Significantly for our study, the Shandong peninsula forms a relic area of retention for Type II (*I tall-surpass you*), a dialect island within the Northern territory of Type I, in addition to parts of Northeastern Fujian, exemplified by the Fuzhou dialect. Type I is encroaching on the native Type II in the Central China transitional zone, for example, in the Gan and Xiang dialect groups, located in Jiangxi and Hunan provinces.



MAP 2.4. Comparative constructions in Sinitic languages

Historically, the Type I *Compare* comparative began to appear only in the late Tang period from the tenth century onwards (see Chappell and Peyraube, this volume). The prestige ascribed to the language of the imperial court—and today to Standard Mandarin—has seen this Type I or *比* 比 comparative move slowly into all dialect areas, either borrowed in its entirety to replace the native structure, or used side by side with the native form as an alternative comparative strategy, when not hybridized with local strategies as in Southern Min, Gan, Hakka, and Xianghua.

It was observed that the Type II *Surpass* comparatives are very common across South-east Asia. This feature appears to dovetail once again neatly with the fact that the *Surpass* schema—and its transitive structure—are also found right across Southern and South-western China. Type III, the zero-marked comparative belonging to the Action schema, is a pan-Sinitic feature that codes in the main dimensions, age, and physical features.

The remaining four types of comparative structure are more restricted in their distribution. Type IV, the *Adverbial* comparative, is representative of Southern Min, while Type V is a hybrid form that combines the syntax of the Type I *Compare* comparative with the marker of the native comparative of the region. This may be a Type IV *Adverbial* comparative, as in the case of Hakka and Southern Min dialects, or a plural classifier that serves as the marker in a *Polarity* comparative, as in Xianghua. Type VI, the *Topicalized* comparative reverses the typical comparative word order for the comparee and the standard, placing the standard NP in topic position. This construction has to date only been identified for a handful of Wu and Hui dialects in Central Eastern China (Li Lan 2003). Type VII, the *Polarity* comparative, is formed with two conjoined clauses whose predicates contain antonyms. The *Polarity* type deserves more attention in future research to pinpoint its exact extent of use in Sinitic languages, as too does the *Topic* comparative.

More broadly speaking, however, in terms of linguistic areas, the two main types of comparative that divide China into the North versus the Centre and the South, are the Type I *Compare* comparative and the Type II *Surpass* comparative. Contrasting to this, the area of greatest variation appears to be along the central eastern coast where the Type VI *Topic* comparative is found mainly in certain Wu dialects of Zhejiang province, but also in some Hui dialects in Anhui province. A little further to the south, Type IV, the *Adverbial Comparative*, and Type V, the *Hybrid Comparative* are widespread in the Min dialects of Fujian province, and secondarily, in the case of Type V, in certain Hakka and Gan dialects as well. Type IV can be seen as a characteristic feature of Min dialects.

2.5 Synthesis

On the basis of earlier studies on typology and typo-geographical features, and the three constructions examined in this chapter, I would like to propose that there are at least five linguistic areas that can be discerned for Sinitic languages in China.

These are:

- (i) the Northern,
- (ii) the Central Transitional,
- (iii) the Southwestern,
- (iv) the Far Southern,
- (v) the Southeastern.

This classification is not at all incompatible with the broad divisions put forward in earlier typological studies by Hashimoto and Norman, yet better accounts for the greater diversity and variation found in the south and the southeast of China.

2.5.1 Northern China

Northern China is constituted by just one large area unlike the South which has splintered into three. Nonetheless, it is complex in harbouring a non-conformist enclave in the form of the Shandong peninsula and another area in the Northwest of contact with Altaic languages (Peyraube forthcoming).

2.5.2 Northern Area

Not surprisingly, this new areal research on grammaticalization patterns and syntactic structures confirms that there is a distinct Northern area characterized by three key features which are the native use of TAKE verbs as the source of object markers, causative verbs as the source of agent markers in the passive and the use of the dependent-marked, Type I *Compare* comparatives with *bì* 比. The core languages of this area largely comprise most of the Northern Mandarin subgroups (Lanyin (Northwestern), Central Plains, Beijing, Northeastern) and the Jin dialect group of Shanxi and Inner Mongolia. Hence, geographically it largely coincides with the area to the north of the Yangtze River corridor.

Furthermore, the borderline largely corresponds with that for the use of the interrogative pronoun SHÉI 誰 ‘who’, identified in a study by Wang and Chappell (2012) of over 300 Chinese languages, where it is overwhelmingly preponderant in the six dialect subgroups of Mandarin located in this northern region, as well as in the Jin group. It also roughly coincides with the use of several features that were early remarked upon by Norman (1988) and used as his criteria for identification of Mandarin dialects. These are (i) *tā* 他 as the third person singular pronoun (Cao et al., 2008, Map 3), (ii) the general negative *bù* 不 (Cao et al., 2008, Map 28), the negative adverb *méi* 沒 used in past perfective contexts (Cao et al., 2008, Map 29), and (iii) the diminutive suffix -ér /ə/ 兒 (Cao et al., 2008, Map 52).¹⁹

¹⁹ The Northern area is nonetheless subsumed by a larger area extending over the Yangtze River border for postverbal [CL-N] phrases which code indefiniteness, and which includes most Mandarin, Jin, Gan, Hakka,

This notwithstanding, there is one enclave within the Northern area consisting of two Mandarin subgroups which do not consistently behave in a ‘Northern’ fashion: these are Jiaoliao and Jilu, both located in the eastern peninsula area of Shandong province. Jilu Mandarin is also spoken in parts of Hebei province while Jiaoliao extends to the Liaoning peninsula on the northern side of the Bó Hǎi 渤海 sea, the outcome of migration (see Map 2.1). Here we find the use of Type II *Surpass* comparatives is common, based on QI 起 ‘arise’, and so too is the SUFFER passive with ZHUÓ 着, both features being equally characteristic of Southwestern Mandarin in the ‘deep’ Southwest (the provinces of Sichuan, Yunnan, Guizhou).

We have suggested that the Shandong peninsula is a relic zone, manifesting features that reflect earlier periods in the development of Mandarin, where these syntactic changes have not yet reached. This fits in with the *Wellentheorie* or ‘wave model’ approach in which linguistic changes gradually move out from a central dialect or variety of language, for example, one with social prestige, and in this case, the language of the imperial court in China, to the more isolated, peripheral areas. This model was adopted in fact by Hashimoto (1988) to account for the distribution of agent markers of the passive in Sinitic languages (see also Wang 1969; Lien 1994).

2.5.3 Southern China

The southern half of China can be divided into three main linguistic areas: Southwestern, Far Southern, and Southeastern areas. They do not linguistically pattern as a unified Southern area and need to be treated separately, albeit sharing a number of features that set them apart from the Northern area. Each is next discussed in turn in geographical order from west to east.

2.5.3.1 Southwestern area

This area consists of the provinces of Sichuan, Guizhou, and Yunnan but extends to parts of Hubei and also to western Hunan. In other words, it equates to the core zone of Southwestern Mandarin dialects. In a nutshell, the Southwestern China area makes use of Type II *Surpass* comparatives, SUFFER-passives formed with the markers ZHUÓ 着 and ÁI 挨 and a variety of object markers, based on mainly the TAKE verb BA 把, but also GIVE and HELP, the latter aligning it with the Central Transitional area, but just for this particular feature in our study.

The Southwestern area is a zone where Sinitic languages intermingle with Tibeto-Burman, Zhuang (Tai-Kadai), and Hmong and whose borders are shared with Tibet, Burma (Myanmar), Laos, and Vietnam. We saw that it also shares the characteristic use of SUFFER-passives, derived from CONTACT class verbs, with Thai, Lao, and Khmer, as well as the Type II *Surpass* comparatives.

northern Xiang dialects, and sporadic cases of Wu and Hui (see Chapter 5). Like comparatives of inequality, the core area of use remains the Northern area incorporating Mandarin and Jin.

2.5.3.2 *Far Southern area* Linguistically speaking, Far Southern China is represented above all by the Yue dialect group whose heartland is in southern Guangdong province, overflowing into the Guangxi Autonomous Region, but also by Hakka languages concentrated in northeastern Guangdong and scattered across Guangxi. Finally, less central than either Yue or Hakka, the Pinghua group of Guangxi may be judiciously affiliated with the Far Southern area. Pinghua is sandwiched between Yue and Southwestern Mandarin languages, its geographical location being revealed in the resultant syntactic patterning.

This area represents first of all the use of the Type II *Surpass Comparative par excellence* and secondly, the use of Give-passives which it shares with Southeastern (see §2.5.3.3). Far Southern stands out from the rest of Sinitic in the reduced incidence of the object-marking construction. This is complemented by a preference for the use of serial verb constructions where TAKE class verbs act as the first verb but have not (yet) grammaticalized into an object marker. Note, however, that certain Hakka dialects, close to or located in Fujian province, pattern with another area, the Southeastern, which includes Southern Min, due to their use of a comitative preposition for object-marking constructions. One example of this would be the Meixian dialect of Hakka which uses *lau*¹¹ 拢 < ‘with’²⁰.

Pinghua acts like a linguistic pivot or revolving door between the Far Southern and Southwestern areas, patterning sometimes like Southwestern and sometimes like Far Southern (to continue the metaphor, depending on where the door stops). For example, many dialects in this group use SUFFER-passives formed with AI 挨, as in Southwestern, while on the other hand, they disprefer the use of grammaticalized object-marking constructions, just as in Yue. Many more of the special typological features of Southern Pinghua are discussed in de Sousa (this volume) in relation to Standard Hong Kong Cantonese and the Guangxi dialects of Yue, not to overlook, importantly, the relation to Zhuang, a Tai-Kadai (or Kra-Dai) language which is the major language in Guangxi. The Yue dialects share further features with Zhuang with respect to the extreme polyfunctionality of classifiers, the use of gender suffixes, a set of postverbal adverbs, and *Verb-Theme-Recipient* (that is, DO-IO) as one of the main ditransitive word orders. Again, Pinghua patterns differently in these respects, showing its more marginal status as a member of this area.

We also noted in the body of this discussion that the Far Southern area patterns in a very similar way to the Southeast Asian linguistic zone for the use of both Type II *Surpass comparatives* and TAKE SVCs, but not in its passive structure, since Far Southern uses GIVE, a development which is rare for GIVE verbs in Southeast Asia.

²⁰ All these dialects may of course use the formal marker, JIĀNG 将, or even BĀ 把. As remarked upon earlier, this is a borrowing from Mandarin and does not represent the colloquial register of these dialects. Regarding Hakka, the same situation applies for Hakka dialects spoken in Taiwan, doubtlessly due to the fact that the communities are in contact with Taiwanese Southern Min.

2.5.3.3 Southeastern area The Southeastern area is typified by the Min dialects of coastal Fujian province and Taiwan, stretching a little further north to encompass certain Southern Wu dialects from whose ancestral language, it is claimed by some scholars, they have evolved (You Rujie 1992). Min and Southern Wu thus present two very special groups of dialects that tend to show many conservative features.

This linguistic area represents a very special source of object markers based on comitative prepositions, such as *ka⁷~kang⁷* 共 which is quite generalized in Min, derived from an earlier verb meaning ‘to accompany’ (see Chappell et al. 2011), and *tse⁷⁴⁵* < 着, a comitative derived from the verbal meaning ‘stick together’ in the Shaoxing dialect of Southern Wu. The use of comitatives is not as widespread in Southern Wu, however, as the GIVE~HELP verbs (Chappell 2013a).

This area latches onto both Far Southern and some parts of the Central Transitional area for the use of GIVE-passives. In fact, a swathe of coastal dialects use agent markers derived from different lexical verbs of giving, encompassing the area for Jianghuai Mandarin located further north in Jiangsu province, moving down through the Wu dialects in Zhejiang, to the Min dialects of Fujian, and finally to the Hakka and Yue dialects of Guangdong. As Table 2.4 showed, the particular GIVE-verb usually identifies the dialect group: *pə⁷* 拨 for Wu, *khit⁴* 乞 for most of the Min area, *hou⁷* 與 for Southern Min dialects in the Xiamen area and Taiwan, *pun* 分 for Hakka and *pei²* 眇 for Yue.

For the comparative, we find that this area is originally a Type II *Surpass* area. At its northern end, Type II is being replaced by the Type I *Compare* from Mandarin, while in Fujian, for the Min dialects, the use of the Type IV *Adverbial* and Type V *Hybridized* strategies is endemic. From historical documentation, however, we know that in addition to Type IV, the Type II *Surpass* comparative was also possible in earlier stages of Southern Min. The *Adverbial* Comparative presents an interesting variation on the Action schema for comparatives, being transitive in structure and semantics.

Min dialects represent the core of this linguistic area, repeatedly demonstrating original syntactic behaviour that is not found elsewhere in Sinitic. For example, Wang Jian (this volume) has shown that Min dialects constitute a type on their own with respect to bare classifier phrases of the form CL-N. They do not permit this structure in any position of the clause, unlike the majority of Sinitic languages which allow the postverbal use at the minimum. S. Ngai (this volume) similarly observes that the numeral for ‘one’ SOK 翁 is an exclusive Min characteristic, not seen elsewhere in Sinitic. In a study on experiential aspect markers as evidentials, Min dialects prove, once again, to form a group on their own due to their use of BAT 別 derived from a verb meaning ‘to know’ whereas most other Sinitic languages use GUO 過 < ‘to cross, pass (through)’ (Chappell 2001c). The use of special interrogative particles in Min has also been observed in several studies on this topic (Zhu Dexi 1985; A. Yue-

Hashimoto 1991; Zhang Min 2000), not to mention distinctive interrogative pronouns for WHO (Wang and Chappell 2012) based on the composite form WHAT +PERSON OR WHICH+PERSON.

2.5.4 Central China

We discuss the Central China Transitional Area last of all, albeit quite evidently out of any geographical order. This is simply because its highly mixed features only make sense in the light of the other four areas found in Northern and Southern China. As it lies between the North and the South, this linguistic area shows characteristics of those found on either side, yet also ones which are not at all predictable.

2.5.4.1 Central China Transitional Area It comes as no surprise to find that the typological features of the ‘buffer zone’ between the North and the South of China necessarily fluctuate where different dialect groups have come into contact over the millennia. This is the ultimate outcome, linguistically-speaking, of the continual migrations which have taken place in the direction from north to south, particularly during the first millennia AD and up to the time of the Song dynasty (960–1279) (see You 1992, Chappell 2001a on dialect history). A typical example (and here we are considering the geographical distribution for one feature at a time) is the Northern Wu dialects which make use of the TAKE source for object markers, while simultaneously patterning with the Southeastern area, when it comes to the source of agent markers in the passive, since they use GIVE. Put differently, they do not use the typical Northern source of a causative verb, or even the SUFFER-verb BĒI for the passive. In general, they also show the native use of Type II *Surpass* comparatives.

A further interesting example is the Jianghuai or Southern Mandarin subgroup, spoken mainly in the part of Jiangsu province, north of the Yangtze River, and so shares a frontier with the Wu dialect group. It contrasts greatly with the Mandarin subgroups located in the North of China: for object markers, it generally patterns like the Central Transitional area with GIVE~HELP verbs as the common source, while for agent markers in the passive, it patterns with the Type III GIVE ~ SUFFER area. Type I *Compare* comparatives are however the norm, as in the North.

The mixed patterning of languages in this intermediate zone is particularly clear for the case of agent markers of the passive where the greatest diversity and variation is evident for Hui, Gan, Xiang, and Jianghuai Mandarin (Types III and IV, where causative, SUFFER verbs including BĒI, GIVE, TAKE, and WAIT verbs intermingle). Hashimoto (1976a) presaged the idea of a transitional area in Central China for Xiang, Gan, and Southern Wu which was later formalized in the work of Norman (1988). He remarked that both Northern and Southern types can be found in this middle area of China. For example, gender affixes on animal terms are typically prefixes in the North and suffixes in the South while in this transitional area one and the same language may show both (cf. Wu Yunji 1995 on this topic in Xianghua).

In fact, the situation is more complex than believed, given that, first, sources found neither in the North, nor in the South, turn up in this central area, such as the use of TAKE and WAIT verbs as markers of passive constructions. Second, a variety of combinations of features can be found for the markers of the three construction types examined in the present chapter. That is, in this transitional area, it is not a simple case of the pairing of features for construction types, one being a Northern type and the other, the corresponding Southern type of feature.

This kind of linguistic turbulence found in the Central Transitional area, with almost a new realignment for each combination of features, is responsible for the lack of any clearly defined characteristics. However, this is only to be expected where languages with different evolutionary pathways meet geographically, if not ‘physically’. The languages situated at the centre of this turbulence are those belonging to Wu, Jianghuai Mandarin, Hui, Gan, Xiang, and some of the Southwestern Mandarin subgroups (in geographical order from east to west) without overlooking co-territorial Hakka communities, a language group which is scattered over the southern extremes of this central zone in China.

2.6 Conclusion

This analysis has shown that there are at least four main sources for object markers found in the basic form of the differential object marking construction, seven different sources for agent markers in the passive, as well as seven different structural types for the comparative construction of inequality, each possessing its own specific source of markers. Compared with the relevant Standard Mandarin constructions, the markedly contrasting, and in certain areas crosslinguistically atypical, semantic sources for the markers of differential object marking, passive and comparative constructions point to the importance of exploring the grammatical diversity of Sinitic languages in a more detailed and systematic way. This would unreservedly include the non-standard Mandarin dialect groups, for which data are similarly insufficient for these key areas. The findings of such research can in turn contribute to enriching our knowledge of how linguistic areas are set up within a language group such as Sinitic as well as add to the stock of grammaticalization pathways in the world’s languages, and consequently a better understanding of diachronic change, if not the extent of variation possible.

In sum, much more work is needed to gather together and match up morphosyntactic features associated with grammatical structures and grammaticalization pathways to see if they correspond to the formation of true linguistic areas within China. Certainly, the five areas proposed here will need to be much more finely drawn, as the task of searching for a consistent matching of features progresses. In spite of such caveats, a broad overlapping can nevertheless be detected at this early stage, as shown, even though the borders for each feature do not coincide with any

mathematical precision. In spite of this, the correspondence and overlapping of the geographical areas for key linguistic features, approximate as it may be at present, points to the existence of the five linguistic areas within China that cross-cut the traditional boundaries made on the basis of phonological and lexical criteria for Sinitic languages. This regrouping according to geographical areas has been defined by a specific set of syntactic features based on both structural types and grammaticalization pathways.

Acknowledgements

Apart from acknowledging the funding received from the European Research Council and the Agence Nationale de la Recherche (see front covers), I would like to thank the SINOTYPE and DIAMIN team members in Paris for many enjoyable hours of critical discussion on a large number of the topics which are presented in this chapter: Manjun Chen 陈曼君, Weirong Chen 陈伟蓉, Yujie Chen 陈玉洁, Hilario de Sousa 苏沙, XuPing Li 李旭平, Sing Sing Ngai 倪星星, Alain Peyraube 贝罗贝, Wang Jian 王健, and Jianming Wu 吴健明.

3

Grammatical change in Sinitic languages and its relation to typology

ALAIN PEYRAUDE

3.1 Introduction

Historical linguistics and typology are two domains which have always been linked in China as elsewhere, at least as far as grammar goes. However, if the renewal of research into diachronic syntax from the 1980s onwards in the West has essentially depended on advances in typology, in particular, in the wake of Greenberg's (1963) work on universals, the contrary situation has held sway in China: it is rather research on historical grammar which has played an important role for the very recent domain of typology of Sinitic languages, as distinct from Chinese dialectal studies which have mainly been devoted to phonological problems.

In spite of this, the association between these two domains of research, currently claimed to be close, is actually quite indirect. The three basic mechanisms which govern grammatical change, that is, the processes of reanalysis (including grammaticalization and exaptation), analogy (including the phenomena of lexicalization and/or degrammaticalization), and external borrowing through language contact, do not involve, or only minimally, typological research strictly defined. External borrowing is the only exception which could be included in the definition of linguistic areas, insofar as the latter's relevance for typology.

The same applies to the recognized motivations for grammatical change: semantico-pragmatic change, structural requirements, etc. They do not directly involve typology to any great extent.

What connects these two domains then? Quite simply the fact that diachronic linguistics often enables us to provide, if not explanations, at least grounded hypotheses about the common properties which Sinitic languages share with other languages (the dream of every typologist being to find implicational universals, if not universal, at least relative ones), or more often, the basic differences which are revealed between them.

After drawing an outline of a model of grammatical change (based on Peyraube 2005, 2013) and providing some general typological principles and characteristics for the Chinese language—or Sinitic languages (based on Greenberg 2001, and on Chappell 2001a, this volume), examples will be selected in this chapter from the following topics to illustrate the points made above: passives and causatives, post-verbal and preverbal adverbs, double object constructions, and verbs of saying.

3.2 A model of grammatical change

Scholars working on diachronic syntax try to find answers to the following four questions: (i) ‘What motivates grammaticalization in the first place? (ii) What mechanisms lead to it? (iii) What are its probable paths of progression through time? (iv) What are its end results?’ (Hopper and Traugott 1993: 32).

The two essential mechanisms said to account for the appearance of new grammatical forms remain roughly those identified by Meillet (1912) a hundred years ago: analogy and grammaticalization. A third and external mechanism has been added to these two internal ones: external borrowing. The second of these two internal mechanisms, i.e. grammaticalization, has been extensively discussed and commented upon over the past thirty years.

The following issues have been much debated: (i) Does grammaticalization have any theoretical value?¹ (ii) Can the unidirectionality principle in grammaticalization be upheld, and, if so, is this unidirectionality of theoretical importance? (iii) Do degrammaticalization cases exist which contradict the unidirectionality principle, ‘degrammaticalization’ being a term used widely today (Norde 2002; Heine 2003)? (iv) Is the notion of exaptation, borrowed from biology, relevant for linguistics, as suggested by Lass in 1990? (v) Do pragmatic inferencing (metonymization), metaphorical extension, and subjectification represent the only main motivations behind syntactic change, and hence the major mechanisms for semantic change?

I would first like to propose a coherent model of grammatical change, by providing more solidly based definitions of such notions as grammaticalization, degrammaticalization, exaptation, but also of reanalysis and analogy, re-grammaticalization (Greenberg 1991), functional renewal, re-functionalization (Giacalone-Ramat 1998), and hypoanalysis or under-analysis (Croft 2000), redefining their role and function in grammatical change.

This model of grammatical change, developed in Peyraube (2005, 2013), has only two internal mechanisms of syntactic change: analogy and reanalysis. In this model, grammaticalization, degrammaticalization, and exaptation are secondary processes,

¹ Newmeyer (1998: 226, chapter on ‘Deconstructing grammaticalization’) states: ‘there is no such thing as grammaticalization, at least in so far as it might be regarded as a distinct grammatical phenomenon requiring a distinct set of principles for explanation.’

while degrammaticalization and exaptation have to be distinguished, as the first one belongs to analogy and the second one to reanalysis. There is a third external mechanism: borrowing. Analogy (or generalization) thus comprises degrammaticalization (with this referring to cases of lexicalization most of the time) while reanalysis comprises both grammaticalization and exaptation.

But what are precisely analogy, reanalysis, grammaticalization, degrammaticalization, and exaptation?

3.2.1 Mechanisms of grammatical change

3.2.1.1 Analogy McMahon (1994: 71) defines analogical extension as follows: ‘generalization of a morpheme or relation which already exists in the language into new situations or forms’. It only modifies the surface structure and does not modify the underlying structure. Very recently, Kiparsky (2005, 2012), rethinking analogy, distinguishes two types of analogy: (i) exemplar-based analogy, and (ii) non-exemplar-based analogy. Cases of degrammaticalization (see the definition below) are ordinary analogical changes of the exemplar-based type. Examples of non-exemplar-based analogical changes are fusions of two words into one, which occur spontaneously without any particular model. The latter is driven by a language-independent preference for structural economy: other things being equal, one word is always better than two. Among the contractions of this kind in Classical Chinese (Late Archaic, fifth–second centuries BC) are to be found: *zhī* 之 ‘third personal pronoun’ + *yú* 于 ‘to, at’ > *zhū* 諸; *bù* 不 ‘not’ + *zhī* 之 ‘third personal pronoun’ > *fú* 弗.

The opposite process, fission of one word into two words is always exemplar-based, it occurs by analogy with specific existing constructions. Examples are *zhū* 諸 > *zhī* 之 + *yú* 于; *fú* 弗 > *bù* 不 + *zhī* 之.

3.2.1.2 Reanalysis Harris and Campbell (1995: 61) were the first to consider reanalysis, instead of grammaticalization, as one of the two major internal mechanisms of syntactic change, for which they adopted the definition of Langacker (1977): ‘Change in the structure of an expression or class of expressions that does not involve any immediate or intrinsic modification of its surface manifestation.’ As this could also very well characterize grammaticalization, by opposition to analogy, which is a ‘change in the surface manifestation of a syntactic pattern that does not involve immediate or intrinsic modification of underlying structure’ (also from Langacker 1977), the two notions of grammaticalization and reanalysis have sometimes been confused, especially by Chinese linguists. It is necessary to distinguish them.

Even if most grammaticalization cases are also reanalyses, the reverse is not always true. Consequently, I have adopted the following definition by Hagège (1993: 62): ‘An operation by which language builders cease to analyse a given structure as they did previously, and introduce a new distribution of, and new relations between, the syntactic units that constitute this structure.’ As a matter of fact, only Hagège’s

definition allows us to consider major typological shifts such as word order change (OV > VO) as cases of reanalysis. Grammaticalization and exaptation are processes belonging to reanalysis.

3.2.1.3 Grammaticalization Meillet (1912)'s definition of grammaticalization ('*l'attribution du caractère grammatical à un mot jadis autonome*'—‘the shift of an independent word to the status of a grammatical element’) has been improved upon, first by Kuryłowicz (1965) with the characterization: ['(It) consists in the increase of the range of a morpheme advancing from a lexical to a grammatical or from a less grammatical to a more grammatical status'], then later by Hopper and Traugott (2003: 231) with the refinement: ‘It is a robust tendency for lexical items and constructions to be used in certain linguistic contexts to serve grammatical functions, and once grammaticalized, to be used to further develop new grammatical functions.’ An example from a Sinitic language would be the different stages of grammaticalization of the word *gòng* 共 from a lexical verb ‘to share with’ to an adverb ‘together’, thence to a preposition ‘with’, and finally to a conjunction ‘and’: V > Adverb > Preposition > Conjunction (see Liu and Peyraube 1994).

For Greenberg (1991: 303), however, grammaticalization is not only a shift from lexical to grammatical. It is ‘a development of grammatical elements from all sources’. Many Chinese linguists have taken up this broad definition. However, in my viewpoint, it is better not to see grammaticalization as equivalent to grammatical change, and therefore it should not be considered as ‘a development of grammatical elements from all sources’. Actually, confined to diachronic studies, grammaticalization is a concept familiar to all Chinese linguistics specialists, since, from at least the time of the Yuan dynasty onwards (1279–1368), Chinese scholars have observed that empty (grammatical) words were full (lexical) items in ancient times (see Sun C. 1996: 11; Hong B. 1998). Zhou Boqi 周伯琦, a scholar of the Yuan dynasty 元朝 (thirteenth–fourteenth centuries), wrote: ‘today’s empty words are all former full words: *jīn zhī xū zì jiē gǔ zhī shí zì* 今之虚字皆古之实字’.

Much work has been done on grammaticalization since the 1990s and several principles (or tendencies) have been suggested, such as the four typical ‘heuristic’ principles (Layering, Divergence, Specialization, and Persistence) proposed by Hopper (1991), or the MGMF (‘more general, more frequent’) principle put forward by Hagège (1993: 212), or the notion of cline, or the hypothesis of cyclic change, etc. (see Peyraube 2013).

One of these principles has been particularly discussed: the unidirectionality principle, said to be the main characterization of grammaticalization, associated with it since the beginning. Haspelmath (1999) has claimed: ‘Grammaticalization is irreversible’: from content words to empty words or from lexical item > grammatical element (full word > empty word) or from less grammatical > more grammatical. Given the theory of unidirectionality, it has been hypothesized that diachronically all minor categories (prepositions, conjunctions, auxiliary verbs, pronouns,

demonstratives, etc., i.e. the relatively ‘closed’ categories) have their origins in major categories (those relatively ‘open’ lexically, such as nouns and verbs). This assumption has been proved to be valid in Chinese languages. It is however probably better to treat unidirectionality as a defining characteristic of grammaticalization (grammaticalization is unidirectional by definition) rather than a principle.

Note that claiming that grammatical items originate in lexical ones does not entail by any means hypothesizing a language stage in which everything is lexical.

3.2.1.4 Degrammaticalization Counterexamples to unidirectionality began to be discussed in the 1990s with well-documented instances of what has been called ‘degrammaticalization’. Some argue that such counterexamples are not damaging to the unidirectionality hypothesis because they are sporadic (for example, Haspelmath 2004, Hopper and Traugott 2003). Others (a growing number) draw the opposite conclusion and conclude that no special type of change such as grammaticalization even exists. Degrammaticalization involves the following process:

- (i) Grammatical element > Lexical Item; or
- (ii) Empty word > Full word (‘Lexicalization’, Ramat 2001) or More grammatical > Less grammatical.

An example of this for English is: *text* [+ Noun] > *to text* [+ Verb] as in: *Just text me!* and for Chinese is: demonstrative pronoun *shì* 是 ‘this’ > copular verb *shi* 是 ‘to be’.

As seen above, these cases of degrammaticalization (or lexicalization) are subsumed by the mechanism of analogy.

3.2.1.5 Exaptation The term ‘exaptation’ is widely used in work on evolution of language, but has now also been applied to historical morphosyntax. It was apparently first used in evolutionary biology by Gould and Vrba (1982): ‘We wish to restrict the term *adaptation* only to those structures that evolved for their current utility; those useful structures that arose for other reasons, or for no conventional reasons at all, and were fortuitously available for other changes, we call *exaptations*’ (Gould 1983, cited in Lass 1990: 80). Lass gave the following definition: ‘Exaptation... is the opportunistic co-optation of a feature whose origin is unrelated or only marginally related to its later use.’ It is a marginal morpheme with an old function that becomes a more central morpheme with a new function (formulation by Traugott, 2004). An example in Chinese is the re-use of the Classical Chinese (eleventh–second century BC) modal particle *yě* 也 in sentence-final position, as an adverb in Medieval Chinese (second–fourteenth century AD), where *yě* 也 means ‘also’.

Exaptation is a special case of reanalysis. It has nothing to do with extension or analogy, or reformulation of paradigms in accordance with a target or a model. It can neither be classified as grammaticalization, nor as degrammaticalization. It is something different, as it is the re-use of an old form (A) for something else completely new (B), with no direct or indirect connection between A and B.

3.2.1.6 External borrowing Analogy and reanalysis (including grammaticalization) are internal mechanisms of change. There is a third and external one: external borrowing through language contact. Borrowing, contrary to analogy, but like reanalysis, can introduce an entirely new structure into a language, and in this sense can produce radical change. It is an ‘attempted reproduction in one language of patterns previously found in another’ (Haugen 1950, quoted in McMahon, 1994: 200).

Some of the debatable universals, principles, and constraints of external borrowing (see Moravcsik 1978) are the following:

- (i) Borrowing moves from the more to the less prestigious language. This condition is not absolute. We know many examples of Chinese structures borrowed from Altaic languages (Khitan, Jurchen, Mongolian, Manchu), languages that have never been more prestigious than Chinese within the Sinosphere. Mei T. (1988) has thus convincingly shown that the opposition between *zánmen* 咱们 ‘we, inclusive’ versus *wǒmen* 我们 ‘we, exclusive’, which appeared in Chinese under the Jin (twelfth century), has been borrowed from Altaic languages, either Khitan or more probably Jurchen.
- (ii) Basic vocabulary is only affected infrequently. We nonetheless know that English borrowed a good deal of basic vocabulary from Norse: *sky*, *skin*, and even the pronouns *they*, *them*, *their*.
- (iii) Structural compatibility is supposed to be required. Weinreich (1953: 25), after Jakobson, has stressed that a language ‘accepts foreign elements only when they correspond to its tendencies of development’. However, any insistence that grammatical borrowing happens only in situations of shared structural similarities is simply wrong. Many examples involve grammatical borrowing from typologically divergent languages. See example above of Altaic–Chinese contact.
- (iv) Some categories (lexical elements) are said to rank highest in terms of borrowability, others lowest, if borrowable at all, such as the case of inflectional affixes. This claim is also debatable. An absolute ranking of this nature provides little real satisfaction.
- (v) Basic patterns are very difficult to borrow. Nonetheless, a rather large number of cases have been reported in which basic word order patterns have been borrowed. Faarlund (1990: 84) further claims that ‘all known instances of a change from VO to OV are due to contact with OV languages’. Thomason and Kaufman (1988: 55) also believe that the word order of a language is one of the syntactic features that is very easily borrowable.

To sum up, if borrowing must indeed be accorded a more significant position among the three mechanisms of syntactic change, it is safer to consider all the proposed universals and principles of borrowing as general tendencies, instead of absolute constraints. And it may prove to be the case that language contact and the

outcome of borrowing could have, at best, a trigger effect, launching or accelerating grammatical phenomena which then develop independently.

3.2.2 Motivations for grammatical change

Analogical change has two main motivating factors: (i) A > B motivated by the abnormality or complexity of A, or by the generality or simplicity of B (pull/push model); and (ii) semantic-pragmatic change, especially metaphorical extensions. The motivations for reanalysis are also (i) semantic-pragmatic change, but that which involves pragmatic inferencing or metonymization rather than metaphorical extension, which is more closely related to analogy; (ii) subjectification and/or intersubjectification (Traugott 2004); and (iii) other motivations, such as phonological change. The main motivation for external borrowing is obviously language contact.

3.3 Typology

Apart from what is often called genetic classification, or historical linguistic classification, which is usually considered classification *par excellence* and is based almost exclusively on phonological criteria, Greenberg (2001) distinguishes two other kinds of linguistic classification: areal classification and typological classification.

3.3.1 Areal typology

The areal approach to classification privileges considerations of a geographical kind. Parallel facts observed from within the diverse nature of languages cannot always be explained by a common origin. They are simply due to the ongoing relationship of contiguity. There is a diffusion of traits, normally understood as not being inherited. The clearest case of diffusion and convergence thus occurs between unrelated languages.

This approach has led some linguists to discuss the possibility of identifying large linguistic areas for East Asian and Southeast Asian languages, comprising Sino-Tibetan, Austronesian, Austroasiatic, Tai-Kadai, and Hmong-Mien languages (see Chappell, Chapter 2, this volume).

But what do we call a ‘linguistic area’? Several definitions have been provided, such as the ones by Emeneau (1956) or by Thomason (2001a: 99), but they are not sufficiently restricted.² The one by Enfield (2005: 190) is preferable: ‘A linguistic area is defined as a geographical region in which *neighboring languages belonging to different language families* show a *significant set of structural properties in common*

² Emeneau (1956): ‘(A linguistic area is an) area which includes languages belonging to more than one family but showing traits in common which are found not to belong to the other members of (at least) one of the families.’ Thomason (2001a: 99): ‘A geographical region containing a group of three or more languages that share some structural features as a result of contact rather than as a result of accident or inheritance from a common ancestor.’

[italics my own], where the commonality in structure is due to contact and where the shared structural properties are not found in languages immediately outside the area (ideally where these include languages belonging to the same families as those spoken inside the area).’

To have a real ‘linguistic area’, we thus need (i) a geographical area; (ii) a set of languages belonging to different families in this area; (iii) the situation that these various languages share some linguistic features; and (iv) the fact that these linguistic features are unknown to the languages immediately outside the area, be they related or not.

Such a restricted definition does not, however, allow the identification of a large number of linguistic areas. Typical examples are the Balkans linguistic area and the Northeastern Amazonian area. In the Balkans area, a *Sprachbund*, a confederation of languages has been proposed, including Modern Greek, Bulgarian (Slavic), Rumanian (Romance), and Albanian, as they share a dozen linguistic features unknown to the other languages belonging to these families (Comrie 1989: 205). In the Vaupez-Içana basin (Northeast of Amazonia), five linguistic features are shared by North Arawak, East Tucano, Maku, all languages belonging to different families (Aikhenvald 1999).

Passive markers having their origin in a verb of giving in some East Asian and Southeast Asian languages could be another example of ‘linguistic area’ (see §3.4 on ‘Passives and causatives’).

The fact remains that it is often difficult to distinguish, for the features which involve resemblances between languages, those which are the result of borrowing through contact, from those which could be explained by the same source. In the absence of historical features, as is often the case, common origin and diffusion are two hypotheses which we can envisage. For Sinitic languages, in any case, historical facts do exist and certainly allow us to justify many typological hypotheses, such as the ones formulated by Chappell (this volume) who, in the context of areal linguistics of Sinitic languages of China, has examined the feasibility of a classification into five micro-areas as a refinement of the traditional North–South division originally proposed. See also the following sections; Aikhenvald and Dixon (2001), Chappell (2001b).

3.3.2 *Typological classification*

In the framework of typological classification, languages are grouped into types, according to their common linguistic (most of the time syntactico-semantic) structures and properties. One can thus distinguish according to different parameters: (i) languages with a verb–noun distinction (such as Indo-European languages); (ii) languages with no verb–noun distinction at all (Santali, in India); (iii) languages with a weak verb–noun distinction (such as Sino-Tibetan), or (a) languages with an ergative construction (such as Basque or several languages from the Southern Caucasus); (b) languages with an accusative construction (French or Malay); (c) mixed languages (Australian or Oceanic languages); see Hagège (1982: 39–40).

The classification in the European tradition into inflectional languages (Latin, Russian, etc.), agglutinative languages (Turkish), and isolating languages (said to be monosyllabic, such as Chinese) is a typological classification. Shaped during the nineteenth century by scholars including the von Schlegel brothers—Johann and Friedrich, by August Schleicher, and by Wilhelm von Humboldt, this approach has since considerably evolved.

One of the issues currently discussed for such a typological classification is certainly the order of constituents for a language. Correlations have been recognized by Greenberg (1963), which allow distinguishing more precisely different types of languages. Thus, if a language displays an SOV (subject-object-verb) order, it has postpositions and the modifiers precede the heads in endocentric constructions; if, on the contrary, it displays a SVO order, it has prepositions and the ‘Head + Modifier’ construction.

The most interesting and the most productive typological classifications need to be established on the basis of correlations between different parameters and suggest, by this means, that there can be causal links of a universal nature between the properties under consideration (that is, the existence of implicational universals). It is for this reason that the most recent typological work has been closely associated with research into linguistic universals and that, correspondingly, multidimensional typologies have become more and more frequent.

3.4 Passives and causatives

There are strong links, crosslinguistically, between verbs of giving and dative prepositions, between verbs of giving and causative verbs (Heine and Kuteva 2002; Comrie and Polinsky 1993; Lord et al. 2002), but it is absolutely atypical, from a typological point of view, to have verbs of giving developing into passive markers.

In many Sinitic languages, however, verbs from the lexical domain of giving, such as Mandarin *gěi* ‘to give’, are often used as a passive marker. This appears to be a unique characteristic of Sinitic languages (Wu, Yue, Hakka, Min, and Jianghuai Mandarin, see §2.3 in Chappell this volume and Map 2.3). An example follows.³

- (1) 鱼给貌吃了。

Yú gěi māo chī le.
fish PASS cat eat ASP
'The fish has been eaten by the cat.'

³ The romanizations used in this chapter are *pīnyīn* for Standard Mandarin and for examples from earlier periods of Chinese, Church Romanization for Taiwanese Southern Min, Yale for Cantonese, and the IPA for other languages. For examples cited from other sources, a reproduction is provided according to the original source.

A very common usage of *gěi* is also to introduce the indirect object in a dative construction. *Gěi* is then a dative preposition.

- (2) 我送了一本书给他。

Wǒ song le yīběn shū gěi tā.
1SG offer ASP one.CL book to him
'I offered a book to him.'

Another meaning of *gěi* is a causative one.

- (3) 他才十五岁了，不给他开车吧！

Tā cái shíwǔ suì le, bù gěi tā kāi chē ba.
3SG just fifteen age PRT NEG let 3SG drive car PRTIMP
'He is just fifteen, don't let him drive the car!'

In Archaic Chinese (Classical Chinese, eleventh–second century BC) and in Medieval Chinese (second–fourteenth century BC), several verbs are used to form analytic causative constructions: *shǐ* 使 ‘to send, to cause’, *lìng* 令 ‘to command’, and to a lesser extent, *wèi* 遣 ‘to give’ (Jiang S. 2002). In Medieval Chinese we also find *jiào* 教 ‘to instruct’, *jiāo* 交 ‘to exchange’ and later *jiào* 叫 ‘to call’ in the Modern period; *yǔ* 与 ‘to give’, *yào* 要 ‘to want’, *zhāo/zhuó* 著 ‘to use, to place’, *qǐ* 乞 ‘to give’, *ràng* 让 ‘to yield’.

- (4) 又教弟坐。 (祖堂集)

Yòu jiào dì zuò.
in.addition CAUS disciple sit
'(And) furthermore, (he) let his disciples sit.' (Zǔ táng jí, dated 952)

- (5) 乞我惶了推門推不開。 (金瓶梅词话)

Qǐ wǒ huáng le tuī mén tuībùkāi.
CAUS 1SG frighten ASP push door push.NEG.open
'(It) made me so frightened (that I) could not open the door.'
(Jin Píng Méi Cíhuà, 16th c.)

- (6) 我著孩子們 做與你吃。 (老乞大諺解)

Wǒ zhuó háizimen zuò yǔ nǐ chī.
1SG CAUS child.PL make to 2SG eat
'I'll get my children to make you (something) to eat.'
(Lǎo Qǐ Dà Yánjiě, dated 14th c.)

Most of these causative verbs are or have also been used as passive markers: *yǔ*, *qǐ*, *gěi*, *jiào*, *ràng*, and even *zhuó*. Some of these verbs are also dative markers, grammaticalized from verbs meaning ‘to give’: *yǔ*, *qǐ*, *gěi*. To account for the links between the verbs of giving, datives, causatives, and passives, two main solutions can be proposed for consideration:

Grammatical change in Sinitic languages and its relation to typology 63

1. Several different pathways implicating a polygrammaticalization process:
V [+ give] > dative marker (for *yǔ*, *qǐ*, and *gěi*);
V [+ give] > V [+ causative] (for *yǔ*, *qǐ*, *gěi*, and *jiào*);
V [+ give] > passive marker (for *yǔ*, *qǐ*, *gěi*, and *jiào*).

2. One single grammaticalization chain: The different meanings (dative, passive, causative) are all linked and derived from one another. In this case, two solutions would then emerge as possibilities:
(a) V [+ give] > dative > causative > passive;
(b) V [+ give] > dative > passive > causative.

Since we can show that the causative and passive functions are historically related, the first proposal of polygrammaticalization is not economic. Nor is the second proposal of just one grammaticalization chain a desirable hypothesis, since it is not supported by either the Sinitic data or data from surrounding language families, and it furthermore violates the principle of unidirectionality in that it involves a preposition (the dative marker in (a) and the passive marker in (b)) grammaticalizing into a verb (the causative).

Chappell and Peyraube (2006a) hypothesized that there have been at least two different pathways of grammaticalization for verbs of giving:

- (i) V [+ give] > dative marker;
- (ii) V [+ give] > causative > passive.

The second pathway [+ give] > causative occurred ca. 800 AD for both *jiào* and *qǐ* and ca. 1800 AD for *gěi*. This hypothesis conforms to the available dating: the causative use appeared not long before the passive one.

Reinforcing this analysis is the fact that the syntactic configuration for dative (ditransitive) constructions is different from that for the passives and causatives: dative markers follow the main verb in most Sinitic languages, whereas the causative and passive exponents precede it.⁴ Finally, causative verbs from sources other than verbs of giving similarly develop into passive markers (but never into dative prepositions). This applies, for example, to *jiào* 訂 or *ràng* 諉 used as passive markers in contemporary Chinese.

What is most important to notice is that all the passive markers that have their source in verbs of giving in the investigated languages are not directly derived from these verbs of giving, but come from causative verbs, themselves derived from verbs of giving (see also Jiang S. 2002; Jiang L. 1999):

⁴ Note that the dative is to be distinguished from the benefactive marker ‘for’ and prepositions with the meaning of ‘instead of’ which also tend to be placed in preverbal prepositional phrases in Sinitic.

- (7) V [+ give] > V [+ causative] > passive marker.

Since the development from a verb of giving directly into a passive marker is atypical, while the development of passives directly grammaticalized from causatives is a common phenomenon, Chappell and Peyraube (2006a) have proposed the following implicational universal:

If a language has a passive marker whose origin is a verb of giving, then it necessarily has a causative verb realised by the same form and having its source in a verb of giving.

A second example illustrating the importance of diachronic studies in resolving problems in linguistic typology concerns constructions with a postverbal adverb.

3.5 Postverbal adverbs

3.5.1 Cantonese adverbs

In Standard Mandarin, adverbs are necessarily preverbal, while in contemporary Cantonese they are either preverbal or postverbal. Moreover, in the beginning of the twentieth century, they were basically postverbal, as noticed by Leblanc (1910: 144) in this regard: ‘the adverb is generally placed after the verb’. Examples are found below for the following Cantonese postverbal adverbs *sin* 先 ‘first’, *jyuh* 住 ‘for the moment’, *gwo* 过 ‘again’, *tim* 添 ‘also, more’, *màaih* 埋 ‘also, more, again’, *saai* ‘all, completely’ [expression of universal quantification], *jaih* ‘too’.

- (8) 我去先。

Ngóh heui sin

1SG go first

‘I go first.’

- (9) 眇郁住。

Máih yūk jyuh!

NEG_{IMP} move now

‘Don’t move now!’

- (10) 放心啦我请过你哋。

Fongsàm lā ngóh chéng gwo néihdeih.

be-assured PRT 1SG invite again 2PL

‘Don’t worry, I will invite you again.’

- (11) 重可以游水添。

Juhng hóyíh yàuhséui tim.

moreover can swim also

‘Moreover, (you) can also swim.’

Grammatical change in Sinitic languages and its relation to typology 65

(12) 我哋係翻埋一齊。

Ngóhdeih haih fāan màaih yātchāih.

1PL be return again together

‘We are back together again.’

(13) 佢哋去晒西藏。

Kéuihdeih heui saai Sàijohng.

3PL go all Tibet

‘They have all gone to Tibet.’

(14) 呢件衫你著细得滯。

Nī gihn sāam néih jeuk sai dak jaih.

this CL dress you wear small PRT too

‘This dress is too small for you.’

Where do these seven adverbs come from? Did similar constructions exist in earlier stages of the Chinese language that will allow us to propose a hypothesis of internal change? In fact, postverbal adverbs never existed in the whole history of the Chinese language. Adverbs have always been preverbal, in Archaic, Medieval, as well as in Modern and Contemporary Chinese. It is therefore quite impossible to propose such a hypothesis of internal change.

If one assumes that Cantonese is a Sinitic language, the only remaining possible hypothesis is that of ‘external borrowing’. Peyraube (1997) has shown that this hypothesis can be taken seriously, as the Yue populations have been in permanent contact all along their history, with those speaking non-Sinitic languages, notably Miao-Yao (Hmong-Mien), Kam-Tai (Tai-Kadai), and Austroasiatic, or even perhaps Austronesian languages. Meecham (1993) already suggested that ‘leaving aside the Austronesian question, it seems highly likely that the peoples called Yueh at various times spoke Austroasiatic languages, early forms of Miao/Yao (Hmong-Mien), Tai-Kadai languages, and perhaps other language families now extinct’.

If we look carefully at the placement of adverbs in non-Sinitic languages with which Cantonese has always been in contact, namely the two linguistic families of Kam-Tai and Miao-Yao (Hmong-Mien), it can be seen that numerous adverbs are postverbal, allowing us to identify these as likely sources of borrowing (see also §7.4.3 in de Sousa, this volume, for Pinghua).

3.5.2 Kam-Tai languages

The Kam-Tai family of languages includes two branches: Tai and Kam-Sui. Examples of sentences with clause-final adverbs belonging to these two branches are given here.

(15) Zhuang (Tai)

Mun³¹ pai²⁴ ko:n³⁵.

2SG go first

‘You go first.’ (compare with [8]) (Li J. 1990)

- (16) Sui (Kam-Sui)
*Ta:p*⁷ *kon*⁵.
jump first
'Jump first!' (Zhang J. 1980)

There are also examples where two synonymous adverbs are present, one is preverbal, probably borrowed from Chinese, whereas the other is postverbal, leading to a hybridized structure.

- (17) *ŋa*² *sjen*³ *fan*² *ha:i*¹ *man*¹ *kon*⁵.
2SG first say to 3SG first
'Tell him first.' (Zhang J. 1980)

Hmong-Mien languages also have postverbal adverbs, as in:

- (18) Qian dong dialect (Hmong-Mien)
*Ken*⁵⁵ *va*⁴⁴.
Cry a.lot
'Cry a lot.' (Wang F. 1985)
- (19) Bunu (Yao, Hmong-Mien)
*Cuŋ*³ *nau*² *i*¹ *py*³ *thiŋ*¹ *zau*⁴.
1SG eat one bowl again
'I eat another bowl (of rice).' (Mao Z. 1982)
- (20) *Kau*² *muŋ*⁴ *te*².
2SG go first
'You go first.' (Mao Z. 1982)

These correspondences, between Cantonese on the one hand, and Kam-Tai and Hmong-Mien languages on the other, are probably not a coincidence. As these languages have been in contact for a long period, it is probable that a borrowing mechanism has taken place. Two competing hypotheses can nonetheless be put forward:

- (i) Kam-Tai and/or Hmong-Mien languages might have borrowed their post-verbal adverbs from Cantonese;
- (ii) Cantonese might have borrowed postverbal adverbs from Kam-Tai or Hmong-Mien languages.⁵

In the absence of historical evidence which would allow the fixing of a specific date of borrowing, it is obviously difficult to favour one hypothesis over the other. If the

⁵ Stating that the contacts between Cantonese and Yao have been more conspicuous than the ones with Hmong, Dai Q. et al. (1992) believes that borrowing between Cantonese and Yao is more in evidence. He cites several syntactic structures to support his viewpoint.

first hypothesis were correct, however, the origin of Cantonese postverbal adverbs would remain unexplained, given that the hypothesis of internal change cannot be supported. In contrast, the second hypothesis can however be retained as a tentative proposal, though be it nigh impossible at the current stage of research to verify if these postverbal adverbs are to be found in earlier stages of the Kam-Tai or Hmong-Mien languages.

One can also, of course, go one step further and claim that Cantonese and Kam-Tai and/or Hmong-Mien are genetically related. Such an hypothesis, with the implication that Cantonese be distinguished from not only Chinese but from the Sino-Tibetan family, should of course be tested by similar analyses of other syntactic constructions.

3.6 Dative constructions

For the feature of ditransitive constructions, Sinitic languages are mixed languages according to Malchukov et al. (2010), and Haspelmath's classification (2005), based on Mandarin ditransitives: on the one hand, there are indirective or prepositional object constructions with the recipient [R] flagged by an adposition, typically with its source in a verb of giving, and the theme [T] zero marked, while, on the other hand, there are double object constructions in which neither recipient nor theme is marked.

3.6.1 Standard Mandarin

It turns out that in Standard Mandarin, there are not just two main types of word orders but six (see Zhu D. 1979; Peyraube 1986).

- A. V + R + T [R = Recipient/ indirect object (IO); T = Theme/ direct object (DO)]

(21) 他送了我一本书

Tā sòng-le wǒ yǐběn shū.
3SG offer-ASP 1SG one.CL book
'He offered me a book.'

- B. V + T + *gěi* + R [R is flagged by *gěi* = dative preposition]

(22) 他送了一本书给我。

Tā sòng-le yǐběn shū gěi wǒ.
3SG offer-ASP one.CL book to 1SG
'He offered me a book.'

- C. V + *gěi* + R + T

(23) 他送给我一本书。

Tā sòng gěi wǒ yǐběn shū.
3SG offer to 1SG one.CL book
'He offered me a book.'

D. *gěi* + R + V + T

- (24) 他给我写了一封信。

Tā gěi wǒ xiě-le yifēng xìn.
3SG to 1SG write-ASP one.CL letter
'He wrote a letter to me.'

E. *Bǎ* + T + V (+*gěi*) + R [T is flagged by *bǎ*, a direct object or 'pretransitive' marker]

- (25) 他把书送(给)我。

Tā bǎ shū sòng(gei) wǒ.
3SG BA book offer (to) 1SG
'He offered me a book.'

A sixth type of ditransitive construction of the double object or 'neutral' type is also found in other Sinitic languages, particularly of the Southern group, which M. Hashimoto (1976) has identified as a key parameter in distinguishing Northern from Southern Sinitic (Wu, Hakka, Cantonese Yue, Min, etc.). See also Chappell and Peyraube (2007) and §3.6.2 in this chapter on the exceptional status of Min.

F. Verb + DO + IO

- (26) Sin-on dialect of Hakka

人客一到就愛分茶佢。

Nyin²hak⁶ yit⁶ tau⁴ tshyu⁴ oi⁴ pin¹ tsha² ki².
guest as:soon:as arrive then must give tea 3SG
'As soon as the guests arrive, give them tea.'

What was the situation in earlier stages of Chinese and what is it today in Sinitic languages other than Standard Mandarin? There were three main dative structures available in Classical (Late Archaic) Chinese (fifth–second century BC):

- (i) V+IO+DO
- (ii) V+DO+ *yú*于+ IO [where *yú* is a dative preposition introducing the IO]
- (iii) *yǐ*以+DO+ V+ IO (or V+IO+ *yǐ*+ DO) [where *yǐ* is a preposition introducing the DO].

Examples are given below.

- (27) 公賜之食。 (左传) V+IO+DO

Gōng cì zhī shí.
prince offer 3SG food
'The prince offered him food.' (Zuò zhuàn, 5th c. BC)

- (28) 尧让天下与许由。 (庄子) V+DO+ *yú*于+ IO

Yáo ràng tiānxià yǔ Xǔ Yóu.
Yao leave Empire to Xu You
'Yao left the Empire to Xu You.' (Zhuāngzǐ, 4th c. BC)

Grammatical change in Sinitic languages and its relation to typology 69

- (29) 今买诸商人 (左传) yǐ以+DO+ V+ IO (N.B. zhū諸 = zhī 之 + yú于)

Jīn mǎi zhū shāng rén.
now buy it.from merchant people
'(He) bought it now from the merchant.' (Zuō zhuàn)

- (30) 孔子以其兄之子妻之 (论语) V+IO+ yǐ + DO

Kǒngzǐ yǐ qí xiōng zhī zǐ qī zhī.
Confucius OM 3SG brother GEN daughter marry 3SG
'Confucius gave him his niece in marriage.' (Lun yu, 5th c. BC)

A new construction emerged in the Pre-Medieval period with two verbs in series: V₁ + V₂ + IO + DO. The verbs that were able to fill the V₁ slot are all specific verbs of giving such as 'transmit, offer, sell, distribute', etc. Verbs in V₂ position are restricted to three different verbs that express only a general sense of giving. These are: yǔ 与 (與), yú 予 and wèi 遺.

The complex verb construction is obviously redundant, as the meaning 'to give' expressed by the V₂ is already included in that of the V₁. An example follows.

- (31) ...而厚分与其女财 (史记)

Ér hòu fēn yǔ qí nǚ cái.
and generously share give 3SG daughter property
'And (he) generously gave parts of his property to his daughter'.
(Shǐ jì 117, 1st c. BC)

Significantly, beginning in the Late Han period, there was a process of lexical unification for the V₂. The verb yǔ 与 (與) gained ascendancy over the two others: { yǔ 与, yú 予, wèi 遺} > { yǔ 与 (與)} (see Peyraube 1986, 1988).

In the following example, yǔ 予 in the Shǐ jì (1st c. BC) has been replaced by yǔ 与 in the Hán shū (1st c. AD):

- (32) 分与文君僮白人。 (汉书)

fēn yǔ Wénjūn tóng bái rén.
distribute give Wenjun slave hundred people
'(He) distributed a hundred slaves to Wenjun.' (Hán shū 57, 1st c. AD)

Another new structure emerged under the Early Medieval period: V1+DO +V2+IO, where the V₂ position is filled by yǔ 与 or sometimes (more rarely) by wèi 遺.

- (33) 时跋跋提国送狮子两头与乾陀罗王。 (洛阳伽蓝记)

Shí Bábátí guó sòng shī zǐ liǎng tóu yǔ
that-time Bactria kingdom offer lion cub two CL give
Qiántuólóuó wáng
Gandhara king

'At that time, the kingdom of Bactria offered two lion cubs to the king of Gandhara.' (Luòyáng qiélán jì, 6th c. AD)

Peyraube (1986) claims that the right derivation is the following one: $V_1+V_2+IO+DO > V_1+DO+V_2+IO$. Later on, circa the eighth–ninth centuries, V_2 *yǔ* 与 ‘to give’ grammaticalizes into a dative preposition ‘to’ ($[yǔ \text{ 与}, +V] > [yǔ \text{ 与}, + \text{Prep}]$), as in the following example.

- (34) 说与他道。 (朱子语类)

<i>shuō</i>	<i>yǔ</i>	<i>tā</i>	<i>dào.</i>
speak	to _{<give}	3SG	<i>dao</i>
'(He) spoke of <i>dao</i> to him.' (<i>Zhūzǐ yǔlèi</i> , 11 th –12 th c. AD)			

Around the same period, i.e. the ninth century, a new structure emerged where *yǔ* 与 + IO is found before the verb: *yǔ* + IO+V+DO.

- (35) 与老僧过净瓶水。 (祖堂集)

<i>yǔ</i>	<i>lǎo</i>	<i>sēng</i>	<i>guò</i>	<i>jìng</i>	<i>píng</i>	<i>shuǐ.</i>
to _{<give}	old	monk	pass	drinkable	bottle	water
'Pass me (Old Monk) a bottle of drinkable water.' (<i>Zǔ táng jí</i> , 952)						

Finally, as soon as the disposal construction using the prepositions *jiāng* 将 or *bǎ* 把 for introducing the preverbal DO became widespread around the eleventh–twelfth centuries, this latter structure, BA+DO+V+PREP+IO, completed the set of all the ditransitive constructions found in contemporary Standard Mandarin Chinese. In fact, the final development occurred in the eighteenth century with the lexical replacement of the dative preposition *yǔ* by the dative preposition *gěi*, also arising from a verb meaning ‘to give’ through a grammaticalization process.

What can be appropriately observed is that the structure V + DO + IO, in current use in certain Southern Sinitic languages (see construction F above), has never been historically attested in any stages of the history of the construction, except in the Pre-Archaic period (fourteenth–eleventh century BC), where it represented less than 20% of all the dative constructions (Peyraube 1988).

3.6.2 Early Southern Min (ESM)

Apart from Mandarin, Southern Min is the only other Sinitic language for which we have a reasonable corpus of historical materials, in this case, dating back to the early modern period of the sixteenth century. We have four verbs of giving in Early Southern Min (ESM): *khit*⁴ 乞 ‘to give’, *thou*³ 度 ‘to give’, *u*³ 與 ‘to give’, and *hou*⁷—‘to give’. (See C. Lien 2005; Chappell 2013).

These four verbs in the ESM period (from the sixteenth century onwards) have also been grammaticalized into prepositions marking the IO. We find the following three structures in representative documents for ESM: Verb+IO+DO, Verb+PREP+IO+DO, Verb+DO+PREP+IO, where the different prepositions used are all grammaticalized from verbs of giving. Examples follow with 乞 *khit*⁴ (in the structure V + DO+PREP+IO), with 度 *thou*³ (in V+ PREP+IO), which is certainly less frequent

Grammatical change in Sinitic languages and its relation to typology 71

and probably less grammaticalized than *khit*⁴, and with *u*³ 與, similar to Late Medieval Chinese *yǔ*, where one finds only the Verb+PREP+IO(+DO) structure, and never the Verb+DO+PREP+IO.⁶

(36) V+DO+PREP+IO

你撻落荔枝乞阮為記。

*Li*² *tan*³*loh*⁸ *nai*⁷*chi*¹ *khit*⁴ *gun*² *ui*⁵ *ki*³.

2SG throw.DIR litchi to<give 1SG as token

‘Throw down the litchi branch to me as a token (of your love).’

(*Lì Jìng Jì* [Romance of the Litchi Mirror], 1566)

(37) V+ PREP+IO

我送度汝。

Gua *sang* *tou* *lu*.⁷

1SG offer to<give thee

‘I offer (it up) to thee.’

(Mysteries of the Rosary, 24a, p.177, *Doctrina Christiana en letra y lengua china*, ca.1607)

(38) Verb+PREP+IO(+DO)

我今求汝。

Gua — *kiu* *lu*

1SG now beg 2SG

汝賜與我汝子來探我神魂。

lu *su* *u* *gua* *lu* *kia* *lay* *tam* *gua* *sinhun*.

2SG grant to<give 1SG 2SG son come search 1SG soul

‘I beseech thee, grant to me the coming of thy son to search my soul (and forgive me my sins).’

(Mysteries of the Rosary, 15a, p. 165, *Doctrina Christiana en letra y lengua china*)

These two postverbal prepositional object construction types found in ESM are the continuation of the Medieval Chinese ones. One can nevertheless notice that the preverbal position for the dative prepositions, i.e. the PREP+IO Verb DO, which

⁶ There are several available documents for Early Modern Southern Min: the *Doctrina Christiana en letra y lengua china* (ca. 1607); the operas *Lì Jìng Jì* 荔鏡記 (*The Romance of the Litchis and the Mirror*, dated 1566, but probably reflecting the language of the fifteenth or even the fourteenth century), supposedly written in both Chaozhou and Quanzhou dialects, and with different later versions known as the *Lì Zhi Jì* 荔枝记, dated 1581, 1638, 1885; *Arte de la lengua Chiō Chiu* (*Grammar of the Chiō Chiu language*, 1620); *Bocabulario de la lengua sangleya* (ca. 1617); *Dictionarium Sino-Hispanicum* (*Chinese-Spanish Dictionary*, 1604). I use the collated version of the *Doctrina Christiana* with both Chinese characters and the Spanish romanization which is provided in an appendix to Van der Loon (1967) but add my glossing and translations. See Chappell and Peyraube (2006a) for more details.

⁷ I reproduce here the original romanization used in the *Doctrina Christiana*, which does not indicate the tones.

appeared in Late Medieval Chinese around the ninth century, is never attested in ESM. Why? It could be because this development in Late Medieval Chinese occurred long after the Min dialects had split off (see Chappell and Peyraube 2007).

Moreover, in terms of word order typologies, Southern Min does not use the Cantonese type of double object construction Verb+DO+IO ('I give book you') and therefore is a counterexample to one of the parameters used in Hashimoto's (1976) North–South typological classification for Sinitic languages. It therefore aligns with the northern strategy for the double object construction in terms of word order while lacking the preverbal position for the dative preposition that is found in Mandarin.

3.6.3 Cantonese

As mentioned before, in Southern Sinitic languages such as Cantonese, Hakka, Wu, and some of the Xiang, Gan, and Pinghua dialect groups (but not Southern Min), the preferred word order is V+DO+IO instead of V+IO+DO.

- (39) 我畀一本书你。

Ngóh béis yāt búñ syù néih.
1SG give one CL book 2SG
'I give you a book.'

Where does this V + DO + IO structure come from, given that it is not attested in the different stages of Mandarin (Archaic, Medieval, Modern), except for a short period during the Pre-Archaic epoch (fourteenth–eleventh century BC)? Two different hypotheses have been put forward. The first one is an 'External Borrowing hypothesis' suggested by M. Hashimoto (1976) and later developed by Peyraube (1981), and by Yue-Hashimoto (1993), which puts forward the view that the structure has been borrowed from non-Sinitic languages with this word order and with which Cantonese has been in close contact, in the same manner as for the Cantonese postverbal adverbs (see §3.5). The source language could be either some Tai-Kadai language, or possibly even an Austroasiatic language. However, it is not possible to identify precisely which one specifically nor the date, even approximately, when the borrowing took place. Nonetheless, different advocates of this hypothesis would agree that the source language for the borrowing is a Tai language, most likely Zhuang.

Contra this 'External Borrowing hypothesis', it can be observed that there are a large number of other dialects which also display this V+ DO+IO word order and yet do not belong to the Cantonese or Yue group, lying in fact far to the north of this southern border area. In addition to the branches of Sinitic mentioned earlier (Hakka, Wu, Gan, Xiang, Pinghua etc.), it is striking to find that even dialects of Southwestern Mandarin, including Enshi 恩施, Badong 巴东, Dangyang 当阳, Jingmen 荆门, Jiangling 江陵, Yidu 宜都, Hankou 汉口, Hanyang 汉阳, Tianmen 天门, and Jingshan 京山, all located in Hubei province, use the V+ DO+IO word order as

do the Jianghuai Mandarin dialects of Tongcheng 桐城, Anqing 安庆, Wuhu 芜湖 all located in the province of Anhui (data from Yue-Hashimoto 1993:111). See also Cao et al. (2008, Map 96).

Furthermore, if contemporary Zhuang and Thai languages (all Tai-Kadai) have V+ DO + IO order, we know that Ancient Thai (thirteenth century) had the order of V + DO + PREP + IO. Therefore, it could not have been the source for the borrowing. As de Sousa points out (this volume), even amongst the Zhuang languages of Guangxi province, intermingled with Western Yue dialects, both word orders can be found. This effectively means the contact situation is much more complex than believed.

The second hypothesis is an ‘Internal Derivation’ hypothesis. Two different scenarios could have been in action:

- (i) V + IO + DO (which has been in existence from the Pre-Archaic period, 14th–11th c. BC) > V + DO + IO; or
- (ii) V + DO + PREP + IO (which has been in existence from the Pre-Archaic period) > V + DO + IO.

The first one, involving constituent movement, is however unlikely to be the right one. There is no evidence to argue the case that either the DO has moved backwards over the IO to be inserted immediately after the verb, or for the IO to move into a position following the DO.

In contrast to this, Xu L. and Peyraube (1997) argued in favour of the second scenario, involving the deletion of the preposition which introduces the IO. They proposed the following four arguments:

- (i) almost all the V + DO + IO structures can also appear under the V + DO + PREP + IO constituent order;
- (ii) there is often a pause between the DO and the IO of the V + DO + IO form, showing that there was probably an element between the two objects that has been deleted;
- (iii) the verbs involved in both structures V + DO + PREP + IO and V + DO + IO share the same semantic constraints, while the verbs involved in V + IO + DO do not (Shen Pei 1992; Peyraube 1987, 1988);
- (iv) new information is carried by the IO in both V + DO + PREP + IO and V + DO + IO structures, while it is carried by the DO in the V + IO + DO form.

There is one question still unsolved. As the V + DO + IO structure comes from the V + DO + PREP + IO one, through deletion of the dative preposition, why is this deletion process allowed in Cantonese, the two southern subgroups of Mandarin—Jianghuai and Southwestern Mandarin—as well as in some other Central and Southern dialects such as Wu, Gan, Xiang, and Hakka, not to mention also in Pre-Archaic Chinese (fourteenth–eleventh century BC), while it is not allowed in later stages of Chinese, nor in contemporary Standard Mandarin?

3.7 Verbs of saying

Grammaticalization of verbs of saying into complementizers is a well-known phenomenon for many African, East and Southeast Asian languages (see Saxena 1988; Lord 1993; Heine and Kuteva 2002; Fang M. 2006; Yeung K. 2006; Xu H. and Matthews 2007; Chappell 2008; Chen I. and C. Lien 2008; Güldemann 2008; C.F. Lien 2011; Chappell and Peyraube 2013 and others)

A complementizer (also known as a subordinating conjunction) is a grammatical word that enables a dependent clause to be embedded under the main clause. Example: 'I heard THAT there was a terrible train accident in China'. Fang M. (2006) was one of the first scholars to identify such complementizers in Pekinese, which have arisen from a *say* verb, as in:

- (40) 有很多人，他们就认为说这得政府给我们解决，....

Yǒu hěn duō rén, tāmen jiù rènwéi shuō
there.be very many people 3PL then consider that_{<say}
zhè déi zhèngfu gěi wǒmen jiějué.
this need government for 1PL resolve

'There are lots of people who consider that this has to be resolved for us by the government.'

Chappell and Peyraube (2006b) found similar phenomena in earlier stages of Chinese, and in some Southern dialects, such as Southern Min. The next example is from contemporary Taiwanese Southern Min.

- (41) 這事感覺講這足危險。

Che sū kám-kak kóng che chiok gùi-hiàm.
this situation feel that_{<say} this very dangerous
'[Toyotomi said: I] feel that the situation is very dangerous.' (Japanese tales 1081–3, Chappell 2008)

The grammaticalization process of *say* verbs into complementizers on the basis of quotative verbs is a two-stage process, as shown in Chappell (2008):

- (i) V [+ SAY] > Semi-complementizer;
- (ii) Semi-complementizer > Complementizer.

According to Chappell (2008), the Sinitic languages that have reached the first but not (yet) the second one are: Cantonese (the variety spoken in Hong Kong) and the Hakka of Sixian (四县话). The ones that have reached the second stage and thus have real complementizers arising from verbs of saying are: Pekinese, Taiwanese Mandarin, Taiwanese Southern Min, and the Chaozhou dialect in Fujian province. In other Chinese dialects and Sinitic languages, such as *pǔtōnghuà* (Standard Mandarin) and the dialects of Changsha (a variety of Xiang, in the Hunan province), Shanghai

Grammatical change in Sinitic languages and its relation to typology 75

(belonging to the Wu group of dialects), Nanchang (a Gan dialect in Jiangxi), Huojia (a Jin dialect in Shanxi), etc., the verbs of saying have not even reached the first stage of grammaticalization, still being used as quotative verbs. See also Chappell and Peyraube (2013).

Historically, in Pre-Archaic and Early Archaic (fourteenth–sixth century BC), we have the following *say* verbs: 言 *yán*, 云 *yún*, 曰 *yuè* (in oracle bone inscriptions), and 语 *yǔ* (in bronze inscriptions) (cf. Wang W. 2003). In Late Archaic, we also have 谓 *wèi*, 说 *shuō*, and 道 *dào*.

These different verbs of saying can be classified into two groups. In the first one, *yǔ* 语 and *yán* 言 are used as main verbs with their full lexical meaning; they are rarely used to introduce direct quotations, a function which will appear later, at the end of the Archaic period. In the second group, *yún* 云 and *yuè* 曰 are used as quotative verbs most of the time (both direct, especially in the case of *yuè*, and indirect). As for *wèi* 谓, it is often used as a verb for introducing indirect quotations. Finally, *shuō* 说 and *dào* 道 are actually very rarely used as *say* verbs in Classical Chinese. This is because *shuō* principally means ‘to explain’, and *dào*, ‘to discuss’. None of these verbs are complementizers or semi-complementizers. They all have their full lexical meaning but may also be used as quotative verbs in some cases.

In Medieval Chinese, *huà* 话 appears during the Tang period (ca. seventh–eighth century AD) and *shuō* 说 as a *say* verb appears between the Tang and the Song dynasties (ninth–eleventh century AD), where it is often used as a verb which introduces direct quotations. In the *Xiào jīng zhí jiě* 孝经直解 (fourteenth century), there are already 55 *shuō*, out of which 27 are used for direct quotations (see Wang, W. 2003). As for *dào*, it is also widespread during the Northern Song period (tenth–twelfth century AD) as a lexical *say* verb, but most of the time it is used for indirect quotations. Among the 41 *dào* in the *Sāncháo bēiméng huibiān* 三朝北盟汇编 (twelfth century), 26 are indirect quotation verbs (Wang W. 2003). *Shuōdào* also appears under the Song, but becomes widely used only under the Yuan (thirteenth–fourteenth century AD):

(42) 说道韩家有几万。 (三朝北盟汇编 5)

Shuōdào Hán jiā yǒu jǐ wàn.
Say+say Han family have several ten-thousand
(They) said the Han had several thousands.
(*Sāncháo bēiméng huibiān* 5, 12th c.)

Finally, *jiāng* 讲 appears under the Yuan (1206–1368).

Liu D. (2004) hypothesized that the ‘V+dào’ structure comes from the ‘*shuō+dào*’ compounds, after the verb *shuō* 说 has been extended to other verbs of saying through process of analogical extension (*lèituī kuòzhǎn* 类推扩展 in Liu D.’s terms), and later, *dào* 道 grammaticalizes into a complementizer : V₁ + V₂-*dào* > V₁ + Complementizer. This grammaticalization process, according to Liu D. (2004),

could have occurred at the end of the Tang or at the beginning of the Song (tenth–eleventh century AD).

Liu D.'s hypothesis is difficult to uphold, as will be next explained. Chappell and Peyraube (2006b, 2013) propose another scenario that is more firmly grounded in diachrony. They argue (i) that *dào* never became a real complementizer, as the process stopped at the first stage (semi-complementizer); and (ii) that the grammaticalization process for the other *say* verbs occurred much later than the Tang-Song period. Relying on Li M. (2003), they also observed that many *say* verbs have or have had the cognitive meaning of *yǐwéi* 以为 'to believe, to think'. Among them, one can find *yán* 言 (whose meaning of 'believe' is already attested under the Six Dynasties period, seventh–eighth century AD), *yún* 云 (also under the Six Dynasties and the Tang), *dào* 道 (later, around the fourteenth century). The semantic change has been the following one:

SAY verbs (factive) > to consider (*rènwéi* 认为, non factive) > to think (*yǐwéi* 以为, contrafactive) (cf. Li M. 2003).

Examples of some *say* verbs having a cognitive meaning are:

- (43) 时论乃云胜山公。 (世说新语: 方正)

Shí lùn nǎi yún shèng Shān Gōng.
at.that.time evaluate then consider surpass Shan Gong
'The evaluators of that time then considered (that he was) superior to Shan Gong' [*yún* = 'consider']. (*Shi shuō xīn yǔ: Fāng Zhèng* 5th c.)

- (44) 周秀便理会得道是个使钱的勤儿。 (大宋宣和遗事: 享集)

Zhōu Xiù biàn lìhuide dào shì ge shǐ qián de qínr.
Zhou Xiu then understand think be CL spend money PRT pimp
'Zhou Xiu then understood (the situation) and thought (he was) the pimp who had spent the money' [*dào* = 'to think'].
(*Dà Sòng Xuǎn hé Yì Shi: Hēng Jí* 13th c.)

In other words, the complementizer probably does not evolve directly from any verb [+ SAY], but from a cognitive verb meaning 'to think', 'to believe'. The semantic change could have been as follows:

SAY > CONSIDER > THINK > COMPLEMENTIZER

The change from 'say' to 'consider' to 'think' is a typical example of what Traugott (1989: 35) calls 'subjectification'. Moreover, the last derivation (THINK > COMPLEMENTIZER) and only this particular one, represents a grammaticalization process which could not have taken place before the second half of the twentieth century.

Chappell and Peyraube (2013) use evidence from Early Modern Southern Min (fourteenth–seventeenth century, see above, note 4), the contemporary spoken Beijing dialect and contemporary Taiwanese Mandarin.

Grammatical change in Sinitic languages and its relation to typology 77

In Early Modern Southern Min, the main *say* verb in the *Doctrina Christiana* (1607) is *yuè* 曰; in the *Lì Jīng Jì* 荔鏡記 and *Lì Zhī Jì* 荔枝记, we also find: 呀 *tann3*, 說 *seh4*, and 叫 *kio3*. None of these verbs is used as a complementizer. They are full lexical verbs, quotative verbs or, at the best, semi-complementizers (first stage). For example:

- (45) 汝子西士氏差一天人名曰山呀臘迷降下。報說僚氏要援汝身與神魂。
Lu kia Jesu che tien yīn — *San Galabe cang he*
2SG son Jesus send angel called Saint Gabriel go down
po sue Diosi ay uan lu sin cang sinhun
announce-say God want hold 2SG body and soul
'Your son Jesus sent the angel Gabriel down to announce to you that God wants to bring your body and soul (up to heaven to be blessed).' (see also Note 5) (*Doctrina Christiana in letra y lengua china*, in Van der Loon, 1967: 179–80)

The grammaticalization of *shuō* 說 in the Beijing dialect is also a late phenomenon. It is only attested as a semi-complementizer in the second half of the nineteenth century, and not any earlier (Fang M. 2006). Examples of this first stage are found in the *Guānhuà zhǐ nán* 官话指南 (*The Mandarin guide*), reflecting the language of the late nineteenth century, as in:

- (46) 是, 我來是和您打聽一件事情, 是我聽見說您這西院裏那處房要出租, 是真的麼?

Shì, wǒ lái shì hé nín dǎtīng yī jiàn shìqíng,
yes 1SG come be with 2SG enquire one CL matter
shì wǒ tīngjiàn shuō nín zhè xī yuàn-li nà
FOC 1SG hear say 2SG this west courtyard-in that
chūfáng yào chūzū shì zhēnde ma?
kitchen want rent be true PRT

'Yes, I've come to enquire about a matter with you. I heard it said you wanted to rent out the kitchen in the west courtyard. Is that true?'

The use of *shuō* 說 in this last example represents the first stage of the semi-complementizer. To reach the second stage of grammaticalization as a real complementizer, we have to wait until the second half of the twentieth century for this category to emerge in the Beijing dialect.

- (47) 我总是觉得说生活里缺了点儿什么。

Wǒ zǒngshì juéde shuō shēnghuó-lì quē le diǎnr shénme
1SG always feel say<that life-in lack ASP little something
'I've always felt that there is something a little lacking in my life.'

(Fang M. 2006: 109)

A similar case exists for contemporary Taiwanese Southern Min and Taiwanese Mandarin: complementizers are only attested from the late twentieth century.

- (48) 伊 干單唸 講, 阿和趁足濟 錢... (Taiwan Southern Min)
I kan-na liām kóng, á-hô thàn chiok-chē chin...
3SG just nag say<that (name) earn much money
'He kept insisting that A-Ho earns a lot of money.' (Fate 455–6, from Chappell 2008)

- (49) 他们不知道说系里面最近有什么样的事情。
Tāmen bù zhīdào shuō xì-lǐmiàn zuijìn yǒu shénmeyàng
3PL NEG know say<thatdepartment-in recently have what
de shìqíng
PRT matter
'They didn't know what was happening in the department.'
(Taiwanese Mandarin, data from Huang S. 2003: 440)

To sum up, the grammaticalization of *say* verbs into real complementizers, which is attested in only some of the Sinitic languages and dialects, though far from all, is a relatively late phenomenon. It probably did not occur before the second half of the twentieth century. Two stages for the grammaticalization process have been identified: (i) from a quotative verb > semi-complementizer; (ii) from a semi-complementizer > complementizer.

3.8 Conclusion

I have tried to show that, in spite of quite indirect links between diachronic syntax and typology, historical grammar can help us test well-grounded hypotheses in the domain of linguistic typology, at least in the case of Sinitic languages, as the history of Chinese languages goes back to the second millennium BC. Furthermore, data extracted from Chinese dialects, that is from the non-Mandarin Sinitic languages, may also serve in future to provide insights for solving historical issues.

We should nevertheless be aware, when doing research on the historical evolution of Chinese or on the diversity of Sinitic languages and dialects, of the importance of taking into consideration the phenomena of external borrowing and contact-induced change, and not only of the internal mechanisms of change. The situation in Chinese linguistics today is about the same as it was ten years ago, when Traugott (2001) wrote: 'Most of analyses up to now have been conducted largely in the context of putative homogeneous developments. When we look at contact situations, complications arise.' This has been shown above in particular for the case of Cantonese.

Part II

Typological Studies of Sinitic Languages

OUP UNCORRECTED PROOF – REVISES, 27/10/2015, SP1

4

The semantic differentiation of demonstratives in Sinitic languages

YUJIE CHEN

4.1 Introduction

This analysis explores the common categories found in the semantic differentiation of demonstratives in Sinitic languages. While two-term systems are very common in Sinitic languages, there are also one-term, three-term, four-term, even five-term systems, based on the scale of distance and visibility. Neutral demonstratives form a one-term system, notably which can co-occur with other types of systems in the same language.

Systems greater than three-term can be single-layered or multi-layered, determined by semantic properties other than distance such as visibility in Sinitic. Special morpho-phonological processes, such as lengthening, reduplication, and tone change, have been identified in systems greater than the three-term type to denote an additional semantic differentiation within the system.

As the result of my investigation of a representative sample of 303 Sinitic languages, distance is found to be more easily encoded by demonstrative systems in Sinitic languages than any other semantic properties.

4.1.1 *Nature of demonstratives*

Demonstratives are closely related more specifically to pointing/indexical gestures and, in a sense, they can be regarded as the lexical correspondence of gestures. In many situations, they have the same function as gestures: both of them are used to identify particular objects. The most effective way to identify an object is by the use of gestures which can directly single out an intended object from many entities of the same kind. Demonstratives as a linguistic expression have natural disadvantages compared with gestures for pointing, because they refer to referents by means of phonological forms which are not primarily visual, that is, by words. For identifying

objects more accurately and speedily, many semantic properties can be encoded by demonstrative systems, such as distance, animacy, altitude, and so on, as Diessel (1999) points out. These semantic properties can help an addressee recognize the exact location of an object and then identify it.

Many linguists regard distance as a default and essential property of a demonstrative. Diessel (2005) considers that demonstratives indicate the relative distance of a referent to its reference centre (also called its deictic centre). While he recognizes that distance is just a typical property of demonstratives, other contrastive pairs, such as ‘visible’ versus ‘out-of-sight’, ‘higher elevation’ versus ‘lower elevation’, ‘uphill’ versus ‘downhill’, ‘upriver’ versus ‘downriver’, and ‘moving toward’ versus ‘moving away from’ a deictic centre, can all be encoded by demonstratives. Compared with other semantic items, it is distance, however, which forms the basic core of all the properties related to demonstratives, i.e. it appears to be the most easily encoded by demonstratives. This conclusion can similarly be justified for Sinitic languages and a sample of other Southeast Asian languages, as will be shown in this analysis.

Nonetheless, it must be pointed out that distance may not always be the essential property of demonstratives in all languages. In some languages, there are demonstratives which are not marked by distance, i.e. they do not indicate any distance from the reference centre to the intended referents. This type of demonstrative is called a ‘neutral demonstrative’, which we will also discuss in one of the following sections, §4.2.1.

4.1.2 Previous studies on semantic properties encoded by demonstratives

Demonstratives are used to identify a particular referent in extra-linguistic contexts or in discourse mainly by their pragmatic features. The first deictic use is also called the ‘straightforward situational use’ in Himmelmann (1996: 205), and the ‘exophoric use’ in Diessel (1999: 6). Other aspects are their discourse deictic use, anaphoric function, and recognitional function, etc. In addition to these pragmatic functions, a referent can be more easily recognized by some semantic properties, such as distance, visibility, animacy, or even topographical factors. These semantic features can be selectively encoded by demonstrative systems in different languages, as they are supplementary devices to the pragmatic functions of indexing: they enable more rapid and accurate identification of intended referents than only through pragmatic functions.

Many linguists distinguish between distance features and the qualitative features coded by the semantic properties of demonstratives (Lyons 1977: 648; Diessel 1999: 35). The former indicates the distance of a referent to its reference centre, while the latter indicates the characteristics of a referent. From the definition we can see that only the former can be used for the differentiation of demonstratives, the latter itself is not used to divide the deictic range based on distance from reference centre to the intended referents. Nonetheless, ontological category, as one of the

qualitative features, is indeed related to the hierarchy of demonstratives. This entails that the distinctions made for deictic range in a language may vary according to the ontological category of referents, as proposed by Chu Zexiang and Deng Yunhua (2003). For example, in the Wujiang dialect (Wu dialect), demonstratives referring to ordinary objects have a two-way distinction, while those for location have a three-way distinction (Liu Danqing 1999). So to describe a demonstrative system of a language, it is necessary to first differentiate the ontological categories of referents for which demonstratives are used.

Chu and Deng (2003) find a hierarchy for the differentiation of demonstratives concerning different ontological categories, which can be reformulated as the following implicational universal (Liu and Liu 2005):

Degree/manner ⊃ time ⊃ ordinary object ⊃ location

This universal means that for any given language, if one category of demonstratives has a N-way distinction, the category on its right-hand side has at least a N-way distinction. That is to say, if the demonstrative system for degree or manner is two-term in a given language, the demonstrative systems for other categories on the right-hand side of the degree/manner feature shown in the above diagram are also at least two-term. This phenomenon shows that the differentiation of demonstratives is related to the semantic categories of referents, while it is also possible that syntactic position can influence the differentiation of demonstratives.

As Diessel (2005) points out, it is possible for the division of adnominal demonstratives to be different from that of pronominal demonstratives in the same language, and he uses Tongan (Polynesian) to support this viewpoint. In Tongan, the adnominal demonstratives include two terms which distinguish between proximal and distal, whereas, the pronominal demonstratives are comprised of three distance-marked demonstratives and a distance-neutral one. Diessel concludes that there may be different distance contrasts in different syntactic categories for one and the same language, without providing further discussion on this question.

Chu and Deng (2003) study the differentiation of demonstratives from an ontological perspective, while Diessel (2005) mainly focuses on the relation between syntactic properties and distance contrasts. The discussions above show that it is possible that distance contrasts vary depending on syntactic or semantic categories of demonstratives in a language. So it is clearly necessary to distinguish between different semantic and syntactic categories to completely describe a demonstrative system. Taking into account all these factors, demonstrative systems may prove to be quite complex, even though they form a closed type of word class.

Ontology and syntactic position are always closely related to the differentiation of demonstratives. It seems that some non-parallel divisions of demonstratives are decided by syntactic properties, such as in German, where there is one term system for adnominal demonstratives versus a two-way contrast for adverbial

demonstratives. In modern German, as Diessel (2005) points out, the adnominal demonstratives *dieser* and stressed *der*, *die*, *das* are distance-neutral, while the adverbial forms *hier* ‘proximal’ and *da* ‘distal’ are used to express the distance contrast. This phenomenon in fact can be explained from an ontological viewpoint: adnominal demonstratives are generally used to refer to ordinary entities while adverbial ones are used for locations.

4.1.3 Objectives of this analysis

Diessel (2005) gives a basic description of demonstratives crosslinguistically, pointing out that 54.4% of all languages have a two-way contrast between proximal and distal in adnominal demonstratives, whereas only 37.4% have a three-way distinction. The majority of the latter have a distance-oriented system, that is to say, they distinguish three points for referents located in different positions on a distance scale. About two-thirds of the three-term languages are distance-oriented¹ and only one third is person-oriented.

When Diessel discusses the geographical distribution of adnominal demonstratives, he draws a conclusion that two-term and three-term systems are the most prevalent throughout the world, while the occurrence of smaller and larger systems is restricted to certain areas: one-term systems occur quite frequently in Africa, and they are also found in Europe (German, French) and Meso-America (Mam). Four- and five-term systems occur in three regions: North America, Africa, and the Pacific region. The data on which his conclusions are based include some Asian languages, but he fails to mention that there are one-term, three-term, and systems with even more than three-terms found in Asian languages as well.

This chapter will discuss the semantic differentiation which underlies the different types of demonstrative systems found in Sinitic languages, with a brief description of the situation in Tibeto-Burman languages at the end. Except for the most common type of two-term system, there are both one-term demonstrative systems and those with greater than three terms to be found in Sinitic languages, thus presenting a great variety of types. Systems with or exceeding three-terms may be single-layered or multi-layered. Qualitative features, such as ontological category, animacy, sex, humanness, number, etc., have different effects on demonstrative systems in the various Sinitic languages, which we will not be able to discuss in detail in this chapter (see however Chen 2010 for an in-depth study).

¹ Fillmore (1982: 49–50) divided demonstratives into two classes: distance-oriented systems which take the speaker as the only deictic centre, and person-oriented systems, which take both speaker and hearer as the deictic centre. Anderson and Keenan (1985: 282) point out, in person-oriented systems, the middle term indicates objects close to or identifiable by the addressee, while in distance-oriented systems, the middle term indicates that the referent is farther than a proximal referent and closer than a distal one.

The research focus of this chapter will in fact be on semantic features which present great diversity in these languages, and on which the differentiation of demonstrative systems is ultimately based. A sample of 303 Sinitic languages is used to investigate this topic from a crosslinguistic perspective (these data all come from Chen 2010: 178–205). Hence, we mainly focus on spatial indexing, which is the typical function of demonstratives, and do not discuss other types of functions of demonstratives in this chapter.

4.2 Common categories in the differentiation of demonstratives in Sinitic languages

4.2.1 One-term systems and neutral demonstratives

One-term systems consist of only one member, which can function as either a demonstrative pronoun or an adnominal. As Anderson and Keenan (1985) point out, demonstratives in a one-term system simply indicate something like ‘present to the speaker’ or ‘present in the extralinguistic context of the utterance’, and they do not commit to the distance from the speaker or the visibility for the speaker. From this viewpoint, it can be said that the demonstrative in a one-term system is distance-neutral and can be called a neutral demonstrative.

While Anderson and Keenan (1985: 280) argue that a distance-neutral demonstrative differs very little from a definite article or a third person pronoun, Himmelmann (1996) takes a totally different standpoint. As he points out, in colloquial German there are two sets of adnominal demonstratives preceding head nouns: *dieser* (*diese*, *dieses*, *die* etc.) and stressed *der* (*die*, *das*, *die* etc.),² all of which are distance-neutral. Diessel (2005), like Himmelmann, treats these as demonstratives rather than definite articles, for the reason that they are normally used to focus the hearer’s attention on entities in the surrounding situation, which is not the function of definite articles. In order to indicate the distance from a reference centre to an intended referent, these neutral forms must be accompanied by a demonstrative adverb *hier* ‘proximal’ or *da* ‘distal’. Diessel’s analyses (1999, 2005) clearly address the issue of neutral demonstratives as a linguistic category, but provide few examples and do not discuss their characteristics. In this part, we will provide a detailed explanation of neutral demonstratives based on the sample of more than 30 Sinitic languages.

4.2.1.1 [Gəʔ⁴⁴] 焉 in the Suzhou dialect 苏州 (Wu group 吴语)

As mentioned, neutral demonstratives have nothing to do with distance and their function is just to point out an intended referent without indicating any distance meaning from the deictic centre to the referent. In the following paragraphs, we take

² The choice of the form depends, of course, on the gender, number, and case of the head noun.

as an instance [gəʔ⁴⁴] in the Suzhou dialect (Wu) to illustrate the function of neutral demonstratives.

The lexeme [gəʔ⁴⁴] has first been described by Xiaochuan Huanshu (1981) who regards it as a medial demonstrative corresponding to the medial demonstrative *so-* in Japanese. In Japanese, the demonstrative system is person-oriented, and the *so-*-series of medial demonstratives is used for referents close to hearers. We do not agree with Xiaochuan's analysis because the meaning of [gəʔ⁴⁴] is relatively complicated, compared with the *so*-series of demonstratives.

In the Suzhou dialect, [gəʔ⁴⁴] is a very special demonstrative which is used alongside a separate set of proximal and distal demonstratives, [E⁴⁴/KE⁴⁴] and [UE⁴⁴/KUE⁴⁴]. According to Liu Danqing (1995), [gəʔ⁴⁴] can indicate referents appearing in a speech situation, which may be close to or far away from the deictic centre, and under these conditions, [gəʔ⁴⁴] is always used in combination with gestures. Liu therefore claims that [gəʔ⁴⁴] can serve both proximal and distal functions, as shown in example (1).³

[Context: a tour guide introduces a stone stele to her tourists with gestures:]

(1) 碑块碑有五百年历史哉。

gəʔ⁴⁴ kʰuE⁵²³ pE⁴⁴ Y²³¹ n⁵²pəʔ⁴³ nil²²³ liəʔ²³sɪ⁵¹ tsE³.

DEM_{NEUT} CL tablet have five.hundred year history PRF

'This/that stone stele has five hundred years of history.'

[Gəʔ⁴⁴] in sentence (1) can be used under the condition that the stele is visible. Furthermore, it may be either close to or far away from the speaker. Shi Rujie (1999) calls [gəʔ⁴⁴] a specific demonstrative whose distance meaning varies according to the contexts in which it occurs, giving further elaboration, as follows:

(i) [gəʔ⁴⁴] can refer to objects in the vicinity of a deictic centre when it is in contrast to the distal demonstratives [UE⁴⁴/KUE⁴⁴];

(2) 我欢喜碑种样子个，勿欢喜弯种。

n̩əu²³¹ huø⁴⁴ci⁵¹ gəʔ⁴⁴ tson⁵¹ ia²³¹tsi³ kəʔ⁴³, fəʔ⁴³ huø⁴⁴ci⁵¹ uE⁴⁴

1SG like DEM_{NEUT} type style MOD NEG like DEM_{DIST}

tson⁵¹.

type

'I like this kind of style, but don't like that kind (of style).'

³ I thank Prof. Wang Jian who provided the IPA transcription of the examples in the Suzhou dialect. In this chapter, examples have been transcribed in IPA as far as possible. In cases where the transcription uses the Chinese *pīnyīn* romanization in capitals, according to the standard pronunciation of the corresponding Chinese characters in Standard Chinese, this has been adopted to overcome the lack of transcription in the source materials.

The semantic differentiation of demonstratives in Sinitic languages 87

- (ii) [gəʔ⁴⁴] has a distal meaning when it contrasts to the proximal demonstratives [E⁴⁴/KE⁴⁴].
- (3) 哀个勿好，痴个好一点。
E⁴⁴ kəʔ² fəʔ⁴³ hæ⁵¹, gəʔ⁴⁴ kəʔ⁴ hæ⁵¹ iəʔ⁴³tI⁵¹.
DEM_{DIST} CL NEG good DEM_{NEUT} CL good a.little
'This is not good, but that is a little better.'

Nonetheless, I disagree with Shi Rujie: From the examples above, we can see that the meaning of [gəʔ⁴⁴] is in fact flexible: it can be explained as either proximal or distal in meaning, according to the other demonstrative forms which co-occur with it. Furthermore, according to Jin Youjing (1962), when co-occurring with proximal and distal demonstratives, [gəʔ⁴⁴] can apparently also function as a medial demonstrative, which indicates a referent which is closer than a distal object and further away than a proximal object, as shown in this example, with which I will also take issue below:

- (4) 哀搭是苏州，弯搭是杭州，痴搭是上海。

E ⁴⁴ ta?	Zi ²³¹	səu ⁴⁴ tsøY ⁴⁴
DEM _{PROX}	COP	Suzhou
uE ⁴⁴ ta?	Zi ²³¹	fī ²²³ tsøY ⁴⁴
DEM _{DIST}	COP	Hangzhou
gəʔ ⁴⁴ ta?	Zi ²³¹	Zā ²³¹ hE ⁵¹ .
DEM _{NEUT}	COP	Shanghai

'This is Suzhou, that over there is Hangzhou, and that (in the middle) is Shanghai.'

(Context: the speaker is in Suzhou, and (s)he points out in turn Suzhou, where she is located, the town of Hangzhou which is furthest away, and then Shanghai which is somewhere in between the two.)

However, Li Xiaofan (1998) presents a different analysis for the medial meaning of [gəʔ⁴⁴]. As he points out, if there are three contrastive referents, [gəʔ⁴⁴] is never used as a medial demonstrative. Instead, the three-way contrast is expressed by the proximal [E⁴⁴/KE⁴⁴] and the distal [UE⁴⁴/KUE⁴⁴] in combination with gestures. My synthesis is thus as follows.

[Gəʔ⁴⁴] as a basic demonstrative morpheme is clearly a neutral demonstrative whose meaning is inferred from the particular context, and which can serve as a pronoun or an adnominal element for many types of semantic functions: it can refer to ordinary objects as in example (1), to location in combination with [ta?] 搭 or [mil] 面 as in example (4), and to time in combination with [ciə²³] 歇. Its meaning is inferred from the particular context.

4.2.1.2 [kəʔ⁵⁵] 葛 as a neutral demonstrative in the Chongming dialect (Wu group 吴语)

The demonstrative system in the Chongming dialect is very complex, as Liu and Liu (2005) have observed. The lexeme [kəʔ⁵⁵], as a member of the complex

demonstrative system, is distance-neutral and can be used flexibly, like [gə?⁴⁴] in the Suzhou dialect, to denote referents nearby or far away from a given deictic centre.

- (5) 吉本书我咯，葛本书是我图书馆里借咯。

<i>tɕia?</i> ⁵⁵	BEN	SHU	WO	GE,
DEM _{PROX}	CL	book	1SG	MOD
<i>kə?</i> ⁵⁵	BEN	SHU	SHI	WO TUSHUGUAN LI JIE GE.
DEM _{NEUT}	CL	book	COP	1SG library LOC borrow MOD

‘This book is mine, and that book is borrowed from the library.’

Liu and Liu (2005) consider that [kə?⁵⁵] in example (5) has a distal meaning.

By way of contrast, Liu and Liu (2005) consider that [kə?⁵⁵] in sentence (6) has a proximal meaning.

- (6) 葛瓶香水是我新买咯，送拨尔。

<i>kə?</i> ⁵⁵	PING	XIANGSHUI	SHI	WO	XIN	MAI	GE,
DEM _{PROX}	bottle	perfume		COP	1SG	new	buy PRT
SONG	BO	ER.					
send	give	2SG					

‘This bottle of perfume was just bought by me, and I would like to give it to you (as a present).’

Once again, the distance meaning of [kə?⁵⁵] seems to be inferred from the particular contexts in which it appears, rather than being an intrinsic semantic feature of the demonstrative itself. [Kə?⁵⁵] can also be used to refer to several objects that co-occur in the same speech situation. Moreover, these objects may be close or far away, according to their actual distance from the deictic centre. A case in point is given in example (7) where none of the instances of [kə?⁵⁵] has any contrastive meaning whatsoever, while being completely neutral for the distance feature.

- (7) 葛个我里爸爸，葛个我里姆妈，葛个我里妹妹。

<i>kə?</i> ⁵⁵	GE	WO	LI	BABA
DEM _{NEUT}	CL	1SG	MOD	father
<i>kə?</i> ⁵⁵	GE	WO	LI	MUMA
DEM _{NEUT}	CL	1SG	MOD	mother
<i>kə?</i> ⁵⁵	GE	WO	LI	MEIMEI.
DEM _{NEUT}	CL	1SG	MOD	younger.sister

‘This/that is my father, this/that is my mother, and this/that is my younger sister.’

(Context: all relevant members, including the speaker, their father, their mother, and their sister are all in the speech situation, and the speaker introduces them one by one to the hearer.)

The semantic differentiation of demonstratives in Sinitic languages 89

The demonstratives in example (7) are interpreted according to the speech situation in which they occur. While with every indexing action, the actual distance from the referents to the deictic centres can be considered an objective ‘real world’ one, and as a possible semantic feature for expression, it is obligatorily suppressed under the condition where distance-neutral demonstratives are used. Distance-neutral demonstratives function purely to identify objects.

4.2.1.3 [ny35] as a neutral demonstrative in the Wuhan dialect 武汉 (Jianghuai Mandarin 江淮官话)

Similarly to the Suzhou and Chongming examples above, in the Wuhan dialect (Jianghuai Mandarin), as Zhu Jiansong (1992) observes, the neutral demonstrative [ny35] is always used to point out objects whose distance to the deictic centre is irrelevant.

- (8) 口是生的, 口是熟的。

ny ³⁵	SHI	SHENG	DE,	ny ³⁵	SHI	SHU	DE.
DEM _{NEUT}	COP	unripe	MOD	DEM _{NEUT}	COP	ripe	MOD

‘This one is unripe, and this one is ripe.’

- (9) 我在口里, 他在口里。

WO	ZAI	ny ³⁵ LI	TA	ZAI	ny ³⁵ LI.
1SG	at	DEM _{NEUT}	3SG	at	DEM _{NEUT}

‘I am here, and he is here.’

4.2.1.4 Neutral demonstratives referring to locations

Diessel (2005) only discusses neutral demonstratives as adnominal elements. In fact, as shown already, neutral demonstratives can serve many types of syntactic and semantic functions, determined by the context of use and their overt features for distance. From a syntactic perspective, neutral demonstratives can serve as pronouns (including those referring to ordinary objects and those referring to location)⁴ or adnominal elements, while from a semantic perspective, neutral demonstratives are used to refer to ordinary objects, location, manner, degree, and time.

For example, in the Qingliu 清流 dialect (a sub-branch of Hakka), the term [t^hain]²¹[li²⁴] 田里 can be used independently as a demonstrative pronoun or can directly follow a noun or a pronoun. It is described as a demonstrative ‘unrelated to distance’ by Xiang (1999). Specifically, it may be used to refer to a particular location without indicating any distance meaning, and can thus be viewed as a neutral demonstrative. There are in addition three distance-marked demonstratives for indicating location: the proximal 这里 [ti²⁴li²⁴] ‘here’; the medial 扁里 [pair²⁴li²⁴] ‘there’; and the distal 解里 [ka²⁴li²⁴] ‘over there (yonder)’, as shown in the following examples.

⁴ In Sinitic languages, demonstratives referring to location generally function as pronominal elements rather than as adverbial elements, as in English.

- (10) 这里有一只青口[kuain]⁵⁵。
ti²⁴li²⁴ YOU YI ZHI QINGkuain⁵⁵.
DEM_{PROX} have one CL frog
'There is a frog here.'
- (11) 扁里出香菇木耳弄事子。
pain²⁴li²⁴ CHU XIANGGU MUER NONGSHIZI.
DEM_{MED} yield mushroom agaric others
'That place yields mushrooms, agarics, and other things as well.'
- (12) 我也要去解里。
WO YE YAO QU ka²⁴li²⁴.
1SG also want go DEM_{DIST}
'I want to go over there as well.'

4.2.1.5 Ways to express distance contrasts with neutral demonstratives

In some languages, the neutral demonstrative, while being a member of the whole demonstrative system, is complementary to a separate set of distance-marked demonstratives, as is the case of [gə?⁴⁴] in the Suzhou dialect: [gə?⁴⁴] is a neutral demonstrative, while distance-marked [E⁴⁴/KE⁴⁴] and [UE⁴⁴/KUE⁴⁴] respectively indicate proximal and distal meaning.

According to Dai (1999), in the Taihe 泰和 dialect (a sub-branch of Gan), the neutral demonstrative 格 [GE] is the sole member of the demonstrative system for ordinary objects, and it can function as a pronominal or an adnominal element, while there are two demonstratives for location. Therefore, the distance contrast for two ordinary objects between distality and proximity is realized by attaching the distal location demonstrative [ko⁴²lo³³] to [GE], while the neutral demonstrative itself is used to express proximal meaning, displayed in the following sentences.

- (13) 格是一堆橘子。
GE SHI YI DUI JUZI.
DEM_{NEUT} COP one CL orange
'This is a pile of oranges.'
- (14) 格栋屋是叔叔个，葛罗格栋屋是伯伯个。
GE DONG WU SHI SHUSHU GE
DEM_{NEUT} CL house COP father's.younger.brother MOD
ko⁴²lo³³ GE DONG WU SHI BOBO GE.
DEM_{DIST} LOC DEM_{NEUT} CL house COP father's.elder.brother MOD
'This house belongs to my father's little brother, and that house belongs to my father's elder brother.'

This is just one way to express distance contrast with neutral demonstratives.

The semantic differentiation of demonstratives in Sinitic languages 91

Another way to express distance contrasts with neutral demonstratives depends on the use of different construction types, such as those in the Ningbo 宁波 dialect (a member of the Wu dialect group, data from Pan and Tao 1999) and the Shengze 盛泽 dialect (also a member of Wu, data from Liu 1999), both exemplified below:

Ningbo dialect:

- (15) 个个是我个，个个是渠个。

GE	GE	SHI	WO	GE,	GE	GE	SHI	QU	GE.
DEM _{NEUT}	CL	COP	1SG	MOD	DEM _{NEUT}	CL	COP	3SG	MOD

‘This is mine, and this is his.’

In example (15), the two neutral [GE] demonstratives are used for two referents which exist in the speech situation, and they do not indicate any distance contrast.

- (16) 个个是我个，个面一个渠个。

GE	GE	SHI	WO	GE,		
DEM _{NEUT}	CL	COP	1SG	MOD		
GE	MIAN	YI	GE	SHI	QU	GE.
DEM _{NEUT}	side	one	CL	COP	3SG	MOD

‘This is mine, and that is his.’

In contrast to (15), in example (16), the neutral [GE] 个 is used to identify a referent nearby, while the combination of [GE] 个 and the noun [MIAN] 面 ‘side’ is used for a referent far away from the deictic centre.

Shengze dialect:

- (17) 葛只是茶杯，葛只是酒杯。

GE	ZHI	SHI	CHABEI,	GE	ZHI	SHI	JIUBEI.
DEM _{NEUT}	CL	COP	cup	DEM _{NEUT}	CL	COP	wine.glass

‘This is a cup, and this is a wine glass.’

In example (17), the two instances of the neutral demonstrative, 葛[GE], are used to indicate two referents which appear in the speech situation without indicating any distance contrast, just as 个 in example (15) in the Ningbo dialect.

- (18) 葛只是茶杯，葛面只是酒杯。

GE	ZHI	SHI	CHABEI,	GE	MIAN	ZHI	SHI	JIUBEI.
DEM _{NEUT}	CL	COP	cup	DEM _{NEUT}	side	CL	COP	wine. glass

‘This is a cup, and that is a wine glass.’

As example (18) shows, to express the distance contrast, the neutral [GE] 葛 is used to identify a referent nearby, while the combination of [GE] 葛 and the noun [MIAN] 面 ‘side’ is used for a referent far away from the deictic centre, similarly to the Ningbo dialect.

In some languages, a neutral demonstrative is the only member of a demonstrative system for one specific category, such as in the Taihe dialect, where [GE] is the only demonstrative referring to ordinary objects. While in other languages, the neutral demonstrative may be complementary to a set of distance-marked demonstratives, such as in the Suzhou, Chongming, and Wuhan dialects.

4.2.1.6 *Nature of neutral demonstratives*

As we have shown in the preceding sections, the function of a neutral demonstrative is to point out a specific object. Neutral demonstratives are essentially not related to distance, but they can nonetheless implicate a distance meaning when they are used in a contrastive context. In such contexts, their distance meaning consequently varies.

In other words, neutral demonstratives are not concerned with distance when used alone. Because their main function is to point out a specific object (see definition in §4.2.1.1), that is, because they have the same basic pragmatic function as distance-indicating demonstratives, they can co-occur with those distance-indicating demonstratives, and thus obtain a temporary distance meaning derived from the particular context. However, this is not stable, as shown clearly in the examples above.

If a distance contrast is expressed by different linguistic forms in a language, two possible referents in the same situation can be identified simply by these linguistic devices which, consequently, do not rely on gestures any more. It is therefore possible for distance to be encoded by demonstratives as an intrinsic element in the process of the formation of demonstrative paradigms. The more semantic properties are encoded by demonstratives, the higher the identifiability of the referents which are denoted by such demonstratives.

The gradual incorporation of semantic properties into demonstrative paradigms makes the system more and more ‘hands-free’ from the pointing function, both literally and linguistically-speaking. This makes it possible for demonstratives to no longer depend on the concomitant use of gestures for their correct interpretation. Demonstratives without accompaniment by any gestures resemble pure language symbols more closely than those that are obligatorily used in combination with such gestures. This is because they indicate referents only by their form and meaning, and not by any kind of bodily postures.

Neutral demonstratives have a much higher dependency on bodily postures than those distance-marked ones, and from this point of view, the latter type is more grammaticalized than the former. It is reasonable to propose that neutral demonstratives may appear earlier than distance-marked ones. Nevertheless, this remains an open question to be explored.

When there is a neutral demonstrative morpheme in a language, it can indicate the complete domain of indexing by itself, and can be used to refer to any referent appearing in this domain, when used in combination with bodily postures. Neutral demonstratives constitute a one-term system by themselves, while distance-marked

demonstratives can also form a parallel range of indexing, as we have seen. Consequently, if there are neutral demonstratives and distance-marked demonstratives found in the same language or dialect, the neutral ones should not be viewed as part of any hierarchy of demonstratives on a distance scale.

Neutral demonstratives as an independent demonstrative system thus co-occur with distance-marked systems in the Suzhou dialect, where two sets of demonstrative systems are used: one is distance-neutral, and is a one-term system, while the other is distance-marked, and is a two-term system.

In a binary proximal-distal system, the proximal and the distal have the whole domain of indexing in their shared scope, while the neutral demonstrative is also able to index objects across the entire domain on its own. Hence, it can be used to identify several objects co-occurring in the same situation, as does [kəʔ⁵⁵] in the Chongming dialect (see §4.2.1.2).

4.2.2 Two-term systems

As mentioned already, Diessel (2005) points out, the two-way distinction is the most common of all types of contrasts: one is used for proximal referents, while the other is used for distal referents. In many Sinitic languages, demonstrative systems present precisely this proximal-distal contrast, just as in Standard Mandarin, where there are a proximal [tʂə⁵¹] 这 ‘this’ and a distal [na⁵¹] 那 ‘that’.

Here one important observation should be made: in some languages, such as in the Suzhou dialect, there are three demonstrative morphemes: one neutral demonstrative and two distance-marked demonstratives. Generally speaking, this system can be called a two-way distinction rather than a typical three-way distinction, because the indexing function of the neutral demonstrative overlaps to some extent with that of proximal and distal demonstratives, i.e. the neutral demonstrative can always replace the proximal or distal ones when used for referents in particularly well-defined speech situations. The neutral demonstrative therefore does not occupy the same demonstrative territory as the proximal and distal ones, and it is obviously independent of those proximal and distal ones.

It can be concluded that if a demonstrative system consists of one neutral demonstrative and two distance-marked ones, it involves a two-term system in the form of the latter. As we have argued, the neutral demonstrative carries out a separate function which can however interact with the distance-marked demonstratives. Nonetheless, from a higher level perspective, this type of situation can also be called a three-term system which can be represented as a multi-layered one, as in Table 4.1.

4.2.3 Three-term systems

According to Diessel (2005), three-term systems only rank second to two-term systems in distribution. He does not mention Sinitic languages when he talks about

TABLE 4.1. Representation of systems with neutral demonstratives

	Primary layer	Secondary layer
Three-term system	Distance unmarked (Neutral demonstratives)	
	Distance marked	Proximal Distal

TABLE 4.2. Three-term demonstrative systems in Sinitic languages

Type	Semantic differentiation			Examples
Person-oriented	proximal	medial	distal	Guangshan 光山 (Central Plains Mandarin)
Distance-oriented	proximal close proximal	medial proximal distal	distal distal yonder	Lengshuijiang 冷水江 (Xiang) Xinyu 新余 (Gan Dialect) Lichuan 黎川 (Gan dialect)

three-term systems, but there are indeed many three-term configurations in Sinitic languages, which present complex and diverse systems as displayed in Table 4.2.

In Sinitic languages, three-term systems are relatively common. The demonstrative systems which have been identified and described in this language group are overwhelmingly distance-oriented, while person-oriented systems are quite rare, one example being the Guangshan 光山 dialect (Central Plains Mandarin). As Zhang Xianmin (2011) points out, in the Guangshan dialect, the proximal [te³¹] and the distal [la³¹] locate referents by reference to the speaker, while the medial [ní³¹] has both the speaker and the addressee as its reference centre, which accordingly presents the typical properties of person-oriented demonstratives.

If only distance was concerned, all three-term distance-oriented demonstrative systems could be viewed as systems consisting of proximal, medial, and distal demonstratives. If this conclusion were acceptable, we would be obliged to call all the following three different categories of demonstratives ‘medial demonstratives’ from a relativistic point of view: medial demonstratives in a proximal-medial-distal system, distal demonstratives serving as medials in a proximal-distal-yonder system, and proximal demonstratives serving as medials in a close-proximal-distal system. But in fact, this is not the case. These three types of three-term systems are quite different in their semantic essence.

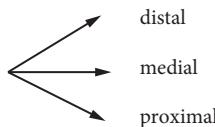


FIGURE 4.1 Basic single-layered demonstrative system with three unmarked terms

4.2.3.1 Proximal-medial-distal system

In this type of system, all demonstratives are distance-marked, and they have a relatively fixed meaning. When these three demonstratives co-occur in the same situation, they are successively used to identify the referents ranged from near to far in location. Each of them can be used independently, i.e. they are not necessarily used in contrast to other demonstratives in a given situation. For example, when a medial demonstrative is used independently, it indexes an object which is neither close to nor far away from the reference centre, and it can be used in this way even when it is not in contrast to a proximal or a distal demonstrative. In this type of system, each member is an unmarked form, which means that they have the same opportunity to be selected (Figure 4.1).

Representative dialects of this type of system are next discussed:

4.2.3.1.1 Case one: Weifang dialect 潍坊 (*Central Plains Mandarin* 中原官话)

In the Weifang dialect, there is a three-way distinction of demonstratives: proximal 这[*tʃə²¹*], medial 訂[*n̩iə²¹*]/乜[MIE]⁵, and distal 那[*nA²¹*]. When used to distinguish between two contrastive referents, any two of these three terms may be selected according to the actual distance from a reference centre to the referents in question.

As Zhao (1996) points out, if neither of two contrastive referents are far away from the reference centre, proximal and medial are used, while if neither of two contrastive referents are close to this reference centre, medial and distal demonstratives are selected, and finally if one is obviously closer to the reference centre and the other far away from it, proximal and distal are selected.

The medial demonstrative [*n̩iə²¹*] has a very clear use. In general, it is used to identify a referent which is neither close to nor far away from the reference centre. Hence, this type of three-term system with proximal, medial, and distal demonstratives forms a continuum with a complete range of indexing possible, but as such, there are no clear-cut boundaries between members of the set. Therefore, medial demonstratives are able to always show a certain degree of flexibility in their actual uses, i.e. in some situations, a medial demonstrative can be interchangeable with a proximal or a distal one. As Qian and Luo (1992) point out, when there is a pair of proximal-distal contrasts, both the medial [MIE] 乜 and the distal [*nA²¹*] 那 can be used to refer to distal objects.

⁵ The medial in the Weifang dialect has two free variants: 訂 [*n̩iə²¹*] and 乜 [MIE].

- (19) 我不要这个，我要也/那个。

WO BU YAO *tʃə²¹* GE, WO YAO MIE/nA²¹ GE.
1SG NEG want DEM_{PROX} CL 1SG want DEM_{MED/DIST} CL
'I don't want this one, I want that one.'

When there are three contrastive referents, however, [MIE] 也 and [nA²¹] 那 are respectively used for medial objects and distal objects, and cannot be interchanged.

- (20) 你是要这个啊，是也个啊，还是老娘家的那个啊？

NI SHI YAO *tʃə²¹* GE A, SHI MIE GE A,
2SG COP want DEM_{PROX} CL PRT, COP DEM_{MED} CL PRT,
HAISHI LAONIANG JIA DE nA²¹ GE A?
or grandmother home MOD DEM_{DIST} CL PRT?

'Do you want this one, that one, or the one in grandmother's home over there?'

4.2.3.1.2 Case two: Lengshuijiang dialect 冷水江 (Xiang group 湘语)

In the Lengshuijiang dialect (Chen 1995), there is a three-term system consisting of a proximal [YI]/[GE]以/咯,⁶ a medial [Nü₁], 女₁ and a distal [Nü₂].⁷ The proximal forms [YI]/[GE] are used for referents close to the speaker, the medial form [Nü₁] refers to objects which are not near the speaker but are closer to the hearer, and the distal form [Nü₂] is used for all other situations. The indexing domains of these three demonstratives are thus very clear.

- (i) [YI]/[GE], [Nü₁] and [Nü₂] can co-occur, referring to three contrastive objects from near to far respectively:

- (21) 以丘田五分，女₁丘田三分，女₂丘大滴，六分八。

YI QIU TIAN WU FEN,
DEM_{PROX} CL land five CL_{UNIT.FOR.LAND}
Nü₁ QIU TIAN SAN FEN,
DEM_{MED} CL land three CL_{UNIT.FOR.LAND}
Nü₂ QIU DA DI LIU FEN BA.
DEM_{DIST} CL big MOD, six CL_{UNIT.FOR.LAND} eight.

'This land measures five *fen*, that is three *fen*, while the big one over there is 6.8 *fen*.'

- (ii) If there are two objects which are in contrast to each other, there are three possible ways of expression that can be selected, depending on the actual distance from reference centre to referent.

⁶ [I] and [GE] are two free variants of the same morpheme.

⁷ [Nü₁] and [Nü₂] differ only in tone pitch, for which the author does not provide any details.

The semantic differentiation of demonstratives in Sinitic languages 97

- a. If a referent is close to the reference centre and the other is a little further away from it, the combination proximal-medial is selected.

(22) 以只鸡是我屋里咯，女₁只是其屋里咯。

YI	ZHI	JI	SHI	WO	WU	LI	GE
DEM _{PROX}	CL	chicken	COP	1SG	home	inside	MOD
Nü ₁	ZHI	SHI	QI	WU	LI	GE.	
DEM _{MED}	CL	COP	3SG	home	inside	MOD	

‘This chicken is mine, and that one is his.’

- b. If two objects are far away from the reference centre and they are in a contrastive situation, the combination medial-distal is used.

(23) 女₁个是我个大崽，女₂个是老二。

Nü ₁	GE	SHI	WO	GE	DAZAI
DEM _{MED}	CL	COP	1SG	MOD	eldest.son
Nü ₂	GE	SHI	LAOER.		
DEM _{DIST}	CL	COP	second.eldest		

‘That is my eldest son, and that one over there is my second eldest.’

- c. If one referent is close to the deictic centre, while the other is very far away from it, the combination proximal-distal is used.

(24) 女₁拿到以個两件东西去，其拿到女₂個两件去。

NU	NADAO	YI	GE	LIANG	JIAN	DONGXI	QU
2SG	take.away	DEM _{PROX}	CL	two	CL	thing	go
QI	NADAO	Nü ₂	GE	LIANG	JIAN	QU.	
3SG	take.away	DEM _{DIST}	CL	two	CL	go	

‘You can take away these two things, and he can take away those two over there.’

- (iii) All demonstratives can be used in non-contrastive situations, according to the distance from deictic centre to referent. The medial form [Nü₁] allows a little flexibility in use, but the range of these three demonstratives in general remains quite distinct.

In proximal-medial-distal systems, when two contrastive objects need to be pointed out, there is no rigid criterion allowing the speaker to choose between proximal-medial and medial-distal demonstrative forms. Hence, the deictic domain of a medial demonstrative may appear to partially overlap with a proximal or a distal one. Nonetheless, these three terms do not overlap with each other when used together in the same utterance or sentence.

4.2.3.2 *Proximal-distal-yonder and close-proximal-distal: Marked versus unmarked demonstratives and multi-layered systems in Sinitic languages*

According to Chen Minyan et al. (2003), in many sub-branches of the Gan dialect group, close-proximal-distal systems are to be found, such as those spoken in Xiajiang county 峡江县, and in Jianghuang 姜黃 town in Zhangshu county 樟树县. This type of system presents a common property: the occurrence of a close form invariantly implies the occurrence of a proximal form, i.e. if a close form is selected to refer to an intended referent, it must be in contrast with a proximal form, but the occurrence of a proximal demonstrative does not necessarily imply the existence of a close one.

Furthermore, Chen Minyan et al. (2003) go on to describe that in still other sub-branches of the Gan dialect group, three-term proximal-distal-yonder systems can also be found, such as those spoken in Zixi County 资溪县, Pengze County 彭泽县, and Linquan town 林泉 in De'an County 德安县. In these systems, the occurrence of a yonder form invariantly implies a contrast with a distal referent.

So in these two types of three-term systems, the close and the yonder are the marked forms respectively, compared to the proximal and the distal in function. The close and the yonder demonstratives can be viewed as secondary layers or branching nodes that serve as marked variants under either the first distal or the first proximal layer. Based on the fact that the close and the yonder forms never appear independently when there is no distance contrast, i.e. the intended referent is not in contrast with other referents on the same segment of a distance scale, this means the marked forms—that is, the close and the yonder demonstratives—are never selected. In other words, only under contrastive conditions can the marked close and yonder forms be used, and their appearance must be accompanied by the corresponding unmarked forms, as shown in Table 4.3.

In some Sinitic languages, the close and yonder demonstratives are morphologically or phonologically derived from the corresponding unmarked forms, which reflects the relation of dependence between the forms and meanings of demonstratives and their markedness, as Liu and Chen (2008, 2009) propose. Some special devices, such as the lengthening of the main vowel or rhyme of a syllable, placement of stress on a syllable, the reduplication of a syllable, or tone sandhi to rising tone are used to express finer differentiations along the distance scale for demonstratives in certain Sinitic languages. Therefore, marked demonstratives may be derived from the corresponding unmarked demonstratives with one or more of these special morpho-phonological devices.

In the Anyi 安义 dialect (Gan), rising tone is used to distinguish the yonder from the distal demonstrative:

- (i) distal demonstrative: [he (*Yinshang* tone)]
- yonder demonstrative: [he ↗ Rising tone]

TABLE 4.3. The uses of marked demonstrative forms in three-term systems

Three-term systems	Permissible uses of marked and unmarked demonstratives
Gan dialects	
Type I:	Close, proximal (co-occurring)
(i) Xiajiang county,	*Close
(ii) Jianghuang town in Zhangshu county	Proximal
	Distal
Type II:	Proximal
(i) Zixi County,	Distal
(ii) Pengze County,	Distal, yonder (co-occurring)
(iii) Linquan town in De'an County	*Yonder

Legend: '*' means unable to be used independently

In the Andong district of Pingxiang city 萍乡市安东区 (Gan), reduplication is used:

- (ii) distal form: [len (*Yinping tone*)]
yonder form: [len (*Yinping tone*) DE len (*Yinping tone*)]

In Hutang town of Fengcheng county 丰城湖塘 (Gan), the device of lengthening is used:

- (iii) proximal form: [ko (*Yinping tone*)]
close form: [k^ø (*Yinping tone*)]

Based on the materials available, it is clear that marked close demonstratives are only to be found in the Gan dialect group for Sinitic languages, while the marked

yonder demonstratives are more widespread, not being exclusive to Gan dialects. The latter is a type which I will introduce below.

4.2.4 Three-term demonstrative systems involving semantic properties other than distance

4.2.4.1 Single-layered and multi-layered proximal-distal-yonder systems

Up to this point we have only considered the division of demonstratives along a single distance dimension. In fact, according to numerous studies, there are many other dimensions which can be encoded by a demonstrative system from a typological perspective (see §4.1.1 for more details on these features). In Sinitic languages, except for distance, visibility appears, however, to be the only other parameter which can be encoded.

In dialects spoken in Yugan county town 余干县城 and Nanfeng town 南丰乡 in Duchuan county 都昌县, the distal is used for visible referents and the yonder, for referents out of sight.

Yugan dialect:

(25) (导游:)该仂是东山岭, 该个是干越亭, 口仂是琵琶洲。

(DAOYOU:) *oi¹-LE* SHI DONGSHANLING,

(tour guide:) DEM_{PROX}-location COP Dongshanling

oi¹-GE SHI GANYUETING,

DEM_{PROX}-CL COP Ganyueting

ŋ⁵-LE SHI PIBAZHOU.

DEM_{DIST}-location COP Pibazhou

(游客:)康山忠王庙在哪儿?

(YOUKE:) KANGSHAN ZHONGWANG MIAO ZAI NAR?

(A tourist:) Kangshan Zhongwang temple at where

(导游:)口仂! 要撑船去!

(DAOYOU:) *hoi⁵-LE!* YAO CHENG CHUAN QU!

(The tour guide:) DEM_{YOND}-location! need pole boat go

[Context: a tour guide is pointing out directions]

‘(The tour guide:) ‘This is Dongshanling, this is Ganyueting, and that is Pibazhou.’

(A tourist:) ‘Where is the Kangshan Zhongwang temple?’

(The tour guide:) ‘Over there! You need to take a boat to get there.’

As shown in example (25), the yonder demonstrative has a fixed meaning, so it can appear independently without needing to be contrasted with other demonstratives. Nor can it be considered a morphologically marked, derived demonstrative. Therefore, the demonstrative system to which it belongs is not a multi-layered one with two nodes of branching.

The semantic differentiation of demonstratives in Sinitic languages 101

Crosslinguistically, the closeness value to a deictic centre can also be encoded by a semantically specific demonstrative, distinguished by the possibility of its independent use. In this case, the system to which it belongs is also a single-layered one. But this type of system has not been found in Sinitic languages up to now, as the presence of a close demonstrative, where attested in Sinitic languages, implies a proximal one, as we have argued.

Consequently, a three-term system with a *yonder* demonstrative may be multi-layered or single-layered according to the nature of the *yonder* demonstrative, whereas for the close demonstrative, only the multi-layered system is possible in Sinitic languages, as will be seen in §4.2.4.

Three criteria can be used to judge whether a *yonder* demonstrative is a marked form or an unmarked form. These are:

Criterion 1: Can it be used in non-contrastive situations? (that is, it is not in contrast with other demonstratives).

Criterion 2: Is it in contrast with the proximal form in the same demonstrative system?

Criterion 3: Is it in contrast with the distal form in the same system.

If the answers to these three questions are all ‘yes’, the *yonder* form is then an unmarked form in the demonstrative system, and the system is thus a single-layered system (Figure 4.2).

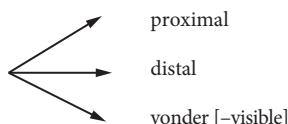


FIGURE 4.2 Single-layered unmarked three-term system with a ‘yonder’ demonstrative

Otherwise, the *yonder* is a marked term and the system is multi-layered, as represented in Figure 4.3.

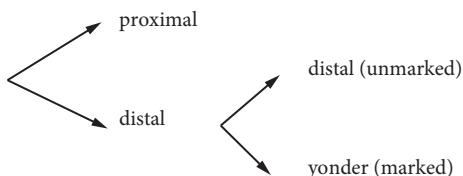


FIGURE 4.3 Multi-layered three-term system with a marked ‘yonder’ demonstrative

For Sinitic languages, although the nature of single-layered proximal-distal-yonder systems has much in common with that of the single-layered proximal-medial-distal

distinction, they nonetheless remain distinct because the former always has a non-visible [-visible] property encoded by its demonstratives. This is the case argued by Tang (2000) who points out that in the Susong 宿松 dialect (Gan) there is a yonder form [uei²¹] 畏 which she claims can be used to refer to invisible objects when the term does not contrast with the distal [n²¹]. I will take issue with this viewpoint and show that it is an inference in many cases. Let us consider the following sentence:

- (26) 三个人只回来哆一个，畏两个人不晓得哪里去在。

SAN GE REN ZHI HUILAI DUO YI GE,
three CL person only come.back PFV one CL
uei²¹ LIANG GE REN BU XIAODE NALI QU ZAI.
DEM_{YOND} two CL person NEG know where go PRT
'Only one of three people came back; no one knows where the other two are.'

- (27) 今朝中时就煮这些鱼，畏些留着过年。

JINZHAO ZHONGSHI JIU ZHU ti²¹XIE YU
today noon ADV boil DEM_{PROX-PL} fish
uei²¹XIE LIUZHE GUO NIAN
DEM_{YOND-PL} leave celebrate spring festival
'This noon only cook these fish, and leave those/(the others) for the Spring Festival.'

In these two sentences, [uei²¹] in fact does not serve a deictic function in my view. In sentence (26), three persons are viewed as a group, but only one of them has come back: [uei²¹ LIANG GE REN] 畏两个人 here refers to the two others and assumes a definite meaning from the context. This is not the deictic use of demonstratives because it does not index objects in a speech situation. As for sentence (27), the explanation is similar to that for sentence (26): in these two sentences, as even Tang points out, [uei²¹] is the equivalent of 'other'.

- (28) 这是桃树，那是橘子树，畏是枇杷树。

ti²¹ SHI TAOSHU, n²¹ SHI JUZISHU, uei²¹ SHI
DEM_{PROX} COP peach.tree, DEM_{DIST} COP orange.tree, DEM_{YOND} COP
PIPASHU.
loquat.tree.

'This is a peach tree, that is an orange tree, and the other yonder (not visible) is a loquat tree.'

The so-called 'invisibility' feature is in fact an inference available as the outcome of semantic change for the yonder demonstrative [uei²¹] from a concrete deictic meaning to an abstract non-deictic meaning, that of 'definiteness', as in (28). Generally speaking, the term referring to the greatest distance always becomes the

The semantic differentiation of demonstratives in Sinitic languages 103

most grammaticalized, specifically undergoing semantic generalization. Tang further claims that *yonder* [uei²¹] can be in contrast with the distal [n²¹] or the proximal [ti²¹], as shown in example (29) and (30).

- (29) 恰不要在那里嬉，快滴在大伯畏里去帮忙。

NEN	BUYAO	ZAI	n ²¹	-LI	XI
2PL	NEG _{IMP}	at	DEM _{DIST}	-location	play
KUAI	DI	ZAI	DABO	uei ²¹ LI	QU BANGMANG.
quick	a.little	at	uncle	DEM _{YOND} -location	go help

'Don't play here, and please go and help your uncle at his place.'

- (30) 冬下天，北京畏里比者这里冷好些。

DONGXIATIAN,	BEIJING	uet ²¹ LI	BI	WOZHE	ti ²¹ LI
winter,	Beijing	DEM _{YOND} -location	compare	1PL DEM _{PROX} -location	
LENG	HAOXIE.				
cold	much				

'In winter, Beijing is much colder than here.'

Hence, there does not seem to be an invariant visibility distinction for the distal-yonder end of the continuum, since distal [n²¹] and yonder [uei²¹] can be used for either visible or invisible objects.

In conclusion, I do not believe that these data show that there is any [-visible] feature encoded by the demonstrative system in the Susong dialect. Rather, it is used to mark definiteness, when there is a two-way contrast, or a yonder distinction, when there is a three-way contrast. Consequently, it is reasonable to consider that there is a proximal-medial-distal distinction of the demonstrative system in the Susong dialect.

4.2.4.2 Single-layered and multi-layered close-proximal-distal system

As for close demonstratives, the criteria to judge if they can be viewed as secondary layers with two branching nodes are as follows:

Criterion 1: Can the close demonstrative be used in non-contrastive situations, i.e. when it is not in contrast with other demonstratives?

Criterion 2: Is it in contrast with the proximal form in the same demonstrative system?

Criterion 3: Is it in contrast with the distal form in the same system?

If the answers are all 'yes', the close, proximal, and distal demonstratives are located in the same layer with just one branching node (Figure 4.4).

Otherwise, if the close demonstrative is located in a secondary layer under the first proximal layer, it will be marked in comparison with the proximal form in use. In this case, we have two branching nodes.

In Sinitic languages, we do not find any instances in which a close demonstrative is located on the same node or layer with the proximal and distal ones. This type of

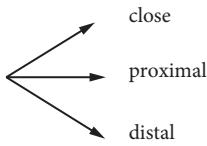


FIGURE 4.4 Single-layered unmarked three-term system with a 'close' demonstrative

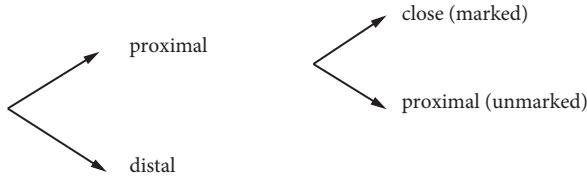


FIGURE 4.5 Multi-layered three-term system with a marked 'close' demonstrative

three-term system always has a secondary bifurcation for the proximal type, as shown in Figure 4.5.

Through this investigation into the demonstrative systems of 303 Sinitic languages, we obtain the following conclusions:

- (i) If the division of a demonstrative system in a language is only based on distance, those systems with more than a three-way distinction can be multi-layered; that is, the yonder and close forms are marked and always form a secondary layer or bifurcation (see §4.2.5 below).
- (ii) In demonstrative systems, morphologically-derived forms, using processes such as stress or reduplication, must co-occur with their corresponding unmarked forms.
- (iii) If semantic properties such as [+/-visibility] are encoded by a demonstrative system, the demonstratives coding them can occur independently, and are not marked, derived forms.

For example, if visibility is encoded as part of a three-term demonstrative system, this system is always single-layered, and it can be viewed as a proximal-distal-yonder system. The Yungan 余干 dialect (Gan) is a representative of this type of system, using three terms, [oi¹], [iŋ⁵], and [hoi⁵] respectively to locate referents on a scale from near to far (see example 25). The proximal one is used to indicate a referent close to the speaker, the distal one is used for visible referents relatively far from the speaker, while the yonder demonstrative is used for invisible referents.

4.2.5 Systems with more than three terms

In Sinitic languages, four-term paradigms are less common than three-term paradigms, but they are certainly attested and present more diversity in their features than

The semantic differentiation of demonstratives in Sinitic languages 105

TABLE 4.4. Four-term demonstrative systems in Sinitic languages

	Primary layer	Secondary layer	
Systems with a four-term distinction on a distance scale	Proximal	marked	close
	Distal	unmarked unmarked marked	proximal distal yonder

TABLE 4.5. Four-term demonstrative systems in Dongxiang county 东乡县 (Gan dialect)

Meaning	Form
Close	ko (<i>Yinping</i> tone)
Proximal	ko (<i>Yangping</i> tone)
Distal	e (<i>Yinping</i> tone)
Yonder	e (<i>Yangping</i> tone)

TABLE 4.6. Four-term demonstrative systems in Zhangshu city 樟树市 (Gan dialect)

Meaning	Form
Close	kø (<i>Yinru</i> tone)
Proximal	ko (<i>Yinru</i> tone)
Distal	he (<i>Yinshang</i> tone)
Yonder	hɛ (<i>Yinshang</i> tone)

two- or three-term systems. Systems with four-terms are generally multi-layered systems with a major binary distinction on the distance scale, as shown in Table 4.4.

Certain Gan dialects are representatives of this type of distinction for demonstratives, such as those spoken in Dongxiang county 东乡县, which differentiates by means of tone change (Table 4.5), and Zhangshu city 樟树市 which differentiates by means of vowel lengthening (Table 4.6).

As I have discussed above, some special devices, such as lengthening, stress, and reduplication, are often used to obtain a finer semantic differentiation for demonstratives. These prosodic and morphological processes are consistently applied to a distal or a proximal demonstrative to indicate further distance or a closer degree of proximity. Some linguists do not treat these lengthened, stressed, or reduplicated forms as demonstratives, but only view them as variants of distal or proximal demonstratives, for example, Wang Huayun (2008), Liu and Chen (2008, 2009).

Their viewpoints can be accepted as reasonable to a certain degree. Strictly speaking, lengthened, stressed, or reduplicated forms are not grammatical features since there is no fixed correspondence between form and meaning for them. The degree of lengthening and the meaning of a lengthened demonstrative form are decided by context, such as [WU] 兀 in the Linxia dialect 临夏 (Wang Sen, 1996): the greater the lengthening of [WU], the further away the referent will be.

I would like to put forward a hypothesis that lengthening, stressed, or reduplicated strategies can be viewed as morphological devices to create a new semantic distinction for demonstratives. They can occur independently, which shows that the demonstratives with these devices occupy a point on a scale. Theoretically, the degree of lengthening decides the distance of the referent from the deictic reference point. Demonstratives could therefore be hypothetically divided into numerous finely differentiated types, but in fact the possibilities for lengthening are restricted physiologically by the articulatory mechanism of human beings, so the differentiation cannot be limitless. Moreover, the reduplicated forms have fixed form and code fixed meaning, as I have shown. So these derived forms are not simple allomorphs of demonstratives that occupy a position on a continuum: they can indeed be viewed as demonstratives in their own right.

Five-term demonstrative systems also exist in Sinitic languages. One such paradigm for location is possible in the dialect spoken in Tailing village in Xixiang town of Yongxin county 永新县西乡台岭 (a sub-branch of Gan), just for the distance scale, which is achieved by the change of tone and reduplication as shown in Table 4.7.

The close and the yonder forms are respectively different from the proximal and the distal in tone pitch, while the further-than-yonder form shows a partial reduplication of the yonder.

In Sinitic languages, just as in other languages in the world, distance—as compared with other properties—is encoded with the most facility by demonstratives. In some dialects, visibility is an additional differentiating feature integrated into the demonstrative system. No other qualitative properties are found in the demonstrative systems of Sinitic languages.

TABLE 4.7. Demonstrative system in the Tailing town dialect

Meaning	Form
Close	ke (<i>Yinshang</i> tone) tō
Proximal	ke (<i>Yinru</i> tone) tō
Distal	kā (<i>Yinping</i> tone) u (<i>Yinru</i> tone) tō
Yonder	kā (<i>Yinshang</i> tone) u (<i>Yinru</i> tone) tō
Further than yonder	kā (<i>Yinshang</i> tone) u (<i>Yinru</i> tone) kā (<i>Yinshang</i> tone) u (<i>Yinru</i> tone) tō

4.3 Two implicational universals on the basis of a sample of Southeast Asian languages

As far as ontological properties are concerned, it is found that only distance and visibility are encoded by demonstrative systems in Sinitic languages, while in other Southeast Asian languages, the situation is more complex. Apart from distance and visibility, there is a large variety of semantic properties which can be encoded, such as elevation and animacy among others.

Based on 303 Sinitic languages (for details and the complete data, see Chen 2010: 178–205), in addition to a further 20 languages from the Tibeto-Burman branch of Sino-Tibetan and the Hmong-Mien, Austroasiatic, and Tai-Kadai language families (listed in the appendix at the end of this chapter), we propose the two following implicational universals for demonstratives:

Implicational Universal One:

The coding of any relevant semantic property by a demonstrative system

- ▷ the property of distance will be coded by this demonstrative system

This implicational universal is explained by the fact that distance is more easily encoded by demonstrative systems than other semantic properties such as visibility, as attested for Sinitic languages, in which the parameter of distance has dominant status for the semantic differentiation of demonstratives.

Implicational Universal Two:

For any one set of semantic properties that applies to proximal demonstratives

- ▷ the same set of properties also applies to distal demonstratives.

This universal can be interpreted in the following way: in a demonstrative system, a demonstrative referring to a more distant referent is always more easily marked by semantic properties other than distance, such as visibility, or altitude, than a demonstrative referring to a relatively more proximal referent.

This can be explained by the fact that it is necessary to use additional semantic properties to identify a more distant object, because it is far away from the reference centre and can thus be potentially confused with many other objects. On the contrary, as one gets closer to a given reference centre, the number of objects diminishes, compared with all those that become possible sources of confusion as one progressively moves out to more distant areas. Objects closer to the reference centre can thus be more easily identified by demonstratives with or without supplementary gestures. The demonstrative system in Langsu (Ngwi-Burmese, Tibeto-Burman) can represent this implicational universal for Southeast Asian languages (Dai 2005) (Table 4.8).

TABLE 4.8. Demonstrative system in Langsu

Distance	Proximal	Distal	Yonder
Altitude			
Horizontal		thø ³¹	thø ⁵⁵
High	tʃe ³¹	xu ³¹	xu ⁵⁵
Low		mɔ ³¹	mɔ ⁵⁵

4.4 Conclusion

There are four main conclusions concerning demonstrative systems in Sinitic languages:

Based on a representative sample of 303 Chinese dialects from different branches of Sinitic, we obtain the following conclusions:

- (i) If there is a one-term demonstrative system in a language for a particular category, the demonstrative in question is neutral, such as 格 [GE] in the Taihe dialect (Gan).
- (ii) The two-term system is the most common system in Sinitic languages and this conforms to the situation crosslinguistically (Diessel 2005). In Sinitic, it is distance-oriented, with proximal and distal forming a pair of oppositions.
- (iii) The three-term system presents a greater diversity in the combination of features, as displayed in Table 4.2.

Generally speaking, in Sinitic languages, the close and yonder demonstratives in this type of system always represent the marked or derived forms compared with the proximal and distal forms in use. Hence, if there is a marked close or a marked yonder form in a dialect, they can be regarded as a sub-layer or secondary branching node under the basic binary opposition of proximal and distal, and the system is multi-layered. Sometimes, however, the yonder demonstratives, unlike the close demonstratives, may be independent in both form and function, and in some dialects, such as the Yugan dialect (余干县城) described in §4.2.4.1, the yonder demonstratives encode the feature of visibility, specifically [-visible], in contrast with the other two terms. In this case, we have a single-layered three-term system.

The semantic differentiation of demonstratives in Sinitic languages 109

- (iv) There are also four and five-term demonstrative systems attested in Sinitic languages, based only on the distance scale, and these are always multi-layered systems with secondary branching nodes.

In these systems, the semantic factors by which the differentiation of demonstratives is established in Sinitic languages (or Chinese dialects) always involves distance and visibility, and does not concern animacy, geographical features, or any other parameters which may be widespread, for example, in certain other of the Sino-Tibetan, Hmong-Mien, Austroasiatic, and Tai-Kadai languages. Some prosodic properties, such as lengthening and stress, if not morphological properties such as reduplicated forms, are common devices to construct demonstrative systems in the language families of Southeast Asia, including the main focus of our analysis, the Sinitic languages.

**Appendix: Additional languages used as reference data from
Tibeto-Burman Hmong-Mien, Austroasiatic, and Tai-Kadai**

- Jianghua Mian (Jianghua Mian [mjəŋ² wa⁶] 江华勉语) (Yao, Hmong-Mien)
Sangkong (Sangkong [san⁵⁵ qhon⁵⁵] 桑孔语) (Yi, Tibeto-Burman)
Zauzuo (Rouruo [zao⁵⁵ zuo³³] 柔若语) (Yi, Tibeto-Burman)
Anung (Anong [a³¹ nun³¹] 阿依语) (Tibeto-Burman)
Wuse (Wusehua [nju⁵⁵ çek⁵⁴ wa²⁴] 无色话) (Unclassified mixed language)
Zhaba (Draba) (Zhaba [ndža⁵⁵ pa⁵⁵] 扎巴语) (Qiangic, Tibeto-Burman)
Lavrung (Khroskyabs) of Yelong (Yelong Lawurong 业隆拉坞戎语) (Qiangic, Tibeto-Burman)
Pumi (Pumi [phzɔ⁵⁵ mi⁵⁵] 普米语) (Qiangic, Tibeto-Burman)
Geman (Geman [kω³¹ man³⁵] 格曼语) (Tibeto-Burman)
Monba (Menba [mø³⁵ kɔ⁵³] 门巴语) (Tibetan, Tibeto-Burman)
Wa (Va) (Wayu 佤语) (Mon-Khmer, Austroasiatic)
Li (Hlai) (Lai [la:i³³] 傣语) (Tai-Kadai)
Jiongnai Bunu (Jiongnai 焰奈语) (Hmong-Mien)
Younuo Bunu (Younuo [zou¹³ no¹³] 尤诺语) (Hmong-Mien)
Bumang (Bumang 布芒语) (Mon-Khmer, Austroasiatic)
Khem (Kemie [khy³¹ met⁵³] 克蔑语) (Mon-Khmer, Austroasiatic)
Buyang (Buyang 布央语) (Kra, Tai-Kadai)
Lhao Vo (Langsu 浪速语) (Yi-Burmese, Tibeto-Burman)
Pala (Bola [po³¹ no⁵¹] 波拉语) (Yi-Burmese, Tibeto-Burman)

5

Bare classifier phrases in Sinitic languages: A typological perspective

WANG JIAN

5.1 Introduction

This chapter re-examines bare classifier phrases in a typological perspective and puts forward three implicational universals with respect to their syntactic distribution and semantic interpretation.

A bare classifier phrase refers to one composed of just a classifier and its head noun without any numerals or demonstratives preceding the classifier. Bare classifier phrases ([CL-N] phrases) can be found in most Sinitic languages. Notably, their syntactic distribution and semantic interpretation of bare classifier phrases may, however, vary according to the language. The three implicational universals for which I will argue are the following:

- (1) preverbal [CL-N] phrases \supset postverbal [CL-N] phrases;
- (2) preverbal indefinite [CL-N] phrases \supset preverbal definite [CL-N] phrases, and
- (3) postverbal definite [CL-NP] phrases \supset postverbal indefinite [CL-N] phrases.

The analysis subsequently proposes that Sinitic languages can be categorized into one of seven types precisely according to the different patterns of syntactic distribution and semantic interpretation of bare classifier phrases that can be discerned, based on a sample of 120 Sinitic languages. Finally, some tentative explanations are offered for the three implicational universals.

5.1.1 *Classifier phrases*

Sinitic languages are classifier languages. The use of classifiers with numerals or demonstratives is common and highly developed throughout the Sinitic linguistic area. In Standard Mandarin the basic syntactic configurations in which classifiers occur are the following:

Bare classifier phrases in Sinitic languages: A typological perspective 111

- (a) Demonstrative-Classifier-NP;
- (b) Numeral-Classifier-NP;
- (c) Demonstrative-Numeral-Classifier-NP.

The following examples illustrate this well.

- (1) a Demonstrative-Classifier-NP

这本书
tsə⁵¹ pən⁰ su⁵⁵
this CL book
'this book'

- b Numeral-Classifier-NP

一本书
i⁵⁵ pən⁰ su⁵⁵
one CL book
'one book'

- c Demonstrative-Numeral-Classifier-NP

这两本书
tsə⁵¹ lian²¹⁴⁻³⁵ pən⁰ su⁵⁵.
this two CL book
'these two books'

In addition, Standard Mandarin has bare classifier phrases. By 'bare classifier phrase' (hereafter [CL-N] phrase), we intend to refer to phrases which are formed by a classifier and a noun without any numerals or demonstratives preceding the classifier. In Standard Mandarin a [CL-N] phrase can only appear in postverbal position. For example:

- (2) 我买了本书。

uo²¹⁴⁻³⁵ mai²¹⁴⁻²¹ -lə⁰ pən⁰ su⁵⁵.
1SG buy-ASP CL book
'I have bought a book.'

- (3) *本书好看。

*pən²¹⁴ su⁵⁵ xau²¹⁴⁻²¹ k^han⁵¹.
CL book interesting
(Attempted meaning: 'A/The book is interesting.')

The [CL-N] phrase 'pən⁰ su⁵⁵本书' in example (2) is postverbal and codes indefiniteness. Example (3) is unacceptable because the [CL-N] phrase 'pən²¹⁴su⁵⁵本书' cannot occur in preverbal position in Standard Mandarin.

It is generally agreed that the omission of the numeral ‘one’ accounts for the occurrence of [CL-N] phrases in this postverbal position in Standard Mandarin. Furthermore, [CL-N] phrases can only have the indefinite reading in Standard Mandarin (Zhu 1982). The difference between definiteness and indefiniteness can be illustrated by the following examples (P. Chen 2004: 1131).

- (4) a George finally bought a house.
b George finally bought the house.

By using ‘a house’ in (4a), the speaker assumes the addressee is not in a position to identify which particular house George bought; in uttering (4b) he assumes that the addressee knows which house he is talking about. The entity of a house is presented as definite in (4b), and indefinite in (4a).

It has also been reported that in some other Sinitic languages, especially those in the south of China, [CL-N] phrases can appear not only in postverbal position but also in preverbal position and have either a definite or indefinite reading. Most of the previous research on [CL-N] phrases focuses on the usage of these phrases in one single dialect, and thus lacks any crosslinguistic comparison. This analysis re-examines the [CL-N] phrases of 120 Sinitic languages in a typological perspective.

This chapter consists of three main sections intervening between the introduction and the concluding words: §5.2 discusses the proposed three implicational universals with respect to the syntactic distribution and semantic interpretation of [CL-N] phrases, while §5.3 examines seven types of Sinitic languages, and §5.4 provides the explanation for the three implicational universals.

5.2 Three implicational universals with respect to the syntactic distribution and semantic interpretations of the [CL-N] phrases

The traditional linguistic definition of typology refers to a classification of structural types across languages. The contemporary definition of typology is the study of the patterns that occur systematically across languages. Croft (2003: 1) refers to the second definition of typology as concerning typological generalization. The patterns found in typological generalization can be viewed as different kinds of language universals. The classic example of a typological universal is the implicational universal which takes the form of ‘if *p*, then *q*’. Thus the presence of one property (i.e. the *implicans*) implies that of another (i.e. the *implicatum*) (Song 2001). Typology in this sense began with Joseph H. Greenberg’s discovery of such patterns in the form of implicational universals

TABLE 5.1. A tetrachoric table

	prepositions	postpositions
Verb-initial	+	—
Non-verb-initial	+	+

TABLE 5.2. Implicational Universal 1

A: ($-P +Q$)	C: ($+P +Q$)
B: ($-P -Q$)	*D: ($+P -Q$)

(P and Q refer to preverbal [CL-N] and postverbal [CL-N] phrases respectively.)

for features of both morphology and word order (Greenberg 1966a). A good example of this type of universal is the following: verb-initial languages are nearly always found to be equipped with prepositions. This can be rewritten as follows: if a language is verb-initial (p), then it is also prepositional (q). A tetrachoric table may be used to indicate clearly which of the logically possible combinations of two (or more) properties is allowed or disallowed. Table 5.1 shows one such example.

The tetrachoric table shows that the combination of verb-initial word order and postpositions is highly unlikely in human languages. Thus the implicational universal effectively serves as a constraint on possible variation within human languages.

As to [CL-N] phrases, according to the outcomes of my research, Sinitic languages can be typologized into seven different types by using two parameters: (i) syntactic position: preverbal vs. postverbal; and (ii) semantic interpretation: definite vs. indefinite. In terms of these two parameters, three implicational universals can be drawn up:

Universal 1: preverbal [CL-N] phrases (P) \supset postverbal [CL-N] phrases (Q), i.e. in a given language, if a preverbal [CL-N] phrase is acceptable, a postverbal one is also acceptable; but the opposite is not always true.

Table 5.2 shows the four logical possibilities. Type A languages only allow postverbal [CL-N] phrases but disallow preverbal ones. Neither preverbal nor postverbal [CL-N] phrases are accepted in Type B languages whereas both preverbal and postverbal [CL-N] phrases are acceptable in Type C languages. In contrast to this, type D languages which allow preverbal [CL-N] phrases, but disallow postverbal ones, do not exist in our sample. This is formulated as Universal 1.

TABLE 5.3. Implicational Universal 2

A: $(-\text{P}_{\text{INDEF}} + \text{P}_{\text{DEF}})$	C: $(+\text{P}_{\text{INDEF}} + \text{P}_{\text{DEF}})$
B: $(-\text{P}_{\text{INDEF}} - \text{P}_{\text{DEF}})$	*D: $(+\text{P}_{\text{INDEF}} - \text{P}_{\text{DEF}})$

(P_{INDEF} and P_{DEF} refer to preverbal indefinite [CL-N] and preverbal definite [CL-N] phrases respectively.)

TABLE 5.4. Implicational Universal 3

A: $(-\text{Q}_{\text{DEF}} + \text{Q}_{\text{INDEF}})$	C: $(+\text{Q}_{\text{DEF}} + \text{Q}_{\text{INDEF}})$
B: $(-\text{Q}_{\text{DEF}} - \text{Q}_{\text{INDEF}})$	*D: $(+\text{Q}_{\text{DEF}} - \text{Q}_{\text{INDEF}})$

(Q_{DEF} and Q_{INDEF} refer to postverbal definite [CL-N] phrases and postverbal indefinite ones respectively)

Universal 2: preverbal indefinite [CL-N] phrases (P_{INDEF}) \supset preverbal definite [CL-N] phrases (P_{DEF}), i.e. in a given language, if preverbal [CL-NP] phrases can have an indefinite reading, it must also be possible for them to have a definite reading; but the opposite is not always true.

Table 5.3 shows the four possibilities. Type A languages allow preverbal definite [CL-N] phrases but disallow indefinite ones. Type B languages disallow both preverbal indefinite [CL-N] phrases and preverbal definite ones. Both preverbal indefinite [CL-N] phrases and preverbal definite ones are acceptable in Type C languages. Type D languages which allow preverbal indefinite [CL-N] phrases, but disallow preverbal definite ones, do not exist in our sample, as we predicted according to our formulation of Universal 2.

Universal 3: postverbal definite [CL-N] phrases (Q_{DEF}) \supset postverbal indefinite [CL-N] phrases (Q_{INDEF}), i.e. in a given language, if a postverbal [CL-N] phrase has a definite reading, then it must also have an indefinite reading; but the opposite is not always true.

Table 5.4 illustrates the four possibilities. The postverbal indefinite [CL-N] phrases are acceptable while the postverbal definite ones are unacceptable in Type A languages. Neither postverbal definite [CL-N] phrases nor postverbal indefinite ones are allowed in Type B languages, while Type C languages allow both postverbal definite [CL-N] phrases and postverbal indefinite ones. The fourth logical possibility, Type D, which allows postverbal definite [CL-N] phrases but disallows postverbal indefinite ones is excluded, as predicted by Universal 3.

The combination of these three universals makes it possible to predict the existence of seven types of Sinitic languages, and at the same time to exclude nine other logical possibilities. Table 5.5 summarizes this patterning concisely.

(Note: The nine combinations that are not possible are subsumed under three main category labels given below while they are fully listed under Note 1.)

TABLE 5.5. Seven possible types of bare classifier configurations in Sinitic languages

Type I: [(+P _{INDEF} +P _{DEF}) (+Q _{DEF} +Q _{INDEF})]	Type VI: [(-P _{INDEF} -P _{DEF}) (-Q _{DEF} -Q _{INDEF})]
Type II :[(+P _{INDEF} +P _{DEF}) (-Q _{DEF} +Q _{INDEF})]	Type VII: [(-P _{INDEF} -P _{DEF}) (-Q _{DEF} +Q _{INDEF})]
Type III: [(-P _{INDEF} +P _{DEF}) (+Q _{DEF} +Q _{INDEF})]	Not attested : *(+P -Q) ¹
Type IV: [(-P _{INDEF} +P _{DEF}) (-Q _{DEF} +Q _{INDEF})]	*(+P _{INDEF} -P _{DEF})
Type V: [(-P _{INDEF} -P _{DEF}) (+Q _{DEF} +Q _{INDEF})]	* (+Q _{DEF} -Q _{INDEF})

5.3 Seven types of Sinitic languages

Next, I discuss each of the seven types in turn, providing examples.

5.3.1 Type I languages: All [CL-N] phrases allowed in both pre- and postverbal positions [(+P_{INDEF} +P_{DEF}) (+Q_{DEF} +Q_{INDEF})]

From the perspective of syntactic distribution, Type I languages allow not only postverbal but also preverbal [CL-N] phrases. These bare classifier phrases, no matter whether they are preverbal or postverbal, can have either a definite or indefinite reading determined by the context. There are ten dialects belonging to Type I languages in my sample of 120 different Sinitic languages.¹ These are the Jixi 绩溪 (Hui group), Lianshui 涟水 (Jianghuai Mandarin), Yiwu 义乌 (Wu group), Liyanuan 淈源 (Xiang group), Xinhua 新化 (Xiang group), Xintang 新塘 (Xiang group), Yining 义宁 (Guibei Pinghua), Luchuan 陆川 (Hakka), Chenghai 澄海 (Min group), and Tongcheng 通城 (Gan group) dialects. The following examples are from the Lianshui 涟水 dialect.

Preverbal [CL-N] phrases in the Lianshui 涟水 dialect

- (5) 条牛捱偷得了。
tʰiɔ³⁵ əw³⁵ ε⁵⁵ tʰəw³¹ təʔ⁰ lie⁰.
CL OX PAS steal RVC PRT
‘The/an ox was stolen.’

¹ The list of languages and dialects included in this database is presented in the appendix at the end of this chapter.

Postverbal [CL-N] phrases in the Lianshui 淮水 dialect

- (6) 我骑挂脚踏车来的。

ŋ²¹³⁻²¹ tç^hi³⁵ kua⁰ tçia?̄³⁴⁻³¹ t^hæ?̄³⁴⁻⁵ ts^həi³¹ lE³⁵ tie⁰.
1SG ride CL bike come PRT
'I rode a/the bike here.'

In the Lianshui 淮水 dialect, the [CL-N] phrases can be interpreted as either definite or indefinite no matter whether they are preverbal or postverbal. Thus, examples (5) and (6) are ambiguous because both the preverbal [CL-N] phrase 't^hi³⁵ əw³⁵ (CL ox)' in example (5) and the postverbal one 'kua⁰ tçia?̄³⁴⁻³¹ t^hæ?̄³⁴⁻⁵ ts^həi³¹ (CL bike)' in example (6) have two possible readings in the absence of any context: definite or indefinite. The bare classifier phrase can be seen to be semantically indeterminate for this feature.

The Jixi 绩溪 dialect also belongs to Type I languages. In the Jixi 绩溪 dialect, both preverbal and postverbal [CL-N] phrases can similarly code either indefiniteness or definiteness. The following examples illustrate this well.

Preverbal [CL-N] phrase in the Jixi 绩溪 dialect

- (7) 老张，个人是尔家门口转。

nə⁵⁵⁻⁵³ tçio²¹ ko³²⁴⁻³⁵ zā³² se⁵⁵ ð⁵⁵ ko²¹ mā³² k^he⁵⁵ tçyē³²⁴.
Lao Zhang, CL man at 2SG home door.mouth walk.around

- a. 'Lao Zhang, there is a man walking around at the entry to your house.'
b. 'Lao Zhang, the man is walking around at the entry to your house.'

Postverbal [CL-N] phrase in the Jixi 绩溪 dialect

- (8) 小王搬只沙发去哩

çia⁵⁵⁻⁵³ ð³² pā²¹ tse?̄³²⁻³⁵ so²¹⁻²² -fa?̄³² k^he³²⁴. ni.
Xiao Wang move CL sofa go PRT
'Xiao Wang took away the/a sofa.'

In examples (7) and (8), both the preverbal [CL-N] phrase 'ko³²⁴⁻³⁵ zā³² (CL man)' and the postverbal one 'tse?̄³²⁻³⁵ so²¹⁻²² fa?̄³² (CL sofa)' can be interpreted as either definite or indefinite. Thus, both examples (7) and (8) are ambiguous. It should be noted that the use of definite [CL-N] phrases in the Jixi 绩溪 dialect must be anaphoric and cannot convey new information. This constraint does not exist in the Lianshui 淮水 dialect. The following sentence, given in example (9), can be used in the Lianshui 淮水 dialect when the speaker points to one of two books, present in the speech context, but its counterpart in the Jixi 绩溪 dialect cannot be used in the same kind of exophoric, deictic context.

Postverbal [CL-N] phrase in the Lianshui 淮水 dialect

- (9) 我要本书。

ŋ²¹³⁻²¹ i⁵⁵ pən²¹³⁻²¹ su³¹.
1SG want CL book
'I want this book.'

Bare classifier phrases in Sinitic languages: A typological perspective 117

5.3.2 Type II languages: Sole exclusion of postverbal definite [CL-N] phrases

$[(+P_{INDEF}+P_{DEF}) (-Q_{DEF}+Q_{INDEF})]$

Distributionally, Type II languages allow both preverbal and postverbal [CL-N] phrases. The preverbal [CL-N] phrases can have two readings: definiteness and indefiniteness, while the postverbal one can only have an indefinite reading. The Loudi 娄底 and Ningxiang 宁乡 dialects, both of which belong to the Xiang group, are the only Type II languages in my sample. The following examples are from the Loudi 娄底 dialect (Liu L. 2001).

Preverbal [CL-N] phrase in the Loudi 娄底 dialect

- (10) 咯多巴多箇书，只箱子怕
ko⁴² to³³ ·pa to³³ ·ko cy⁴⁴, tcio¹³ siɔŋ⁴⁴⁻³³.tsi p^hɔ³⁵⁻⁵⁵
this many.SUF many PRT book CL box.SUF afraid
放 不 落 吧?
xɔŋ³⁵⁻⁵⁵ p^hu³⁵⁻⁵⁵ lo³⁵ ·pa.
put NEG fall PRT
'(I'm) afraid that the/one box cannot contain so many books.'

Postverbal [CL-N] phrase in the Loudi 娄底 dialect

- (11) 我吃碗饭就行。
ŋ⁴²⁻¹¹ tç^hiɔ¹³ ue⁻⁴²⁻¹¹ Yuã¹¹ dzio¹¹ Vɔ¹³.
1SG eat CL rice ADV fine
'For me, just a bowl of rice is enough.'

In example (10), the preverbal [CL-N] phrase 'tcio¹³ siɔŋ⁴⁴⁻³³.tsi (CL box)' can be interpreted as either indefinite or definite, while in example (11), the postverbal [CL-N] phrase 'ue⁻⁴²⁻¹¹ Yuã¹¹ (CL rice)' can only code indefiniteness.

The Ningxiang 宁乡 dialect (Xiang group) also belongs to Type II languages, shown by examples (12) and (13) (Zhang X. 2008).

Preverbal [CL-N] phrase in the Ningxiang 宁乡 dialect

- (12) 部车子停在路中间。
pu²¹ t^hie³³-tsi²¹ tin¹³ tsai⁴⁵ ləu²¹-tsən³³kan³³.
CL car-SUF stop in road-middle
'A/The car stops in the middle of the road.'

Postverbal [CL-N] phrase in the Ningxiang 宁乡 dialect

- (13) 换身衣服出门去。
z²¹ sən³³ i³³fu¹³ tcy²⁴ mən¹³ k^hə⁴⁵.
change CL clothes out door go
'Change some clothes and go out.'

Without its context, example (12) is ambiguous because the preverbal [CL-N] phrase ‘*pu²¹ t^{hie³³}* ts²¹ (CL car)’ can express either indefiniteness or definiteness. By way of contrast, the postverbal [CL-N] phrase ‘*sən³³ i³³ fu¹³* (CL clothes)’ in example (13) is not ambiguous, since it can only code indefiniteness.

5.3.3 Type III languages: Sole exclusion of preverbal indefinite [CL-N] phrases

$$[(-P_{INDEF} + P_{DEF}) (+Q_{DEF} + Q_{INDEF})]$$

From the perspective of syntactic distribution, Type III languages not only allow postverbal [CL-N] phrases but also allow preverbal ones. From the perspective of semantic interpretation, the postverbal [CL-N] phrases of Type III languages code either indefiniteness or definiteness, while the preverbal ones can only have a definite reading. The Guangzhou 广州 dialect (Yue group) is a typical Type III language.²

Preverbal [CL-N] phrase in the Guangzhou 广州 dialect

- (14) 本书好睇。
pun³⁵ sy⁵⁵ hou³⁵t^hei³⁵.
CL book interesting
'The book is interesting.'

Postverbal [CL-N] phrase in the Guangzhou 广州 dialect

- (15) 界杯茶我饮。
pei³⁵ pui⁵⁵ tj^ha¹¹ n³²³ iem³⁵.
give CL tea 1SG drink
'Give me the/a cup of tea to drink.'

In the Guangzhou dialect, preverbal [CL-N] phrases can only have definite reading, as example (14) shows, while postverbal [CL-N] phrases can code either definiteness or indefiniteness. Thus, example (15) is ambiguous without any context.

The Nanning Baihua 南宁白话 (Yue group) (also known as Nanning Cantonese, see Chapter 7) similarly belongs to Type III languages, as the following examples show (Lin and Qin 2008).

Preverbal [CL-N] phrase in Nanning Baihua 南宁白话

- (16) 张凳放喺边哒定?
tʃœŋ⁵⁵ təŋ³³ fɔŋ³³ hei³⁵ pin⁵⁵ tet³ teŋ²²?
CL chair put at which CL place
'Where should (I) put the chair?'

² Special thanks go to Sing Sing Ngai and Hilário de Sousa who offered me the data on the Guangzhou dialect.

Bare classifier phrases in Sinitic languages: A typological perspective 119

Postverbal [CL-N] phrases in Nanning Baihua 南宁白话

- (17) 系细王整只电脑烂嘅。
hei²² t̩ai³³ wɔŋ²¹ tʃen³⁵ tʃɛk³ tin²²nu²⁴ lan²² kɛ³³.
be little Wang make CL computer broken MOD
'It is little Wang who broke the computer.'
- (18) 讲只好消息给你哋听。
kɔŋ³⁵ tʃɛk³³ hu³⁵ t̩iu⁵⁵t̩ek⁵ kəi⁵⁵ ni²⁴ti²² thɛŋ⁵⁵.
say CL good news give 2PL hear
'(I) will tell you apiece of good news.'

The preverbal [CL-N] phrase 'tʃoen⁵⁵ t̩əŋ³³ (CL chair)' in example (16) can only code definiteness, while the postverbal [CL-N] phrases 'tʃɛk³ tin²²nu²⁴ (CL computer)' in example (17) and 'tʃɛk³³ hu³⁵ t̩iu⁵⁵t̩ek⁵ (CL good news)' in example (18) express definiteness and indefiniteness respectively, determined by their different contexts.

The Suzhou 苏州 dialect is also one of the Type III languages. There are, however, some further interesting constraints on definite postverbal [CL-N] phrases in the Suzhou 苏州 dialect. In the Suzhou 苏州 dialect, definite postverbal [CL-N] phrases can only occur under two circumstances. The first one is when they are modified by an attributive phrase or relative clause, as example (19) shows. The second one is when they act as the objects of reduplicated verbs, as example (20) shows (Shi and Liu 1985).

Postverbal [CL-N] phrases in the Suzhou 苏州 dialect

- (19) 我欢喜僚上趨件衣裳。
ŋəu³¹⁻³³ huø⁴⁴⁻⁵⁵ ci⁵³⁻²¹ nE³¹⁻³³ zã³¹⁻²² t̩a⁴¹²⁻⁵⁵ dzil³¹⁻²¹ i⁴⁴⁻⁵⁵zã²³⁻²¹.
1SG like 2SG above time CL clothes
'I like the clothes you wore last time.'
- (20) 我等等个客人。
ŋəu³¹⁻³³ tən⁵³-tən⁵³⁻²¹ kə?̩⁵⁻⁴ kʰa?̩⁵n̩in²³⁻³⁵.
1SG wait-RDP CL guest
'I will wait for the guest.'

Besides the three dialects mentioned above, the following six belong to Type III languages: Shanghai 上海 (Wu group), Wenzhou 温州 (Wu group), Tunchang 屯昌 (Min group), Yantai 烟台 (Jiaoliao Mandarin), Xiangxiang Hutian 湘乡壶天 (Xiang group), and Hezhou 贺州 (Yue group). In other words, there is a total of nine dialects in my sample belonging to this type.

5.3.4 Type IV languages: Mirror-image bare classifiers—preverbal definite and postverbal indefinite [CL-N] phrases

$$[(-P_{INDEF} + P_{DEF}) (-Q_{DEF} + Q_{INDEF})]$$

Type IV languages allow both postverbal and preverbal [CL-N] phrases. Interestingly, preverbal [CL-N] phrases in Type IV languages can only be interpreted as definite and postverbal ones can only express indefiniteness. The Fuyang 富阳 dialect (Wu group) represents a dialect which belongs to Type IV, as the following example shows (Li 2013).

Preverbal and postverbal [CL-N] phrases in the Fuyang 富阳 dialect

- (21) 个老板买了部车子。
kəʔ⁵⁻³ lɔ¹³⁻²⁴ pɛ̃⁴⁴⁻²¹ ma¹³.ləʔ bu⁰ tsʰo⁴⁴tsi⁰.
CL boss buy ASP CL car.SUF
'The boss bought a car.'

The preverbal [CL-N] phrase 'kəʔ⁵⁻³ lɔ¹³⁻²⁴ pɛ̃⁴⁴⁻²¹ (CL boss)' in example (21) codes definiteness and the postverbal one 'bu⁰ tsʰo⁴⁴tsi⁰' (CL car) expresses indefiniteness. Besides the Fuyang 富阳 dialect, the Jieyang 揭阳 dialect (Min group) also belongs to Type IV languages, as examples (22) and (23) show (H. Xu 2007).

Preverbal [CL-N] phrase in the Jieyang 揭阳 dialect

- (22) ko²¹³⁻⁵³ hui⁵⁵ hǎ³³ aĩ²¹³ me¹¹?
CL fish still want Q-PRT
'Do you still want to keep the fish?'

Postverbal [CL-N] phrase in the Jieyang 揭阳 dialect

- (23) lu⁵³ na¹¹ aĩ²¹³, ua⁵³ tsu⁵⁵⁻¹¹ kʰioʔ⁵⁻² kai⁵⁵⁻¹¹ siō³³ lu⁵³⁻⁵⁵ pan²¹³.
2SG if want 1SG then take CL box 2SG put
'If you want (to buy them), I will get you a box to put (the goods in).'

According to H. Xu (2007), the preverbal and postverbal [CL-N] phrases in the Jieyang 揭阳 dialect can similarly only express definiteness and indefiniteness respectively. In my sample there are only two dialects which belong to this type.

5.3.5 Type V languages: Exclusively postverbal but bivalent [CL-N] phrases

$$[(-P_{INDEF} - P_{DEF}) (+Q_{DEF} + Q_{INDEF})]$$

Type V languages only allow postverbal [CL-N] phrases and these may express either definiteness or indefiniteness. Up until now, we have found just one

Bare classifier phrases in Sinitic languages: A typological perspective 121

Type V language: the Nanning Shangyao 南宁上尧 dialect (Pinghua group) (de Sousa 2013).³

Postverbal [CL-N] phrases in the Nanning Shangyao 南宁上尧 dialect

- (24) 个只香蕉呢，就系只外甥吃。
kə⁵⁵ tʃət³ tsiu⁵³tʃi³³ nə⁵⁵ tʃəu²² hei²⁵ tʃət³ wai²²tʃəŋ⁵³ hət³.
DEM CL banana PRT ADV give CL grandson eat
'This banana is for the grandson to eat.'
- (25) 我想介绍个女人系你。
ŋa¹³ tʃəŋ³³ kai⁵⁵tʃiu²² kə⁵⁵ nəi¹³ jənən¹¹ hei²⁵ nəi¹³.
1SG want introduce CL woman give 2SG
'I want to introduce a woman to you.'

The postverbal [CL-N] phrases 'tʃət³ wai²²tʃəŋ⁵³ (CL grandson)' and 'kə⁵⁵ nəi¹³jənən¹¹ (CL woman)' in example (24) and (25) express definiteness and indefiniteness respectively. The Nanning Shangyao 南宁上尧 dialect does not allow preverbal [CL-N] phrases.

5.3.6 Type VI languages: All [CL-N] phrases excluded

$$[(-P_{INDEF} - P_{DEF}) (-Q_{DEF} - Q_{INDEF})]$$

Type VI languages do not allow any kind of [CL-N] phrases, no matter whether they are preverbal or postverbal. Most Min dialects which are distributed in Fujian and Taiwan belong to Type VI languages. I have six representative dialects in my sample. These Min dialects make use of [DEM-CL-N] and [ONE-CL-N] phrases to denote definiteness and indefiniteness respectively. Examples (26)–(31) are from the Fuzhou 福州, Xiamen 厦门, Fuding 福鼎, Yong'an 永安, Shipi 石陂, and Taipei 台北 dialects respectively.⁴ I use the notational convention '*' (x) in the following set of examples to show that the demonstrative or numeral for 'one' may not be omitted to serve as evidence of the widespread restriction on the use of bare classifier phrases in Min. The first two examples are taken from the Fuzhou 福州 dialect.

- (26) a *许头爪飞去了。
*(xi²¹) lau⁵² tseu³¹ pui⁴⁴ ɔ²¹ lau³¹.
DEM CL bird fly go PRT
'The bird flew away.'

³ As there are many hundreds of dialects belonging to each of the ten branches of Sinitic, we predict that there will be others belonging to Type V which our small random sample cannot proportionally represent. Their discovery awaits a larger sample of dialects.

⁴ Special thanks go to Lin Jangling and Professor Li Rulong who gave me the data on Min dialects.

- b 我送*(蜀)本书乞伊。
ŋuai³¹ sɔŋŋ²¹ *(so³¹) βuon³¹ tsy⁴⁴ khøy²³ ·i.
1SG give one CL book give 3SG
'I gave him a book.'

The next two examples are from the Xiamen 厦门 dialect.

- (27) a (许)只爪飞去嘒。
*(hit⁵) tsia?⁵ tsiau⁴⁴-a⁵³ pe⁴⁴ ·khi ·lo.
DEM CL bird-SUF fly go PRT
'The bird flew away.'

- b 伊共我讲*(蜀)项代志。
i⁴⁴ kan²² gua⁴⁴ kɔŋ⁴⁴ *(tshit³²) han²² tai²¹tsi²¹.
3SG with 1SG tell one CL thing
'He told me something.'

The two following examples are from the Fuding 福鼎 dialect.

- (28) a (许)头爪飞走了。
*(hi⁵⁵) lau³¹ tseu⁵⁵ pui³⁴ tsau⁵⁵ ·lo.
DEM CL bird fly go PRT
'The bird flew away.'

- b 我送伊*(蜀)本书册。
ua⁵⁵ θen⁵² i³⁴ *(θi?²) puon⁵⁵ tsi³³za⁵².
1SG give 3SG one CL book
'I gave him a book.'

The next examples are from the Yong'an 永安 dialect.

- (29) a (兀)只爪子飞行了。
*(uo²¹³) tʃin⁴⁴ tsuɔ⁵³-tsã²¹ pue⁵² kio³³ ·no.
DEM CL bird-SUF fly go PRT
'The bird flew away.'

- b 我送*(寡)本书欠渠。
ŋuo⁵² san²⁴ *(kuɔ²¹) puã³³ ſy⁵² khëi²⁴ ny⁵².
1SG give one CL book give 3SG
'I gave him a book.'

Two examples are also provided from the Shipi 石陂 dialect.

- (30) a *(兀)个只仔飞掉了。
*(u²¹³) ko³³ tsia²¹³ tsia²¹³-te fye⁴⁴ ·thau ·lo.
DEM CL bird fly away PRT
'The bird flew away.'

Bare classifier phrases in Sinitic languages: A typological perspective 123

b 床腹底倒*(个)只人。

tshyɔŋ³³ pu²¹³ti²¹ tɔ²¹ *(ko⁵³) tsia²¹³ nein³³.
bed abdomen_{LOC}bottom_{LOC} lay one CL person
'A man is lying on the bed.'

The final two examples are from the Taipei 台北 dialect.

(31) a [(这)个东西有多重?

*(tsit³²⁻⁴⁴) le²⁴⁻¹¹ bī?⁴⁴⁻³² kiā³³ u³³⁻¹¹ gua³³⁻¹¹ tan³³?
DEM CL thing have much weight
'How much does it weigh? (literally: How much does *this thing* weigh?)'

b 与我*(一)本书。

ho³³⁻¹¹ gua⁵³⁻⁴⁴ *(tsit³²⁻⁴⁴) pun⁵³⁻⁴⁴ tsu⁴⁴
give 1SG one CL book
'Give me a book.'

In the sentences above, the numerals and demonstratives in [ONE-CL-N] and [DEM-CL-N] phrases cannot be deleted, no matter whether they are postverbal or preverbal.

Hence, most Min dialects in Fujian and Taiwan do not allow [CL-N] phrases. The exceptions are the Chaoshan 潮汕 (including the Type IV Jieyang dialect) and Tunchang 屯昌 dialects (Min group) which are however located in Guangdong 广东 and Hainan 海南 provinces respectively, outside the Min heartland. It is certainly possible that the patterning for [CL-N] phrases in the Chaoshan 潮汕 and Tunchang 屯昌 dialects could be the result of language contact with Cantonese and Hlai (one of Tai-Kadai languages) respectively, an issue which would evidently need further research, outside the scope of this description.

5.3.7 Type VII languages: Exclusively postverbal indefinite [CL-N]

$$[(-P_{INDEF} - P_{DEF}) (-Q_{DEF} + Q_{INDEF})]$$

Type VII languages only allow postverbal [CL-N] phrases which code indefiniteness. Standard Mandarin is typical of the Type VII languages, as examples (2) and (3) show. According to the data I have collected, there are 90 dialects belonging to this type which include most Mandarin, Jin, Gan, Hakka, and northern Xiang dialects.

5.4 The explanations for the three implicational universals

The first implicational universal formulated as

preverbal [CL-NP] phrase (P) ⊃ postverbal [CL-NP] phrase (Q)

can be shown to reflect a more general tendency, namely, grammatical material that is postposed tends to undergo more reduction and fusing and it will grammaticalize

in both form and semantic content to a greater extent than preposed material (Bybee et al. 1990: 23).

Let me take an example from the Beijing 北京 dialect. The postverbal preposition 在 [tsai⁵¹] ‘at’ is often read as [·tə] or deleted in the Beijing dialect, while the preverbal one is never weakened or deleted (Liu D. 2003a: 337). Another example is from the Lianshui 淄水 dialect of Jianghuai Mandarin. The prepositional phrases 蹲这块 [tən³¹ tsə?³⁴ k^huə?⁰] ‘to be here’ or 蹲那块 [tən³¹ na?³⁴ k^huə?⁰] ‘to be there’ have grammaticalized into quasi-durative markers when they are postverbal. The preposition 蹲 [tən³¹] and the demonstrative in this quasi-durative marker are often omitted and only the locative 块 [khuə?⁰] is left. The prepositional phrases 蹲这块 [tən³¹ tsə?³⁴ k^huə?⁰] ‘to be here’ or 蹲那块 [tən³¹ na?³⁴ k^huə?⁰] ‘to be there’ have also grammaticalized into quasi-progressive markers when preverbal. But the omission of preposition 蹲[tən³¹] and the demonstrative in the quasi-progressive marker is not possible under most circumstances.

Returning to the subject at hand, the definite and indefinite [CL-N] phrases have their source in either [DEM-CL-N] or [ONE-CL-N] phrases, resulting from the omission of the demonstratives or numeral ‘ONE’ respectively. That the demonstratives and numeral ‘ONE’ in postverbal position are much more prone to omission than those in preverbal position reflects the general tendency proposed by Bybee et al. (1990). Thus postverbal [CL-N] phrases are more frequent than preverbal ones in Sinitic languages, as my sample has shown, given the predominance of Type VII languages.

The phrases with demonstratives or numeral ‘ONE’ are labelled as definite or indefinite respectively in Standard Mandarin (Chen P. 1987). But how can we determine the kind of referentiality that a [CL-N] phrase has when it is not connected intrinsically with the features of definiteness and indefiniteness? Chao (1968: 76) claims that there is a very strong tendency for the subject to have a definite reference and the object to have an indefinite reference in Chinese. Thus, syntactic role and distribution play a significant role in determining whether a [CL-N] phrase is definite or indefinite. From this point of view, only postverbal indefinite [CL-N] phrases are unmarked and all the other possibilities are marked. According to the data I have collected, 90 out of 120 dialects allow only postverbal indefinite [CL-N] phrases. Consequently, it is not surprising to find that the classifiers in postverbal indefinite [CL-N] phrases are nearly always exceptionlessly realized as default forms without special phonological changes in my sample of 120 Sinitic languages. We can also predict that the classifiers in marked [CL-N] phrases will be realized as non-default forms that undergo special phonological changes. I explain this below with examples.

According to You (2003: 174), if a monosyllabic classifier in the Wenzhou 温州 dialect (Wu group; Type III language) changes its basic tone to a *rù* (entering) tone, it can act as a proximal demonstrative. There are eight basic tones in the Wenzhou 温州 dialect: 阴平 *yīnpíng* (33), 阳平 *yángpíng* (31), 阴上 *yīnshǎng* (35), 阳上 *yángshǎng* (24), 阴去 *yīnqù* (42), 阳去 *yángqù* (11), 阴入 *yīnrù* (313), and 阳入 *yángrù* (212). The following are the rules of tone sandhi in the Wenzhou dialect.

Bare classifier phrases in Sinitic languages: A typological perspective 125

a) If the tone of a monosyllabic classifier with a voiceless initial is not *rù*, it will change to *yīnrù* (upper entering tone), as the following examples show:

(32) 个[kai⁴²]——个[kai³¹³]人 ‘this man’ 本[pan]³⁵——本[pan]³¹³书 ‘this book’

b) If the tone of a monosyllabic classifier with a voiced initial is not *rù*, it will change to *yángrù* (lower entering tone), as the following examples show:

(33) 头[deu³¹]——头[deu²¹²]鱼 ‘this fish’ 枚[mai³¹]——枚[mai²¹²]镜 ‘this mirror’

c) If the tone of a monosyllabic classifier is *rù*, it will keep its basic tone, regardless of the voicing feature, as the following examples show:

(34) 粒[lø²¹²]——粒[lø²¹²]糖儿 ‘this piece of candy’ 荸[ka³¹³]——荸[ka³¹³]香蕉 ‘this banana’

You Ruijie 游汝杰 (pers. comm.) points out that the definite [CL-N] phrases in the Wenzhou dialect derive from [proximal demonstrative-CL-NP] phrases resulting from the omission of proximal demonstrative 居[ke²¹] ‘this’. This can be seen in the fact that the tone of the classifier in a definite [CL-N] phrase keeps its tone sandhi contour invariant as it occurred in the original [ke²¹-CL-N] phrase. The following examples illustrate this clearly. Compare example (35a) with example (35b), bearing in mind that the tone sandhi rules given above are in operation: rule (c) for example (35) and rule (a) for example (36).

- (35) a 居只笔忒贵。
ke²¹ tsei³¹³ pi³¹³ thru²¹ tçy⁴²
This CL pen very expensive
‘This pen is very expensive.’
- b 只笔忒贵。
tsei³¹³ pi³¹³ thru²¹ tçy⁴².
CL pen very expensive
‘This pen is very expensive.’

In the Wenzhou 温州 dialect, we have a clear case of definite bare classifier [CL-N] phrases being used to make a contrast with [distal demonstrative-CL-NP] phrases, as the following examples show.

- (36) 张凳匱你坐，我坐许张凳。
tçi³³⁻³¹³ tan⁴² ha⁴² n*j*²⁴ zuɔ²⁴, iŋ²⁴ zuɔ²⁴ ci³⁵ tçi³³ tan⁴².
CL stool give 2SG sit 1SG sit that CL stool
‘You sit on this stool, I sit on that stool.’

- (37) 许碗饭匱你吃，我吃碗饭。
ci³⁵ y³⁵ va¹¹ ha⁴² n*j*²⁴ ts^hi³¹³, iŋ²⁴ ts^hi³¹³ y³⁵⁻³¹³ va⁶.
That bowl(CL) rice give 2SG eat 1SG eat bowl(CL) rice
‘You eat that bowl of rice, I eat this one.’

According to Pan and Tao (1999: 33–4), a classifier with a special tone sandhi in a [CL-N] phrase in the Shanghai 上海 dialect (Wu group; Type III language) codes not

only definiteness but also proximal deixis. Compare example (38a) with example (38b) for the tone sandhi of the classifier [pəŋ⁵³⁻⁵⁵]. 本

- (38) a 个本书好, 哀本勿灵。

gəʔ¹² pəŋ⁵³⁻⁵⁵ s1⁵³ hɔ³⁴, E⁵³ pəŋ⁵³⁻⁵⁵ vəʔ¹² lin²³.
this CL book good that CL NEG good
'This book is good, that one is not good.'

- b 本书好, 哀本勿灵。

pəŋ⁵³⁻⁵⁵ s1⁵³ hɔ³⁴, E⁵³ pəŋ⁵³⁻⁵⁵ vəʔ¹² lin²³.
CL book good that CL NEG good
'This book is good and that one is not.'

The [CL-N] phrase [pəŋ⁵³⁻⁵⁵ s1⁵³] 本书 in which classifier [pəŋ⁵³] 本 undergoes tone sandhi in example (38b) takes on the same function as its counterpart [gəʔ¹² pəŋ⁵³⁻⁵⁵ s1⁵³] 个本书 in example (38a) with the [DEM-CL] phrase. If the classifier in a definite [CL-N] phrase is realized simply as a neutral tone, the phrase cannot express proximal deixis in the exophoric deictic context of pointing. Thus the following sentence is not acceptable in the Shanghai 上海 dialect and can be usefully compared with (38b).

- (39) *本书好, 哀本勿灵。

*pəŋ⁵³⁻³ s1⁵³ hɔ³⁴, E⁵³ pəŋ⁵³⁻⁵⁵ vəʔ¹² lin²³.
CL book good that CL NEG good
(Attempted meaning: 'This book is good and that one is not.')

The [CL-N] phrases with a neutral tone in the Shanghai 上海 dialect can express either indefiniteness, if they are postverbal, or definiteness if they are preverbal. The following example illustrates this neatly.

- (40) 我昨日买了本书, 本书勿灵。

ŋu³³ zɔʔ¹²-nii¹²⁻²³ ma⁴⁴ ləʔ³ pəŋ⁵³⁻³ s1⁵³ pəŋ⁵³⁻³ s1⁵³ vəʔ¹² lin²³.
1SG yesterday buy ASP CL book CL book NEG good
'I bought a book yesterday and the book is not good.'

Obviously, the definite [CL-N] phrases with a neutral tone on the classifier in the Shanghai 上海 dialect cannot code proximal deixis. This unmarked form shows in its turn a greater degree of grammatical reanalysis and semantic/pragmatic generalization than those undergoing special tone sandhi, since the latter forms reflect in fact their original position, meaning and consequent tone sandhi when following a demonstrative (see Chapter 3 for more discussion on reanalysis and grammaticalization).

In the Xiangxiang 湘乡 dialect (Xiang group; Type III language), if a [CL-N] phrase is encoded as definite, the classifiers must undergo tone sandhi, as the following example shows (He 2006: 21).

Bare classifier phrases in Sinitic languages: A typological perspective 127

(41) 隻猪杀哩三百多斤。

tçio^{13/33} cy⁴⁴⁻³³ sa¹³⁻³³ .li sã⁴⁴⁻³³ pe¹³⁻³³ to⁴⁴⁻³³ tçen⁴⁴⁻³³.

CL pig kill ASP three hundred more jin

‘A [tçio¹³]/The [tçio³³] pig was killed and it is more than three hundred jin in weight.’

The classifier 隻[tçio¹³] in the Xiangxiang 湘乡 dialect must change its basic tone from 13 to 33 in order to express definiteness. If 隻[tçio¹³] does not undergo tone sandhi, it will code indefiniteness.

Another piece of evidence in favour of the markedness properties of definite bare classifier phrases is constituted by the following three properties which all involve phonological attrition. According to Liu and Shi (2012), the pronunciation of a definite classifier in the Yantai 烟台 dialect (Jiaoliao Mandarin; Type III language) has three characteristics: (i) the final of a definite classifier is changed to a schwa, or even deleted;⁵ (ii) the syllable of the classifier is shortened with a glottal stop coda; and (iii) the tone of the classifier is neutralized.

Similarly, in the Lianshui 淮水 dialect (Jianghuai Mandarin; Type I language), only the classifier in an indefinite postverbal [CL-N] phrase, which has the unmarked value, can be realized as the default form, i.e. in neutral tone. In contrast to this, the syllable of a classifier in a marked [CL-N] phrase is preceded by an unreleased glottal stop and is shortened. Thus, the final of the general classifier 个 [kəw⁵⁵] in a definite [CL-N] phrase, no matter whether it is preverbal or postverbal, is changed from [əw] to [ə?].

An even more complicated situation arises in the Hezhou 贺州 dialect (Yue group; Type III language) where a preverbal [CL-N] phrase can only be definite; while a postverbal one can code either definiteness or indefiniteness, depending on whether the classifier undergoes tone sandhi or not. If a disyllabic postverbal [CL-N] phrase expresses indefiniteness, the classifier undergoes tone sandhi in accordance with the sandhi rule for a disyllabic word. If a disyllabic phrase is definite, no matter whether it is preverbal or postverbal, the classifier is pronounced in its basic tone. This presents somewhat of a conundrum in that it appears to be a counterexample to the markedness theory outlined above (Zhang and Tang 2010). Compare first of all (42a) with (42b).

(42) a 头人

tou¹³²⁻²¹ nuon¹³²

CL person

‘a person (indefinite)’

(with tone sandhi)

b 头人

tou¹³² nuon¹³²

CL person

‘the person (definite)’

(without tone sandhi)

⁵ In the Yantai 烟台 dialect, a syllable is formed by an initial, a final, and a tone, as is the typical phonological segmentation for most Sinitic languages. Initials consist of consonants or semi-vowels; finals consist of vowels or vowels plus one of the two nasal sounds such as [n] or [ŋ]. The tone is superimposed over the entire syllable.

How then can this contradictory data be explained? As it turns out, however, the classifier in its basic tone has to be the marked form (and not the sandhi form), because all disyllabic words across the board undergo tone sandhi in the Hezhou 贺州 dialect. According to Zhang and Tang (2010), no tone change is however observed for the first syllable of a trisyllabic word or phrase in the Hezhou dialect.

- (43) 戴眼镜
lai³⁵ njan¹³²⁻²¹ kən³⁵
wear glasses
'to wear glasses'

When a classifier constitutes a trisyllabic phrase with a disyllabic noun, it can be either stressed or unstressed, depending on the interpretation of the [CL-N] phrase. If the phrase is postverbal and denotes indefiniteness, the classifier remains unstressed. If the phrase is definite, the classifier will be stressed. For example, (bold characters indicate stress).

- (44) 头人客头人客
tou¹³² njuon¹³² hak³⁴ tou¹³² njuon¹³² hak³⁴
CL guest CL guest
'a guest (indefinite)' 'the guest (definite)'

Apparently, the stressed form is thus the marked one and the unstressed form is the unmarked one, corresponding to the definite and indefinite readings (Zhang and Tang 2010).

We also find that in some dialects, such as the Jixi 绩溪 dialect (Hui group; Type I language), the classifier in a [CL-N] phrase is always realized as its default form, i.e. tone sandhi is applied in accordance with the rules, no matter where it occurs and whether it is definite or indefinite. This similarly shows that the [CL-N] phrases in these dialects have been subject to a greater degree of reanalysis and generalization compared with their counterparts in other dialects, since there is no longer any correspondence with a semantic or discourse-level distinction.

Table 5.6 illustrates the relations between the pronunciation of the classifiers and the semantic interpretation and syntactic distribution of [CL-N] phrases in 19 Sinitic languages.

5.5 Concluding words

In this chapter, three implicational universals have been proposed, based on the syntactic distribution (preverbal vs. postverbal) and semantic interpretation (definite vs. indefinite) of bare classifier [CL-N] phrases in Sinitic languages. I have set out to explain the following three implicational universals: 1) preverbal [CL-NP] phrases ⊃

Bare classifier phrases in Sinitic languages: A typological perspective 129

TABLE 5.6. Bare classifier phrases in 19 Sinitic languages with their corresponding phonological properties

	Preverbal indefinite	Preverbal definite	Postverbal definite	Postverbal indefinite
Type I: All [CL-N] allowed in both pre- and postverbal positions [(+P _{INDEF} +P _{DEF}) (+Q _{DEF} +Q _{INDEF})]				
Jixi 绩溪 (Hui group)	Tone sandhi in accordance with the rules No specific tone sandhi	Tone sandhi in accordance with the rules No specific tone sandhi	Tone sandhi in accordance with the rules No specific tone sandhi	Tone sandhi in accordance with the rules No specific tone sandhi
Lianshui 涟水 (Jianghuai Mandarin)	Preceded by unreleased glottal stop, Syllable shortening	Preceded by unreleased glottal stop Syllable shortening	Preceded by unreleased glottal stop Syllable shortening	Neutral tone
Xinhua 新化 (Xiang group)	Basic tone	Basic tone	Basic tone	Basic tone
Lianyuan 淊源 (Xiang group)	Basic tone	Basic tone	Basic tone	Basic tone
Xintang 新塘 (Xiang group)	Basic tone	Basic tone	Basic tone	Basic tone
Type II: Sole exclusion of postverbal definite [CL-N] [(+P _{INDEF} +P _{DEF}) (-Q _{DEF} Q _{INDEF})]				
Loudi 娄底 (Xiang group)	Basic tone	Basic tone		Basic tone
Type III: Sole exclusion of preverbal indefinite [CL-N] [(-P _{INDEF} +P _{DEF}) (+Q _{DEF} +Q _{INDEF})]				
Guangzhou 广州 (Yue group)	Basic tone	Basic tone	Basic tone	Basic tone
Nanning baihua 南宁白话 (Yue group)	Basic tone	Basic tone	Basic tone	Basic tone
Wenzhou 温州 (Wu group)	Change to <i>rì</i> tone	Change to <i>rì</i> tone	Neutral tone; Basic tone	
Suzhou 苏州 (Wu group)	Special tone sandhi	Special tone sandhi	Neutral tone	Neutral tone
Shanghai 上海 (Wu group)	Neutral tone	Neutral tone	Neutral tone	Neutral tone
Yantai 烟台 (Jiaoliao Mandarin)	Special tone sandhi	Special tone sandhi	Neutral tone	Neutral tone
	Neutral tone	Neutral tone		
	Rhyme is changed to schwa	Rhyme is changed to schwa	Neutral tone	

(continued)

TABLE 5.6. Continued

	Preverbal indefinite	Preverbal definite	Postverbal definite	Postverbal indefinite
Xiangxiang 湘乡 (Xiang group)	Syllable shortening Change to neutral tone	Special tone sandhi	Special tone sandhi	Basic tone
Hezhou 贺州 (Yue group)	Basic tone Stressed	Basic tone Stressed	Tone sandhi in accordance with the rules No specific tone sandhi Unstressed	
Tunchang 屯昌 (Min group)	Special tone sandhi	Special tone sandhi	Tone sandhi in accordance with the rules No specific tone sandhi Unstressed	

Type IV: Mirror image bare classifiers – preverbal definite and postverbal indefinite
 $[CL-N] [(-P_{INDEF} + P_{DEF}) (-Q_{DEF} + Q_{INDEF})]$

Jieyang 揭阳 (Min group)	Tone sandhi in accordance with the rules No specific tone sandhi Unstressed	Tone sandhi in accordance with the rules No specific tone sandhi Unstressed
Fuyang 富阳 (Wu group)	Tone sandhi in accordance with the rules No specific tone sandhi Unstressed	Neutral tone

Type V: Exclusively postverbal but bivalent $[CL-N] [(-P_{INDEF} - P_{DEF}) (+Q_{DEF} + Q_{INDEF})]$

Shangyao 上尧 (Pinghua group)	Basic tone	Basic tone
--------------------------------	------------	------------

Type VI: All bare $[CL-N]$ excluded

Type VII: Exclusively postverbal indefinite $[CL-N] [(-P_{INDEF} - P_{DEF}) (-Q_{DEF} + Q_{INDEF})]$
 90/120

Beijing 北京 (Mandarin)	Neutral tone
--------------------------	--------------

postverbal [CL-NP] phrases; 2) preverbal indefinite [CL-NP] phrases ⊃ preverbal definite [CL-NP] phrases; 3) postverbal definite [CL-NP] phrases ⊃ postverbal indefinite [CL-NP] phrases. The three implicational universals predict that only seven types of Sinitic languages are possible. Data from 120 Sinitic languages support these predictions.

The first implicational universal reveals a more general tendency, i.e. grammatical material that is postposed will undergo more reduction and fusing and it will undergo reanalysis and generalization of both its form and semantic content to a greater extent than preposed material (Bybee et al. 1990). The second and the third implicational universals corroborate the tendency which was first claimed by Chao (1968) that preverbal noun phrases tend to be definite and postverbal ones tend to be indefinite in Sinitic languages.

Appendix: Sinitic languages found in the sample and their classification

See Chapter 2 for a map of Sinitic languages and more details on their classification into ten branches.

1. MANDARIN 官話 :

- (i) *Beijing Mandarin* 北京官话: Beijing 北京话, Langfang 廊坊, Chifeng 赤峰
- (ii) *Northeastern Mandarin* 东北官话: Jinzhou 锦州, Shenyang 沈阳, Tonghua 通化, Changchun 长春, Daqing 大庆, Ha'erbin 哈尔滨, and Haila'er 海拉尔.
- (iii) *Jiaoliao Mandarin* 胶辽官话: Yantai 烟台(Liu and Shi 2012), Dalian Jinzhou 大连(金州), Qingdao 青岛, Juxian 莒县, and Tonghua 通化.
- (iv) *Jilu Mandarin* 冀鲁官话: Baoding 保定, Tianjin 天津, Tangshan 唐山, Qianxi 迁西, Changli 昌黎, Wuqiao 吴桥, Zibo 淄博, and Shouguang 寿光.
- (v) *Zhongyuan Central Plains Mandarin* 中原官话: Suining 睢宁, Xuzhou 徐州, Suqian 宿迁, Ganyu 赣榆, Xi'an 西安, Guyuan 固原, Lingbao 灵宝, Zhumadian 驻马店, and Xinyang 信阳.
- (vi) *Lanyin Northwestern Mandarin* 兰银官话: Yinchuan 银川, Wuzhong 吴中, and Urumqi 乌鲁木齐.
- (vii) *Southwestern Mandarin* 西南官话: Xichong 西充, Chengdu 成都, Nanchong 南充, Bijie 毕节, Zunyi 遵义(Hu 2010), Kunming 昆明, Zhaotong 昭通, Wuhan 武汉, and Zhijiang 枝江.
- (viii) *Jianghuai Southern Mandarin* 江淮官话: Lianshui 涟水, Guannan 灌南, Yandu 盐都, Xinghua 兴化, Dongtai 东台, Taizhou 泰州, Hai'an 海安 (Zhang Y. 2008; Wang R. 2012), Nanjing 南京, Yangzhou 扬州, Nantong 南通, Heifei 合肥, Liu'an 六安, Zongyang 枞阳, Anqing 安庆, Xuancheng 宣城, and Huangzhou 黄州(Wang H. 2005).

2. JIN 晋语:

Taiyuan 太原, Pingyao 平遥, Jincheng 晋城, Dingxiang 定襄, and Shenmu 神木 (Xing 2002).

3. WU 吴语:

Suzhou 苏州 (Shi and Liu 1985), Shanghai 上海 (Qian N. 1997; Yang 1998), Changshu 常熟, Changzhou 常州, Wuxi 无锡, Wuhu County 芜湖县, Fuyang 富阳 (Li 2010), Yiwu 义乌 (Chen X. 1992), Shaoxing 绍兴, and Wenzhou 温州 (You 2003).

4. MIN 闽语:

Fuzhou 福州, Xiamen 厦门, Fuding 福鼎, Yong'an 永安, Shipi 石陂, Quanzhou 泉州 (Lin H. 2008), Taipei 台北, Tunchang 屯昌 (Qian D. 2002), Jieyang 揭阳 (H. Xu 2007), Chenghai 澄海 (Chen and Lin 2003).

5. GAN 赣语:

Yichun 宜春, Nanchang 南昌, Fuzhou 抚州, Xingzi 星子, Duchang 都昌, Ji'an 吉安, Tongcheng 通城, Susong 宿松 (Tang 2005).

6. HAKKA (Kejia) 客家话:

Meixian 梅县, Luchuan 陆川, Shicheng 石城, Liancheng Xinquan 连城新泉 (Xiang 1997).

7. YUE 粤语:

Hongkong 香港, Guangzhou 广州, Nanning Baihua 南宁白话 (Lin and Qin 2008), Xinhui 新会(Gan 2010), Hezhou 贺州 (Zhang and Tang 2010).

8. XIANG 湘语:

Changsha 长沙, Shaoshan 韶山, Shaoyang 邵阳, Changning 常宁, Xiangxiang 湘乡 (He 2006), Loudi 娄底 (Liu L. 2001), Xinhua 新化 (Luo 1998), Lianyuan 淊源 (Chen H. 1999), Hengdong Xintang 衡东新塘 (Xu Q. 2007), Ningxiang 宁乡.

9. PINGHUA 平话:

Nanning Jiangxi 南宁江西镇, Nanning Shangyao 南宁上尧平话 (Su 2010), Chongzuo Xinhe Zheyuan 崇左新和蔗园话 (Liang and Lin 2009), Guinan Pinghua 桂南平话.

10. HUI 徽语:

Jixi Lingbei 绩溪岭北, Shexian 歙县, Qimen 祁门.

UNCLASSIFIED DIALECT: Waxiang (Xianghua 乡话) 瓦乡话

Acknowledgements

I am grateful to Prof. Hilary Chappell for her insightful comments, which led to great improvements in the organization and argumentation of the chapter. Errors and infelicities which remain are of course my own. Apart from receiving funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013): ERC Advanced Grant agreement No. 230388:

Bare classifier phrases in Sinitic languages: A typological perspective 133

‘The hybrid syntactic typology of Sinitic languages’ (2009–12), the research reported on in this chapter has also been sponsored by the National Philosophy and Social Sciences foundation of China (Project No. 13BYY046), MOE (Ministry of Education in China) Project in the Humanities and Social Sciences (Project No. 11YJC740102), Jiangsu Provincial Project in Philosophy and Social Sciences (Project No. 11YYB005), Qinglan Project and 333 Project.

6

The comparative construction in Sinitic languages: Synchronic and diachronic variation

HILARY M. CHAPPELL AND ALAIN PEYRAUBE

6.1 Introduction

Synchronously, to express a comparative of inequality, two comparative construction types predominate in Sinitic languages: Type I *Compare* and Type II *Surpass*.¹

While the Type I *Compare* comparative is synchronically the predominant one in Sinitic languages (see discussion in §6.3), the distribution of the Type II *Surpass* comparative is much more widespread than has been previously supposed.

Diachronically, the comparative construction of inequality in Archaic and Pre-Medieval Chinese belongs to Type II *Surpass* in terms of its structure, with the comparative marker being YÙ 于 ‘to, at’. So too the constructions formed by RŪ 如 ‘be like’ and sī 似 ‘be like’, which were subsequently used as comparative markers of inequality in the periods of Early and Late Medieval Chinese.² Starting in the Late Medieval period (seventh–thirteenth centuries), we also find a comparative construction belonging to Type I, with BÌ 比 < ‘to compare’ as the comparative marker. This construction then became the dominant one in Northern Chinese and finally in Standard Mandarin today, totally replacing all the Type II constructions.

In Southern Sinitic, however, where Type I *Compare* had not yet been adopted, verbs such as GUÒ 過 (过) < ‘to cross, to surpass’ had time to grammaticalize into

¹ *Compare* type, because the standard comparative marker of this Type I is BÌ 比 whose original meaning is ‘to compare’; *Surpass* type, because the origin of the standard comparative marker used in Type II is a verb, GUÒ 過(过), meaning ‘to surpass’. In this chapter, we adopt the classification for the main general types of linguistic comparison as used in Huddleston and Pullum (2002: 1100).

² Note that here we are referring to the syntactic structure as being a head-marking comparative. As will be explained below, the cognitive schemas of *Location* and *Similarity* are associated diachronically with the Type II *Surpass* structure. To be succinct, we shall refer to Type II mainly as the *Surpass* comparative, using this as a cover term for the structure.

comparative markers and to replace the other comparative markers used in these Type II head-marking structures (*RÚ* 如 and *sì* 似). In this analysis, we argue for the position that this is a language-internal development which would be difficult to attribute to any contact-induced grammaticalization.

6.1.1 Background to the issue

This section provides the necessary background material to the specific issue at hand, namely the synchronic distribution of comparative construction types in China and the diachronic developments which may have led to this patterning.

As foreshadowed above, two dominant structural types have been identified in Sinitic languages for the comparative of inequality:³

- (i) *Compare* type – Type I dependent-marking $NP_A - CM - NP_B - VP$
- (ii) *Surpass* type – Type II head-marking $NP_A - VERB - CM - NP_B$
(where *CM* = comparative marker, NP_A = comparee, NP_B = NP acting as standard)

In Standard Mandarin, the language generally used as the representative for Sinitic, the comparative construction has the form, *Marker – Standard of comparison – Verb*, the ordering which crosslinguistically shows a strong correlation with OV, rather than VO languages (Dryer 1992: 91). In other words, it is disharmonic with VO ordering. Hence, a question relevant for the typology of the whole Sinitic taxon is to identify whether this pattern is in general shared across Chinese languages, or if it is restricted to just certain subgroups.

In fact, we find that in the case of the comparative construction, a large number of Sinitic languages do not use this strategy at all natively, but instead employ a *Surpass* or Type II comparative with the configuration *Verb – Marker – Standard of comparison*, a construction which aligns exceptionlessly with SVO languages in the sample established by Stassen (1985) of 110 languages. This pairing similarly shows a strong correlation in a larger sample used by Dryer (1992: 91–2) of 543 languages, grouped by genera and geographical area.

In terms of geography, the *Compare* comparative is essentially found ‘natively’ in northern China, while the *Surpass* comparative is found in central, southwestern and southern China. The areas using this pattern include many Hakka and Yue dialects, and even several subgroups of Mandarin. In these cases, the ordering is in perfect harmony with VO constituency.

After establishing the synchronic distribution of the *Surpass* construction, a diachronic sketch of the development of the comparative construction is presented in order to address the issue of whether the *Surpass* comparative is an internal

³ Chappell (this volume) and Li Lan (2003) discuss the further types of comparative constructions, distributed across Sinitic languages, which are, however, less common. These are the Topic, Transitive Action, Adverbial, Hybridized and Polarity types.

development within Chinese or is a case of contact-induced grammaticalization on the model of languages from surrounding families. Significant for this study, the relevant linguistic area of Mainland Southeast Asia also uses in the main the Type II *Surpass* strategy.

6.2 Comparative constructions: Definition

Comparative constructions involve a grading process and typically contain two NPs representing the ‘standard’ and the ‘comparee’, a morphological marker of the comparative and a stative predicate denoting the dimension or quality (Stassen 1985; Heine 1997). The comparative marking may also involve a second, degree marker (Heine 1997) which modifies the adjective (if not adverb), such as either the English suffix *-er* or ‘more+ADJ’, and similarly, French *plus+ADJ*, as well as *khah*⁴ 较 ‘more’ in many Min dialects of China.⁴

In using one object (A, the standard) as a benchmark against which to judge another (B, the comparee), different kinds of comparative constructions are possible depending on whether they express that an entity A has either (i) more, (ii) less, or (iii) an equal degree of the given dimension or quality. The first two belong to the comparative of inequality which is divided in this way into superiority and inferiority subtypes. The third is known as the equative and in English uses the form *as ADJ as B* (Huddleston and Pullum 2002: 1100).

More specifically, the comparative of superiority expresses that an entity A possesses a greater quantity or degree than an entity B along a scale for the given dimension, the latter being coded by its predicate. In English, it has the following structure and conforms to the definition of a *Particle* comparative:

NP_A[Comparee] – Stative predicate (ADJ + DEGR._{-er}) – Comparative marker_{than} – NP_B[Standard]

- (1) Carla is taller than Nicholas.

In contrast to this, the comparative of inferiority uses the adverb ‘less’ to express a smaller quantity or degree along a scale or continuum for the given quality or dimension:

NP_A[Comparee] – *less* – Stative predicate – Comparative marker_{than} – NP_B[Standard]

- (2) Richard is less tolerant about these matters than Erica.

6.2.1 Target construction

We single out the comparative of inequality as our target construction in this analysis of Sinitic comparatives. In general, all Chinese languages possess a comparative of

⁴ English, like Chinese languages, also allows a verbal predicate in the comparative, in which case it is the adverb which takes the degree marker: *She runs more quickly/faster than I.*

inequality of the superiority subtype, known as the *chàbījù* 差比句, in Standard Chinese. By way of contrast to English and other European languages, there is, however, no special construction for the comparative of inferiority, that is, of the ‘less than’ subtype. Thus, comparative constructions (*bǐjiào jù* 比较句) are divided into just two categories for Sinitic: the comparative of superiority or *chàbījù* 差比句 and the equative or *děngbījù* 等比句 which will not, however, be treated in this analysis.

6.2.2 Elements in a comparative of inequality

Up to five elements can be identified in comparative structures: These are the following, accompanied by the abbreviations we will use for them in this chapter:

CM	=	comparative marker
NP _A	=	noun phrase referring to the comparee
NP _B	=	noun phrase referring to the person or thing acting as the standard or benchmark
VP	=	verb phrase filled principally by an adjective or a verb as well as other predicative elements
DEGR	=	degree marker

These elements are exemplified in (3) for a basic comparative clause in Standard Mandarin (*pǔtōnghuà* 普通话) with a stative predicate formed by the adjective *gāo* ‘tall’.

- (3) Standard Mandarin: NP_A – CM – NP_B – VP

小王**比**小李高。

Xiǎo Wáng bǐ Xiǎo Lǐ gāo.⁵

NP_A CM NP_B tall

‘Xiao Wang is taller than Xiao Li.’

In comparative constructions, even if the dimension of comparison is predominantly expressed by an adjective, this is not always the case. Hence, in Sinitic languages, the VP may be filled by either a verb or a predicative adjective, the latter being classified by some linguists as a ‘quality’ or ‘stative’ verb. Thus, the label VP for ‘verb phrase’ is intended to include both verbs and adjectives in our syntactic formula for comparatives of inequality. For example, the transitive clause *pà làoshǔ* 怕老鼠 ‘to fear mice’ is possible in both Mandarin and Cantonese.

⁵ For the transcriptions of examples, we use *pīnyīn* romanization for Standard Mandarin and the Yale system for Hong Kong Cantonese. When quoting examples from published sources, the transcriptions have been faithfully reproduced in either IPA, or in a romanization with widespread currency such as that used for Hmong.

- (4) Standard Chinese (Mandarin):

我比你更怕老鼠。

Wǒ bǐ nǐ gèng pà lǎoshǔ.

1SG CM 2SG DEGR fear mice

‘I’m even more afraid of mice than you.’

- (5) Cantonese, Yue: Hong Kong (香港粤语)

我怕老鼠多过你。

ngóh pa louhsyú dō gwo néih.

1SG fear mice DEGR CM 2SG

‘I’m even more afraid of mice than you.’

6.2.3 Crosslinguistic research: Strategies and cognitive event schemas for comparatives

Crosslinguistic research on the major comparative strategies or cognitive schemas shows the following main categories and, importantly, that variation is not without certain limits. Stassen (1985, 2005, 2011) proposes six different categories according to structural criteria, while Heine (1997) puts forward eight separate categories which he defines cognitively (see Table 6.1). For this reason, we have adopted Heine’s framework (1997) for the schemas described in this analysis. However, in §6.3.2, we propose the existence of an additional cognitive schema called ‘Compare’, which represents the most common type in Sinitic, but is not accounted for by those given in Table 6.1 (for which see Table 6.2).

The first three types of comparatives in both the approaches of Stassen and Heine involve case-marking or the use of adpositions as morphological markers. These are dependent-marked comparative constructions which use the separative or ablative ‘from’, the allative ‘to’, or the locative ‘at’ strategies. The fourth type involves a transitive structure and verbs with the meaning of ‘exceed’ or ‘surpass’. The fifth type represents a complex sentence structure of the conjoined or coordinate type

TABLE 6.1. Comparison of analytic approaches for comparative schemas of inequality

	Stassen 1985, 2005	Heine 1997
1 :	Separative	Source ‘from’
2 :	Allative	Goal ‘to’
3 :	Locative	Location ‘at’
4 :	Exceed	Action ‘surpass, defeat’
5 :	Conjoined	Polarity ‘X is A, Y is not A’
6 :	Particle	Sequence e.g. Germanic ‘than’
7 :	–	Similarity ‘as, like’
8 :	–	Topic ‘X and Y, Y is A’

with polarity semantics. In the sixth type, conjunctions such as ‘than’ and its equivalent in many European languages, if not other kinds of particles, are used as markers of the comparative (see Example (1)). The seventh type involves the cognitive schema of similarity, while the eighth represents a topicalization strategy.

In the contemporary situation, Sinitic languages make use of at least four of these cognitive schemas: *Transitive Action* (our Type II), *Polarity*, *Topic*, and also the *Compare* schema (Type I) to be introduced in §6.3.2 (for more details, see the chapter by Chappell, this volume). Archaic Chinese also made use of the *Location* schema, and Medieval Chinese, the *Similarity* schema, both discussed in §6.4. Before proceeding into the main analysis, we provide an example of the unusual topicalization strategy from the Tiantai dialect 天台, a Wu dialect of southern Zhejiang province (see Zhao Jinming 2002b).

- (6) Tiantai,Wu: Zhejiang (浙 江天台吴语)

小王是小李高。

XIĀO WÁNG SHÌ XIĀO LÍ GĀO.⁶

Xiao Wang be Xiao Li tall

‘Xiao Li is taller than Xiao Wang.’ (literally: As for Xiao Wang, Xiao Li is taller.)

6.3 Comparative constructions of inequality in Sinitic languages

As mentioned in the introduction, two types of comparative construction predominate synchronically in Sinitic languages. These are the *Compare* type and the *Surpass* type. Type I, *Compare*, is dependent-marked whereas Type II, *Surpass*, is head-marked.

Note that while the lexical source and forms for the comparative markers may vary widely in Sinitic languages, the structures and associated cognitive schemas remain essentially the same.

6.3.1 ‘Compare’ as a comparative marker in Sinitic

The Type I comparative in Sinitic languages refers to the *Compare* schema which has the syntactic configuration: NP_A – CM – NP_B – VP. As mentioned above, the label ‘VP’

TABLE 6.2. Types versus schemas used in the history of the Chinese language for the comparative of inequality

Type I: <i>Compare</i> NP _A – CM – NP _B – VP	Type II: <i>Surpass</i> NP _A – VERB – CM – NP _B
<i>Compare</i> schema (with BÌ 比)	<i>Locative</i> schema (with YÙ 于) <i>Similarity</i> schema (with RÚ 如, RUÒ 若, sì 似) <i>Surpass</i> schema (with GUÒ 过, SHÈNG 胜)

⁶ We use small capitals for the Mandarin *pīnyīn* romanization, here, and wherever the original article does not provide any kind of phonetic transcription.

typically represents an adjective or stative verb that codes the dimension in question for the operation of comparison. Synchronously, this construction is the most common one to be found in Sinitic languages (for details, see the survey in Li Lan 2003).

(7) Type I: *Compare*

NP_A-CM-NP_B-VP Standard Mandarin

他**比**我高。

Tā_A bǐ wǒ_B gāo_{VP.}
3SG CM 1SG tall

‘S/He is taller than me.’

(8) Standard Mandarin

我去**比**你去方便。

Wǒ_A qù bǐ nǐ_B qù fāngbiàn_{VP.}
1SG go CM 2SG go convenient

‘It’s more convenient for me to go than you.’

The main marker for the Type I *Compare* comparative is the preposition **bì** 比 which is derived from the verb ‘to compare’ as used in a serial verb construction. Its use is attested from Early Medieval Chinese, that is, from third–fourth centuries AD. As **bì** 比 ‘compared to’ cannot, strictly speaking, be treated as a *Particle (Sequence)* or a *Similarity* comparative, we have decided to propose a ninth cognitive schema and thereby to adopt the label of ‘*Compare* comparative’. Specifically, the comparative marker **bì** 比 cannot be functionally equated with particles or conjunctions, such as English *than*, French *que*, or German *als*, found in European comparatives, nor does it have the meaning of ‘be similar to’ in these constructions, contrary to the classification given by Ansaldi (1999), as a *Similarity* comparative.

Furthermore, **bì** 比 cannot be classified as belonging to any of allative, locative, or source categories, even though it involves dependent-marking. Finally, it certainly does not mean ‘exceed’, *pace* Stassen (1985, 2005, 2011) who erroneously describes Mandarin **bì** 比 as an ‘exceed’ comparative (see Feature 121 in the World Atlas of Language Structures – WALS), as also does Ultan (1972).

This is the standard situation for Northern Chinese. As earlier stated, in this Type I, with the comparative marker in a preverbal position, the source and forms of the comparative marker may vary widely in different Northern Sinitic languages or dialects. The following examples, with the comparative markers **GĀN** 超 ‘catch up, overtake’, and the compound form **GĒNG-BÌ** 更 比 are reproduced from Li Lan (2003: 217).⁷

⁷ Li Lan (2003: 217) also gives examples of Type I from Shandong Jiaoliao Mandarin where the marker is apparently either **BĀ** 把 (which is the preverbal direct object marker in Standard Chinese), or **BÉI** 被 (a passive marker). These, however, are most likely cases of Chinese characters being used for

- (9) Tongxin, Ningxia: Lanyin Mandarin (宁夏同心兰银官话)
今儿赶昨儿强多了。

JÍNR GĂN ZUÓR QIÁNG DUŌ LE.
today CM yesterday good_{<strong} DEGR PRT
'Today is much better than yesterday.'

- (10) Dafang, Guizhou: Southwestern Mandarin (贵州大方 西南官话)
我更比你高。

WǑ GÈNG-BÌ NǏ GĀO.
1SG CM 2SG tall
'I'm taller than you.'

Other dependent markers in use are BÌNG 并, GĒN 跟, Bō 拨, ZHÀO 照, PÁNG 傍 etc. for which the distribution is given in Map 99 on comparative markers in Cao Zhiyun et al. (2008, *Grammar volume*), noting that the semantic accuracy of these sources is not guaranteed in many of these cases.

6.3.2 'Surpass' as a comparative marker in Sinitic

The second structural type, Type II, is commonly represented by the *Surpass* schema in Sinitic languages, because the comparative marker has its origin in a verb meaning 'surpass', 'exceed' or 'defeat', 'win' (see § 6.4.2). This is the type called an 'Action schema' in Heine (1997) due to its transitivity. In fact, the verb GUÒ 過 (过) 'to cross, to surpass' frequently turns out to be the source of the comparative marker for Type II in many dialect groups.

- (11) Type II: Surpass

NP_A – VERB – CM – NP_B
Hong Kong Cantonese (香港粤语)
佢 高 過 我。
Kéuih gōu gwo ngóh.
3SG tall CM 1SG
'S/He is taller than me.'

their homophone values and not the original source morphemes (the IPA values are not given in the original source).

- (a) Daye, Gan: Hubei (湖北大冶赣语)
我把渠长。

WǑ BĀ QÚ CHÁNG.
1SG CM 3SG tall
'I am taller than him.'

- (b) Yishui, Jiaoliao Mandarin: Shandong (沂水, 山东中部 胶辽官话)
你被他大。

NǏ BĒI TĀ DÀ.
2SG CM 3SG old
'You are older than him.' (Qian Zengyi et al. 2001: 293)

142 *Hilary M. Chappell & Alain Peyraube*

- (12) Hong Kong Cantonese (香港粤语)
我去方便過你去。
ngóh heui fòngbihn gwo néih heui.
1SG go convenient CM 2SG go
'It's more convenient for me to go than you.'

Comparative markers other than GUÒ can also be found in this Type II *Surpass* construction: KĀ 嘬, Ā 啊, Q I 起, etc. (examples reproduced from Li Lan 2003: 217).

- (13) Yiyang dialect, Xiang: Hunan (湖南益阳湘语)
他高嚜你蛮多。
TĀ GĀO KĀ NĪ MĀN DUŌ.
3SG tall CM 2SG very much
'He is very much taller than you'.
- (14) Fuzhou dialect, Northeastern Min: Fujian (福建福州闽语)
我好啊汝。
WŌ HĀO Ā RŪ.
1SG good CM 2SG
'I am better than you'.
- (15) Jimo, Jiaoliao Mandarin : Shandong (山东即墨胶辽官话)
他高起你。
TĀ GĀO QI NĪ.
3SG tall CM 2SG
'He's taller than you'. (Qian Zengyi et al. 2001: 292)

6.4 Distribution of comparative types in Sinitic languages

For Sinitic languages, we take the following ten subgroups of this branch of the Sino-Tibetan family as a basis for discussion of the distribution of the two main types of comparative strategies: Mandarin or Northern Chinese 北方話, Jin 晉, Xiang 湘, Gan 贛, Hui 徽, Wu 吳, Min 閩, Kejia 客家, Yue 粵, and Pinghua 平話. The most prominent language in the Sinitic taxon is Mandarin or 'Northern Chinese', which incorporates eight dialect subgroups and includes the standard official language known as *pǔtōnghuà* 普通話, literally 'the common language' (see §2.1 in Chappell, this volume, for more details on classification issues).

6.4.1 Sinitic Type I: Compare comparatives in areal perspective

The Type I *Compare* structure is used almost exclusively as the comparative in 5/8 subgroups of the vast Mandarin supergroup of dialects, including the major Northern groups of Beijing 北京, Northeastern (Manchuria) 东北, Northwestern Lanyin 兰银, Central Plains Zhongyuan 中原 and Southern Jianghuai 江淮 (Li Lan 2003). It

is also the only strategy in the northern Jin dialects 晉語 (see also Chen and Li 1996 for data on 63 Mandarin dialects).

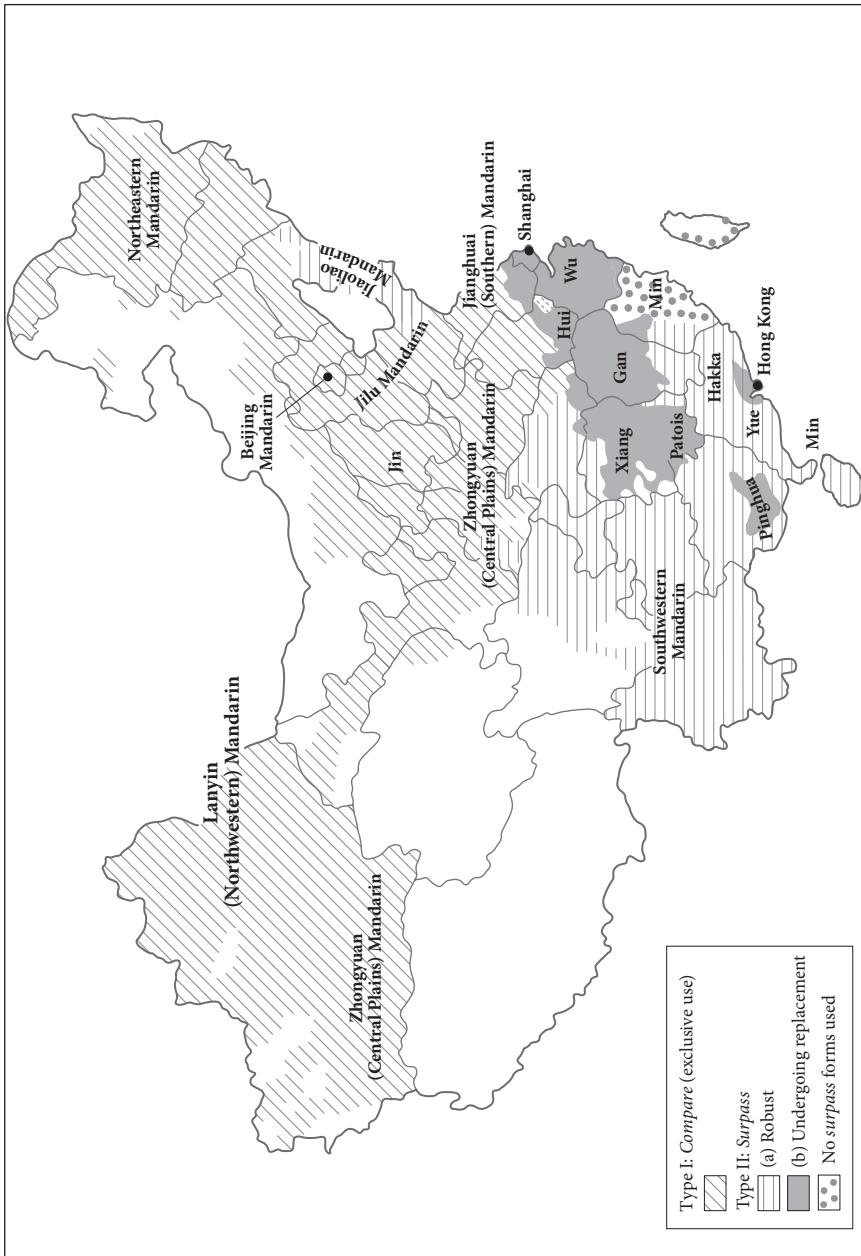
Type I *Compare* is not only represented by Mandarin and the Jin dialect groups found all over the North of China from Xinjiang and Gansu in the northwest across to Beijing and up to the northeast in Manchuria, but also in the contiguous central Hui dialects 徽 of Anhui province. This strongly suggests a distinct *Northern strategy* that was early remarked upon by Hashimoto (1976a, 1978) and more recently by Ansaldi (2010).

6.4.2 Distribution of Type II Surpass comparatives in Sinitic

The dominant position of Type I *Compare* notwithstanding, the distribution of the Type II comparative (*Surpass schema*) is much more widespread than has been previously supposed. In terms of geographical distribution, the southwestern and southern dialect groups of China use it, including particularly the Yue dialects (Cantonese) and the Hakka and Min (or Hokkien) dialects located in Guangdong province. Southwestern Mandarin (parts of Hubei, Sichuan, Guizhou, and Yunnan) and Mandarin dialects of Shandong (Jiaoliao, Jilu subgroups) also use the Type II *Surpass* structure. In the central, transitional zone, certain Xiang dialects in Southern Hunan, Gan dialects in Jiangxi, and many Wu dialects also make use of Type II *Surpass* alongside Type I *Compare* (see examples in §6.3.2). The reader is referred to Map 6.1 for the distribution of these two main types of comparatives.

Therefore, we do not agree with Wu Fuxiang (2010) who claims that Type II *Surpass* is very limited in distribution, being mainly located in the provinces of Guangdong, Guangxi, and Hainan, and also of course in Hong Kong. In other words, in Wu's view, Type II would be mainly limited to Cantonese and to Southern Sinitic languages and dialects closely in contact with Cantonese, from which they have borrowed this structure. We have seen that this is far from being the case, since Type II *Surpass* is the dominant form far to the north in the non-contiguous area of the Shandong peninsula where approximately two-thirds of the Jiaoliao Mandarin dialects employ it (see Qian Zengyi et al. 2001: 291). In another peripheral area, isolated dialects in the west of Shaanxi province also make use of *Surpass* (Map 98, *The Comparative*, in Cao Zhiyun et al. 2008, *Grammar* volume).

In terms of source lexical fields, Type II markers vary considerably, including GUÒ 过 'pass', Qǐ 起 'rise', QU 去 'go', SHÈNG 胜 'win, defeat', sì 似 'similar to'... (See several examples in §6.3.2, and, in particular, Zhang Cheng (2004), Zhao Jinming (2002a, 2002b), and Li Lan (2003)). Despite the different forms for the comparative marker, all make use of the transitive action cognitive schema, as in the Hakka example in (16).



MAP 6.1. Comparative constructions in Sinitic languages: The two main types

The comparative construction in Sinitic languages 145

- (16) Hakka *Surpass* comparative, Bao'an: Guangdong (广东宝安 客家话)

口惹隻山高過个隻山。

nya³ tšak⁶ san¹ kau¹ ko⁴ kai⁴ tšak⁶ (san¹).
this CL mountain_A high CM that CL (mountain_B)

‘This mountain is higher than that one.’

(Chappell and Lamarre 2005: 72)

In Southwestern Mandarin (e.g. Lipu 荔浦 and Liuzhou 柳州, both located in the Guangxi Autonomous Region), we find the following examples which, like many Hakka and Cantonese Yue dialects, use the comparative marker GUÒ 過 (过).

- (17) Lipu, Southwestern Mandarin: Guangdong (广东荔浦西南官话)

他大过我。

TĀ DÀ GUÒ WŌ.

3SG_A big CM 1SG_B

‘He is bigger than me.’ (Wu Hezhong 1998)

- (18) Liuzhou, Southwestern Mandarin: Guangdong (广东柳州西南官话)

坐火车快过坐汽车。

ZUÒ HUŌCHĒ KUÀI GUÒ ZUÒ QÌCHĒ.

sit train_A fast CM sit car_B

‘It’s faster to go by train than by car.’ (Li and Huang 1995)

There are further differences to be noted that involve syntactic constraints. Unlike Standard Mandarin and Mandarin dialect groups such as the Southwestern, exemplified in (18), in Hong Kong Cantonese, the main verb does not need to be repeated in the case of a comparison of complex, non-identical predicates; for example:

- (19) Hong Kong Cantonese (香港粤语)

我食肉多過魚。

ngóh sihk yuhk dō gwo yú.

1SG eat meat more CM fish

‘I eat more meat than fish.’

Compare this example with its equivalent in Standard Mandarin, where the main verb needs to be repeated so that the two full clauses in question can undergo comparison:⁸

- (20) Standard Mandarin

我吃肉比⁹我吃鱼多。

Wǒ chī ròu bǐ nǐ chī yú duō.

1SG eat meat CM 2SG eat fish more

‘I eat more meat than you eat fish.’⁹

⁸ Some speakers find this acceptable but not fully natural and preferred the strategy:

我吃的肉比你吃的魚多。

Wǒ chī de ròu bǐ nǐ chī de yú duō.

1SG eat MOD meat CM 2SG eat MOD fish more

‘I eat more meat than you eat fish.’

146 Hilary M. Chappell & Alain Peyraube

It is not possible to use the more succinct Cantonese *Surpass* structure:

- (21) *我吃肉**比**魚多。
*Wǒ chī ròu **bǐ** yú duō.
1SG eat meat CM fish more

This lack of necessity to repeat the verb is certainly special syntactic feature of Cantonese comparatives (Yue-Hashimoto 1997) but it is not unique to the Yue dialects since it is possible too in the Hui'an dialect of Southern Min, although in a different structural type, the zero-marked comparative (see §10.5 in W. Chen, this volume).

In spite of the fact that the geographical coverage of Type II *Surpass* comparatives is broader than initially believed on the basis of earlier studies, nearly all the dialect groups other than Mandarin may in fact *additionally* use the Type I *Compare* comparative with *bì* 比. In most cases this is, however, a borrowed form. Quite clearly, the *bì* 比 construction is steadily encroaching on the Type II *Surpass* comparative, particularly in the central transitional zone of China for Xiang, Gan, Hakka, and Wu dialects (Li Lan 2003; Cao Zhiyun et al. 2008, *Grammar* volume, Maps 98 and 99).

To be more specific, in these areas, we find either the coexistence of two forms (Type I and Type II) or cases of hybridization of the two, as in Southern Min and Hakka (see Chappell 2001b, this volume; W. Chen, this volume), or even of functional specialization. The latter situation is discussed in Chang Song-hing and Kwok Bit-chee (2005: 232–8) who show, following Yue-Hashimoto (1997), that not all the constructions with GUÒ 過(过) may be replaced by *bì* 比 in Cantonese, and vice-versa.

6.4.3 Geographical areas for Type II Surpass in the world's languages

To place Sinitic languages in a crosslinguistic perspective, it is not surprising to find that the Type II *Surpass* schema common in Southern Sinitic languages is equally predominant in Southeast Asia. Further afield, Bantu languages as well as Afro-Asiatic languages of sub-Saharan Africa and the Middle East make use of this strategy (Stassen 1985; Heine 1997; Ansaldi 1999, 2010).

Some examples of Southeast Asian languages using the Type II *Surpass* comparative schema are Lao (Tai-Kadai), Vietnamese and Khmer (Austroasiatic), Hmong (Hmong-Mien), and Burmese (Tibeto-Burman). The following examples from Lao, Khmer, and Hmong all use markers that are derived from verbs meaning ‘to exceed’ or ‘to surpass’, according to the sources consulted.

- (22) Lao (Tai-Kadai)
qaa⁴ khòoj⁵ suung³ kuaø qaa⁴ caw⁴.
eBr 1SG.P tall MORE.THAN eBr 2SG.P
'My brother is taller than your brother.' (Enfield 2007: 249; eBr = elder brother,
2SG.P = second person singular pronoun)

The comparative construction in Sinitic languages 147

- (23) Khmer (Austroasiatic)
reut tae awn **ciang** kee:aeng (tiat).
get weak **exceed** everybody:else (other)
'[I] keep getting weaker than all the others.' (Haiman 2011: 190)

- (24) White Hmong (Hmong-Mien)
Tus noog no loj dua tus noog ntawd.
clf bird this big **comp.** clf bird that
'This bird is bigger than that bird.' (Jaisser et al., 1995: 182; **clf** = classifier)

Similarly to the case for many Southern and Central Sinitic languages, Wu Fuxiang (2013) has noticed that in certain, though not all, Tai-Kadai, Hmong-Mien, Austroasiatic, and even Austronesian languages spoken in China, the two types, I and II, co-exist. He puts forward the hypothesis that the Type I *Compare* comparative structure in these languages has been borrowed from Northern Chinese. Some examples of languages in China where the Type I *Compare* co-exists with Type II *Surpass* are given here.

Pana (Bānà) 巴那语 (Hmong-Mien, Southwestern Hunan)

- (25) Type II *Surpass*: NP_A–VERB–CM–NP_B
Le²⁴ te³¹ ljou²⁴ kua³⁵ le²⁴ tçi⁴⁴ u²⁴ tçø⁵⁵.
CL eBr old CM CL yBr two year
'Our elder brother is two years older than our younger brother.'
- (26) Type I *Compare*: NP_A–CM–NP_B–VP
va²² za⁴⁴ pi⁴⁴ ni²² za⁴⁴ nɛ̃³⁵.
1SG MOD CM 3SG MOD good
'Mine is better than his.'

Bolyu (Palyu) 傣语 (Austroasiatic, Guangxi and Yunnan)

- (27) Type II *Surpass*: NP_A–VERB–CM–NP_B
tçə² ciu⁵ ?ɔ⁵tsɔ³ kəŋ⁵ nja:n² tçə² zu¹ ta:i⁵.
time harvest cereal cold CM time plant corn
'Autumn is colder than spring.'
- (28) Type I *Compare*: NP_A–CM–NP_B–VP
tçə² ciu⁵ ?ɔ⁵tsɔ³ pi³ tçə² zu¹ ta:i⁵ kəŋ⁵.
time harvest cereal CM time plant corn cold
'Autumn is colder than spring.'

Huihui 回辉语 (Chamic, Austronesian, Hainan Island)

- (29) Type II *Surpass*: NP_A–VERB–CM–NP_B
lu⁴³ kau³³ pion³² la:u³² lu⁴³ ha³³.
bowl 1SG big CM bowl 2SG
'My bowl is bigger than your bowl.'

148 Hilary M. Chappell & Alain Peyraube

- (30) Type I *Compare*: NP_A–CM–NP_B–VP
kau³³ pi¹¹ ha³³ tsat²⁴ tso³³ kiɔ³³ sun³³.
1SG CM 2SG small three inch
'I am three inches shorter than you.'

(All the examples above are reproduced from Wu (2013); our glossing and translations into English).

Hence, the same language contact phenomenon has indifferently affected both Central and Southern Sinitic languages as well as non-Sinitic languages located in China. The 'native' Type II *Surpass* comparative is used alongside the borrowed Type I *Compare* comparative. In the next section, we consider the diachrony of comparative constructions in Chinese to demonstrate that, as far as the earliest written records reveal, syntactically speaking, Chinese began as a Type II comparative language.

6.5 Historical sketch of the comparative constructions

The comparative construction of inequality found in both Archaic (eleventh–third centuries BC) and Pre-Medieval Chinese (second century BC–second century AD) clearly belongs to Type II and has the following structure using the comparative marker YÚ 所: NP_A–VERB–CM (YÚ 所) –NP_B. Two examples are provided.

- (31) Comparative of inequality with Type II structure and *Location* schema as source
季氏富于周公。 (论语：先进)
Jì shí fù yú Zhōu gōng.
Ji family rich CM Zhou Duke
'The Ji family was richer than the Duke of Zhou.' (*Analects*, 5th c. BC)
- (32) 一少于二。 (墨子：经下 41)
yī shǎo yú èr.
one less CM two
'One is less than two.' (*Mòzǐ*, 4th c. BC)

This Type II structure is associated with a *Location* schema, as the main use of YÚ 所 is as a locative preposition 'at, to' and originally a verb 'to go', according to Guo Xiliang (1997). This is why we classify it as the Type II – *Location* schema (see also Table 6.2). Significant in providing the preconditions for a later diachronic change, the Type II syntactic structure also codes the comparative construction of equality: however, it involves a different set of comparative markers. These are: RÚ 如, RUÒ 若 or SÌ 似, all meaning 'to be like, to be similar'.

The comparative construction in Sinitic languages 149

- (33) Comparative of equality using a Type II structure with *Similarity* verbs

NP_A – VERB – CM (RÚ 如) – NP_B

猛如虎很**如**羊贪**如**狼。 (史记 : 项羽本纪)

měng rú hǔ
powerful CM tiger
hěn rú yáng tān rú láng.
ferocious CM ram greedy CM wolf

‘(Be) as powerful as a tiger, as ferocious as a ram, and as greedy as a wolf.’ (*Shi jì*, 1st c. BC)

- (34) 君子之交淡若水,

Jūnzi zhī jiāo dàn ruò shuǐ
gentlemen MOD friendship insipid CM water
小人之交甘若醴。 (庄子: 山木)

xiǎo rén zhī jiāo gān ruò lǐ.
small people MOD friendship rich CM sweet-wine

‘Friendship between gentlemen is as insipid as water, friendship between small-minded people is as rich as sweet wine.’ (*Zhuāngzǐ*, 4th c. BC)

As for the modern Type I *Compare* structure, in Late Archaic Chinese (fifth–second centuries BC) BÌ 比 can only be used as a verb meaning ‘to compare’ in simple S-V-O sentences, while it still does not occur in serial verb constructions.

The same form as the one found in (31) (Type II structure and *Location* schema) continues to be used in Medieval Chinese during the Six Dynasties (third–sixth centuries) and Tang periods (seventh–tenth centuries). Peyraube (1989b) claims that YÚ 于 ‘at’ remains the most common comparative marker, but that one can also find GUÒ 過(过) < ‘surpass’ used as a real grammatical morpheme marking the comparative (see also Zhang Cheng 2005):⁹

- (35) Type II structure associated with both *Location* and *Surpass* schemas

贫于杨子两三倍老过荣公六七年。 (白居易诗)

pín yú Yángzi liǎng sān bèi
poor CM Yangzi two three times
lǎo guò Róng gōng liù qí nián.
old CM Rong Master six seven year

‘Two or three times poorer than Master Yang, six or seven years older than Mr. Rong.’ (*Bái Jūyì shī*, 9th c.)

⁹ Zhao Jinming (2004a) cites a still earlier example of GUÒ from Archaic Chinese, but the example is controversial.

However, what is more interesting is that, beginning in the Late Medieval period (seventh–twelfth centuries), other comparative markers can also be found for expressing the comparative of inequality in the Type II structure: these are *rú* 如 < ‘be like’ and *sì* 似 also < ‘be like’, that were in fact earlier used, in Late Archaic Chinese, to express the comparative of equality (or equative construction) (see example 33). In the following example, we find the comparative markers of inequality *yú* 于 and *sì* 似 occurring in parallel clauses.

- (36) Type II structure associated with both *Location* and *Similarity* schemas

本寺远于日 新诗高似云。 (姚合诗)

běn	sì	yuǎn	yú	rì
this	temple	far	CM	sun
xīn	shī	gāo	sì	yún.
new	poem	high	CM	cloud

‘This temple is farther away than the sun, the new poems are higher than the clouds.’ (*Yáo Hé shī*, 9th c.)

Hence, we clearly have a Type II structure in the second clause of (36) that uses a *Similarity* rather than a *Location* schema to express the comparative of inequality. Hence, we will call it the ‘Type II *Similarity* schema’. This signals an important diachronic change in progress whereby *sì* 似 begins to no longer exclusively code the comparative of equality, as it did earlier in its history. In fact, during two or three centuries (ninth–twelfth centuries), the Type II *Similarity* construction ($NP_A - VERB - CM - NP_B$) with the comparative markers *rú* 如 and *sì* 似 was ambiguous, expressing both a comparative of inequality and a comparative of equality.

In contrast to these two comparative markers, *bǐ* 比, which was essentially used as a verb ‘to compare’ in simple sentences in Late Archaic Chinese, began to be used as V_1 in ‘ $V_1 \dots V_2$ ’ serial verb constructions under the Six Dynasties period (third–sixth centuries) before being grammaticalized into a comparative marker probably by the end of the Tang dynasty (ca. ninth–tenth centuries), according to Ōta (1958), Peyraube (1989b), and also Zhang Cheng (2004). Zhang nevertheless noticed that the structure only appeared in Tang poetry and not in the prose texts of the *Dūnhuáng Biànwén* 敦煌变文 (Buddhist Transformation texts) of the same period. This led to the appearance in the Late Tang period (ninth–tenth centuries) of a new comparative structure, belonging to Type I.

- (37) Comparative of inequality using Type I structure and the *Compare* schema:

$NP_A - CM - NP_B - VP$, comparative marker = *bǐ* 比 :

若比李三又自胜…(白居易诗)

ruò	bǐ	Lǐ Sān	yòu	zì	shèng.
if	CM	Li San	still	myself	better

‘If I am still better than Li San...’ (poem of *Bái Jūyì*, 9th c.)

Starting in the Song-Yuan period (tenth–fourteenth centuries), another Type I construction ($NP_A-CM_{(RÚ\text{如},\ sì\text{似})}-NP_B-VP$) appeared, which used the *Similarity* schema for expressing the comparative of equality, with the comparative markers RÚ 如 and sì 似 now found in preverbal position. This development arose probably because Type II ($NP_A-VERB-CM_{(RÚ\text{如},\ sì\text{似})}-NP_B$) had become confusingly ambiguous, expressing both meanings of inequality and equality.

- (38) Comparative of equality using Type I structure and *Similarity* schema:

$NP_A-CM-NP_B-VP$, comparative marker = RÚ 如:

脸如红杏鲜妍。 (小孙屠)

liǎn rú hóng xìng xiān yán.
face CM red apricot fresh beautiful

‘(Her) face is as fresh and beautiful as a red apricot.’ (*Xiǎo Sūn tú*, 14th c.)

During the Yuan dynasty (thirteenth–fourteenth centuries), the Type I *Compare* schema and Type II *Similarity* schema for expressing the comparative of inequality nonetheless continued to co-exist, as shown by the following two contemporaneous examples.

- (39) Comparative of inequality using Type I *Compare* schema: $NP_A-CM-NP_B-VP$
(这桥)比在前十分好。 (老乞大)

(zhè qiáo) bǐ zài qián shífèn hǎo.
(this bridge) CM at before very good

‘(This bridge) is much better than before.’ (*Lǎo Qīdà*, 14th c.)

- (40) Comparative of inequality using Type II *Similarity* schema:

$NP_A-VERB-CM_{(RÚ\text{如},\ sì\text{似})}-NP_B$

这但轻如你底。 (任风子)

zhè dàn qīng rú nǐ dǐ.
this load light CM 2SG MOD

‘This load is lighter than yours.’ (*Rèn fēngzi*, 14th c.)

It is only at the end of the Yuan dynasty (fourteenth century) that we see a decrease in the Type II comparatives of inequality using the markers YÚ 和 and RÚ/sì (that is,

TABLE 6.3. Occurrence of three comparative structures

Period	Centuries	II : YÚ 和	II : RÚ/sì 如/似	I : BÌ 比
Six Dynasties–Tang	3rd–9th c.	68.1%	14.4%	17.5%
Song	10 th –13th c.	49.1	20.7	30.2
Jin–Yuan	13 th –14th c.	3.4	61	35.6
Ming	14 th –17th c.	1.5	35.1	63.4
Qing	17 th –20th c.	1.3	6.3	92.4

Location and *Similarity* schemas, respectively) in favour of Type I with *bì* (*Compare* schema). This diachronic change is reflected in the data displayed in Table 6.3 (adapted from Huang 1992: 221).

During the 17 centuries represented by the data in Table 6.3, the use of the *Location* schema dwindles to just 1.3%; the two markers of the *Similarity* schema, which saw their heyday in the Jin-Yuan periods, decrease to 6.3%, while the *Compare* schema increases to 92.4%.

We still have not discussed the historical development of Type II comparatives with *GUÒ* 過(过), using the *Surpass* schema, which we have pointed out, is extremely common in Sinitic languages of southern and central China. Nonetheless, only one real example has been found in a Tang dynasty poem (see (35)).

Zhang Cheng (2005) agrees with Peyraube (1989b) that this example (35) should be considered as a comparative construction which makes use of a Type II *Surpass* schema for a comparison of inequality in its second clause. She also noticed that examples of this type involving the comparative marker *GUÒ* 過(过) are very rare in both the Tang and later periods. With respect to the period of the Tang dynasty, she found only 13 examples (12 in the Tang poems and only one in the Buddhist *Dunhuang Bianwen* prose texts). Significantly, all these uses of *GUÒ*, except for the one example in (35) can be considered as full verbs meaning *chāoguò* 超过 ‘to surpass, to exceed’ or *shèngguò* 勝过 ‘to win, to defeat’.

She concludes, and we agree with her, that *GUÒ* is not a morphological marker of the comparative during the Tang period, and that in fact it never grammaticalized from a verb into a comparative morpheme in Medieval or even in Modern Chinese (thirteenth–eighteenth centuries). Significantly, the documents from these periods reflect in the main varieties of Northern Chinese. There are simply no examples of *GUÒ* found in the texts from these periods where *GUÒ* is used as a real comparative marker.

This is also the viewpoint of Wei Pei-chuan (2007) and Wu Fuxiang (2010), who similarly confirm that the few examples of Type II *Surpass* which could be interpreted as a real comparative construction are found only in the Tang poems, and never in prose texts of the same period.

6.6 Hypotheses concerning the *Surpass* schema

We will consider two possible explanations as to the development of *Surpass* comparatives across Central, Southern and Southwestern China in relation to their distribution in Southeast Asia, an area which includes the related Tibeto-Burman languages, as well as those from the unrelated Tai-Kadai (Kra-Dai), Hmong-Mien, Austroasiatic and Austronesian (see §6.4.3). The two hypotheses are the following:

- (i) It is an independent language-internal development involving a native process of grammaticalization
- (ii) It is a contact-induced change resulting in diffusion of the *Surpass* schema—a structural feature which helps define the linguistic area of Southeast Asia.

6.6.1 Hypothesis (i)

Suppose we adopt the following hypothesis:

There is a direct internal derivation of the Surpass model either from Archaic or Medieval Chinese into the modern Sinitic languages.

There is a problem, however, with this first hypothesis: the *Surpass* Type (Type II) in Southern Sinitic languages with GUÒ as the comparative marker cannot be directly internally derived from Medieval Chinese (third–thirteenth centuries), and nor, consequently, from Archaic Chinese (eleventh century BC–third century BC), as the transitive *Surpass* construction is clearly not attested as a model in this period, as concluded in §6.5, which outlines the historical development of comparatives.

The Type II structure which uses locative prepositions (YÙ 于) or *Similarity* verbs (RÚ 如, SI 似) as markers in postverbal position is in fact clearly the more common structure in Medieval Chinese to code the comparative of inequality (as shown in Table 6.3).

6.6.1.1 A proposed scenario for guò GUÒ 過(过) meaning ‘to surpass’, or even SHÈNG 胜 ‘to defeat’, would have indeed been—for obvious semantic reasons—very good candidates to replace YÙ 于< LOC as the comparative marker of inequality in Type II structures. Moreover, *Surpass* is a model associated with a cognitive schema which is attested in many other language families in the world, external to the linguistic area of East and Southeast Asia (see Heine 1997; Heine and Kuteva 2002: 123–6). It did not however have sufficient time to grammaticalize and impose itself upon the *Similarity* comparatives using the markers RÚ 如 and SI 似 of Northern Chinese, since Type I *Compare* (with BÌ 比) had already arisen and was gradually replacing all the variants of Type II.

In contrast to this, in Southern China, where the Type I *Compare* structure had not yet been adopted, GUÒ 過(过) had ample time to grammaticalize into a comparative marker and to replace the other markers used in these Type II *Similarity* comparatives, namely RÚ 如 and SI 似.

6.6.2 Hypothesis (ii) on linguistic areas and contact-induced change

The *Surpass* comparative is one of the structural features (and cognitive schemas) which identifies a linguistic area comprising Tai-Kadai, Austroasiatic, Hmong-Mien, Tibeto-Burman, and many Sinitic languages (see Ansaldi 2010; Bisang 1992; Enfield 2003; Hashimoto 1976a; Matisoff 1991 *inter alia*). This second possible hypothesis would thus appear to be eminently reasonable from the point of view of language contact. If we adopt this explanation, is it possible to demonstrate, however, any such contact-induced change in order to establish the direction of borrowing and

diffusion? To be specific, and by taking only the Tai-Kadai languages as one example among the several different language families of Southeast Asia, we would need to decide, on the basis of adequate data and argumentation, whether we have a case of sinicization, for example, of the Tai-Kadai languages or of taïcization of the Southern Sinitic languages, as has been suggested by Bennet (1979).

Nothing allows us in fact to decide which is the source of the borrowing and which is the target. It could equally be the case that the Type II *Surpass* schema of the languages of Southeast Asia has been borrowed from Sinitic languages, rather than the reverse. This is moreover the hypothesis which both Ansaldi (2010) and Wu Fuxiang (2010) have put forward. In the absence of any historical documents which go back to the early Medieval period for either Southeast Asian languages or for Sinitic languages in Southern China (including at least Yue, Hakka, and South-western Mandarin), we remain in complete ignorance. Therefore, given these conditions, would it not be better to simply state that the *Surpass* comparative type is common to different language families of Southeast Asia as well as to Southern and Central Sinitic languages? This fact can then aptly serve as another piece of evidence enabling us to identify a linguistic area, without needing, for the time being, to locate the source, nor the associated direction of contact-induced change for this feature.

6.7 Conclusion

In our view, until further data is uncovered on Southeast Asian languages, the more conservative hypothesis of an internal development is the only one which can be justified at present. Thus for Sinitic languages, the grammaticalization of *Surpass* class verbs into comparative markers is based on an entirely natural cognitive schema whereby a verb meaning ‘surpass’ grammaticalizes into a comparative marker of inequality, a process which happens independently and repeatedly in many different languages of the world. As far as the diachronic evidence is concerned, a case for contact-induced grammaticalization cannot be established, along the lines of the powerful model of Heine and Kuteva (2005).

This grammaticalization process (Verb ‘to surpass’ > Comparative morpheme) occurred in only some Sinitic languages—being those located mostly in southern and central China but also in the Shandong peninsula in the northeast. The word GUÒ ‘to surpass’ was, in fact, a better-adapted candidate for a comparative morpheme than the *Similarity* verbs RÚ, SÌ, etc. already attested in Medieval Chinese. It also was able to develop in these regions remote from the central area of the empire, where Sinitic languages were not directly in contact with the language of the imperial court, nor consequently with the development of a Type I comparative construction with BÌ (*Compare* schema) that became the standard form in Northern Sinitic, beginning from the Yuan period (thirteenth–fourteenth centuries).

Part III

Individual Studies of Linguistic Micro-Areas in China

OUP UNCORRECTED PROOF – REVISES, 27/10/2015, SP1

7

Language contact in Nanning: Nanning Pinghua and Nanning Cantonese

HILÁRIO DE SOUSA

7.1 Introduction

Five languages/dialects are traditionally spoken in the Nanning area: the Sinitic languages of Nanning Pinghua, Nanning Cantonese, and Old Nanning Mandarin, and the indigenous Tai languages of Northern Zhuang and Southern Zhuang. They have all influenced each other in various ways, and equally interestingly, they have remained distinct in a number of ways. In this analysis, we will discuss aspects of this complex language contact situation from the viewpoint of Nanning Pinghua and Nanning Cantonese. A selection of similarities and differences in their phonology, lexicon, and grammar will be presented during the course of this discussion.

Curiously, Nanning Cantonese, which has been spoken in the area for about 150 years, in some respects resembles the indigenous Zhuang languages more in its phonology and grammar than does Nanning Pinghua, which has been spoken in the area for a millennium. In the last section of the analysis, we will discuss some of the possible sociolinguistic factors that might have caused this unusual outcome for language contact.

7.1.1 Linguistic background

Nanning (南寧 *Nánlíng*) is the capital of Guangxi Zhuang Autonomous Region (廣西壯族自治區 *Guāngxī Zhuàngzú Zìzhìqū*) in Southern China. Situated on the Yong River (邕江 *Yōngjiāng*), a tributary of the Western Branch of the Pearl River, Nanning is the largest city upriver from the Pearl River Delta. Historically, the Nanning region was of great military importance to China due to its proximity to the Vietnamese border, and also its location within the territory of the indigenous Tai-speaking Zhuang people. Nowadays, this geographical importance has

transformed into economic importance due to Nanning's proximity to Vietnam and the Tai-speaking nations in ASEAN. This geographically strategic city has long attracted Han immigrants from various linguistic backgrounds. As a result, the linguistic environment in Nanning is complex, with a multiplicity of Sinitic languages spoken alongside the Tai languages indigenous to the area.

The distribution of languages in the Nanning area is somewhat concentric. The urban area is primarily Sinitic-speaking, and there are three Sinitic languages traditionally spoken in Nanning: Nanning Cantonese (南寧白話 *Nánníng Báihuà*) is primarily spoken in the city centre, and Nanning Pinghua (南寧平話 *Nánníng Pínghuà*) is primarily spoken in the surrounding suburbs and nearby rural areas. Within the city centre, there used to be an enclave of speakers of Old Nanning Mandarin (邕州官話 *Yōngzhōu Guānhuà*). However, most of the remaining speakers of this moribund language are now found in further-away villages/suburbs like 菠蘿嶺 *Bōluólǐng* (most of the original speakers of Old Nanning Mandarin left in Nanning city centre shifted to speaking Nanning Cantonese). Rural villages further away are mostly Tai-speaking, and the Zhuang dialects are commonly divided into Northern Zhuang and Southern Zhuang, each belonging to a different branch of the Tai language family (Northern Tai and Central Tai respectively). In addition to these five languages native to Nanning, nowadays there is also New Nanning Mandarin (南寧普通話 *Nánníng Pǔtōnghuà*, or 南普 *NánPǔ*), which is Nanning's version of modern Standard Mandarin, and these days the *lingua franca* of Nanning. All these languages and dialects in Nanning have mutually influenced each other in different ways.

In this analysis, we will discuss some aspects of this complex language contact situation from the viewpoint of Nanning Pinghua and Nanning Cantonese.¹ Many interesting aspects of this language contact situation can be seen when Nanning Pinghua and Nanning Cantonese are compared with each other, and also when they are compared with Zhuang, Standard Cantonese, and Standard Mandarin. As for Cantonese, Standard Cantonese spoken in the Pearl River Delta is nowadays spoken far away from the Tai languages. However, even Standard Cantonese has a strong Tai substratum (Ouyang Jueya 1989; Li Jingzhong 1994; Bauer 1996; Huang Yuanwei 1997; Li Jinfang 2002: 100–41), a witness to the existence of Tai-speaking people in Guangdong with whom later Chinese settlers from the north mingled.²

¹ Zhuang has, unsurprisingly, also been under the influence of Sinitic languages, but this is not discussed in this analysis. For examples of Sinitic influence on Zhuang see Wang Jun (1962), Dai Qingxia (1992), Qin Xiaohang (2004), Sybesma (2008), Zhao Jing (2008), amongst many others.

² Genetically, it is known that Cantonese people are more than 50% Northern Chinese on their male line of descent and primarily 'native' (Kra-Dai, Hmong Mien, or Austroasiatic) on their female line of descent (Wen et al. 2004; see also Gan et al. 2008 on the genetics of Northern Pinghua people, where the conclusion is that Northern Pinghua people are primarily 'native' on both the male and female lines of descent). In Wen et al.'s (2004) study of the makeup of Sinitic people in general, the 'Northern Chinese' Y haplogroups are (using the nomenclature at the time of publication) O₃ (M122) and O_{3e} (M134), whereas the 'Southern Natives' Y haplogroups are O₁* and O_{1b} (M119-C) and O_{2a1} (M95T).

Language contact in Nanning: Nanning Pinghua and Nanning Cantonese 159

After the arrival of the first groups of Cantonese speakers in the mid-nineteenth century, Nanning Cantonese subsequently evolved under strong influences from local Zhuang languages. As for Pinghua, Nanning Pinghua has been spoken in the Nanning area for at least one millennium. As to be expected, Nanning Pinghua is also influenced by Zhuang. Interestingly, however, with respect to many of its linguistic features, the longer-established Nanning Pinghua is somewhat less like the indigenous Zhuang than the ‘newly-established’ Nanning Cantonese. There are many examples of this, and in §7.2, §7.3, and §7.4 we will see such a selection. In the last section of this chapter, §7.5, we will discuss the possible social-linguistic factors for this less-usual outcome of language contact. In what follows, I give brief introductions to the languages and dialects of the Nanning area and other explanatory notes.

Cantonese is a member of the Yue Chinese subgroup of Sinitic (粵語). (We will discuss briefly the relationship between Yue and Pinghua below.) The first major wave of Cantonese people arrived in Nanning area about 150 years ago, around the time of the Opium Wars (mid-nineteenth century). The majority of them came from the Nanpanshun area (南海 Nánhǎi, 番禺 Pányú, and 順德 Shùndé) just south of Canton in the Pearl River Delta. In the early days of the Republic of China (the 1910s), the numerically dominant Sinitic languages in Nanning were still Nanning Pinghua and Old Nanning Mandarin (Zhou Benliang et al. 2006); it was in the 1930s or the 1940s that Nanning Cantonese became numerically dominant in the city centre, as more and more Cantonese immigrants from the Pearl River Delta settled in the Nanning area due to the Nationalist–Communist Civil Wars and the Second World War. Nanning Cantonese is largely mutually-intelligible with Standard Cantonese; their vocabularies are largely the same, their inventories of tones are the same, and segmentally they are not very different. When compared with Standard Cantonese, it is clear that Nanning Cantonese has been subject to Zhuang influence, to be discussed throughout this chapter. In this analysis, when ‘Cantonese’ is not further qualified, it refers to both Nanning Cantonese and Standard Cantonese.

Nanning Pinghua is a member of the Pinghua Chinese (平話) branch of Sinitic. There are various accents of Nanning Pinghua spoken in the different suburbs of Nanning; their phonological differences are obvious, but they are mutually

The ‘Northern Chinese’ mitochondrial (mt) haplogroups are A, C, D, G, M8a, Y, and Z, whereas the ‘Southern Natives’ mt haplogroups are B, F, R9a, R9b, and N9a. In Gan et al.’s (2008) study of the genetics of Northern Pinghua people, the ‘Han Chinese’ Y haplogroups are O₃, O_{3a5}, and O_{3a5a}, the ‘Tai’ Y haplogroups are O_{1a} and O_{2a*}, and ‘Hmong-Mien’ Y haplogroups are O_{3*} and O_{2a*}. The majority of the Pinghua population sampled is of the ‘non-Sinitic’ O_{2a} Y haplogroup (42.58%), while the second most frequent Y haplogroup, the ‘Sinitic’ O_{3a5a}, is found only in 14.85% of the sampled population. As for the mitochondrial DNA data, the ‘Sinitic’ mt haplogroups are A, C, D, G, M8a, Y, and Z, the ‘Tai’ mt haplogroups are B_{4a}, B_{5a}, M*, M_{7b}*, C, B_{4b1}, M_{7b1}, F_{1a}, B₄*, and R9b, and the ‘Hmong-Mien’ mt haplogroups are B_{4a}, B_{5a}, M*, M_{7b}*, C, B_{4b1}, M_{7b1}, F_{1a}, B₄*, and R9b. The Pinghua population sampled primarily belongs to the B_{4a}, B_{5a}, M*, F_{1a}, M_{7b1}, and N* mt haplogroups, i.e. coinciding with the Tai and the Hmong-Mien population.

intelligible. For daily interactions, there is no ‘standard’ accent of Nanning Pinghua which speakers uniformly gravitate towards.

Nanning Pinghua is associated with the earliest Sinitic immigrants in the area; they arrived before the Mandarin and Cantonese speakers. There were already small groups of Han Chinese people in Guangxi as early as the Qín Dynasty. (For instance, the canal 灵渠 *Língqú* was excavated in northeastern Guangxi (in modern day 興安 Xīngān country) to link the Yangtze and Pearl River systems in 214 BC during the Qín Dynasty. The canal facilitated the migration of the first wave of Han Chinese immigrants into the modern day Guangxi, and subsequently Guangdong.) The most important immigration event associated with the Pinghua speakers happened in AD 1053: many Pinghua speakers claim that their ancestors were soldiers from the historical Shandong area (which includes parts of modern day Henan) who came in AD 1053 during the Northern Sòng Dynasty to wage a war with the local polity 大南 *Daihnanz* (Mandarin Dànnán) headed by the Zhuang leader 儂智高 *Nungz Ciqgau* (Mandarin Nóng Zhīgāo).³ Nowadays Pinghua speakers are found mainly along the old Guilin-Liuzhou-Nanning road (i.e. the Sòng Dynasty military route that the soldiers travelled along from the north to Nanning), and also along the various tributaries of the Pearl River which cross the Guilin-Liuzhou-Nanning road.

In the *Language Atlas of China* (Wurm et al. 1987), Pinghua is divided into two types: Northern Pinghua and Southern Pinghua. Yu Jin (2007) classifies Pinghua into four types: 永江 *Yōngjiāng* ('Yong River'), 官道 *Guāndào* ('Official Road'), 融江 *Róngjiāng*, and 濱江 *Líjiāng*. The former three are types of Southern Pinghua, whereas Líjiāng Pinghua refers to Northern Pinghua. Map 7.1 shows their distribution.

- Yongjiang Pinghua is the prototypical Southern Pinghua, spoken around Nanning and also in their emigrant communities elsewhere, primarily upriver to the west. Wherever grammar is concerned, ‘Nanning Pinghua’ refers to this type of Southern Pinghua in this analysis;
- Guandao Pinghua is also a type of Southern Pinghua, and it is found to the east of Nanning, primarily along the road towards Liuzhou. The Pinghua dialects of 橫縣 *Héngxiàn*, 寶陽 *Bìnyáng* and 來賓 *Láibīn* are of this type; Binyang is famous for being the only county where Pinghua is the dominant language;
- Rongjiang Pinghua is spoken along the Róng River to the north of Liuzhou (Southwestern Mandarin is spoken in Liuzhou city itself). It is situated geographically in Northern Guangxi. However, its phonology is conservative like Southern Pinghua;

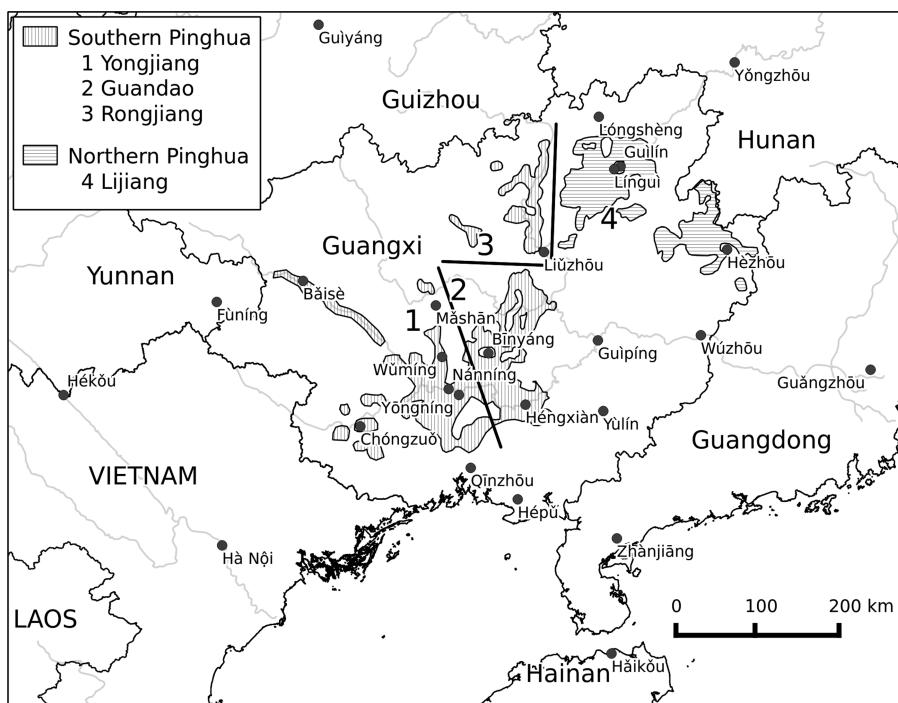
³ The polity headed by 儂智高 *Nungz Ciqgau* and his father 儂全福 *Nungz Cienzfuk* (Mandarin Nóng Quánfù) had numerous name changes. The *Nungz* polity was at various times paying tribute to, or waging wars with, both the 宋 Sòng Dynasty of China to the north and the 李 Lý Dynasty of Đại Việt to the south. In the end, the *Nungz* polity was crushed by the Sòng; many Pinghua people claim that their ancestors were these Sòng soldiers.

Language contact in Nanning: Nanning Pinghua and Nanning Cantonese 161

- Lijiang Pinghua is the prototypical Northern Pinghua, spoken around Guilin (Southwestern Mandarin is spoken in Guilin city itself). Its phonology is significantly Mandarinized. We know that the Mandarinization of the phonology is a later phenomenon, as the older Chinese loanwords in the indigenous languages of the area are Southern Pinghua-like (Qin Fengyu pers.comm.).

Pinghua is a Sinitic language which does not have an influential urban variety. Except in Binyang, all the county towns and city centres in Pinghua-speaking parts of Guangxi are dominated by Cantonese or Mandarin speakers. (However, even Binyang Pinghua is not free from Yue-like features, as will be seen in some examples in the next few sections.)

The affiliation of Pinghua is still a contentious issue within Chinese linguistics, as Pinghua, and Southern Pinghua in particular, share many phonological similarities with Yue. One opinion is that Pinghua is a branch (or branches) of Yue (Liang Jinrong 1997; 2000, Chen Xiaoyan 2007). Another opinion is that just Southern Pinghua is a branch of Yue (Xie Jianyou 2007), while Northern Pinghua is something else (possibly related to a range of Sinitic patois found in Southern Hunan). There is



MAP 7.1 Distribution of Pínghuà dialects in Guǎngxì and their subdivisions

Based on Wurm et al. 1987 and Yú 2007

also Li Lianjin's (2003) opinion that Pinghua and Yue are separate branches of Sinitic, and that Goulou Yue (勾漏 Gōulòu) is in fact a type of Pinghua. Mai Yun (2010) concludes that Yue, together with Pinghua, split off from mainstream Chinese between the Táng and Sòng Dynasties, when Guangxi and Guangdong were part of the 南漢 Southern Hán polity during the Five Dynasties and Ten Kingdoms period.

Later during the Southern Sòng Dynasty, Cantonese received a strong stream of influence from mainstream Chinese again; this new development in Cantonese has in turn spread from the Pearl River Delta, but has not quite reached as far west as the places where Goulou Yue and Pinghua are spoken.⁴ Mai Yun (2010) also correctly points out that Goulou Yue does in fact share many similarities with Pinghua (as also mentioned in Li Lianjin 2003). On the other hand, Zhang Min and Zhou Lieting (2003) observe that Goulou Yue shares more similarities with other Yue dialects than with Pinghua.

The conclusion that can be gathered from these differing viewpoints is that Pinghua and Yue lie on a dialect continuum (ignoring the Cantonese enclaves, including the Nanning Cantonese enclaves, in Guangxi),⁵ and that the isoglosses between Pinghua-like and Cantonese-like features are not bundled along the border between Pinghua and Yue as shown in the *Language Atlas of China* (Wurm et al. 1987).

The reader needs to keep in mind that while I demonstrate many differences between Nanning Pinghua, and Cantonese in this chapter, these are differences between the two opposite ends of the Pinghua-Yue dialect continuum (with Nanning Cantonese being a dialect that was transplanted from the Cantonese end of the dialect continuum directly into the Pinghua end of the continuum, geographically speaking). The linguistic features treated are not necessarily representative of all Pinghua dialects and all Yue dialects respectively. Many traits in Pinghua are also found in neighbouring Yue dialects of the Goulou type and the Lianzhou type (廉州 Liánzhōu, which is the old name of 合浦 Hépǔ),⁶ while the Guandao type of Pinghua, located just to the east of Nanning, sometimes patterns with Goulou Yue

⁴ There is also Li Jinfang's (2002) interesting viewpoint which recognizes there being later Cantonese influences that spread to some of the other Yue dialects. However, for him, Yue is itself not a single genealogical group: the various varieties of Yue are in fact separate (first order) branches of Sinitic; the non-Cantonese branches of Yue received Cantonese influences to various degrees, but otherwise their similarities are due to having a similar Kra-Dai substratum. It is true that it is difficult to identify innovations that are uniquely shared by all the Yue dialects.

⁵ Whether this is the result of Pinghua and Yue belonging to the same branch of Sinitic, or separate branches but with strong convergence along the border, is still difficult to conclude. More detailed studies on the Western Yue dialects and Pinghua dialects are urgently needed to solve this problem.

⁶ The *Language Atlas of China* (Wurm et al. 1987) mentions the 欽廉 Qinlián type of Yue. However, this name is misleading, as there are clearly two types of Yue dialects within this area. 'Qin' stands for 欽州 Qinzhōu; Qinzhōu city and places like 北海 Béihǎi speak a type of Cantonese that is not very different from Standard Cantonese. On the other hand, 'Lián' stands for 廉州 Liánzhōu (i.e. 合浦 Hépǔ), and the Liánzhōu Yue dialect is not mutually intelligible with Cantonese. Many authors (e.g. Mai Yun 2010) use the term Qinlián and assume that the entire region speaks a Cantonese type of Yue. See Chen Xiaojin and Chen Tao (2005) on the Yue dialects of Béihǎi area (including Hépǔ) where both types are demonstrated.

Language contact in Nanning: Nanning Pinghua and Nanning Cantonese 163

to the east and/or Lianzhou Yue to the south rather than with Nanning Pinghua to the west. Some examples will demonstrate these features in the following sections.

The following are brief introductions to the other languages in the Nanning area. Old Nanning Mandarin is regrettably not discussed further in this chapter due to lack of data. Old Nanning Mandarin is phonologically a type of Southwestern Mandarin (similar to Guilin-Liuzhou Mandarin; see Zhou Benliang et al. 2006), but its grammar is primarily Pinghua- or Cantonese-like (Qin Fengyu pers. comm.). Old Nanning Mandarin was brought in by officials and merchants who came to Nanning from Northern Guangxi during the Ming and Qing dynasties. Old Nanning Mandarin was still one of the two dominant languages in Nanning city centre in the 1910s (Zhou Benliang et al. 2006). However, the number of Old Nanning Mandarin speakers dwindled as more and more Cantonese speakers settled in Nanning. The Mandarin enclave has since disappeared from Nanning city centre. Old Nanning Mandarin is now moribund, and its remaining speakers are dispersed in various rural villages.

Northern and Southern Zhuang belong to two separate branches of the Tai language family, and Tai is a branch of the Kra-Dai language family.⁷ The Tai language family is commonly divided into three branches following Li Fang-Kuei (1960, 1977): Northern Tai, Central Tai and Southwestern Tai. The difference between Northern Zhuang and Southern Zhuang can be appreciated through the fact that most of the numerically important Tai languages like Lao, Thai, Shan, and Tai Lü belong to just one branch of Tai (Southwestern Tai), whereas Northern and Southern Zhuang belong to two separate branches of Tai (Northern Tai and Central Tai, respectively). Nevertheless, the grammars of Northern Zhuang and Southern Zhuang are (supposedly) not strikingly different; the differences are mainly phonological and lexical (Zhang Junru et al. 1999: 393). The Tai languages are primarily SVO, but unlike Sinitic languages, most noun modifiers follow the head noun. In this chapter, only examples from Northern Zhuang are given, as data from Southern Zhuang are comparatively rare. (Northern Zhuang data are plentiful as the Northern Zhuang dialect of Wuming (武鳴 *Wǔmíng*), the county neighbouring Nanning to the north, was chosen as the basis for Standard Zhuang by the Chinese government).⁸ The absence of Southern Zhuang data in this chapter does not imply that the Zhuang influences on Pinghua and Yue are only from Northern Zhuang; Southern Zhuang has inevitably also played a role, but the quantity of influence from Northern vs. Southern Zhuang (and influences from other Kra-Dai languages) on Pinghua and Yue dialects is difficult to assess at the moment, for the reasons given.

⁷ The Kra-Dai family is also known as Tai-Kadai. The name Kra-Dai, following Ostapirat (2000, 2004), is gaining in popularity.

⁸ Standard Zhuang is now written in Roman script with no diacritics; the tones are indicated by letters at the end of an orthographic syllable. In the past, Zhuang was written in Chinese or Chinese-inspired characters, similar to how Vietnamese was first written with Chinese-like characters, similar to how Vietnamese was written in characters that followed Chinese principles.

The language contact situation in Nanning area is part of the larger Guangxi language area (Qin Dongsheng 2012). In the wider Southern Guangxi area, other than the language varieties which exist in Nanning, there are also languages like Hakka, Goulou Yue, and Lianzhou Yue with larger numbers of speakers, , and languages like Southern Min, Mien, Bunu, and Vietnamese with smaller numbers of speakers.

Statistics on language usage are difficult to obtain. In terms of official ethnicity, in the 2010 national census, 50.90% of the population was Zhuang and 46.91% of the population was Han within the prefecture-level city of Nanning (including the six counties to the north and east governed by Nanning), which has a population of 6.66 million. Even within the six urban districts of Nanning (i.e. excluding the six counties to the north and east), which has a population of 3.44 million, the Han people were in the minority (41.97%) at the end of 2009.⁹ It is difficult to estimate how much these ethnicity figures correspond to the level of language usage of the various languages. Impressionistically, many Nanning Cantonese speakers claim Zhuang ethnicity, whereas Nanning Pinghua speakers are nearly always Han.

The Nanning Cantonese data are primarily from Lin Yi and Qin Fengyu (2008), abbreviated as ‘L&Q (2008)’ hereafter. The Nanning Pinghua data were collected by the author in Nanning. Unless otherwise indicated, the Nanning Pinghua data are in the accent of Weizilu (位子渌 *Wèizǐlù*; *wèi²² tʂi³³ lk²³*) a village/suburb of Nanning to the west of the city centre and north of the river near Shangyao (上堯 *Shàngyáo*; locally *tʂvŋ²² hiu¹¹*), from where the founders of the Weizilu village relocated. The Weizilu and Shangyao dialects are minimally different from the Pinghua dialect of Xinxu (*Xīnxū*; locally *θəm⁵³ həu⁵³*), which is outlined in Zhang Junru (1987). The Northern Zhuang data come from various sources, as indicated in each instance.

Phonology and vocabulary will be briefly discussed in §7.2 and §7.3 respectively, while various grammatical topics will be discussed in §7.4. The conclusion and further discussion are found in §7.5.

7.2 Phonology

The phonologies of Nanning Pinghua, Nanning Cantonese, and Standard Cantonese share many similarities: they all have at least six tonemes, the stop codas of -p -t -k -m -n -ŋ, medial glides which are distributionally more dependent on the preceding consonant, and few contrastive places of articulation in the coronal region. Except for the last trait, these traits are largely absent in modern Mandarin dialects. On the other hand, these traits are fairly common in Far Southern China and Southeast Asia. Southern Pinghua and Yue dialects, in many cases, have all of these traits, and so do most Tai languages and Kam-Sui languages (another branch of Kra-Dai). Some of these traits are also found in a large number of Hakka, Southern Min, Mienic (Mao

⁹ The data are from <www.nanning.gov.cn/n722103/n722135/n722481/n722721/854686.html>, accessed 10 July 2012.

Zongwu 2004) and Viet-Muong dialects in the region. In this section we will discuss three contact-induced phonological traits.

The first trait to be discussed is the distribution of the medial glides in Nanning Pinghua, Nanning Cantonese, and Standard Cantonese; in this respect the three of them resemble Tai languages more than Sinitic languages. Most Sinitic languages have a maximum syllable structure of CGVX^T, where C is a non-glide consonant, G is a glide, V is a vowel, X is a consonant or vowel, and ^T is tone. An example of a CGVX^T syllable in Mandarin is 天 *tiān* ‘sky’, where *i* is the medial glide (the medial glides are commonly analysed or represented as high vowels in Chinese linguistics). Ignoring tones for the moment, in Western linguistic tradition, a syllable of the shape CGVX would normally be analysed as having an ‘onset’, which is the CG combination (provided that G is analysed as a glide and not a high vowel), and a ‘rime’ (if the remainder is to be analysed as one unit), which is the remaining VX combination. In the Chinese linguistic tradition, on the other hand, the syllable is first divided into an ‘initial’ (聲母 *shēngmǔ*), which consists of the initial C, and a ‘final’ (韻母 *yùnmǔ*), which is maximally the GVX combination. The initial–final analysis of syllable structure has to do with the fact that in most Sinitic languages, including Mandarin, the medial glide has a closer relationship with the following rime (the following VX unit) in terms of structural economy: if the syllable is divided into two, it is more economical to list inventories of C vs. GVX rather than CG vs. VX. Take the example of Standard Mandarin. If the syllable is split between C (initial) and GVX (final), e.g. *t* vs. *ian* for 天 *tiān*, then there are 18 initials and 35 finals, a total of 53 categories. If the syllable is split between CG (onset) and VX (rime), e.g. *ti* vs. *an* for 天 *tiān*, then there are 55 onsets and 21 rimes, a total of 76 categories.¹⁰ It is thus more economical to list inventories of C (initial) vs. GVX (final) in Mandarin and most Sinitic languages.

The situation in the vast majority of Southern Pinghua and Yue dialects is different. In Nanning Pinghua, Nanning Cantonese, and Standard Cantonese, the only medial glide is *w* and its occurrence depends more on the preceding consonant; in these dialects, the medial *w* occurs when the preceding consonant is *k*, *kʰ* or zero. (Nanning Shangyao Pinghua also has very marginal cases of *tʃw* and *tʃʰw*.) In fact, the medial glides in these Southern Pinghua and Yue dialects are so restricted in occurrence that it is common to posit a maximum syllable structure of CVX rather than CGVX, with the C onset in CVX being any consonant including the glides of *j* and *w*,¹¹ and the clusters

¹⁰ The figures are from Duanmu San (2011), who actually rejects inventorial economy as a factor in determining syllable structure, and advocates an onset–rime (i.e. CG vs. VX) division of the syllable for all Sinitic languages.

¹¹ The situation with the medial glides in some non-Cantonese Yue dialects sometimes looks different from the situation in Cantonese, on the surface at least. In some Yue dialects, e.g. 東莞 Dōngguǎn and 台山 Táishān, what must have been syllable-initial *j* and *w* have become *z* and *v* respectively (Lau Chun-fat 2007). The Yue of 台山 Táishān is also interesting in that on the surface they have the medial glides of *j* and *w* which are clearly dependent on the following rime. However, on closer inspection, *j* is only followed by *ɛ*, and *w* is only followed by *ɔ*. The opposite is also true: *ɛ* and *ɔ* are always preceded by *j* and *w* respectively. The medial glides can be easily dispensed with in the underlying representation; they are predictable

of *kw* and *k^wh* being analysed as single segments with secondary articulation: *k^w* and *k^wh*. For instance, Nanning Cantonese is described by L&Q (2008: 10) as having the following inventory of onsets: *p p^h m f t t^h n t l t^f t^h s k k^h η h k^w k^wh j w*,¹² while the rimes have the shape of V or VX (there is also a syllabic *ŋ*). The trait of G forming a closer relationship with the preceding C rather than the following rime (VX) in Southern Pinghua and Yue mirrors the behaviour in many Tai languages. In many Tai languages, medial glides and medial liquids clearly form a closer relationship with the preceding consonant. For instance, Wuming Zhuang is described as having the following inventory of onsets and rimes: onsets *p b m f t d n θ l c k η y ? h j w ?j ?w pl ml kj kw*, while rimes have the shape V or VX (Zhāng Jūnrǔ et al. 1999: 51).¹³

The second phonological trait to be discussed is that Nanning Cantonese and Nanning Pinghua have a lateral fricative *t* (or interdental fricative *θ*) in some accents of Nanning Pinghua like Xinxu (心墟 Xinxū) Pinghua), which contrasts phonemically with *f* in Nanning Cantonese and in most Nanning Pinghua accents.¹⁴ In contrast, Standard Cantonese lacks *t* or *θ*. Amongst Yue dialects, Standard Cantonese and most Yue dialects in the Pearl River Delta are in fact in the minority in not having /t/ or /θ/; /t/ can be found as close to the Pearl River Delta as the Siyi (四邑 Siyi) region (e.g. Huang Jianyun (1990) for 台山 Táishān Yue). The *t* or *θ* is also near universal in the Western Yue dialects. Looking at Zhuang dialects, most Zhuang dialects also have *t* or *θ* (e.g. Standard Zhuang has a contrast between *θ* and *c*). Map 7.2 shows the locations of some Sinitic languages with a lateral fricative [t]

Having *t* or *θ* is an areal phenomenon. However, the directionality of borrowing is not necessarily easy to determine. Linguists in China tend to attribute *t/θ* as a Kra-Dai trait that has diffused into Sinitic languages (Li Jinfang 2002: 110; Mai Yun 2010). It is true that most Zhuang dialects have a *t* or *θ*. In other Kra-Dai languages, *t* or *θ* is also found in most Hlai dialects and Ngao Fon 村話 Cūnhuà in Hainan (Ostapirat 2008). However, this *t* or *θ* in Zhuang dialects is a reflex of what has been reconstructed as *s in Proto-Tai (Li Fang-Kuei 1977; Pittayaporn 2009), given that a reconstructed voiceless lateral onset *l is much more likely to have been realized as [t]. That leaves the other possibility, that is, *s > t/θ is an innovation in Sinitic languages that spread to Kra-Tai languages. With Sinitic languages, *t* or *θ* is found in several non-contiguous

phonetic onglides in /ɛ/ and /ɔ/ respectively. On the other hand, there are some Yue dialects where the medial glide cannot be dispensed with, e.g. Qinzhōu Cantonese (Lin Jinjuan 2008).

¹² The transcription that L&Q (2008) use for Nanning Cantonese has been slightly modified to align with the IPA transcription I have adopted for Nanning Pinghua and Standard Cantonese. Their onsets <kw kw^h> are rendered in this description as *k^w k^wh*, their rimes <en ek un uk> are rendered here as *ŋ ik uŋ uk*, and their tones <21 35 24> are rendered here *11 25 13*. Their simplified Chinese characters have also been substituted with traditional Chinese characters in this chapter. Other romanizations follow those used in quoted source material.

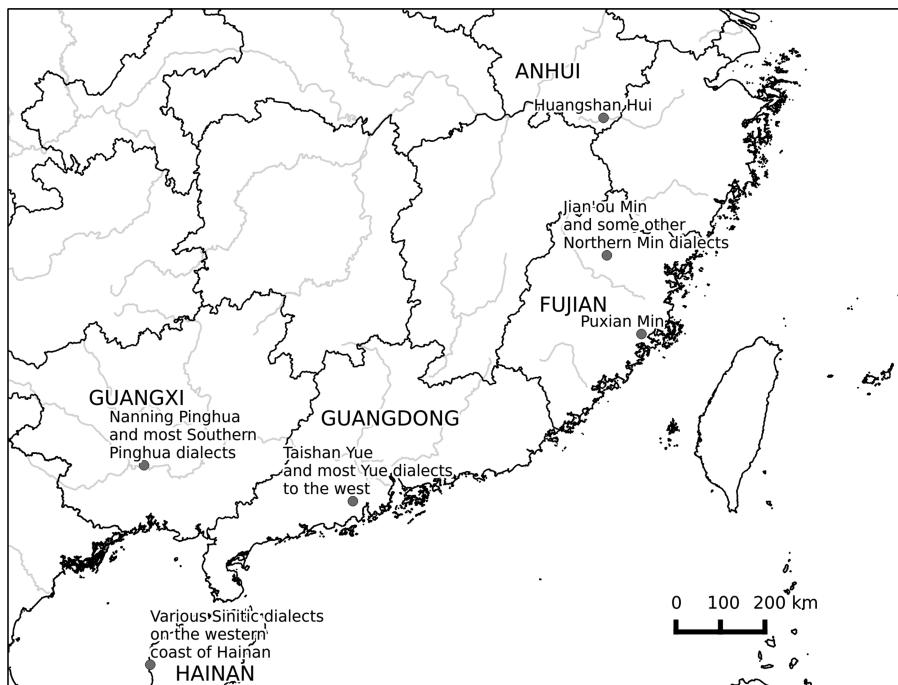
¹³ Underlying long vowels followed by a consonant are realized with a schwa offglide, thus creating rimes with three segments, but this is only a surface phenomenon.

¹⁴ However, Shangyao and Xinxu Pinghua have merged them as *t* and *θ* respectively.

Language contact in Nanning: Nanning Pinghua and Nanning Cantonese 167

areas: Southern Pinghua-plus-Yue, various Sinitic dialects on the western coast of Hainan (Liu Xinzong 2006: 54–5), Puxian Min (Liu Fuzhu 2007), several Northern Min dialects like 建甌 Jiàn'ōu and 政和 Zhènghé (Akitani 2008), and Hui at 黃山 Huángshān (Meng Qinghui 1981). The change of *s > t/θ in all these Sinitic dialects is unlikely to be a retention from an older stage of Chinese, as there is no evidence of *s being pronounced anything like t or θ (judged by, for example, the transliterations between Chinese and foreign languages at various periods of time). Unless we are dealing with multiple independent developments (and putting the Western Hainan case aside, as the direction of borrowing was clearly from Hlai to Sinitic due to the later arrival of Sinitic speakers), the only possibility left is that they have all been influenced by Kra-Dai languages at some point in time, and that t or θ was borrowed from Kra-Dai languages. This loops back to the argument that *s > t/θ was an innovation in Kra-Dai languages, which is problematic as we have seen. Detailed studies are required to explain this phenomenon.

As for Nanning Cantonese, it is unclear whether Nanning Cantonese acquired t after their arrival in Nanning, or whether they started off with t in their inventory when the Cantonese speakers were still in the Pearl River Delta. (Within Pearl River Delta, currently there is only the Cantonese of 佛岡 Fógāng which is reported as



MAP 7.2 Locations of some Sinitic languages with a lateral fricative [ɬ]

having /ɿ/ (Mai Yun 2010: 228).) There are, however, other clearer examples of Sinitic languages having borrowed /ɬ/ or /θ/ in a short period of time after their arrival in Guangxi, e.g. Hakka of 馬山 Mǎshān has a contrast of /θ/ vs. /s/ vs. /ʃ/ (Xie Jianyou 2007: 262), and Southern Min of 平南 Píngnán has a contrast between /ɬ/ and /s/ (Xie Jianyou 2007: 288). Hakka and Southern Min in general are not known to have /ɬ/θ/, and these Sinitic dialects have only been spoken in Guangxi for less than 150 years (but some of these speakers were settled in Western Guangdong before their arrival in Guangxi).

The third phonological feature to be discussed is the historical development of the Entering tones (i.e. tones that syllables with obstruent codas have).¹⁵ In Cantonese, and also in the vast majority of Yue dialects, the Upper and/or Lower Entering tone is further split into two, with the conditioning factor being vowel length (or the related vowel quality difference). For instance, for the Upper Entering tone in Cantonese, (which infers that the syllable in Middle Chinese has a voiceless onset, and a plosive coda), usually syllables with a short vowel have the high tone [⁵], and syllables with a long vowel have the mid tone [³], e.g. 北 [prek⁵] ‘north’, 百 [pa:k³] ‘hundred’ (*pok* and *pæk* respectively in *Middle Chinese* [MC]).¹⁶ Having a tone split with the Entering tone (‘Tone D’ in Kra-Dai studies) based on vowel length is also present in the majority of Kra-Dai languages, including Zhuang, but largely absent in other Sinitic languages and other families in the region, namely Viet-Muong and Hmong-Mien.¹⁷ However, Nanning Pinghua and the other Yongjiang type of Southern Pinghua are different as they also have one of their Entering tones, the Lower Entering tone, split into two. However, the conditioning factor in Nanning Pinghua is different from Yue and Zhuang; the conditioning factor is whether the initial in Middle Chinese is a sonorant (including the 云 *f and 以 *zero initials) or an obstruent. For instance, the *sonorant-initialed 域 *l̥wik* > *wət*²³ ‘region’, 葉 *jep* > *hip*²³ ‘leaf’, 襪 *m̥jwot* > *mat*²³ ‘sock’, vs. the *obstruent-initialed 活 *ywat* > *wət*² ‘live’, 盒 *yop* > *hap*² ‘box’, and 罷 *b̥jwot* > *fat*² ‘punish’. Vowel length plays no role in the splitting of the Entering tone

¹⁵ In traditional Chinese historical phonology, the reconstructed tone system of Middle Chinese has been divided into four tones *píngshēng* 平 Level tone; *shǎngshēng* 上 Rising tone; *qǐshēng* 去Departing tone; and *rùshēng* 入 Entering tone. With the later Sinitic languages, when some or all of these tones developed two allotones (which are often subsequently phonemicized), the allotones are called *yīn* 隹 ‘upper’ and *yáng* 陽 ‘lower’. Correspondences are made in this chapter with the reflexes of these tonal categories in contemporary Sinitic languages.

¹⁶ Baxter (1992)’s transcription system of Middle Chinese is followed here. His ASCII-friendly symbols have been changed into IPA. Baxter’s transcription of Middle Chinese is a transcription of Chinese notation for Middle Chinese phonology, and neither are reconstructions of Middle Chinese in the sense of the Western comparative method. Hence Middle Chinese forms in Baxter’s transcription are not prefixed with an asterisk, thus ‘*’.

¹⁷ Amongst Hmong-Mien languages, the splitting of the Entering tone based on vowel length seems to be only present in some Kimmun dialects (Mienic) in Yunnan (Mao Zongwu 2004: 76, 82, 87) and Hainan (Li Yunbing 2003: 695–6). This may have to do with the fact that these Kimmun dialects are all spoken in the vicinity of Kra-Dai languages.

Language contact in Nanning: Nanning Pinghua and Nanning Cantonese 169

in Nanning Pinghua. This is one feature where Nanning Pinghua is less like Zhuang than Nanning Cantonese. (Despite this, the splitting of the Entering tone based on the sonority of the initial in MC is not a defining feature of Pinghua; Guandao Pinghua dialects like 賓陽 Bīnyáng (Huang Yingfu 2005) are again like Yue in having their Entering tones split according to vowel length.)

In summary, we have seen that the phonology of both Nanning Pinghua and Nanning Cantonese are both conservative with respect to their codas and that they have many tones, like Zhuang languages. Also similar to Zhuang, medial glides have a closer relationship with the preceding onset consonant than with the following rhyme. We have also discussed the *t* or *θ* phoneme as an areal phenomenon in Guangxi and Western Guangdong. We have seen one trait where Cantonese resembles Zhuang more closely than does Nanning Pinghua: in Cantonese and Zhuang, one or both of the Entering tones are split based on vowel length, whereas in Nanning Pinghua, the Entering tone is split based on the sonority of the onset consonant in Middle Chinese.

7.3 Vocabulary

There are some interesting vocabulary differences amongst Nanning Pinghua, Nanning Cantonese, and Standard Cantonese. In Nanning Pinghua, there are forms borrowed from Zhuang, whereas Standard Cantonese has a Sinitic form. Nanning Cantonese sometimes sides with Nanning Pinghua, and sometimes with Standard Cantonese. It is not surprising that Nanning Pinghua has Zhuang loanwords, even for some ‘basic’ vocabulary, as Pinghua has been spoken alongside Zhuang for at least one millennium. The following is a very small sample of vocabulary differences, including some differences in grammatical words, between Nanning Pinghua and Cantonese.

- Nanning Pinghua has a first person inclusive pronoun 伝(隊) *wen*¹¹(*tɔi*²²), c.f. Northern Zhuang *vunz* ‘person’ (also written as 伝 in Old Zhuang characters). (The exclusive pronoun is 我隊 *ŋa*¹³*tɔi*²²). Northern Zhuang dialects tend to have a clusivity distinction (inclusive *raez*, exclusive *dou*), whereas Southern Zhuang dialects tend not to (Wei Jingyun and Qin Xiaohang 2006: 3). Cantonese, and Yue dialects in general, have no clusivity distinction for their first person plural pronoun (e.g. Nanning Cantonese 我哋 *ŋɔ*¹³*ti*²² for both inclusive and exclusive first person). Interestingly, Lianzhou Yue has the form *wen*³⁵ as its first person plural pronoun (with no clusivity distinction; Chen Xiaojin and Chen Tao 2005: 400), which is probably related to Zhuang *vunz* as well;
- Nanning Pinghua *hei*²⁵ (and other forms like *həu*²⁵/ *hui*³⁵/ *hei*⁵⁵) ‘give’,¹⁸ c.f. Northern Zhuang *hawj* ‘give’, Proto Tai *hau^C ‘give’ (Pittayaporn 2009: 356).

¹⁸ In Nanning Shangyao: *hei*²⁵. In Nanning 心墟 Xinxū: *həu*²⁵. In Nanning 石埠 Shíbù Pinghua: *hui*³⁵. In Nanning 亭子 Tíngzi Pinghua: *hei*⁵⁵ ‘give’.

The tones do not match (Tai tone C usually corresponds with the Sinitic Rising tone, but *hvet*²⁵ etc. in Pinghua is in the Departing tone), but the high level pitch in *het*⁵⁵ and the rime in *həui*²⁵ are phonetically similar to the Northern Zhuang *hawj* [həu⁵⁵].¹⁹ Nanning Cantonese has 紿 *ket*⁵⁵ (< Guilin-Liuzhou Mandarin) or 界 *pi*²⁵ ‘give’, and Standard Cantonese has 界 *pei*²⁵ ‘give’;

- Nanning Pinghua *jən*⁵³ ‘cold’, c.f. Proto Tai *?*jen*^A ‘cool’ (Li Fang-Kuei 1977). Cantonese has 淚 *tŋy*³³ ‘cold’; Nanning Pinghua *həŋ*²⁵ ‘too (non-tactile) hot’, c.f. Northern Zhuang *hwngq* ‘hot weather’.²⁰

The following are a few examples where both Nanning Pinghua and Nanning Cantonese have Zhuang loans.

- Nanning Pinghua *ṭek*³ ‘a few’, Nanning Cantonese *ṭek*⁵ ‘a few’, c.f. Northern Zhuang *saek* ‘a few’. Nanning Pinghua also has the Sinitic 幾 *kəi*³³, and Nanning Cantonese 幾 *ki*²⁵ ‘a few’, c.f. Standard Cantonese 幾 *kei*²⁵ ‘a few’;
- Nanning Pinghua *nəm*⁵⁵ ‘stupid’, Nanning Cantonese *nəm*⁵⁵ ‘unsophisticated’, c.f. Northern Zhuang *naem* ‘slow’. Standard Cantonese has other words like 蠢 *tsʰən*²⁵ ‘stupid’;
- Both Nanning Pinghua and Nanning Cantonese have *ṭən*¹¹ ‘shiver’, c.f. Northern Zhuang *saenz* ‘shake, vibrate’. They also have the Sinitic word 震 (Nanning Pinghua *tʃən*⁵⁵, Nanning Cantonese *tʃən*³³), c.f. Standard Cantonese has 震 *tʂən*³³ ‘shiver’;
- Nanning Pinghua *puk*²(-*tfi*³³) ‘pomelo’ (子 -*tfi*³³ is a noun suffix; see §7.4.2), Nanning Cantonese *pok*⁵*lok*⁵ ‘pomelo’, c.f. Northern Zhuang (*mak-*)*bug* ‘pomelo’ (*mak-* is ‘fruit’). Standard Cantonese *luk*⁵*jeu*²⁵ ‘pomelo’, where 柚 *jeu*²⁵ is Sinitic, and *lok*⁵ is perhaps also related to Northern Zhuang *lwg-* ‘melon, child’).

¹⁹ Lin Yi (2010) mentions that ‘有學者 [some scholars]’ consider *hawj* ‘give’ in Zhuang is a loan from Chinese 許 ‘allow, promise, betroth’. It is a possibility that Proto-Tai borrowed this from Chinese; amongst the many arguments raised by Lin Yi that support this is the regular tonal correspondence between Tai Tone C and Sinitic Rising tone. In terms of the tone, the Nanning Pinghua *het*²⁵ is not a regular reflex of Middle Chinese 許 (**xjo*^{RISING}). The verb *het*²⁵ in Nanning Pinghua is likely to be a loan from Zhuang, or at least influenced by Zhuang.

²⁰ Lin Yi (2003) considers the Proto Tai *?*jen*^A ‘cold’ and the Northern Zhuang *hwngq* ‘hot weather’ to be related to Chinese 涼 (*?*jin*^{LEVEL}) and 煙 (**jan*^{LEVEL/DEPART}) respectively. See also de Sousa, Langella and Enfield (2015). Regular reflexes of these in Nanning Shangyao Pinghua would be 涼 *en*⁵³/*in*⁵³ and 煙 *ŋy*^{11/22} respectively, somewhat different from the actual forms in Shangyao Pinghua: *jən*⁵³ ‘cold’ and *hŋy*²⁵ ‘non-tactile hot’ respectively. As for the words for ‘cold’, reflexes of Proto Tai *?*jen*^A ‘cold’ are commonly found in Southwestern Tai languages, but less often encountered in Northern and Central Tai languages (which includes the Zhuang dialects). However, the comparative rarity of reflexes of *?*jen*^A ‘cold’ in modern Zhuang dialects does not preclude the possibility that Southern Pinghua borrowed *jən*⁵³ from historically earlier Zhuang dialects. As for the words for ‘hot’, forms like Northern Zhuang *hwngq* are not very widespread amongst the Tai languages in China as pointed out in Lin Yi (2003), and no one has reconstructed a similar form for Proto-Tai; *hwngq* in Zhuang might be a form borrowed from Chinese.

Not all vocabulary differences between Pinghua and Cantonese are due to Zhuang influence. Nanning Pinghua has many ‘basic’ words which are cognates with those in Mandarin, while Cantonese has distinct forms. With these Mandarin-like words, it is not necessarily the case that Nanning Pinghua has borrowed them from Mandarin; it is usually just the case that Nanning Pinghua has a form that is commonly found in other modern Sinitic languages, whereas Cantonese has a different Sinitic form (i.e. Pinghua and Cantonese retained different forms for the same meaning from older forms of Chinese). These vocabulary differences are not necessarily distinguishing features of Pinghua and Yue; many non-Cantonese Yue dialects also use forms that are cognates with Mandarin rather than with Cantonese. It is simply the case that the Cantonese forms have failed to supplant the corresponding forms in these other Yue dialects, which have otherwise received much influence from Cantonese. The following is a small sample of vocabulary differences between Nanning Pinghua, Nanning Cantonese, and some other Yue dialects that are not the result of by Zhuang influence. (Also quoted are relevant maps in the *Grammar, Lexicon or Phonetics volume of the Linguistic Atlas of Chinese Dialects* (Cao Zhiyun et al. 2008).)²¹

- Nanning Pinghua 是 *ɿi²²* / *θi²²* / *fi²²* ‘be’,²² Mandarin 是 *shì* ‘be’, vs. Cantonese 係 *hei²²* ‘be’ (also Hakka 係 *he⁵³* ‘be’). Shì is as a copula is also found in some Yue dialects in Guangxi, e.g. Lianzhou Yue *si²¹*, Yulin Yue *si²⁴* (see also Grammar Map 038);
- Nanning Pinghua 吃 *het³* ‘eat’, Mandarin 吃 *chī* ‘eat’, vs. Cantonese 食 *sik²* ‘eat’. Actually many Yue dialects also use 吃 rather than 食 for ‘eat’, e.g. Taishan Yue 吃 *hiek³* (see also Lexicon Map 084);
- Nanning Pinghua 看 *han²⁵* / *han⁵⁵* ‘look’,²³ Mandarin 看 *kàn* ‘look’, vs. Cantonese 瞧 *tʰei²⁵* ‘look’. Lianzhou Yue also uses 看 (*hun⁴⁴*) rather than a cognate of 瞧 (see also Lexicon Map 121);
- Nanning Pinghua 鼻 *pet²* ‘nose’ (Entering tone), Mandarin 鼻 *bí* ‘nose’ (descended from an Entering tone syllable), vs. Cantonese 鼻 *pei²²* ‘nose’ (Departing tone). 鼻 ‘nose’ having a -t coda is also found in Lianzhou Yue and some Gouloou Yue dialects (see Xie Jianyou 2007: 188, and Phonetics Map 037);²⁴

²¹ 廉州 Liánzhōu (i.e. 合浦 Hépǔ) Yue data are from Chen Xiaojin and Chen Tao (2005), 玉林 Yùlín Yue data are from Zhang Min and Zhou Liting (2003), and 台山 Taishan Yue data are from Huang Jianyun (1990).

²² In Shangyao Pinghua *ɿi²²*; in Xinxu (心墟 Xinxū) Pinghua *θi²²*; in all other accents of Nanning Pinghua there is a distinction between *f* and *θ/θ̪*: *fi²²*.

²³ ‘Look’: *han²⁵* in Shangyao and Xinxu Pinghua; *han⁵⁵* in most other accents of Nanning Pinghua.

²⁴ Standard Cantonese also has a fossilized form of ‘nose’ with a -t coda: 象拔蚌 *tsəŋŋ²²* *pet²* *pʰɔŋŋ¹³* ‘geoduck clam’ is literally ‘elephant pull clam’, but *pet²* is probably originally ‘nose’, i.e. ‘elephant nose clam’, given that geoduck clams resemble elephant trunks.

- Nanning Pinghua 了 *liu*¹³ PERFECTIVE, Mandarin 了 *le* PERFECTIVE, vs. Standard Cantonese 左 *tsɔ*²⁵ PERFECTIVE, Nanning Cantonese 曬 *tai*³³ PERFECTIVE; 了 as a perfective marker is also quite common in the Yue dialects in Guangxi and far western Guangdong (see also *Grammar Map 063*).

Lastly, there are lexical items that have a different distribution to those outlined above. One example is the passive marker 挨 ‘suffer’ (for an agented or agentless passive construction), which is common across Guangxi, e.g. Nanning Pinghua and Nanning Cantonese 挨 *ŋai*¹¹, Guilin-Liuzhou Mandarin 挨 *ŋæ*³¹, Northern Zhuang *ŋaiz*. Standard Cantonese has the obligatorily agented 罢 *pei*²⁵ ‘give’ passive construction (see also *Grammar Map 095*). Another example is the word for ‘wok’: in Nanning Pinghua 鑄 *tſey*⁵³ vs. Cantonese 鑊 *wɔk*². The word 鑄 *tſey*⁵³ ‘wok’ is found throughout Pinghua dialects, and cognates are also found in some western Yue dialects and also in Xianghua to the north (*tsʰɔŋ*⁵⁵; Hilary Chappell pers. comm.) (see *Lexicon Map 109* and Li Lianjin 2003).

In summary, Nanning Pinghua has more Zhuang loans than Nanning Cantonese. This is to be expected, as Nanning Pinghua has been in continuous contact with Zhuang much longer than has Nanning Cantonese. Nonetheless, Nanning Cantonese has also acquired many Zhuang loanwords during its short existence in the Nanning area. The Sinitic vocabulary of Nanning Pinghua and Nanning Cantonese share many similarities. This notwithstanding, Nanning Pinghua sometimes has words resembling Mandarin rather than Cantonese. This is not necessarily due to Mandarin influence; often these ‘Mandarin-sounding’ words are simply words that are commonly found amongst (non-Cantonese) Sinitic languages, as Pinghua has not been under the influence of Cantonese long enough to acquire these Cantonese words. Lastly, there are some words that are commonly found amongst the various languages in Guangxi, and Nanning Cantonese has acquired them too within its short existence in Guangxi.

7.4 Grammar

It is in the realm of grammar that the most surprising outcome in this complex language contact situation manifests itself. On the whole, the grammars of Nanning Pinghua and Nanning Cantonese are not radically different.²⁵ However, there are some areas where the ‘newly’ established language of Nanning Cantonese resembles Zhuang more than the longer-established Nanning Pinghua, and Nanning Pinghua is in some cases ‘anti-Zhuang’ in its grammar. The following is a selection of

²⁵ Without prolonged exposure, Nanning Pinghua is not very intelligible to speakers of Nanning Cantonese. Nevertheless, Nanning Pinghua is still grammatically, lexically, and phonologically much closer to Cantonese than it is to Mandarin.

Language contact in Nanning: Nanning Pinghua and Nanning Cantonese 173

grammatical differences amongst Nanning Pinghua, Nanning Cantonese, Standard Cantonese, Standard Mandarin, and Northern Zhuang. (Some differences in the forms of the grammatical markers have already been discussed in §7.3.) In §7.4.1 to §7.4.5 we will see examples of how Cantonese resembles Zhuang more than does Nanning Pinghua. In §7.4.6 and §7.4.7 we will see how Nanning Cantonese has become even more Zhuang-like than both Standard Cantonese and Nanning Pinghua. In §7.4.8 we will discuss the difference in the ditransitive word order for ‘give’; the ditransitive word order for ‘give’ in Cantonese is often attributed to Zhuang influence, but the real situation is not so simple.

7.4.1 Gender affixes for animals

Nanning Pinghua is a curiosity amongst Southern Sinitic Languages. Southern Sinitic languages, Yue, Hakka and Min for instance, usually have gender suffixes for animals, e.g. Nanning Cantonese 雞公 *kəi⁵⁵-kʊŋ⁵⁵* ‘cock/ rooster’, 雞乸 *kəi⁵⁵-na²⁵* ‘hen’, 雞項 *kəi⁵⁵-hɔŋ²²* ‘pullet’ (L&Q 2008: 144–5). (Standard Cantonese has the same forms except for 雞項 *kəi⁵⁵ hɔŋ²⁵* ‘pullet’). Tai languages also have gender suffixes, which is not surprising as noun phrases in Tai languages are strongly left-headed, e.g. Northern Zhuang *gaeq-boux* ‘cock/ roosters’, *gaeq-meh* ‘hen’, *gaeq-hanh* ‘pullet’. On the other hand, Nanning Pinghua follows the ‘Northern Chinese’ pattern of using gender prefixes, e.g. Nanning Pinghua 公雞 *kʊŋ⁵³-kəi⁵³* ‘cock/ rooster’, 母雞 *mu¹³-kəi⁵³* ‘hen’, 項雞 *har²²-kəi⁵⁵* ‘pullet’. This is also generally the case in other Yongjiang Pinghua dialects, for instance, in 崇左新和 Xinhé, Chóngzuǒ (Liang Weihua and Lin Yi 2009: 151), 崇左江州 Jiāngzhōu, Chóngzuǒ (Li Lianjin and Zhu Yan'e 2009: 175), and the ones listed in Xie Jianyou (2007: 1258–61), namely 崇左四排 Sípái, Chóngzuǒ, 南寧沙井 Shājǐng, Nánníng, 扶綏龍頭 Lóngtóu, Fúsuí, and 百色那華 Náhuá, Básè.

Looking into languages in the vicinity of Nanning Pinghua, gender prefixes are also used in Southwestern Mandarin, the dominant language in Northern Guangxi. However, the Mandarin of Northern Guangxi is somewhat mixed in having both gender prefixes and suffixes, unlike Nanning Pinghua and other Yongjiang Pinghua dialects, which are more uniformly prefixal. For Mandarin dialects in Northern Guangxi, the *Linguistic Atlas of Chinese Dialects* (Cáo Zhìyún et al. 2008) shows the gender affixes in ‘boar’ and ‘sow’ being primarily suffixal (*Lexicon Maps* 024 and 025). On the other hand, the data in Xie Jianyou (2007) show that the two Mandarin localities of 臨桂 Línguì and 鹿寨雒容 Luòróng, Lùzhài are primarily prefixal for the gender affixes. In the dictionary of 柳州 Liǔzhōu Mandarin (Liu Cunhan 1995), both gender prefixes and suffixes are found.

Having gender prefixes rather than suffixes is one trait where Nanning Pinghua is less Zhuang-like than Nanning Cantonese. However, this cannot serve as a defining feature of Southern Pinghua in general. For instance, to the immediate northeast of

Nanning, Binyang Pinghua (a type of Guandao Pinghua) has gender suffixes rather than prefixes (Lin Yi pers. comm.).

7.4.2 Noun suffixes

The noun suffix 子 *-tſi³³* is ubiquitous in Nanning Pinghua. The word 子 originally meant ‘son’ or ‘child’; in some noun compounds it still has that meaning or something clearly related, e.g. 孝子 *hau²⁵ tſi³³* ‘filial son’, 瓜子 *kʷa⁵³ tſi³³* (melon child) ‘seed’. The noun suffix 子 *-tſi³³* has a nominalizing function, when it is suffixed to verb roots, e.g. 扣子 *kʰeu²⁵-tſi³³* (fasten-NMZ) ‘button’, 鑿子 *tſak²-tſi³³* (chisel-NMZ) ‘chisel (n.)’. However, most of the time it is suffixed to noun roots (most of them monosyllabic), and the only function seems to be nothing other than prosodic; Nanning Pinghua thus appears to have a dispreference for monosyllabic nouns, like Mandarin. The noun suffix 子 *-tſi³³* does not even have a diminutive meaning in many cases such as in the following examples of nouns that take this same suffix: 車子 *tʃe⁵³-tſi³³* ‘car’, 果子 *ku³³-tſi³³* ‘fruit’, 蕉子 *tſiu⁵³-tſi³³* ‘banana’, 薄子 *puk²-tſi³³* ‘pomelo’ (< Zhuang *bwg* ‘pomelo’), 蝦子 *ha⁵³-tſi³³* ‘shrimp’ (not ‘juvenile shrimp’ or ‘shrimp roe’), 星子 *tən⁵³-tſi³³* ‘star’, 鬍子 *hu¹¹-tſi³³* ‘moustache’, 椅子 *zi³³-tſi³³* ‘chair’, 窗子 *tʃʰar⁵³-tſi³³* ‘window’, 亭子 *tən¹¹-tſi³³* ‘pavilion’ and 骰子 *tæk³-tſi³³* ‘dice’. Note that the diminutive suffixes are 兒 *-ni¹¹⁻⁵⁵* and 仔 *-tſai³³*. Having a frequently used noun suffix is a trait more associated with Mandarin and Xiang (Wu Yunji 2005: 89–105) spoken to the north. Both 子 and 仔 (Mandarin *zǐ* and *zǎi* respectively) have long histories; in the Han dynasty (206 BC–AD 220), the word for ‘son’ was 子 *zǐ* in the north and 仔 *zǎi* in the south (Wu Yunji 2005: 100). Presumably the rampant use of 子 *-tſi³³* is a Northern influence in Nanning Pinghua. However, it is difficult to determine how early, or how late, this Northern influence in Pinghua is.

In contrast to Nanning Pinghua, Cantonese is much less concerned with making nouns not monosyllabic. However, the derivative tone sandhi (into the [²⁵] or [⁵⁵] tone) in Cantonese is functionally similar to the noun suffix 子 *-tſi³³* in Nanning Pinghua, and, similarly, it usually does not alter the semantics of the noun root. Derivative tone sandhi is ubiquitous in Standard Cantonese (e.g. 繩 *sŋ¹¹* > *sŋ²⁵* ‘rope’), but rarer in Nanning Cantonese (L&Q 2008: 90). Zhuang languages also do not show any dispreference for monosyllabic nouns.

The rampant use of the noun suffix 子 *-tſi³³* is another trait which makes Nanning Pinghua resemble Zhuang less than does Nanning Cantonese for this feature of avoiding monosyllabic words, where such devices exist but are rarely used.

7.4.3 Position of adverbial 先 ‘first, ahead’

Sinitic languages tend to disprefer adverbials being in a postverbal position (see, e.g., de Sousa 2015). Using the adverbial 先 ‘first, ahead’ as an example, Mandarin has 先

Language contact in Nanning: Nanning Pinghua and Nanning Cantonese 175

xiān which is placed before the verb, e.g. 你先吃蘋果 *nǐ xiān chī píngguǒ* (2SG first eat apple) ‘you eat the apple(s) first’. Standard Cantonese, however, is known to have some adverbials, including 先 *sin⁵⁵* ‘first’,²⁶ which are placed after a postverbal object, countering the general Sinitic trend of not having adverbials postverbally, e.g. 你飲湯先 *nei¹³ jiem²⁵ tɔŋ⁵⁵ sin⁵⁵* (2SG drink soup first) ‘you eat the soup first’, 我走先 *ŋɔ¹³ tseu²⁵ sin⁵⁵* (1SG leave first) ‘I shall leave first’. The following is an example of 先 *tin⁵⁵* ‘first’ in Nanning Cantonese, for which the position is also postverbal.

Nanning Cantonese

- (1) 我 行 阿邊， 你哋 行 噥邊， 瞭 下 邊個 到 先。
ŋɔ¹³ hay¹¹ a⁵⁵-pin⁵⁵, ni¹³ti²² hay¹¹ lu⁵⁵-pin⁵⁵, tʰei²⁵ ha¹³ pin⁵⁵kɔ³³ tu³³ tin⁵⁵.
1SG walk this-way 2PL walk that-way see TENT who arrive first
'I walk this way, you walk that way, see who gets there first.' (L&Q 2008: 336)

Having postverbal adverbials is often considered a Tai trait in Cantonese (e.g. Huang Yuanwei 1997: 71–2; Li Jinfang 2002: 116–7). The following are examples of the postverbal use of ‘first’ in Tai languages.

Northern Zhuang

- (2) *vih maz mbouj hawj gou gong?*
for what NEG give 1SG first
'Why do you not give it to me first?' (Luo Liming et al. 2005: 521)

Thai

- (3) *pāi kɔ̄n tāaj kɔ̄n.*
go first die first
'You go first, you die first.' (said to someone trying new things)²⁷

(c.f. Cantonese 行先死先 *hay¹¹ sin⁵⁵ sei²⁵ sin⁵⁵* ‘one who does something first suffers first’.)

In the *Linguistic Atlas of Chinese Dialects* (Cao Zhiyun et al. 2008), all of Guangxi, except the Xiang speaking corner in the northeast, is indicated as having the adverbial 先 ‘first, ahead’ placed after the verb (*Grammar Map 084*). Nanning Pinghua and other Yongjiang Pinghua dialects are also said to have postverbal 先 ‘first, ahead’ (e.g. Nanning Tingzi Pinghua (Qin Yuanxiong et al. 1997: 204); Chongzuo Jiangzhou Pinghua (Li Lianjin and Zhu Yan'e 2009), Chongzuo Xinhe Pinghua (Liang Weihua and Lin Yi 2009: 320–1)). However, in my Nanning Shangyao

²⁶ Other examples of postverbal adverbials in Cantonese are 多 *tɔ⁵⁵* ‘more’ (pre-object) and 添 *tʰim⁵⁵* ‘as well’ (post-object). The functional equivalents of these adverbials in Mandarin are preverbal.

²⁷ I would like to thank Pittayawat Pittayaporn for providing possible contexts for this expression.

Pinghua data, 先 *tin*⁵³ is usually placed in front of the verb. This is a trait of Shangyao Pinghua which is different from both Zhuang and Cantonese.

Nanning Pinghua

- (4) 佢 就 想 辦法 先 抓 個 筐 果,
kəi¹³ *tʃəu²²* *tɛŋ³³* *pan²²fap³* *tin⁵³* *na⁵³* *kə⁵⁵* *kʷʰaŋ⁵³* *ku³³*,
3SG then think method first take DEM basket fruit
'He thought of a way to take the basket of fruit first.'

Sometimes 先 *tin*⁵³ is placed after the verb; in these cases, 先 *tin*⁵³ has a spatial meaning (usually used in contrast to 後 *heu*²² 'behind'). For instance, contrast 我先行 *ya¹³ tin⁵³ heŋ¹¹* (1SG first go) 'I shall go first', which is something that one might say as one departs, vs. 我行先 *ya¹³ heŋ¹¹ tin⁵³* (1SG go first) 'I shall go in front', which is in contrast to something like 你行後 *nɔi¹³ heŋ¹¹ heu²²* (2SG go after) 'you will go behind'. See also §3.5 in Peyraube, this volume, on word order and adverbs in Sinitic, Hmong-Mien, and Tai.

7.4.4 *CL + N noun phrases*

In isolation or in a preverbal position, Nanning Shangyao Pinghua does not allow noun phrases to begin with a bare classifier ('[CL + N] noun phrases');²⁸ the classifier must be preceded by at least a demonstrative or a quantifier.²⁹ This is similar to Mandarin.

- (5) *(個) 個 細仔仔 有 有 老子 老娘,
*(kə⁵⁵) kə⁵⁵ t̪ei⁵⁵men⁵³tʃai³³ mi¹³ jəu¹³ lau¹³tʃi³³ lau¹³neŋ⁵³,
DEM CL child NEG have father mother
'The child has no father or mother.'

- (6) *(個) 只 新府 呢, 穿 衫 個 陣 時 呢,
*(ə⁵⁵) tʃət³ t̪en⁵³fəu³³ ne⁵⁵, tʃun⁵³ tam⁵³ kə⁵⁵ tʃen²² ti¹¹ nε⁵⁵,
DEM CL daughter-in-law SFP wear clothes DEM CL time SFP
要 佢 幫 扣 扣子。
iu⁵⁵ kəi¹³ paŋ⁵³ kʰəu²⁵ kʰəu²⁵ tʃi³³.
need 3SG help fasten button

'The daughter-in-law, when she put on clothes, she needed her [the servant] to help with doing up the buttons.'

²⁸ Subsections §7.4.4 to §7.4.7 are also discussed in de Sousa (2013).

²⁹ This is the case in the varieties of Nanning Pinghua that I am familiar with. In the *Nanning Tingzi Pinghua Dictionary* (Qin Yuanxiong et al. 1997), there are also no examples of noun phrases that are preverbal classifier-initial. However, Bu Lianzeng gives one such example in the Pinghua of 四堂 Sitáng (2011: 97), a rural town to the northeast of Nanning city centre, and also one example from Binyang Pinghua (2011: 96), spoken further northeast from Sitáng. This is perhaps a feature associated with Guandao Pinghua.

Language contact in Nanning: Nanning Pinghua and Nanning Cantonese 177

On the other hand, [CL + N] noun phrases are ubiquitous in Cantonese and Zhuang. In Cantonese, preverbal [CL + N] noun phrases have a definite interpretation. In Northern Zhuang, preverbal [CL + N] noun phrases can be either definite or generic (see, e.g. Qin Xiaohang (1995) on the generic usage). Not allowing bare preverbal [CL + N] noun phrases is another trait that makes Nanning Pinghua less Zhuang-like than Nanning Cantonese (see also §5.3 in Wang Jian, this volume, on bare classifier phrases in Sinitic languages in general).

Nanning Cantonese

- (7) 張 椅 放 哪 邊 墘 定?
tʃæŋ⁵⁵ tʂɛŋ³³ fɔŋ³³ hɛi²⁵ pɪn⁵⁵ tʂat³³ tʂɛŋ²²?
CL chair put LOC which CL place
'Where should I put the chair?' (L&Q 2008: 270)
- (8) 只 張 小平 啊， 做 嘥 噉 衣□ 嘅!
tʂɛk³ tʂæŋ⁵⁵ tʂiu³⁵ pʰmŋ¹¹ a tʂu³³ jɛ¹³ kem²⁵ ji⁵⁵ jau⁵⁵ kɛ³³!
CL Zhāng Xiǎopíng TOP do thing so careless MOD
'This Zhāng Xiǎopíng, he does things so carelessly.' (L&Q 2008: 276)

Northern Zhuang

- (9) go faex maj ndaej vaiq.
CL tree grow VCOMPL fast
'Trees grow quickly.' (Wei Jingyun and Qin Xiaohang 2006: 223)
- (10) go faex raek dwk hat.
CL tree break VCOMPL ONOM
'The tree went crack ([ha:t³⁵]) and broke.' (Wei Jingyun and Qin Xiaohang 2006: 231)

Nevertheless, in a postverbal position, Nanning Pinghua is the same as Cantonese and Zhuang: it allows [CL + N] noun phrases to occur postverbally; these postverbal [CL + N] noun phrases can be definite or indefinite, and specific or non-specific. Mandarin is different; it only allows [CL + N] noun phrases to occur postverbally, and they must be indefinite. According to Wang Jian (this volume), Nanning Pinghua thus belongs to Type V languages that have bivalent bare classifiers exclusively in the postverbal position. See de Sousa (2013) for examples of bare postverbal [CL + N] noun phrases.

7.4.5 Possession and association

All Sinitic languages have modifier markers (also known as attributive markers, subordinate markers, linkers, amongst many other names); they signify that the preceding constituent is a noun modifier (e.g. 的 *de* in Mandarin). Nanning Pinghua has a modifier marker 個 (*kə*)⁵⁵. One prototypical function of the modifier marker is

to indicate possession or association (in this construction, the number of the possessum is not specified).

Nanning Pinghua

- (11) 我 個 書
ŋa¹³ kə⁵⁵ təi⁵³
1SG MOD book
'My book(s)'
- (12) 老子 個 蛋糕
lau¹³tfi³³ kə⁵⁵ tan²²kau⁵³
father MOD cake
'Father's cake(s)'

Nanning Cantonese whose modifier marker is 嘅 *ke³³* is next exemplified.

Nanning Cantonese

- (13) 門口 嘅 樹 擋 路 多， 剪 開啲 去 啦。
mun¹¹hau²⁵ ke³³ kʰy²² tɕy²⁵ lu²² tʂ⁵⁵ tʃin²⁵ hoi⁵⁵-ti⁵⁵ hy³³ la⁵⁵.
door MOD tree block road much cut open-COMPR go HORT
'The tree in front of the door is too obstructing, let us chop it down so that it is less so.' (L&Q 2008: 318)

Zhuang does not have a native modifier marker which is as multifunctional as the modifier markers in Sinitic languages. However, Zhuang does have a possessive marker, e.g. Northern Zhuang *duh*.

Northern Zhuang

- (14) *doxgaiq duh de*
thing POSS 3SG
'His/her thing(s)' (Wei Jingyun and Qin Xiaohang 2006: 218)

Nevertheless, the modifier marker 嘅 *ke³³* in Cantonese and the possessive marker *duh* in Northern Zhuang are not the most commonly used constructions to indicate possession. In Cantonese, when the possessum is a tangible object, a classifier is used instead of a modifier marker to indicate possession. When a classifier is used to indicate possession, the possessum is singular, unless the non-singular classifier 啟的 *ti⁵⁵* is used.³⁰

³⁰ It takes the form N_{POSSESSOR}–CL–N_{POSSESSUM} and is not dissimilar to the Austronesian possessive classifiers coding alienability, particularly in Oceanic languages where pronominal affixes referring to the possessor are attached to a set of classifier-like free morphemes in combination with the separate possessed noun. These are however limited to a small number of general categories such as edible and potable objects, plants, weapons, etc. (see Chappell and McGregor 1996).

Language contact in Nanning: Nanning Pinghua and Nanning Cantonese 179

Nanning Cantonese

- (15) 佢 隻 崇
k^hy¹³ tʃek³ tʃei²⁵
3SG CL son
'His/her son' (L&Q 2008: 278)
- (16) 我 哟 崇女 個 個 都 好 聽講。
ŋɔ¹³ ti⁵⁵ tʃei²⁵-ny²⁵ kɔ³³ kɔ³³ tu⁵⁵ hu²⁵ t^hɛŋ⁵⁵ kɔŋ²⁵.
1SG CL:NSG son-daughter CL CL all very listen-talk
'All of my children are very obedient.' (L&Q 2008: 319)
- (17) 我 粒 地 貓 屎窟 噉 大，
ŋɔ²⁴ nəp⁵ ti²² mɛu⁵⁵ fi³⁵fet⁵ kem³⁵ tai²²,
1SG CL(small) land cat buttocks like_so big
'My (small) piece of land is as big as a cat's buttocks,' (L&Q 2008: 263)
- (18) 崩 江 哟 水 多 濁 過 柳 江。
jvŋ⁵⁵ kɔŋ⁵⁵ ti⁵⁵ sui²⁵ tɔ⁵⁵ tʃuk² kɔ³³ lɛu¹³ kɔŋ⁵⁵.
Yōng River CL:NSG water more muddy SURP Liǔ River
'The water of the Yōng River is much muddier than that of the Liǔ River.'
(L&Q 2008: 319)

Possession in Northern Zhuang is usually zero-marked. However, since the classifier is most usually present (in front of the noun), on the surface the PossP+CL(+N) possessive construction in Cantonese and the CL+(N)+PossP possessive construction in Zhuang look identical, except for the (expected) difference in the position of the possessor phrase.

Northern Zhuang

- (19) *ponj saw kou*
CL book 1SG
'My book'
(Zhang Junru et al. 1999: 404)
- (20) *aen vanj mwngz*
CL bowl 2SG
'Your bowl'
(Zhang Junru et al. 1999: 404)
- (21) *gij saw youq gwnz daiz cungj dwg gjij mwngz hwj.*
CL:NSG book LOC top table all be CL:NSG 2SG FP
'The books on the table all belong to you.' (Luo Liming et al. 2005: 484)

On the other hand, Nanning Pinghua does not usually allow the Cantonese-like (and Zhuang-like) PossP+CL(+N) possessive construction. A modifier marker must be used, except for some nouns like kin terms which allow possession to be zero-marked. This is another feature that makes Nanning Pinghua less Zhuang-like than Nanning Cantonese.

Nanning Pinghua

(22)	細	蘇	[個	/*	隻	/*	的]	狗兒
	˧˥	˧˥	[kə˥˥]	/*	tʃət ³	/*	tik ⁵]	kəu ³³ -ji ¹¹
	Little	Sū	MOD		CL		CL:NSG	dog-DIM
	'Little Sū's puppy/ puppies'							

7.4.6 ADJ + CL + N

Nanning Cantonese has diverged from Standard Cantonese and other Cantonese dialects in the Pearl River Delta due to strong Zhuang influence. There are many examples of this; in §7.4.6 and §7.4.7 we will discuss just two such examples.

In Standard Cantonese, there are many examples of CL + N noun phrases (as discussed in §7.4.4), e.g. 架車 *ka³³ tɿ^hɛ⁵⁵* (CL car) ‘the car’, 間屋 *kan⁵⁵ ʊk⁵* (CL house) ‘the house’. If an adjective is added, it is most usually placed between the classifier and the noun (CL + ADJ + N), e.g. 架紅車 *ka³³ hʊŋ¹¹ tɿ^hɛ⁵⁵* (CL red car) ‘the red car’, 間空屋 *kan⁵⁵ hʊŋ⁵⁵ ʊk⁵* (CL empty house) ‘the empty house’. Only size adjectives, primarily 大 *tai²²* ‘big’ and 細 *səi³³* ‘small’, can occur immediately in front of the classifier, and they are usually further modified by a degree adverb, e.g. 帕大間空屋 *kəm³³ tai²² kan⁵⁵ hʊŋ⁵⁵ ʊk⁵* (such big CL empty house) ‘such a big empty house’. For other adjectives, the word order of ADJ + CL + N is not possible.

One way to resolve this, while keeping all the constituents in that order, is to insert a distal demonstrative between the adjective and the classifier (ADJ + DIST.DEM + CL + (N)). For instance, 空嘅間屋 *hʊŋ⁵⁵ kɔ²⁵ kan⁵⁵ ʊk⁵* (empty that CL house) ‘the house which is empty’ is grammatical, whereas *空間屋 **hʊŋ⁵⁵ kan⁵⁵ ʊk⁵* (empty CL house) is not grammatical. In fact, the adjective in 空嘅間屋 *hʊŋ⁵⁵ kɔ²⁵ kan⁵⁵ ʊk⁵* (empty that CL house) ‘the house which is empty’ is actually in a relative clause. The most common relativization strategy in Standard Cantonese has the configuration of REL + DIST.DEM + CL + (N) where DIST.DEM is the distal demonstrative 嘅 *kɔ²⁵*. The proximal demonstrative 呢 *ni⁵⁵* cannot be used in this construction, e.g. *空呢間屋 **hʊŋ⁵⁵ ni⁵⁵ kan⁵⁵ ʊk⁵*. In summary, ADJ + CL + N noun phrases are rare in Standard Cantonese.

Nanning Pinghua is the same as in Standard Cantonese in that ADJ + CL + N noun phrases are rare, and one way to resolve this is to insert a demonstrative between the adjective and the classifier: ADJ + DEM + CL + (N), again making the adjective part of a relative clause. The difference with Standard Cantonese is that in Nanning Pinghua, both the neutral demonstrative 個 *kə⁵⁵* and the distal demonstrative 嘅 *ji²²* can be used. In the following examples, it is ungrammatical to leave out the demonstratives.

Language contact in Nanning: Nanning Pinghua and Nanning Cantonese 181

Notice that 的 *tik*⁵ is a comparative suffix of the adjective, e.g. 細的 *tei*⁵⁵-*tik*⁵ (small-COMPR) ‘smaller’, similar to 啟 *ti*⁵ in Cantonese, which also functions as the non-singular classifier, as in Pinghua. It is not the Mandarin modifier marker 的 *de*.

Nanning Pinghua

(23)	細的	个	間	房	我	住,
	<i>tei</i> ⁵⁵ - <i>tik</i> ⁵	<i>kə</i> ⁵⁵	<i>kan</i> ⁵³	<i>fʊŋ</i> ¹¹	<i>ŋa</i> ¹³	<i>tʃəi</i> ²² ,
	small-COMPR	DEM	CL	room	1SG	stay
	大的	个	間	房	你隊	住,
	<i>taɪ</i> ²² <i>tik</i> ⁵	<i>kə</i> ⁵⁵	<i>kan</i> ⁵³	<i>fʊŋ</i> ¹¹	<i>nəi</i> ¹³ <i>tɔɪ</i> ²²	<i>tʃəi</i> ²² ,
	big-COMPR	DEM	CL	room	2PL	stay
	最大	二	間	房	系	老師 住。
	<i>tʃui</i> ⁵⁵ - <i>taɪ</i> ²²	<i>ni</i> ²²	<i>kan</i> ⁵³	<i>fʊŋ</i> ¹¹	<i>hei</i> ²⁵	<i>lau</i> ¹³ <i>tei</i> ⁵³ <i>tʃəi</i> ²² .
	SUPL-big	DIST.DEM	CL	room	give	teacher stay

‘I stay in the smaller room, you stay in the larger room, let the teacher stay in the largest room.’

Mandarin is similar in that ADJ + CL + N noun phrases are rare. In comparison with Nanning Pinghua and Standard Cantonese which have the ADJ + DEM + CL + (N) construction, Mandarin requires an extra modifier marker between the adjective and the demonstrative: ADJ + MOD + DEM + CL + (N). The adjective is again in a relative clause which also requires a modifier marker in Mandarin. For example: 大的那個房間 *dà de nà ge fángjiān* (big MOD that CL room) ‘the room which is big’.

On the other hand, in Nanning Cantonese, ADJ + CL + (N) noun phrases are quite common. The following are two examples. (Recall that the equivalents in Standard Cantonese would require a distal demonstrative between the adjective and the classifier.)

Nanning Cantonese

- (24) 黃 色 支 筆 有 寫 得 咳, 黑 色 支 重 得。
*wɔŋ*¹¹ *fik*⁵ *tfi*⁵⁵ *pət*⁵ *mu*¹³ *te*²⁵ *tek*⁵ *tai*/³³, *hək*⁵ *fik*⁵ *tfi*⁵⁵ *tʃor*²² *tek*⁵.
yellow colour CL pen NEG write can PRF black colour CL still can
‘The yellow pen is unusable, the black one can still be used.’ (L&Q 2008: 278)
- (25) 嫣糊 高 隻 男 恁 好 嘞 唷。
*ma*⁵⁵ *wu*¹¹ *ku*⁵⁵ *tfək*³ *nam*¹¹ *tʃət*²⁵ *hu*²⁵ *lək*⁵ *ke*³³.
quite tall CL male child very capable MOD
‘The rather-tall boy is very capable.’ (L&Q 2008: 277)

Compare the ADJ + CL + N construction in Nanning Cantonese with the CL + N + ADJ construction in Northern Zhuang.

Northern Zhuang

- (26) *gou ndaenj haeuj aen ranz laep-saengsaeng bae.*
1SG squeeze enter CL house dark-IDEO go
'I went into the pitch black house.' (Wei Jingyun and Qin Xiaohang 2006: 226)

The ADJ + CL + N template in Nanning Cantonese can be explained simply as the CL + N + ADJ template in Northern Zhuang with the adjective shifted to the front to fit the Sinitic requirement of having right-headed noun phrases. A question that one might ask is why the adjective is shifted to the front of the classifier (ADJ + CL + N) rather than between the classifier and the noun (i.e. the usual Sinitic word order of CL + ADJ + N). In fact, both the ADJ + CL + N word order and the CL + ADJ + N word order are present in Nanning Cantonese. Given that many Nanning Cantonese speakers are ethnic Zhuang people who shifted to speaking Nanning Cantonese within the last one or two generations, having the ADJ + CL + N word order is in fact understandable: the ADJ + CL + N template in Nanning Cantonese requires less change from the CL + N + ADJ template in Northern Zhuang in terms of surface adjacency of the constituents. It is also worth noting that in Northern Zhuang, the classifier always precedes the noun immediately.³¹ A Nanning Cantonese speaker from a Zhuang background might thus have a preference for the adjective not intervening between the classifier and the noun.

7.4.7 Anaphoric use of lone classifiers

Another interesting feature in Nanning Cantonese is that lone classifiers can be used as discourse anaphors (that is, they refer to previously mentioned referents). By 'lone classifier', I mean classifiers which occur without either a head noun or any modifiers such as a numeral or a demonstrative. This differs from the bare classifier, discussed in §7.4.4 and Chapter 5 in this volume, which forms a noun phrase with its head noun.

Nanning Cantonese

- (27) *啲 狗 我 中意 隻, 有 中意 隻, 隻 難睇 多。*
ti⁵⁵ keu²⁵ ηɔ¹³ tʃuŋ⁵⁵ ji³³ tʃɛk³, mu¹³ tʃuŋ⁵⁵ ji³³ tʃɛk³, tʃɛk³ nan¹¹ tʰpi²⁵ tɔ⁵⁵.
CL:NSG dog 1SG like CL NEG like CL CL ugly too
'The dogs, I like this one, I do not like that one, that one is too ugly.' (L&Q 2008: 277)

This is parallel to the anaphoric use of lone classifiers in Zhuang.

³¹ In Northern Zhuang, within a noun phrase, all noun modifiers follow the noun except for the classifier and numerals other than one, e.g. *aen ranz ndeu* (CL house one), *song aen ranz* (two CL house), *sam aen ranz* (three CL house).

Language contact in Nanning: Nanning Pinghua and Nanning Cantonese 183

Northern Zhuang

- (28) *mwngz bi bi ndaem faex, go baenzlawz ha?*
2SG year year plant tree CL how Q
‘[Y]ou plant trees every year, how are they doing?’
(Sio and Sybesma 2008: 191, quoting Qin Xiaohang 1995: 83)
- (29) *mwngz dawz duz ma de daeuj hawj gou, gou cawz duz.*
2SG take CL dog that come give 1SG 1SG buy CL
‘[Y]ou bring that dog to me, I’ll buy it.[.]’
(Sio and Sybesma 2008: 191, quoting Qin Xiaohang 1995: 83)

In Nanning Pinghua, Standard Cantonese and Mandarin, lone classifiers cannot be used preverbally as an anaphor. In these languages, lone classifiers can only exist in a postverbal position, and they have an indefinite ‘one’ interpretation (a numeral ‘one’ can be inserted in front of the classifier with no change in semantics). This postverbal indefinite ‘one’ use of lone classifiers is also present in Nanning Cantonese and Northern Zhuang.

Nanning Pinghua

- (30) 買 (一) 隻 系 佢 欣 哟!
mai¹³ (y¹t³) tʃət³ hei²⁵ kəi¹³ hen⁵³ la⁵⁵

Standard Cantonese

- (31) 買 (一) 隻 界 佢 玩 哟!
mai¹³ (j¹t⁵) tʃək³ pet²⁵ kʰey¹³ wan²⁵ la⁵⁵

Standard Mandarin

- (32) 買 (一) 隻 紿 他 玩 吧!
mǎi (yī) zhī gěi tā wán ba
buy one CL give 3SG play HORT
‘Buy one for him/her to play with!’

Nanning Cantonese

- (33) 你 打 死 我 隻 貓, 着 賠 翻 隻 我。
ni¹³ ta²⁵ ɿi²⁵ y¹ŋ¹³ tʃək³ məu⁵⁵, tʃək² pʰui¹¹ fan⁵⁵ tʃək³ y¹ŋ¹³.
2SG hit die 1SG CL cat need compensate back CL 1SG
‘You killed my cat, you need to compensate me by getting me another one.’
(L&Q 2008: 351)

Northern Zhuang

- (34) *neix lij miz geij duz bit, gou aeu duz.*
this still exist few CL duck 1SG want CL
‘[T]here are still some ducks here, I want one.[.]’
(Sio and Sybesma 2008: 191, quoting Qin Xiaohang 1995: 84)

7.4.8 *Ditransitive word order for ‘give’*

In the preceding subsections we have seen some syntactic differences between Nanning Pinghua and Nanning Cantonese that were caused by varying degrees of Zhuang influence. In this subsection we will discuss one example where the Zhuang influence is often thought of as clearly evident, but is in fact much less direct.

Cantonese and Nanning Pinghua have different ditransitive word orders for the verb ‘give’. ‘Ditransitive’ here refers to cases where both the objects are unmarked. Nanning Pinghua has the order of *give + recipient + theme*. This word order is often associated with Northern Chinese. Nonetheless, this word order is also found in Southern China and Southeast Asia, e.g. Southern Min (Zhang Min 2011) and Vietnamese (Nguyễn Đình-Hoà 1997: 115). The following is an example from Nanning Pinghua.

Nanning Pinghua

- (35) 系 佢 錢， 佢 就 抓 去 賭。
hei²⁵ kai¹³ tʃin¹¹, kai¹³ tʃəu²² ja⁵³ hei²⁵ tu³³.
give 3SG money 3SG then take go gamble
'If you give him/her money, s/he will take it to gamble.'

On the other hand, Cantonese has the *give + theme + recipient* ditransitive word order, as exemplified by the following example from Nanning Cantonese.

Nanning Cantonese

- (36) [給/ 界] 500 文銀 我 媽 過年。
[kei⁵⁵/ pi²⁵] ɿ¹³ pak³ men⁵⁵-yən¹¹ ɿɔ¹³ ma⁵⁵ kɔ³³nin¹¹.
give give five hundred yuan-money 1SG mother celebrate:New:Year's
'I gave my mother five hundred yuan for New Year's.' (L&Q 2008: 351)

The theme-recipient word order in Cantonese is often attributed to a Tai influence (Huang Yuanwei 1997: 72–3; Li Jinfang 2002: 117). However, Zhuang actually has both the *theme-recipient* and the *recipient-theme* word orders.

Wuming Northern Zhuang

- (37) te²⁴ hau⁵⁵ kau²⁴ θo η²⁴ ɿan²⁴ ma k³⁵.
3SG give 1SG two CL fruit
'S/he gives me two pieces of fruit.' (Zhang Junru et al. 1999: 423)
- (38) muŋ³¹ hau⁵⁵ ciən³¹ho η²⁴ kau²⁴ luət²⁴?
2SG give wage 1SG Q
'Will you give me wages?' (Zhang Junru et al. 1999: 423)

Language contact in Nanning: Nanning Pinghua and Nanning Cantonese 185

Standard Northern Zhuang

- (39) *gou hawj mwngz bonj saw he.*
1SG give 2SG CL book then
'I'm giving you one book (i.e. not two).' (Luo Liming et al. 2005: 623)

Wénsān Hēimò 文山黑末 Southern Zhuang

- (40) *jv³³ kau¹¹² sā¹¹² bryu⁴² sz¹¹².*
give 1SG three CL book
'Give me three books.' (Zhang Junru et al. 1999: 424)
- (41) *jv³³ sz¹¹² kau¹¹² sā¹¹² bryu⁴².*
give book 1SG three CL
'Give me three books.' (Zhang Junru et al. 1999: 424)

Looking at the ditransitive word order in some other Kra-Dai languages, there are languages like Bouyei (i.e. the continuation of Zhuang in Guizhou to the north; Yu Cuirong 2009: 131), and Mulam (Wang Jun et al. 2009: 599) with the Mandarin-like *recipient-theme* word order, whereas languages like Lakkja (Liu Baoyuan 2009: 267) and Thai (Thepkajana 2010: 410) have the Cantonese-like *theme-recipient* word order. Kam is described by Long Yaohung (2003: 164) as having the Mandarin-like *recipient-theme* word order, whereas by Liang Min (2009a: 208) as having the Cantonese-like *theme-recipient* word order. Similarly, Sui is described by Zhang Junru (2009: 523) as having the Mandarin-like *recipient-theme* word order, whereas by Li Jinfang (2002: 117) as having the Cantonese-like *theme-recipient* word order. In addition, many of these Kra-Dai languages and others like Maonan (Liang Min 2009b: 674) and Lao (Enfield 2007: 363–6) have a variety of constructions to convey 'give', with one common strategy being the serial verb construction in the configuration of *give + theme + give + recipient*, which in languages like Lao and Thai is in fact less semantically restricted than the ditransitive *give + theme + recipient* construction. In Nanning Pinghua as well, there is more than one 'give' construction: impressionistically, the serial verb 'give' construction in the word order *give + theme + give + recipient* is more common than the ditransitive 'give' construction (*give + recipient + theme*). The following is an example of the serial verb 'give' construction.

Nanning Pinghua

- (42) *個 隻 男 孩子 呢 系 了 幾 隻 果 系 佢。*
kə⁵⁵ tʃət³ nam¹¹ hai¹¹ tʃi³³ ne⁵⁵ hei²⁵ liu¹³ kəi³³ tʃət³ ku³³ hei²⁵ kəi¹³.
DEM CL male child TOP give PFV few CL fruit give 3SG
'The boy then gave a few pieces of fruit to him.'

As there are a variety of constructions in various word orders to convey 'give' (and other trivalent events) in Kra-Dai languages, and it is possible that either of the

theme + recipient or *recipient + theme* word orders could be due to a Chinese influence, the *theme-recipient* word order in Cantonese cannot be straightforwardly attributed to Kra-Dai influence.

In fact, the ditransitive ‘give’ construction (*give + theme + recipient*) in Cantonese (and many other Southern Sinitic languages) is probably an internal development. Looking at earlier documents of Cantonese from the nineteenth and early twentieth centuries, the serial verb ‘give’ construction (*give + theme + ‘to’ + recipient*) appeared earlier than the ditransitive ‘give’ construction (*give + theme + recipient*). As there were no alternatives to the serial verb ‘give’ construction in competition for expressing ‘give’ in earlier Cantonese, the coverb ‘to’ had the opportunity to be elided to create the ditransitive ‘give’ construction (*give + theme + recipient*). Note that having a serial verb ‘give’ construction does make Cantonese resemble Tai languages. Also notice that in Cantonese, only the ‘giving’ type of ditransitive sentences has the order of *verb + theme + recipient*; for other ditransitive sentences like ‘steal’ or ‘teach’, the order is the ‘usual’ Sinitic word order of *verb + recipient + theme*. (For the development of the ditransitive ‘give’ construction in Cantonese, see, e.g. Peyraube and Xu 1997, Phua 2007, Yiu 2010, and Chin 2010; for this development in Southern Sinitic languages in general, see Zhang Min 2011.)

It is interesting that Nanning Pinghua and Nanning Cantonese have different ditransitive word orders for ‘give’. While, the ditransitive ‘give’ construction (*give + theme + recipient*) in Cantonese cannot be directly attributed to Tai influence, the serial verb ‘give’ construction (*give + theme + give + recipient*) in both Cantonese and Nanning Pinghua is at least partially Tai inspired (see Zhang Min 2011). See also §3.6 in Peyraube, this volume.

A summary of the grammar points discussed in this section is provided in §7.5.

7.5 Conclusion and discussion

Nanning is a multilingual area: the Sinitic languages of Nanning Pinghua, Nanning Cantonese, Old Nanning Mandarin, and New Nanning Mandarin are spoken alongside the indigenous Tai languages of Northern Zhuang and Southern Zhuang. In this analysis, we have looked at some of the outcomes of this language contact situation from the viewpoint of Nanning Pinghua and Nanning Cantonese, the two largest Sinitic languages spoken in the area.

On the whole, Nanning Pinghua and Nanning Cantonese share many similarities. This is particularly true of their phonologies, which has led to the widely held view within China these days that Pinghua, or Southern Pinghua at least, is a branch of Yue (which includes Cantonese) (§7.1). There are also some differences; in the preceding sections, we have discussed some of the differences between Nanning Pinghua and Nanning Cantonese, and also some of the differences that they both have with Standard Cantonese, Northern Zhuang, and Mandarin. Sometimes the differences between Nanning Pinghua and Nanning Cantonese are simply natural

Language contact in Nanning: Nanning Pinghua and Nanning Cantonese 187

variations that exist amongst Sinitic languages; for instance, with respect to vocabulary, where Nanning Pinghua has many words which resemble Mandarin rather than Cantonese (§7.3). Sometimes Nanning Pinghua has Zhuang-like features, which Nanning Cantonese lacks, for instance, the numerous Zhuang loanwords in Nanning Pinghua not found in Cantonese (§7.3). This is predictable, since Nanning Pinghua has been spoken alongside Zhuang for at least one millennium, whereas Nanning Cantonese only arrived in the area about 150 years ago. Nonetheless, in most of this chapter we have shown that both Nanning and Standard Cantonese have many more Zhuang-like features than does Nanning Pinghua. Examples raised are:

- the splitting of the Entering tone based on vowel length (§7.2);
- gender suffixes for animals (§7.4.1);
- lack of dispreference for monosyllabic nouns (§7.4.2);
- adverbials like ‘first’ occurring postverbally (§7.4.3);
- having preverbal [CL + N] noun phrases (§7.4.4); and
- using a classifier to ‘link’ a possessor phrase and the possessum noun (§7.4.5).

In addition, we have seen that Nanning Cantonese has diverged from Standard Cantonese due to further influence from Zhuang (i.e. Nanning Cantonese has Zhuang-like features that are neither found in Nanning Pinghua, Standard Cantonese, nor in Mandarin). The examples we have seen are:

- possessing [ADJ + CL + N] noun phrases (§7.4.6); and
- anaphoric use of lone classifiers (§7.4.7).

We have also seen in §7.4.8 that Nanning Pinghua and Cantonese have different ditransitive word orders for ‘give’. The *give + theme + recipient* ditransitive word order for ‘give’ in Cantonese is often attributed to a Zhuang influence, but we have seen in this discussion that this matter is not as straightforward as commonly thought.

The fact that Nanning Cantonese often resembles Zhuang more than Nanning Pinghua does requires explanation. There are two reasons for this. The first reason is that Cantonese itself already has a strong Tai substratum (Ouyang Jueya 1989; Li Jingzhong 1994; Bauer 1996; Huang Yuanwei 1997; Li Jinfang 2002: 100–41), as mentioned in §7.1.

The second reason is that after the Cantonese people’s arrival in Nanning, there has been massive language shift from Zhuang to Nanning Cantonese. Aspects of this language transferral from Zhuang to Nanning Cantonese have been discussed in, e.g. Kwok Bit-Chee (2010) and Qin Fengyu and Wu Fuxiang (2009). On the other hand, while Pinghua people have gradually accumulated many Zhuang loanwords and customs during the millennium of their presence in Guangxi, e.g. Pinghua shamanism, 師公 *tei⁵³ koy⁵³*, shares many similarities with Zhuang shamanism, Pinghua people have nonetheless traditionally kept a social distance from the Zhuang

population. Pinghua people are known for the conservatism amongst the various ethnic groups in Guangxi (Xu Jieshun 1999). For instance, before the 1950s, inter-marriage between Pinghua and Zhuang people in Nanning area was rare, and Pinghua people have largely excluded Zhuang people from their most important industry, which is the planting and processing of sugarcane (e.g. Zhu Zhiyan 2004 portrays a rather unfriendly-type of relationship between the Sugarcane (i.e. Pinghua) people and the Zhuang people). The relatively distant relationship that the Pinghua communities had with Zhuang people has probably contributed to the slower rate of grammatical influence from Zhuang to Pinghua, at least in areas like Nanning where the concentration of Pinghua speakers is higher. (The implication of having larger Pinghua-speaking communities is that Pinghua people in those areas could afford to maintain larger social distance from the indigenous Zhuang majority.) In addition, the arrival of the prestigious Mandarin language in Nanning during the Ming dynasty has perhaps pulled the grammar of Pinghua somewhat more towards the Mandarin grammatical profile.

Speakers' attitude is an important social factor in contact situations (Thomason 2010: 38–9; Thomason 2001a: ch.4; Thomason 2001b; Fought 2010; Goméz Rendón 2008: ch.2; amongst others). The attitude of speakers towards the languages in contact, or the contact situation itself, often exaggerates the rate of contact-induced linguistic change (or retention). Speakers' attitude also can often produce unusual results in relation to the 'normal' outcomes of language contact caused by other social factors such as the intensity of contact, or linguistic factors such as typological distance. Nonetheless, speakers' attitude is also one of the less explored factors of language contact. This is due to the difficulty in formulating the notion of 'speakers' attitude': the notion of 'speakers' attitude' covers a wide range of phenomena, and many of which are subtle and difficult for outsiders to observe. The linguistic outcome of this 'attitude' is even more difficult to predict than the other linguistic and social factors that are relevant in language contact situations. In this analysis, an attempt has been made to explain the unexpected difference in the kinds of contact-induced change in Nanning Pinghua and Nanning Cantonese, based on the speakers' attitude towards the other languages in the area. It is hoped that a positive contribution can be made towards the study of the role of speakers' attitude towards language contact situations.

In conclusion, we have seen that Nanning Pinghua and Nanning Cantonese, two (of the three) Sinitic languages in Nanning area, are both influenced by the indigenous Zhuang languages. Interestingly, however, the grammar of Nanning Cantonese, which has been spoken in the area for about 150 years, resembles the indigenous Zhuang languages more than Nanning Pinghua, which has been spoken in the area for at least one millennium. This probably has to do with the massive language shift from Zhuang to Nanning Cantonese, and also the Pinghua people's more conservative approach towards interactions with Zhuang people.

Acknowledgements

Various topics discussed in this chapter formed the body of talks given at the following conferences: *Syntax of the World's Languages IV*, Dynamique du Langage, CNRS and Université Lumière-Lyon 2, 2010; *Humanities of the Lesser-Known—New Directions in the Descriptions, Documentation, and Typology of Endangered Languages and Musics*, Lunds Universitet, 2010; 第五屆漢語方言語法國際研討會 [*The Fifth International Conference on the Grammar of Chinese Dialects*], 上海大学 Shanghai University, 2010; *Southeast Asian Linguistics Society Conference 21*, Kasetsart University, 2011; and *Workshop on Ecology, Population Movements, and Language Diversity*, Université Lumière-Lyon 2, 2011.

I would like to thank all who attended my talks on these topics and provided me with valuable comments. I would also like to thank all of my Nanning Pinghua informants, and my main informant 梁世華 *lɛŋ¹¹ t̪i⁵⁵wa¹¹*/ Liáng Shíhuá in particular, for teaching me their language(s). I have learnt a great deal from all the linguists that I met in Guangxi, 林亦 Lin Yi and 覃鳳餘 Qin Fengyu and 田春來 Tian Chunlai of Guangxi University in particular. My colleagues in the SINOTYPE team have given me much support and many valuable comments; I would like to thank all of them, and especially our director, Prof. Hilary Chappell.

On the origin of special numerals for ‘one’ in southeastern China: [kei²¹³] in the northwestern Min dialect of Shaowu

SING SING NGAI

8.1 Introduction

Classifier systems and their extended functions have been the topic of much investigation in recent decades (Greenberg 1977, 1978; Craig 1986; Croft 1994; Aikhenvald 2000 among others). While Sinitic languages are well known for their perplexingly hybrid typological features (see, for example, Chappell 2001; Chappell, Li, and Peyraube 2007), extensive research has nonetheless been carried out on Sinitic noun phrases involving classifiers, in terms of their historical development (cf. e.g. Peyraube 1991, 1998; Wang 1994; Yang-Drocourt 2004) and their multifarious functions (cf. e.g. Sun 1989; Shao 1993; Tai 1994; Ahrens and Huang 1996; Bisang 1999; Xu and Matthews 2011).

In this chapter, we set out to unravel the interrelationship between the Shaowu numeral ‘one’ [kei²¹³] and the general classifier [kei²¹³], represented by the same grapheme 个 in the literature, which is found in the Shaowu dialect (hereafter referred to as ‘Shaowu’) spoken in north-western Fujian province in China.¹ To this end, we first delineate the numeral usage of [kei²¹³] in various Shaowu numeral classifier constructions. Next, a hypothesis for its formation will be put forward, supported by diachronic material and synchronic data. To further our argument, the

¹ The Shaowu data used in this chapter are mostly extracted from natural conversations or stories told by native speakers in their seventies who have lived in the city of Shaowu for most of their lives. Data from dialectal localities other than Shaowu are retrieved from reliable publications, while diachronic data are taken from classical texts and historical documents.

On the origin of special numerals for ‘one’ in southeastern China 191

typo-geographical distribution of similar phenomena occurring in the region will be described.

The presentation is organized in the following way:

- (1) Historical background and typological features of the Shaowu dialect
- (2) Numeral paradigms and constructions involving ‘one’ in Shaowu
- (3) Typological distribution of different Sinitic morphemes for ‘one’
- (4) Hypotheses for the genesis of the numeral function of [kei²¹³]
- (5) Implications: human cognition and origin of basic numerals

8.2 Setting the stage

8.2.1 Historical background

The Shaowu dialect belongs to the Shaojiang² sub-group (邵将区) of the Min supra-group³ (cf. e.g. Pan et al. 1963; Wurm et al. 1987: 18; Zhou and Ouyang 1998) of the Sinitic branch of the Sino-Tibetan family, and is named Far Western Min by Norman (1988). It is spoken by approximately 300 000 speakers, with a varying degree of fluency depending on age groups, in the city of Shaowu and its environs in north-western Fujian province. The city was established as a county in as early as the Three Kingdoms period (c. AD 260). Historically, it has mainly been administered by Jianzhou (建州) prefecture in northern Fujian (Northern Min speaking), as well as being intermittently governed by Fuzhou (福州) prefecture in eastern Jiangxi province (Gan-speaking) for almost two hundred years. Most recently, it has come under the jurisdiction of Nanping (南平) municipality since the 1980s.

Shaowu has over the centuries seen massive waves of migration from the Gan-speaking Jiangxi province and the Wu-speaking Zhejiang province. Historical records show that migration peaked in the Eastern Jin (c. fourth century), Mid Tang (c. ninth century) and Southern Song (c. twelfth century) dynasties, as a consequence of invasions, civil wars, famine, and plagues. The outbreak of the Sino-Japanese war (1937–45) also pushed the coastal inhabitants of Fujian, notably those of Fuzhou (福州), towards this less war-ridden inland area. Dialects brought in by immigrants from all corners of China enriched and diversified Shaowu’s phonological, lexical, and morpho-syntactic make-up.

The diverse group of Min languages spoken in Fujian, Hainan, and Taiwan are represented in Map 8.1 in addition to certain co-territorial Hakka and Mandarin varieties.

² The term ‘Shaojiang Group’ (邵将区) refers to the dialect localities around Shaowu (邵武) and Jiangle (将乐).

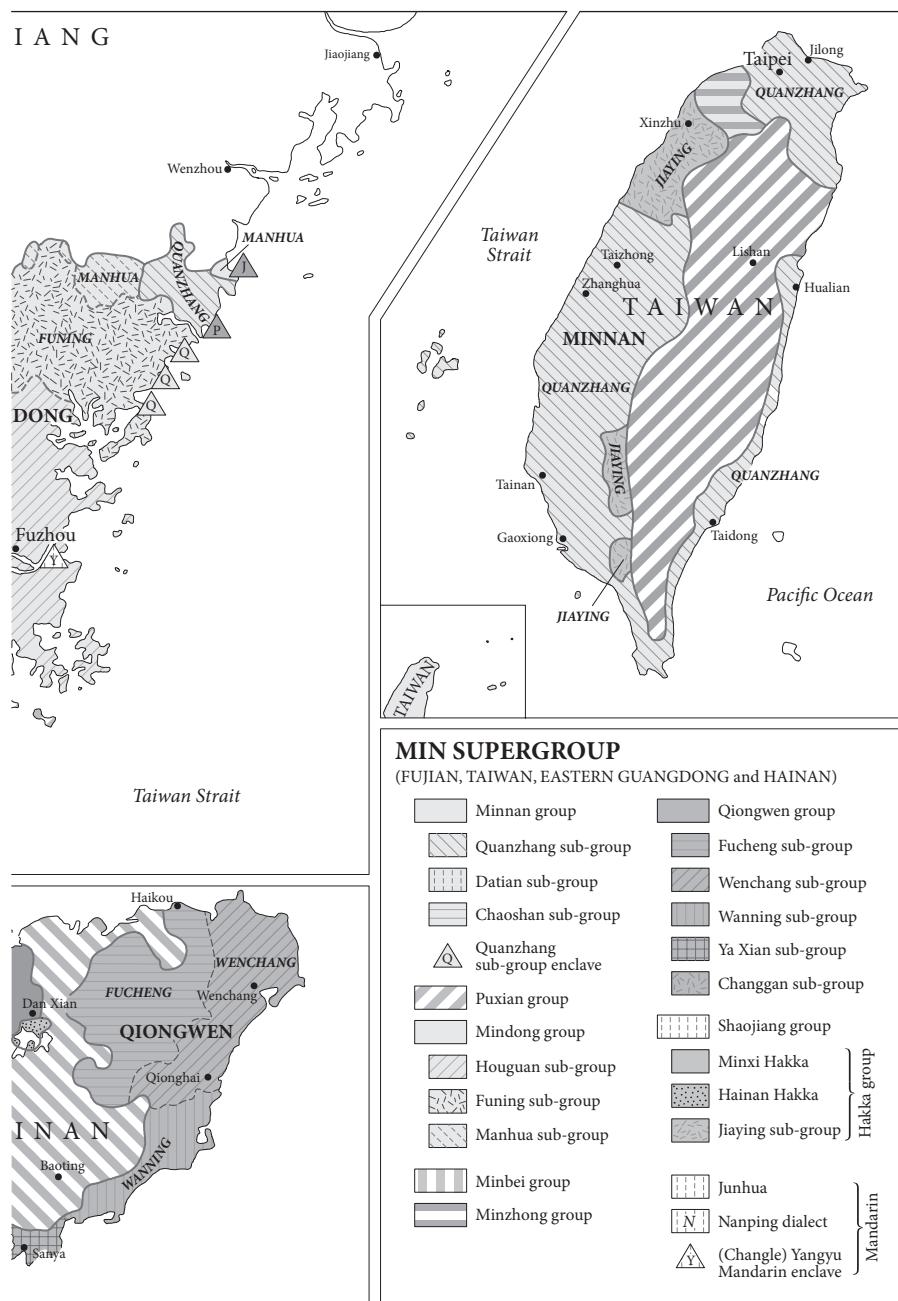
³ The Min supra-group is divided into various sub-groups essentially based on phonological and lexical criteria by Sinitic linguists.



MAP 8.1 Min-speaking areas in southeastern China and location of the city of Shaowu (starred)

Redrawing based on reproduction of Map B-12, Language Atlas of China, Wurm et al., 1987

On the origin of special numerals for ‘one’ in southeastern China 193



8.2.2 Typological features of Shaowu

The Shaowu dialect is a putative transitional language, which possesses a mix of features from the Min and Gan dialect groups in terms of its phonology, lexicon, and morphosyntax (cf. Yuan 1989; Chen and Li 1991, among others). It has a basic word order of SVO. Tone sandhi phenomena are not profuse, and when present, are typically governed by prosodic considerations with no grammatical or lexical value. Morphologically, derivational processes such as affixation, reduplication, and compounding are common, which can serve as word-class changing strategies, including nominalization and adjективization.

Like most Sinitic languages, Shaowu noun phrase structure displays predominantly pre-nominal modification. The linear order of noun phrase elements is shown below:

$$(\text{PRO}) - (\text{GEN}) - (\text{DEM}) - \text{NUM-CL} - (\text{ATTR}) - (\text{REL}) - (\text{LW}) - \text{N}^4$$

8.2.3 Phenomenon under investigation—Shaowu [kei²¹]: numeral and/or classifier?

The Shaowu [kei²¹³] (or [kei²¹] in sandhi form) is idiosyncratic in the fact that not only can it be used as the general classifier for humans and objects, for instance, [san²¹ kei²¹ nin²²] 三个人 (three-CL_{general}-person) ‘three people’, but it can also express the notion of singularity when appearing in the numeral slot in the [Num+CL+N] construction, on condition that the co-occurring classifier is not [kei²¹³] itself. For instance, [kei²¹ pən⁵⁵ cy²¹] 个本书 (one-CL-book) means ‘one book’. As [kei²¹³] is the phonetic realization for both the numeral ‘one’ and the general classifier in Shaowu, the graphemic representation used is simply [个] in the literature, which appears to invoke a strong etymological link between the pan-Sinitic general classifier, including that of Shaowu, and the Shaowu numeral ‘one’ [kei²¹³]. Nonetheless, we will maintain that these two morphemes are synchronically homonyms bearing different grammatical functions, although they arose from the same lexical source as a result of lexical conversion.

In the *Linguistic Atlas of Chinese Dialects –Volume 2: Lexicon* (Cao et al., ed. 2008: 191), only two morphemes for numeral ‘one’ have been identified in over 900 dialect localities spreading across China. They are [IT]⁵ — and [SOK]⁶ 蜀. There is no

⁴ N.B. Given the right syntactic condition, [kei²¹³] can fill any of the slots of [GEN], [NUM], [CL], and [LW].

⁵ [IT] is an abstract phonetic representation for the pan-Sinitic numeral ‘one’, for which we drew inspiration from the different MC reconstructions made by linguists such as: Dong Tonghe’s [*jet], Edwin Pulleyblank’s [*jit], ZhengZhang Shangfang’s [*iIt], and Pan Wuyun’s [*it]. (See also note 12).

⁶ [SOK] is an abstract phonetic representation for the Min-specific numeral ‘one’, for which we drew inspiration from the different MC reconstructions made by linguists such as: Li Rong’s [*üiok], Dong Tonghe’s [*üjuok], Edwin Pulleyblank’s [*düiok], and ZhengZhang Shangfang’s [*düöök]. (See also note 12).

On the origin of special numerals for ‘one’ in southeastern China 195

mention of [kei²¹³] 个, or its cognates, employed as ‘one’ in any linguistic maps of China so far published. The coming to light of other morphemes denoting the concept of ‘one’ in Sinitic thus deserves research attention.

In Sinitic, the numeral ‘one’ is predominantly the [IT] form and its cognates, except in some Min dialects where [SOK] forms are used, likely a shared innovation from Proto-Min (Laurent Sagart pers. comm.), whereas in its sister branch, Tibeto-Burman, the [TSI-] or [TE-] forms (or similar) are pervasive.⁷ The apparent absence of possible cognate forms for [KA] in Sino-Tibetan leads us to propose that the [KA]-form for numeral ‘one’ is a local innovation of the Shaowu area and environs.⁸

8.2.4 On the various morphemes for numeral ‘one’

The numeral ‘one’ [IT] — made its first documented appearance on the oracle bones (Djamouri 1994) some 3500 years ago. Today, it is the pan-Sinitic numeral ‘one’ used in counting and in numeral classifier constructions in most Sinitic dialects (Mandarin, Jin, Gan, Kejia (Hakka), Hui, Xiang, Yue groups, etc.) (Cao et al. 2008: Map 191 in *Lexicon volume*).

The numeral ‘one’ [SOK] 蜀 is a specifically Min feature. It is only found in areas covered by the Min dialect group, being concentrated in the area along the coast of Fujian and northern Guangdong, and to some extent in those Min dialects spoken on Hainan and Taiwan. Some Min dialects allow free alternation between [IT] — and [SOK] 蜀. In counting, however, [SOK] 蜀 has to be used in numeral classifier constructions, as, for instance, in the Fuzhou dialect spoken on the eastern coast of Fujian.

The numeral ‘one’ [KA] 个 has only been identified in a confined, contiguous region encompassing northwest Fujian, northeast Jiangxi, and southwest Zhejiang provinces. The specific phonetic form in Shaowu is realized as [kei²¹³]. We will hereafter refer to it as the indigenous numeral ‘one’ for those [KA]-form numerals found around northwestern Fujian and environs.

⁷ For instance, the numeral ‘one’ in Central Tibetan is [tɕi?]⁵³, while in Burmese, it is [ti?]⁹⁴.

⁸ [KA] is an abstract phonetic representation for the indigenous numeral ‘one’ in Shaowu (and surrounding areas), for which we drew inspiration from the different MC reconstructions made by linguists such as: Bernhard Karlgren’s [*kA], Wang Li’s [*kA], Li Rong’s [*kA], Edwin Pulleyblank’s [*kA], and ZhengZhang Shangfang’s [*kA] (1987). (See also note 12).

8.3 Numeral paradigms and constructions in Shaowu

There are three morphemes for numeral ‘one’ in the Shaowu dialect, namely, the pan-Sinitic [IT] [i⁵³] 一, the Min-specific [SOK] [ci²¹] 蜀 and the indigenous [KA] [kei²¹³] 个. As the phenomenon under investigation is essentially morphemes for numeral ‘one’ in Shaowu, the following discussion will mainly focus on numeral paradigms and constructions that involve the usage of ‘one’.

8.3.1 Cardinal and ordinal numbers

Cardinal numerals are used in counting as a specific process of ascertaining numerosity by enumerating a sequence of lexemes, whereas ordinal numerals identify the position of a given number in a set relative to other members in the set (e.g. ‘the third book he wrote’). Cardinal and ordinal numeral paradigms in the Shaowu dialect are shown in Tables 8.1 and 8.2 respectively.⁹

TABLE 8.1. Some cardinal numbers in Shaowu

Shaowu cardinal numbers	Written form	Gloss
[i ⁵³] / [kei ²¹]	一 / (个)	‘one’
[ni ³⁵] / [lion] ⁵⁵	二 / (两)	‘two’
[san ²¹]	三	‘three’
[cin ³⁵]	十	‘ten’
[kei ²¹ pa ⁵³]	*(个)百	‘a hundred’
[kei ²¹ pa ⁵³ len ²¹ i ⁵³]	*(个)百零一	‘a hundred and one’
[kei ²¹ t ^h ien ²²]	*(个)千	‘a thousand’
[kei ²¹ van ³⁵]	*(个)万	‘ten thousand’

Notice that in counting involving the use of the cardinal numbers of ‘one’ and ‘two’, [kei²¹³] 个 and [lion]⁵⁵ 两 are used on a par with [i⁵³] 一 and [ni³⁵] 二, although the latter are more prevalent today, especially among the younger

⁹ Notes on the transcriptions and conventions used in this chapter:

- (i) IPA symbols together with tonal notations are used for Shaowu data and surrounding dialects for cross-dialectal comparisons.
- (ii) For those morphemes not serving particular comparison purposes but shared by a wide array of dialects across the country, we use a mix of IPA symbols and the Standard Chinese *pīnyīn* romanization in capital letters to give a general or abstract phonetic representation of the morpheme, after taking into account their respective Old Chinese (OC) and Middle Chinese (MC) reconstructed forms that have been proposed by various historical linguists.
- (iii) For sentences retrieved from classical texts whose exact pronunciations are unknown to us, the Standard Chinese *pīnyīn* romanization in small capitals will be used to create an abstract phonetic representation in the main text, or in italics in the notes.

On the origin of special numerals for ‘one’ in southeastern China 197

generations who speak Shaowu. This suggests that [i⁵³] — and [ni³⁵] 二 could be later forms borrowed from other dialects, and most likely from Standard Mandarin which has exerted a great influence in the region since the 1950s. Note also that 个 [kei²¹] is used instead of — [i⁵³] to precede counting units such as ‘hundred’, ‘thousand’, and ‘ten thousand’, and is obligatory in use with these counting units.

This could be attributed to the fact that these counting units (i.e. hundred, thousand) are considered to be measure words, and that 个 [kei²¹³] thus fills the numeral slot in a [NUM-MW-N] noun phrase construction, much like its behaviour, as we are later to see, in [NUM-CL-N] noun phrases. And this further suggests that 个 [kei²¹] is a more basic form than — [i⁵³], as borrowing numbers wholesale for the purpose of counting is a simpler process than replacing a numeral element in a fixed construction. This may in turn suggest that [kei²¹³] 个 and [lioni⁵⁵] 两 are likely to be the more indigenous forms that are slowly being replaced by the pan-Sinitic YI — and ER 二 in counting, but still remain resistant to change in ordinal counting constructions, starting from the number ‘hundred’ onwards, as shown in Table 8.2:

TABLE 8.2. Some ordinal numbers in Shaowu

Shaowu ordinal numbers	TCH>Written form	Gloss
[t ^h i ³⁵ i ⁵³]	第一	‘first’
[t ^h i ³⁵ ni ³⁵]	第二	‘second’
[t ^h i ³⁵ san ²¹]	第三	‘third’
[t ^h i ³⁵ cin ³⁵]	第十	‘the tenth’
[t ^h i ³⁵ kei ²¹ pa ⁵³]	(第) ¹ 个百	‘the hundredth’
[t ^h i ³⁵ kei ²¹ pa ⁵³ len ²¹ i ⁵³]	(第)个百零一	‘the hundred and first’
[t ^h i ³⁵ kei ²¹ t ^h ien ²²]	(第)个千	‘the thousandth’
[t ^h i ³⁵ kei ²¹ van ³⁵]	(第)个万	‘the ten-thousandth’

¹ The ordinal suffix [t^hi³⁵] 第 can be optional for numbers above ‘ten’, if the context spells out information on the ordering, and there may also be prosodic reasons for this.

8.3.2 Numeral classifier constructions involving the use of ‘one’

In this section, we talk about the attributive quantification of nouns (e.g. three books), which involves the use of classifiers in most Sinitic languages including Shaowu, and we sketch the use of numeral ‘one’ in various Shaowu numeral classifier constructions. Although Sinitic nouns alone are bare nouns that bear no grammatical markings for number, case, gender, or definiteness, a syntactic requirement exists whereby a classifier is inserted between the determiner and the head noun. It is a bound morpheme which serves to individuate and categorize the noun,

198 Sing Sing Ngai

and in fact to ‘actualize’ it (see also McGregor 2002: 16–22 for a definition of classifiers in terms of distribution). Most, if not all, Sinitic languages code the quantity and semantic features of their nouns by the conjoined presence of a numeral and a classifier.

Four main construction types involving the use of classifiers in the Shaowu dialect are listed below:

- (i) Numeral + CL + Noun
- (ii) Ordinal prefix + Numeral + CL + Noun
- (iii) Demonstrative + (Numeral ‘one’) + CL + Noun
- (iv) Universal quantifier + (Numeral ‘one’) + CL + Noun

In the following sub-sections, we expose how the choice of numeral ‘one’ changes in interaction with other grammatical elements in a Shaowu classifier noun phrase.

8.3.2.1 *The [Numeral + CL+ Noun] construction*

According to Aikenvald (2000: 335–6, citing Zubin and Shimojo 1993), a general classifier has essentially three complex functions, all of which Shaowu [kei²¹³] 个 fulfills as the CL in the [Num-CL-N] construction (as shown in the following examples):

Function 1 states that a general classifier can have an unspecified referent, typically with a collective meaning. This can be illustrated by the following Shaowu example:

- (1) □ 这 个 唔 好。
tɕiong²¹ kei²¹ n̩⁵⁵ xau⁵⁵
this CL NEG good
'This is no good.'

Function 2 states that a general classifier serves as a fallback for residual nouns that do not conventionally take a sortal classifier, e.g. abstract nouns such as ‘idea’, ‘problem’, ‘wish’. This can be illustrated by the following Shaowu example:

- (2) □ 我 有 蜀 个 想法。
xan³⁵ iɔu⁵⁵ ci²¹ kei²¹ siɔŋ⁵⁵fan⁵³.
1SG have one CL idea
'I have an idea.'

Function 3 states that a general classifier serves as a default classifier (as shown in example (3) for Shaowu), one that can substitute for a sortal classifier (as shown in example (4) for Shaowu):

- (3) □ 他 有 三 个 笔。
xu³⁵ iɔu⁵⁵ san²¹ kei²¹ pi⁵³.
3SG have three CL_{GEN} pen
'He has three pens.'

On the origin of special numerals for ‘one’ in southeastern China 199

- (4) 口他 有 三 管 笔。
 xu³⁵ iou⁵⁵ san²¹ kuɔn⁵⁵ pi⁵³.
 3SG have three CL_{SPEC} pen
 ‘He has three pens.’

Examples (3) and (4) are exactly identical in terms of meaning, which show that [kei²¹³] 个 is in effect a general classifier, which can replace specific sortal classifiers and act as the default classifier.

Table 8.3 gives some examples of the Shaowu numeral classifier construction type (i), i.e. [NUM-CL-N]. The use of the Min-specific numeral ‘one’ [ci²¹] 蜀 is obligatory when preceding the general classifier [kei²¹³] 个 which is used to individuate human nouns, inanimate objects, often round and small, or abstract nouns:

TABLE 8.3. The [ci²¹_{Num} + kei²¹³_{CL} + N] construction in Shaowu

Human nouns			Inanimate objects/Abstract nouns		
蜀 ci ²¹	个 kei ²¹	人 nin ²²	蜀 ci ²¹	个 kei ²¹	想法 sion ⁵⁵ fan ⁵³
one	CL	person	one	CL	idea
‘one/a person’			‘one/an idea’		
蜀 ci ²¹	个 kei ²¹	学生 xo ³⁵ sən ²¹	蜀 ci ²¹	个 kei ²¹	书包 cy ²¹ pau ²¹
one	CL	student	one	CL	schoolbag
‘one/a student’			‘one/a schoolbag’		

Note that the following NUM-CL combinations are ungrammatical in Shaowu, when [ci²¹] 蜀 is not used.

- (5) * — 个 (学生)
 i⁵³ kei²¹ (xo³⁵sən²¹)
 one CL student
 (‘one student’)

- (6) * 个 个 (学生)
 kei²¹ kei²¹ (xo³⁵sən²¹)¹⁰
 one CL student
 (‘one student’)

¹⁰ While it is ungrammatical to use the Shaowu numeral ‘one’ [kei²¹³] 个 before the general classifier [kei²¹³] 个 to instantiate the notion of singularity, it is however grammatical to reduplicate the general classifier 个 [kei²¹³] in order to express universal quantification. Thus a noun phrase such as 个个学生 [kei²¹kei²¹xo³⁵sən²¹] is acceptable when meaning ‘each and every student’, but this is not the intended meaning here.

200 *Sing Sing Ngai*

Examples (7) and (8) illustrate the use of [kɛi²¹³] 个 as a general classifier in full sentences.

- (7) 口这 蜀 个 人 顶 好。
tɕiong⁵³ ci²¹ kɛi²¹ nin²² tin⁵⁵ xau⁵⁵.
DEM one CL_{gen} person ADV good
'This person is very kind.'

- (8) 口我 食 了 了 两 个 桃儿。
xan³⁵ cie³⁵ liau⁵⁵ ə⁰ liɔŋ⁵⁵ kɛi²¹ tʰau⁵³ ə⁰.
1SG eat finish PRF two CL peach
'I have eaten two peaches.'

Table 8.4 gives some examples of the Shaowu numeral classifier construction type (i) involving the obligatory use of the indigenous Shaowu numeral 'one' [kɛi²¹³] 个, preceding any classifiers other than the general classifier [kɛi²¹³] 个:

TABLE 8.4. The [kɛi²¹³_{Num} + CL + N] construction in Shaowu

Animate nouns			Inanimate nouns		
个	头	牛	个	架	车
kɛi ²¹	tʰəu ⁵³	ny ²²	kɛi ²¹	ka ²¹³	tçʰia ²¹
one	CL	ox	one	CL	car
'an ox'			'a car'		
个	只	猫儿	个	本	书
kɛi ²¹	tçia ⁵³	mau ⁵³ ə ⁰	kɛi ²¹	pən ⁵⁵	çy ²¹
one	CL	cat. DIM	one	CL	book
'a cat'			'a book'		

Note that the following NUM-CL combinations are ungrammatical in the Shaowu dialect, when [kɛi²¹³] 个 is not used to code 'one'.

- (9) * — 只 (猫儿)
i⁵³ tçia⁵³ (mau⁵³ ə⁰)
one CL cat
(‘a cat’)

- (10) *蜀 只 (猫儿)
ci²¹ tçia⁵³ (mau⁵³ ə⁰)
one CL cat
(‘a cat’)

Examples (11), (12), and (13) are quantity-specific statements which demonstrate the imperative of using [kɛi²¹³] as the numeral 'one' and not any of the other options.

On the origin of special numerals for ‘one’ in southeastern China 201

- (11) Question: 胡 买 了 几 栋 厝?
xu³⁵ mie⁵⁵ ə⁰ ki⁵⁵ tun²¹³ tç^hio³⁵
3SG buy PRF how.many CL_{block} house
'How many houses has he bought?'

Answer: 个 栋 厝。
kei²¹ tun²¹³ tçhiɔ³⁵
one CL_{block} house
'One house.'

- (12) 𠂇 口我 拿 个 管 笔， 没 是 两 管。
xan³⁵ ni³⁵ na³³ kei²¹ kuɔŋ⁵⁵ pi⁵³ mau³⁵ ci⁵⁵ liɔŋ³⁵ kuɔŋ⁵⁵.
1SG only take one CL_{straw} pen NEG be two CL_{straw}
'I only took one pen, not two.'

- (13) 这 口我 田, 个 头 牛 拉 不 动,
tçioŋ⁵³ ə⁰ t^hən⁵³ kei²¹ t^həu⁵³ ny²² xai²¹ ŋ³ t^huŋ⁵⁵
DEM CL field one CL_{head} ox pull NEG move
还 要 使 个 头 牛 带 个 下。
ai²¹ nun³⁵ sə²² kei²¹ t^həu⁵³ ny²² tai²¹³ kei²¹ xa²².
also must use one CL_{head} ox drag one VCL
'One ox alone cannot pull the plough, – one more is needed to pull it.'

8.3.2.2 The [Ordinal prefix + Numeral + CL + Noun] construction

Insofar as the construction type (ii) is concerned, the pan-Sinitic numeral ‘one’ [i⁵³] — is used after the ordinal prefix [t^hi³⁵] 第 (cf. Table 8.2 in §8.2.1):

- (14) 第 一 个 学生
t^hi³⁵ i⁵³ kei²¹ xo³⁵sən²¹
PREF one CL student
'the first student'

Note that it would be unacceptable to native speakers of Shaowu, if we replaced [i⁵³] — by [ci²¹] 蜀 or by [kei²¹³] 个 in the above construction.

- (15) *第 蜀 个 (学 生)
t^hi³⁵ ci²¹ kei²¹ (xo³⁵ sən²¹)
PREF one CL (student)
(‘the first student’)

- (16) *第 个 个 (学 生)
t^hi³⁵ kei²¹ kei²¹ (xo³⁵ sən²¹)
PREF one CL (student)
(‘the first student’)

8.3.2.3 The [Demonstrative + (Numeral ‘one’) + CL+ Noun] construction

Shaowu has a two-way distinction of demonstratives: proximal [t̪ciɔŋ²¹] and distal [ɔŋ²¹]. For construction type (iii) which involves the use of demonstratives, the general classifier [kei²¹³] 个 is employed, when human nouns, inanimate objects, or abstract nouns are involved. It precedes the head noun and follows the Min-specific numeral one [çɪ²¹] 蜀, although the latter can be omitted (Table 8.5):

TABLE 8.5. [DEM + çɪ²¹_{Num} + kei²¹³_{CL} + N] construction in Shaowu

Human nouns				Inanimate objects/Abstract nouns			
□这 (蜀) t̪ciɔŋ ²¹ DEM 'this person'	个 kei ²¹ CL	人 nin ²² person		□那 (蜀) ɔŋ ²¹ DEM 'that idea'	个 kei ²¹ CL	想法 siɔŋ ⁵⁵ fan ⁵³ idea	
□那 (蜀) ɔŋ ²¹ DEM 'that little boy'	个 kei ²¹ CL	因子 kin ⁵³ tsə ⁰ child.DIM		□这 (蜀) t̪ciɔŋ ²¹ DEM 'this small tin'	个 kei ²¹ CL	罐儿 kuan ²¹³ nə ⁰ tin.DIM	

Note that the following DEM-NUM-CL combinations are ungrammatical in Shaowu, if [çɪ²¹] 蜀 is not used.

- (17) *□这 一 个 (学生)
 t̪ciɔŋ²¹ i⁵³ kei²¹ (xɔ³⁵sən²¹)
 DEM one CL student
 ('this student')

- (18) *□这 个 个 (学生)
 t̪ciɔŋ²¹ kei²¹ kei²¹ (xɔ³⁵sən²¹)
 DEM one CL student
 ('this student')

If classifiers other than the general one [kei²¹³] 个 are employed in construction type (iii), then the indigenous Shaowu numeral ‘one’ [kei²¹³] 个 will be called upon to precede the specific classifier and follow the proximal or distal demonstrative, although the numeral ‘one’ [kei²¹³] can be readily omitted as mentioned above (Table 8.6):

On the origin of special numerals for ‘one’ in southeastern China 203

TABLE 8.6. [DEM + $k\epsilon i^{213}$ Num + CL + N] construction in Shaowu

Animate nouns				Inanimate nouns			
□这	(个)	头	牛	□那	(个)	架	车
t _ç iɔŋ ²¹	(k _ε i ²¹)	t ^h əu ⁵³	ny ²²	ɔŋ ²¹	(k _ε i ²¹)	ka ²¹³	t _ç ^h ia ²¹
DEM	(one)	CL	ox	DEM	(one)	CL	car
‘this ox (or any kind of cattle)’				‘that car’			
□这	(个)	只	猫	□那	(个)	本	书
t _ç iɔŋ ²¹	(k _ε i ²¹)	t _ç ia ⁵³	mau ⁵³	ɔŋ ²¹	(k _ε i ²¹)	pən ⁵⁵	çy ²¹
DEM	(one)	CL	cat.DIM	DIM	DEM	(one)	CL
‘this cat’				‘that book’			

By way of contrast, the following DEM-NUM-CL noun phrases are ungrammatical in the Shaowu dialect, if [$k\epsilon i^{213}$] 个 is not used to code ‘one’; just like the NPs without demonstratives, as shown in sub-section (i).

- (19) *□这 一 只 (猫儿)
 t_çiɔŋ²¹ i⁵³ t_çia⁵³ (mau⁵³ ə⁰)
 DEM one CL cat
 (‘this cat’)
- (20) *□这 蜀 只 (猫儿)
 t_çiɔŋ²¹ çi²¹ t_çia⁵³ (mau⁵³ ə⁰)
 DEM one CL cat
 (‘this cat’)

Note that the use of bare classifiers as definite and/or indefinite markers, such as in some Wu, Hui, and Yue dialects (see Chapter 5), does not exist in Shaowu. For Shaowu, the use of a numeral and/or demonstrative with a classifier remains obligatory, while it still obeys the general rule of NUM_{one}-CL combinations sketched above for [DEM-(NUM)-CL] NP contractions.

8.3.2.4 Quantifier + (Numeral ‘one’) + CL + Noun

Here we focus on construction Type (iv) involving the use of universal quantifiers and the numeral ‘one’. The same rule applies regarding the use of ‘one’ + CL for the two situations, depending on whether the classifier is $k\epsilon i^{213}$, as in (21) and (22), or if it is a specific classifier, as in (25) and (26).

- (21) 每 (蜀) 个 人 皆 是 学生。
 mei²² çi²¹ k_εi²¹ nin²² ka³⁵ çi²² xo³⁵ sən²¹.
 every one CL person all be student
 ‘Everyone is a student.’

204 *Sing Sing Ngai*

- (22) 没 有 *(蜀) 个 人 是 学生
mau³⁵ iɔu⁵⁵ ci²¹ kɛi²¹ nin²² ci²² xɔ³⁵sən²¹.
NEG have one CL person be student
'No one is a student.'

Note that the following NUM-CL combinations are ungrammatical in Shaowu, when [ci²¹] 蜀 is not used.

- (23) *每 一 /个 个 人
mɛi²² i⁵³/ kɛi²¹ kɛi²¹ nin²²
every one CL person
('everyone')

- (24) *没 有 一 /个 个 人
mau³⁵ iɔu⁵⁵ i⁵³/ kɛi²¹ kɛi²¹ nin²²
NEG have one CL person
('no-one')

- (25) 每 (个) 头 牛 皆 食 草。
mɛi²² kɛi²¹ tʰəu⁵³ ny²² ka³⁵ çie³⁵ tʰau⁵⁵.
every single CL cattle all eat grass
'All cattle graze on grass.'

- (26) 没 有 *(个) 头 牛 食 肉。
mau³⁵ iɔu⁵⁵ kɛi²¹ tʰəu⁵³ ny²² çie³⁵ ny⁵³.
NEG have one CL cattle eat meat
'No cattle feed on meat.'

Note that the following NUM-CL combinations are ungrammatical in the Shaowu dialect, if [kɛi²¹³] 个 is not used.

- (27) *每 一 /蜀 头 牛
mɛi²² i⁵³ /ci²¹ tʰəu⁵³ ny²²
'every one CL cow'
('every cow')

- (28) *没 有 一/蜀 头 牛
mau³⁵ iɔu⁵⁵ i⁵³ / ci²¹ tʰəu⁵³ ny²²
NEG have one CL cow
('There is no cow.')

Older Shaowu speakers consider it unacceptable to replace [ci²¹] 蜀 by [i⁵³] — in examples (23) and (24); and likewise [kɛi²¹³] 个 by [i⁵³] — in examples (27) and (28), regarding it as a 'Mandarinism' adopted by younger generations. Likewise, if [i⁵³] —

On the origin of special numerals for ‘one’ in southeastern China 205

appears in Num-CL constructions involving the use of specific classifiers, the NP will sound unauthentic and incongruous to the native ear.

We should like to mention that apart from acting as numeral ‘one’ in NP constructions, [kei²¹³] 个 also acts as numeral ‘one’ in the [V- Num_{one}-VCL] construction, as shown in examples (29) and (30).

- (29) □看 个 下
nian²¹³ kei²¹ xa³⁵
look one VCL
'take a look'
- (30) 去 个 回
k^hɔ²¹³ kei²¹ fei²²
go one VCL_{round}
'go once'

Again, we see that in this VP construction type, the numeral ‘one’ [kei²¹] 个 is used, which suggests that Shaowu numeral ‘one’ [kei²¹] 个 has the stronghold over pan-Sinitic numeral ‘one’ [i] — and will not be easily displaced by it.

To sum up this section, there are in total three numeral morphemes in the Shaowu dialect that express the notion of singularity and they are in complementary distribution (Table 8.7):

TABLE 8.7. Different numeral ‘one’ morphemes in Shaowu

Numeral ‘one’ in Shaowu	Written form	Syntactic positions
[i ⁵³]	—	cardinal and ordinal number ‘one’, which cannot precede counting units such as [pa ⁵³] 百(hundred), [t ^h ien ²²] 千(thousand), or [van ³⁵] 万(ten thousand) etc.
[ci ²¹]	蜀	numeral ‘one’ in numeral classifier constructions when the classifier is [kei ²¹³] 个, as in 蜀个人 [ci ²¹ kei ²¹ nin ²²] (three-CL-person) ‘one person’
[kei ²¹³]	个	numeral ‘one’ in numeral classifier constructions when the classifier is not [kei ²¹³] 个, as in 个本书 [kei ²¹ pən ⁵⁵ cy ²¹] (one-CL-book) ‘a book’
		numeral ‘one’ preceding counting units such as [pa ⁵³] 百(hundred), [t ^h ien ²²] 千(thousand), or [van ³⁵] 万(ten thousand) etc., presumably because these units are regarded as measure terms, hence are treated like classifiers
		numeral ‘one’ in constructions involving the use of verbal classifiers, as in □看个下 [nian ²¹³ kei ²¹ xa ³⁵] (look-one-VCL) ‘take a look’

It is statistically probable that there are more cumulated instances of different classifiers used in Shaowu than instances involving the use of the general classifier [kei²¹³] 个, and since all verbal classifiers call for the use of the numeral ‘one’ [kei²¹³] 个 too, it is very likely that the numeral [kei²¹³] 个 has a higher frequency of usage than its Shaowu counterpart [ci²¹] 蜀. And it is certainly higher than [i⁵³] used in numeral classifier NP constructions. Together this leads us to hypothesize that [kei²¹³] 个 is likely to be the indigenous form for the numeral ‘one’ in Shaowu and environs, because of its higher frequency of usage and its degree of stability in various numeral-classifier constructions, including that with verbal classifiers.¹¹

8.4 Typological and geographical distribution of [KA] 个 and other morphemes meaning ‘unique’, ‘lone’ undergoing semantic change to ‘one’

As mentioned above, Shaowu is not the only dialect locality where [KA] 个 is found to act as the numeral ‘one’ in [Num-CL-N] constructions. In regions around Shaowu, especially in the Shaojiang dialect group nestled in northwestern Fujian, and also in areas at the northeastern corner of Jiangxi province and southwestern tip of Zhejiang province, we find a concentration of dialect localities where [KA] 个 is employed as ‘one’. In §8.4.1, a list of dialect localities where the phenomenon of having [KA] 个 as numeral ‘one’ is drawn up, whereas in §8.4.2, other morphemes etymologically related to ‘unique’ and ‘lone’ are displayed which have grammaticalized and semantically changed into ‘one’. As the transitional or hybrid nature of some Fujian dialects is not the main focus of this analysis, we simply use the classification system established by the *Language Atlas of China* (1987).

8.4.1 Typological and geographical distribution of the numeral 个

This section gives a general overview of the different morphemes for the numeral classifier ‘one’ in the Min speaking areas, extracted from dialect dictionaries and reliable publications.

Table 8.8 represents the Shaojiang Sub-group of the Min Supra-group in north-western Fujian province.

¹¹ If the numeral ‘one’ [kei²¹³] were a later lexical infiltration into the Shaowu dialect, one may as well ask why it did not start by replacing the primary counting unit (whatever that may have been), but went ahead instead to encroach upon the more stable numeral classifier constructions. A more logical assumption would be that [kei²¹³] was once the universal numeral ‘one’ in enumeration and in other constructions, until the numerals for pan-Sinitic [IT] ‘one’ and pan-Min [SOK] ‘one’ gradually supplanted it to a certain extent, including in certain syntactic configurations as exposed above. We can, however, catch a glimpse of its erstwhile ubiquity when, even today, elderly Shaowu speakers still count ‘one, two, three’ as [kei²¹³], [lion⁵⁵], [san²¹].

On the origin of special numerals for ‘one’ in southeastern China 207

TABLE 8.8. [KA] 个 as numeral ‘one’ in the Shaojiang Group in northwestern Fujian province¹

Dialect locality	Num ‘one’	[‘one’+ CL other than [KA] 个 + N]	[‘one’/ ‘other Num+ CL[KA] 个 or equivalent + N]
光泽 Guangze ~30 km to the northwest of Shaowu	[kei ²¹³] or [kə ⁵] 个	个 益 fa ²¹ ka ⁵ tsəŋ ⁴⁴ one CL flower 'one/a flower'	一 个 团 子 œi ⁵ kei ²¹³ kin ⁴¹ tsə one CL child SUFF 'a small boy'
		个 栋 山 ka ⁵ tun ⁴⁴ sɔŋ ²¹ one CL mountain 'one/a mountain'	一 个 桃 儿 œi ⁵ kei ²¹³ xau ⁴¹ ə one CL peach.SUFF 'a peach'
将乐 Jiangle 106 km south of Shaowu	[ka ²¹⁴] or [ka? ⁵] 个	个 口僕 人 ka? ⁵ çə ³³ ɻi ²² one CL person 'one/a person'	
		个 股 花 ka ¹²¹ ku ³³ fa ⁵⁵ one CL flower 'one/a flower'	
		个 口天 ka ²¹⁴ kvnj ⁵⁵ one day 'one/a day'	
		(HYFYDCD 1999: pp. 374, 378, 381)	
顺昌 Shunchang 90 km southeast of Shaowu	[ka ³⁵] or [ka? ⁴] 个	个 只 客人 ka? ⁴ tsia ¹¹ kʰɔ ¹¹ ŋin ¹¹ one CL guest person 'one/a guest'	
		个 口 塔 ka ³⁵⁻¹¹ tʃa ³⁵ tʰɔ ¹¹ one CL tower 'one/a tower'	
		个 从 花 ka? ⁴ tsʰiun ¹¹ xo ⁴⁴ one CL flower 'one/a bunch of flower(s)'	
		(HYFYDCD 1999: p. 375)	

¹ Certain cells in the tables of §8.4.1 are left blank uniquely due to lack of data from related publications, not because the relevant phenomena are ungrammatical.

Table 8.8 shows that these three dialect localities (Guangze, Jiangle, and Shunchang), all located near Shaowu, have the same [KA_{one} + CL+ N] construction as found in Shaowu. However, in Jiangle and Shunchang dialects, their [KA] forms seem to have only assumed the numeral ‘one’ role, as their general classifiers (and we take this as the one that unitizes human nouns such as ‘person’) do not use a [KA]-related form.

Likewise, the Central Min Sub-group shown in Table 8.9 shows the same pattern, as represented by Shaxian, Sanming, and Sanyuan dialects.

TABLE 8.9. [KA] 个 as numeral ‘one’ in the Central Min dialects of central Fujian province

Dialect locality	Num ‘one’ Morpheme	[‘one’+ CL other than [KA] 个 + N]	[‘one’/‘other Num+ CL[KA] 个 or equivalent + N]
沙县 (Shaxian) 160 km southeast of Shaowu	[kei ¹¹] [ko ²⁴] 个	个 本 ko ²⁴ pui ²¹ one CL 'one/a book'	书 cfy ³³ book
		个 从 kei ¹¹ t ^h un ²² one CL 'one/a tree'	树 t ^h iu ²⁴ tree
		(HYFYDCD 1999: p. 375)	
三明 (Sanming) 213 km southeast of Shaowu	[ko ³³] 个	个 只 ko ³³ tʃin ²¹³ one CL 'one/a person'	人 la ⁴¹ Person
		个 只 ko ³³ tʃin ²¹³ one CL 'one/a month'	月 nye ³⁵³ month
		(HYFYDCD 1999: p. 380)	
三元 (Sanyuan) 186 km southeast of Shaowu	[ko ³³] 个	个 只 ko ³³ tʃin ²¹³ one CL 'one/a person'	人 nã ⁴¹ person
		个 只 ko ³³ tʃin ²¹³ one CL 'one/a month'	月 nye ³³ month
		(MYYJ 1991: p. 211)	

On the origin of special numerals for ‘one’ in southeastern China 209

Moving away from Min dialects, we find that the Taining Gan and Mingxi Hakka dialects possess similar configurations (Tables 8.10 and 8.11). They are located in the vicinity of Shaowu.

Table 8.12 represents the Wu dialects spoken in northern Fujian, northeastern Jiangxi, and southwestern Zhejiang provinces.

As more data are available in the aforementioned Wu dialects (Pucheng, Guangfeng, Jiangshan) for [KA] 个 as numeral ‘one’, we attempt to demonstrate how numeral ‘one’ in these dialects interacts with their classifiers, especially in relation to their general classifiers, as compared to Shaowu.

The dialect of Pucheng uses the same strategy as Shaowu, i.e. when the classifier is not [KA] 个 itself, then [KA] 个 in its other function of the numeral ‘one’ can precede this sortal classifier; however when the NP contains the general classifier and a

TABLE 8.10. [KA] 个 as numeral ‘one’ in Gan dialects in western Fujian province

Dialect locality	Num ‘one’ Morpheme	[‘one’+ CL other than [KA] 个 + N]	[‘one’/‘other Num+ CL[KA] 个 or equivalent + N]
泰宁 Taining 75 km to the southwest of Shaowu	[kæ ⁵¹] 个	个 □ _僞 kæ ⁵¹ sa ³³	人 nin ³³
		one CL	person
		个 栋 kæ ⁵¹ tun ⁵¹	厝 tsʰy ⁵¹
		one CL	house
		‘one/a house’	
		个 支 kæ ⁵¹ tsi ³¹	烟 len ³¹
		one CL	cigarette
		‘one/a cigarette’	
		(TNXZH : pp. 657, 658)	

TABLE 8.11. [KA] 个 as numeral ‘one’ in the Hakka dialects of southwestern Fujian province

Dialect locality	Num ‘one’ Morpheme	[‘one’+ CL other than [KA] 个 + N]	[‘one’/‘other Num+ CL[KA] 个 or equivalent + N]
明溪 (Mingxi)	[kʂ ²⁴] 个	个 张 kʂ ²⁴ tion ³³	纸 tʃe
		one CL	paper
		‘one/a piece of paper’	
		(HYFYDCD 1999: p. 377)	

TABLE 8.12. [KA] 个 as numeral ‘one’ in the Wu dialects of N Fujian, NE Jiangxi, and SW Zhejiang

Dialect locality	Num ‘one’ Morpheme	[‘one’+CL other than [KA] 个 + N]	[‘one’/‘other Num+CL[KA] 个 or equivalent + N]
浦城(县城) Pucheng (county) in northern tip of Fujian Province	介 [ka ⁴⁴] for CL _{other than [KA]}	介 ka ⁴⁴ 头 tiao ¹³ 牛 nion ¹³ one CL ox	特 te ⁴² 个 ke ³²³ 客 k ^h a ⁴² one CL guest
	特 [te ⁴²] for CL [KA] or equivalent	介 ka ⁴⁴ 支 tci ³⁵ 笔 pie ⁴² one CL pen 'a pen'	
		打 介 下 na ⁵³ ka ⁴⁴ hp ²¹ na ⁵³ ka ⁴⁴ hp ²¹ hit one VCL 'hit once'	(Zhengzhang 1985: 41)
广丰 (Guangfeng) in northeastern tip of Jiangxi Province	[k ^y ? ⁵] 个	个 口 农 两 口 农 k ^y ? ⁵ lei ⁴⁴⁵ non ³⁴¹ liā ²⁴ non ³⁴¹ non ³⁴¹ one CL person two CL person 'one person'	
		个 口 桃 三 口 桃 k ^y ? ⁵ lei ⁴⁴⁵ daɔ ³⁴¹ sã ⁴⁴⁵ lei ⁴⁴⁵ daɔ ³⁴¹ one CL peach three CL peach 'a peach'	
		个 本 书 四 本 书 k ^y ? ⁵ pon ⁴⁴⁵ cie ⁴⁴⁵ ci ⁵ pon ⁴⁴⁵ cie ⁴⁴⁵ one CL book four CL book 'a book'	
		个 块 碗 五 块 碗 k ^y ? ⁵ k ^h ua ⁴² uã ⁵² ño ⁴ k ^h ua ⁴² uã ⁵² one CL bowl five CL bowl 'a bowl'	
		这 个 口 东 西 ie ⁵ k ^y ? ⁵ lei ⁴⁴⁵ ton ⁴⁴⁵ ci ⁴⁴⁵ DEM one CL thing 'this (one) thing'	
		那 个 口 东西 xy ⁵ k ^y ? ⁵ lei ⁴⁴⁵ ton ⁴⁴⁵ ci ⁴⁴⁵ DEM one CL thing 'that (one) thing'	(Akitani 2001; Hu Songbai, pers.comm.)

On the origin of special numerals for ‘one’ in southeastern China 211

Jiangshan (Jiangshan)	[kə?⁵] or [æ⁵²]	个	□ ^a	农	两	个	农
at the	(the latter being	kə?⁵	lE⁴⁴	nā³¹³	nā̄³³	kæ⁵²	nā³¹³
southwestern	the weakened	one	CL	person	two	CL	person
corner of	phonological	'one person'		'two people'			
Zhejiang Province	form of [kæ⁵²])	个	个	桃	三	个	桃
		kə?⁵	lE⁴⁴	dəw⁹¹¹	sā⁵²	kæ⁵²	dəw⁹¹¹
		one	CL	peach	three	CL	peach
		'a peach'		'three peaches'			
		个	口	瓯	四	本	书
		kə?⁵	lE⁴⁴	u⁴⁴	çɪ⁵	pæ⁷⁴	çɪə⁴⁴
		one	CL	bowl	four	CL	book
		'one/a bowl'		'four books'			
		个	本	书	五	口	瓯
		æ⁵²	pæ⁷⁴	çɪə⁴⁴	ŋuə⁴	lE⁴⁴	u⁴⁴
		one	CL	book	five	CL	bowl
		'a book'		'five bowls'			
		个	桠	花			
		æ⁵²	o⁴⁵	xua⁹¹¹			
		one	CL	flower			
		'a flower'					
		个	句	话事			
		æ⁵²	kya⁵²	yə³³çø⁵²			
		one	CL	sentence			
		'a sentence'					
		这	个	口	东西		
		iE?⁵	kə?⁴	lE⁴⁴	tə?⁴çɪ⁴⁴		
		iE?⁵	kə?⁴	lE⁴⁴	tə?⁴çɪ⁴⁴		
		DEM	one	CL	thing		
		'this (one) thing'					
		那	个	口	东西		
		xa?⁵	kə?⁴	lE⁴⁴	tə?⁴çɪ⁴⁴		
		DEM	one	CL	thing	(Akitani 2001)	
		'this (one) thing'					

^a [ləi⁴⁴⁵] and [lE⁴⁴] are respectively the general classifier in Guangfeng and Jiangshan dialects, which are used to individuate human nouns and small, round objects in particular. Their origin is unknown, although we suspect that they could be related to 粒, whose semantic scope in southern dialects is much wider than in Mandarin, for instance. The dropping of the [-p] coda of 粒 in other Wu dialects is not unusual. For example, in the Wenzhou dialect, it is pronounced as [lə⁹²¹²] (Zhengzhang 2008). However the question remains to be explained as to why the tonal category of [ləi⁴⁴⁵] and [lE⁴⁴] have both shifted from the entering (*rùshēng*) to the high-level (*yīnpīng*) category.

TABLE 8.13. Comparing [Num-CL-N] constructions between the Shaowu (Min-Gan) and Pucheng (Wu) dialects

Dialect locality	when the CL is not [KA] 个 in [Num-CL-N] construction	when the CL is [KA] 个 in [Num-CL-N] construction
Shaowu (Min-Gan)	uses the numeral ‘one’ morpheme [kei ²¹³] 个	uses the Min-specific numeral ‘one’ morpheme [ci ²¹] 蜀
Pucheng (Wu)	uses the numeral ‘one’ morpheme [ka ⁴⁴] 特	uses the numeral ‘one’ morpheme [te ⁴²] 特

numeral ‘one’ is called for, then Pucheng uses the morpheme [te⁴²] 特 (etymologically an adjective meaning ‘special’) to precede classifier [KA] 个; where Shaowu would use the Min-specific numeral [ci²¹] 蜀. In so doing, the reduplication of the morpheme [KA] 个 in potentially both numeral and classifier slots, which could lead to a universal quantification reading, can be avoided, as well as any eventual ‘conflict of role’, as summarized in Table 8.13.

The dialect of Guangfeng has a different story in the sense that it possesses a general classifier [ləi⁴⁴⁵], which is different from the [KA] form. Consequently, it retains its numeral ‘one’ [kə?⁵] (a KA form) in all its numeral classifier constructions, as there is no conflict for its [KA] form between the numeral and classifier roles.

Interestingly, the neighbouring Jiangshan (Wu) dialect, in its [‘one’ + CL+ N] construction, uses the [KA] form [kə?⁵] as numeral ‘one’ and the classifier [lE⁴⁴] □ for human nouns and small, round inanimate objects, whereas in [Num >1 + CL + N] constructions, another [KA] form [kæ⁵²], which appears to be a cognate of [kə?⁵], takes the classifier position. This makes it possible to have a non-conflictual {Num ‘one’ [kə?⁵] + general classifier [lE⁴⁴] □/other CLs + N} and {Num> 1 + general classifier [kæ⁵²] 个 + N} combinations.

Notably, the Jiangshan general classifier [kæ⁵²] 个 has undergone an apparent phonological attrition to be weakened to [æ⁵²] which also acts as numeral ‘one’ in the [Num-CL-N] noun phrase, leading to two instances of ‘one’ morphemes in the Jiangshan numeral repertoire ([æ⁵²] and [kə?⁵]). Due to the lack of written historical evidence, we cannot establish a definite relationship between [kə?⁵], [kæ⁵²], and [æ⁵²], except to say that [æ⁵²] as numeral ‘one’ is likely to have derived from the classifier [kæ⁵²] through phonetic attrition, given that the tonal value and vowel quality are exactly the same. That a numeral may stem from a classifier needs however further evidence and more substantiation. As far as Jiangshan is concerned, more research is needed to tease apart the interrelationship between the classifier [kæ⁵²] 个 and the numeral ‘one’ [kə?⁵], as well as the interaction and a possible complementary distribution between numeral ‘one’ [kə?⁵] and [æ⁵²] in the Jiangshan dialect.

8.4.2 Morphemes etymologically related to ‘unique’ and ‘lone’ that have undergone semantic change into the numeral ‘one’

The following sub-sections intend to show the reader how the numeral ‘one’ may be derived from adjectives historically meaning ‘alone’, or ‘unique’, the etyma for which have been established using traditional comparative methods in historical phonology.¹²

8.4.2.1 Dialects using Min-specific numeral [SOK] 蜀 ‘lone’, ‘unique’ for ‘one’

Etymologically, a common Min numeral for ‘one’ is [SOK] 蜀 which is possibly related to the adjective [DOK] 獨 meaning ‘lone’, ‘unique’.¹³ It is found in a limited, albeit large part, of Min-speaking areas, covering Mindong, Puxian, and some of the Southern Min groups. Traditionally [SOK] 蜀 is regarded as a Min-specific lexical item, and indeed it appears as one of only two morphemes listed for numeral ‘one’ that were elicited for all 930 Chinese dialects surveyed in the *Linguistic Atlas of Chinese Dialects: Lexicon Volume* (Cao et al., 2008: Map 191) (Table 8.14).

8.4.2.2 Dialects using attributive morphemes meaning ‘lone’, ‘special’ KU 孤、KUA 寡、TSI 特 for ‘one’

The coastal and central varieties of Min (such as Youxi, Fuzhou, Putian) tend to use the Min-specific [SOK] 蜀 as the numeral ‘one’, which is not the case for the more inland varieties such as Pucheng, Yong'an, and Jian'an whose numeral ‘one’ derives from adjectives meaning ‘lone’ or ‘special’, as shown in Table 8.14.

8.4.2.3 Dialects using the pan-Sinitic ‘one’ [IT] —

As for some Gan-speaking areas such as Jianning (建宁), as well as most Hakka-speaking areas such as Ninghua (宁化), Qingliu (清流), and Liancheng (连城) within the confines of Fujian province, the pan-Sinitic [IT] is used as numeral ‘one’ (cf. Li and Zhang 1992). Most regions in the neighbouring Jiangxi province (Gan-speaking) and Zhejiang province (Wu-speaking) use the pan-Sinitic ‘one’ [IT], as do most parts of China (cf. Cao et al. 2008: Map 191 in *Lexicon Volume*) (Map 8.2).

¹² The works consulted for these etyma include Bernhard Karlgren (1964), Li Fang-Kuei (1971), Wang Li (1958), Li Rong (1952), Edwin Pulleyblank (1984), Zhou Fagao (1968), Zheng Zhang Shangfang (1987), Dong Tonghe (1967), Pan Wuyun (2000), William Baxter, and Laurent Sagart (2014). The websites also consulted are: <<http://www.eastling.org/oc/oldage.aspx>>, <<http://www.eastling.org/tdfweb/midage.aspx>>, <<http://ocbaxtersagart.lsait.lsa.umich.edu/BaxterSagartOCbyMandarinMC2014-09-20.pdf>>.

¹³ In the first century AD dictionary of dialectal terms, *Fangyan* ('Regional Expressions'), we find this entry: 《方言》：‘一，蜀也，南楚谓之獨。’ (*Fangyan*: *Yi, shu ye, nanchu wei zhi du.* [*Fangyan*: ‘[SOK] is the equivalent of [IT]. In the southern kingdom of Chu, it is called “獨” (special/lone).’]). 郭璞注：‘蜀，猶獨也’ (*Guopu zhu*: ‘*Shu, you du ye.*’ [Annotation by Guo Pu: ‘[SOK] is like 獨.’]). 《爾雅》：‘獨者，蜀’ (*Erya*: *Du zhe, shu.*’ [The grand dictionary of Erya: ‘Lone 獨 signifies one [SOK].’]).

TABLE 8.14. Dialects using attributive morphemes meaning ‘lone’, ‘special’ for the numeral ‘one’

Dialect locality	Num ‘one’ Morpheme	[‘one’+ CL other than [KA] 个 + N]	[‘one’/other Num+ CL[KA] 个 or equivalent + N]
浦城(石陂) Pucheng (Shipi)	[ku ⁵³] 孤	孤 ku ⁵³ 床 ts ^h onj ³³ one CL	席 ciō ⁵⁵ mat
Northern tip of Fujian Province		‘a mat’ (Zhengzhang 1985: 41)	
永安 (Yong'an) Central Fujian	[kup ²¹] 寡 ^a	寡 kup ²¹ 只 tʃip ¹³ one CL	人 nā ³³ person
		寡 kun ²¹ 本 pūi ²¹ one CL	书 cfy book
		‘a book’ (MYYJ, 1991: p. 211)	
建瓯 (Jian'ou) Northern Min group, Northern Fujian	[tsi ⁴] 特	特 tsi ⁴ 两 liɔŋ ³ one MW5og	米 mi ³ ‘one unit of 5og of rice’ (Zhou 1982)

^a The reason we believe this morpheme to be KUA 寡 and not KA 个 is first of all due to the presence of a medial [u] which is not normally reconstructed for the 个 morpheme. Secondly, the tonal category of the Yong'an [kup²¹] is *shǎng* (上声) which corresponds to the tonal category of 寡 and not 个, which belongs to the *qù* tonal category (去声). Notice that the MC reconstruction of 寡 by many historical linguists is the [KUA] form, *inter alia* Bernhard Karlgren [kWa], Wang Li [kwa], Dong Tonghe [kua], Li Rong [kua], Edwin Pulleyblank [kwa], Zhou Fagao [kua], Zhao Rongfen [kua] (cf. <<http://www.eastling.org/tdfweb/midage.aspx>>).

^b Note that we agree with the conclusion by Zhou Changji (1982) that morphemes for numeral ‘one’ cognate with [TSI] have 特 as their etymon, which later evolved to become [ZHI] 植.

Before taking a closer look at the possible explanations for the coming-into-being of the numeral usage of Shaowu 个 [kei²¹³], it will be useful to first sketch the diachronic development of [KA] 个 in Chinese, as well as different views on its origin. As we will see later, the origin of [KA] 个 is in fact closely related to the development of the numeral reading of [kei²¹³] in the Shaowu dialect.

8.5 Etymological origin and historical development of [KA] 个

The origin of the pan-Sinitic classifier [KA] 个 is perhaps amongst the most obscure of all Chinese classifiers. Lexicographers and philologists in

On the origin of special numerals for ‘one’ in southeastern China 215

TABLE 8.15. Dialects using the specific Min numeral ‘one’ [SOK] 蜀

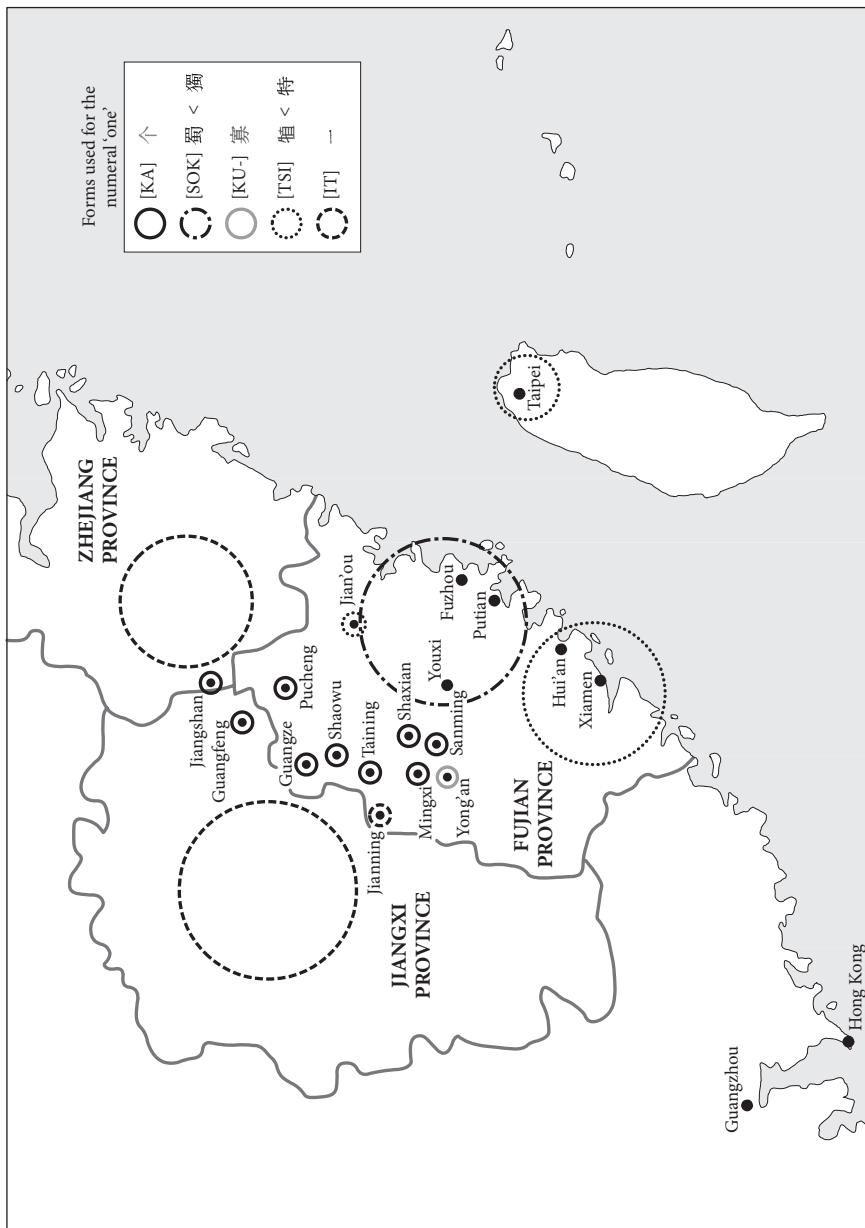
Dialect locality	Num ‘one’ Morpheme	[‘one’+ CL other than [KA] ↑ + N]	[‘one’/‘other Num+ CL _[KA] ↑ or equivalent + N]
尤溪 (Youxi) Houguan Group of Min, at the centre of Fujian province	[cie ³³] 蜀	蜀 本 cie ³³ pon ⁵⁵ tsʰy ³³ one CL book 'a book' (MYYJ, 1991: p. 337)	
福州 (Fuzhou) Houguan Group of Min, in Eastern Fujian	[so?⁵] 蜀	蜀 只 so?⁵ tsie ²⁴ tun ³³ xou?⁵⁵ one CL classmate 'a classmate' (Chen 1998: 23)	
莆田 (Putian) Puxian Group of Min, Southeastern Fujian	[qo?⁷] 蜀	蜀 两 qo?⁷ lieu ⁵ pi ³ one MW:5og rice 'one unit of 5og of rice' (Zhou 1982)	

the past¹⁴ and many Chinese linguists of our days (Liu 1965; Wang 1980; Cao 1994 among others) often take the view that [KA] 个 first appeared as a noun meaning ‘bamboo segments’, which then grammaticalized gradually, under the Han dynasties (206 BC–AD 220), into a specific classifier for items made of bamboo, before turning into a general classifier for human nouns, inanimate objects, and abstract nouns from the Tang dynasty (AD 618–907) onwards. Its path of grammaticalization would be: N > CL_{specific for bamboo} > CL_{general}.

Scholars also seem to have reached a consensus that the classifier [KA] has three written variants in Chinese, i.e. 个, 個, 箇, which are all etymologically linked with ‘bamboo’. [KA] 个 was the oldest written form and served already as a unit of counting from Archaic Chinese (pre-Qin era, c. 400 BC), while 箇 appeared two centuries later in the Western Han (c. 200 BC) and 個 was in use by the end of Eastern Han dynasty (c. AD 200). There was a coalescence of these three written variants in terms of form and function, which took place in the Six Dynasties period circa AD 400. These three forms 个, 個, 箇 have since been freely interchangeable in historical

¹⁴ Amongst others, lexicologists and philologists such as Yang Xiong 杨雄 (50 BC), Xu Shen 许慎 (AD 100), Zheng Xuan 郑玄 (AD 200), Guo Pu 郭璞 (AD 300), Sima Zhen 司马贞(Tang dynasty), Zhang Shoujie 张守节 (AD 700), Dai Tong 戴侗 (AD 950), Wang Jun 王筠 (AD 1800), Duan Yucai 段玉裁 (AD 1800).

216 Sing Sing Ngai



MAP 8.2 Forms for 'one' in Sinitic languages of Southeastern China

texts (cf. e.g. Hong 1963, You 1985, Zhao 1999), but it is not clear whether these three characters ever had distinct meanings, if at all.

The main current notwithstanding, we concur with Hong Cheng (1963) that [KA] 个 has undergone a ‘shortened’ pathway of grammaticalization, i.e. N > CL_{general} without going through the intermediate step of CL_{specific} for bamboo, evidenced by its function almost at the outset as a general classifier to individuate inanimate objects, animals, and humans, long before the specific ‘bamboo classifier’ 節 came into existence. Examples can be found in *Guó Yǔ*【国语】[Chronicles of the Kingdoms] (c. 500 BC) where it is used like a measure word such as LÚ PÍ SÌ GÈ 鹿皮四个 (deer skin four MW) ‘four deer hides’. In the following section, we put forth a hypothesis in an attempt to explain why the numeral usage of [KA] 个 has occurred in Shaowu, and also in certain dialects found in this region of China.

8.6 Hypothesis on the origin of numeral reading of [kei²¹³] 个 in Shaowu

8.6.1 Numeral ‘one’ [kei²¹³] 个 in Shaowu as originating from the adjective [*keajH] 介 meaning ‘unique’

A possible explanation for the existence of numeral [kei²¹³] 个 is that it comes from a different lexical source, and not from any process of reanalysis that would be motivated semantically (i.e. CL_{[kei₂₁₃] > Num_[kei₂₁₃]).}

¹⁵ This may not be surprising for the relatively inaccessible Fujian province, homeland to many an ethnic group in the remote past and cradle of a variegated array of dialect subgroups, often unintelligible to each other. It is also this very region that gave rise to the Min-specific morpheme [SOK] 蜀 meaning ‘one’. Although this does not entail that a ‘numeral’ borrowing or innovation should necessarily take place in this area, it does however suggest a favourable geographic location for this to happen.

If we do not consider the possibility of ‘chance occurrence’ (which is nevertheless considerable in the history of language development), where then, can this [kei²¹³] come from, if not from areal diffusion? The ancient Chinese classics have also been examined and are next discussed.

We find that in these millennia-old texts, the form [KA] 个 was often interchangeable with the form for the morpheme [*keajH] 介, as reconstructed by Baxter and Sagart (2014) for Middle Chinese, which means ‘unique’, ‘different’, or ‘lone’, and moreover that both forms could appear in the same text by the same author (e.g. Zuo Qiuming’s Commentaries *Zuó Zhuàn*【左传】written in c. 500 BC). In effect, the grand Kang Xi Dictionary (*Kāng Xī Zì Diǎn*【康熙字典】) compiled in 1716 during the

¹⁵ If the process of a classifier re-analysed into a numeral were possible, then we might as well ask why, to our knowledge, there are no other instances of CL>NUM in other Sinitic languages, or indeed, within Shaowu itself. Furthermore, this change would violate the principle of unidirectionality in grammaticalization.

218 Sing Sing Ngai

reign of the eponymous Qing emperor, describes 个 as ‘transposable with [JIÈ] 介 (GÈ YÒU TÔNG ZUÒ JIÈ 个又通作 介).¹⁶ The Qing scholar Wang Yinzhī (王引之) referred to 个 as a simplified written version of 介 (which is the older form) simply by omitting a stroke (JIÈ ZÌ LÌ SHÙ ZUÒ, SHÉNG ZÉ WÉI GÈ YÍ 介字隶书作，省则为个矣). Huang Xianfan (2004: 475), among others, has also pointed out that [KA]/ [GÈ] 个 and [KA]/ [JIÈ] 介 are substitutable in ancient texts.¹⁷

Due to space limitations, we only mention that the similarity in semantic content, syntactic position, and phonological makeup of [*keajH] 介 (meaning ‘lonely’ in Old Chinese (OC)) with [KA] 个 may have contributed to their coalescence and later re-analysis from diachronic and synchronic perspectives. We thus contend that [KA]/ [JIÈ] 介 became a free variant of [KA]/ [GÈ] 个 based on their:

- (i) similar graphemic forms (介 / 个);
- (ii) similar semantic content (ADJ ‘single’, ‘lone’ in OC vs. individualising CL from the Tang dynasty onwards);
- (iii) similar syntactic position (both are placed between Num and N);
- (iv) similar phonetic forms reconstructed for OC ([*keajH] ~ [*kaH] > [KA]).

8.6.2 Remnants of [KA]/[JIÈ] 介 carrying its ‘unique’ reading in Modern Chinese dialects

That there was a coalescence of the morphemes 介 and 个 is of general consensus amongst lexicographers. We have suggested how this could have happened,

¹⁶ Online electronic version: <<http://tool.httpcn.com/Html/KangXi/Pic/78.shtml>> accessed on 10th November 2011. ‘康熙字典：【書秦誓】若有一介臣。【大學】作一个臣。即一介，互通。又：一夫曰一介。’ (Kangxi zidian; Shu-qinshi: ruoyou yijie chen; Daxue: zuo yige chen. Ji yijie, hutong. you: yifu yue yijie. [from the Kangxi Dictionary: ‘The Book of the Kingdoms – the Qin Oath: if there were a unique subordinate; *The Book of Great Learning*: this is equivalent of “yīgè chén” meaning ‘one subordinate’. Jiè and Gè are transposable. Furthermore: one man “yī fù” can also be described as “yí jiè”.])

¹⁷ Indeed, annalists frequently interchanged the characters for 介 and 个 in their texts; take for example: 【左傳·襄八年】(Zuo Qiuming’s commentary on the Spring and Autumn Period: In the eighth year of the reign of Xiang Gong)

一 介 行 李
yī jiè xíng lǐ
one JIE diplomat
'a single diplomat'

【左傳·昭二十八年】(Zuo Qiuming’s commentary on the Spring and Autumn Period: On the twenty-eighth year of reign of Zhao Gong)

君 亦 不 使 一 个。
jūn yì bù shǐ yī gè.
2SG but NEG send one GE
'Your Highness did not even send one.'

I use the labels [KA]/ [JIÈ] for the character 介 and [KA]/ [GÈ] for the character 个 in this section to reflect their status as cognates and also their pronunciation in Standard Mandarin today is respectively [JIÈ] 介 and [GÈ] 个.

On the origin of special numerals for 'one' in southeastern China 219

conditioned by resemblance in semantic, syntactic, and phonological features between the two. Perhaps such coalescence can best be demonstrated by the free interchangeability between these two characters and morphemes for [KA] – 介 and 个 – as general classifiers in Hakka (or Kejia) dialects. Indeed, [KA]/ [JIÈ] 介 and [KA]/ [GÈ] 个 are homophonous in most, if not all, Hakka dialects, e.g. Wengyuan (翁源), Liannan (连南), Jiexi (揭西), Xiuzhuan (秀篆), as richly displayed in Li and Zhang (1992: 441, 451) [kai] 介 is used in the Num-CL constructions of these four dialects, among others (respective tone values are not provided).

- (31) 两 三 个 人
 liɔŋ̩ saŋ̩ kai nin
 two three CL people
 ‘two or three people’

(32) 口这 两 介 一 样 大
 ti liɔŋ̩ kai it iɔŋ̩ tʰai.
 DEM two CL one type large
 ‘These two are as large as each other.’

Likewise, in the Fuzhou dialect, [ka²¹³] 介 is used as a general classifier in the Fuzhou (Eastern Min) dialect. According to Chen Zeping (1998: 127), [ka²¹³] 介 is used on a par with [tsie?²⁴] 隻 (or ‘只’ in simplified graphemic form) which is another general classifier, having the function of individuating human and animal nouns, inanimate objects, and abstract concepts. The only difference, according to Chen, is that [ka²¹³] 介 has to precede a head noun [Num- /ka²¹³/ - N], unlike [tsie?²⁴] 隻 which can also follow the head noun [N-Num- /tsie?²⁴/]. The examples he gives are as follows:

- (33) 三 介 哥
 san⁵⁵ ka²¹³ ko⁵⁵
 three CL brother
 ‘three brothers’

(34) 两 介 桶
 nan²⁴² ka²¹³ t^høyŋ³³
 two CL bucket
 ‘two buckets’

The adjectival meaning of ‘sole’, ‘unique’, ‘different’ associated with [KA]/[JIÈ] 介 are retained in the following written (§ 8.6.2.1) and dialectal (§ 8.6.2.2) examples.

8.6.2.1 Lexicalized expressions involving jie 介 in written Chinese

Lexicalized expressions employed especially in modern written Standard Chinese involving the use of *jie* 介 meaning ‘unique’ or ‘different’ are shown here (more examples are included in the notes).

220 *Sing Sing Ngai*

- (35) 介 \dot{M}^{18}
jiè lì
different stand
'unique'
- (36) 耿 介¹⁹
géng jiè
candid distinct
'upright'
- (37) 一 介 书生²⁰
yí jiè shūshēng
one small scholar
'a humble scholar'

8.6.2.2 *Usage of [ka⁵⁵] 介 as a human classifier having a strong semantic coloration of 'alone' or 'sole' in the Shanghai Wu dialect:*

The Shanghai Wu dialect uses the general classifier [gə?¹] 个 (or [fiə¹] in unstressed syllables, as in [i⁴² fiə¹ nin²²] 一个人 'one person') especially to individuate human nouns, possibly as a result of Mandarin influence (Erbaugh and Yang 2006: 196, 199).²¹ In addition to [gə?¹] 个, there is also a special Num-CL construction for human nouns which uses [ka⁵⁵], often preceded by personal pronouns, as shown here.

- (38) 阿拉 一 介 头
a²²la²¹ i⁴² ka⁵⁵ dy²¹
1SG one single SUF<head
'I alone'
- (39) 伊拉 两 介 头
i⁵⁵la²¹ lian²¹³ ka⁵⁵ dy²¹
3PL two single SUF<head
'they (two) alone'

¹⁸ And synonyms such as *jièpi* 介僻 'inflexible', *jièyi* 介异 'different', *jièlin* 介吝 'aloof', *jiètè* 介特 'special', etc. In Beijing Mandarin, 个色 [kɤ⁵³ sɤ⁵³] is used to mean 'peculiar'; in Tianjin Mandarin, 个别 [kɤ⁵³ pie⁴⁵] is used to mean 'different' or 'peculiar' (HYFYDCD: 376). These [kɤ⁵³] 个 morphemes are likely to be etymologically related to [KA]/ [JIÈ] 介. The same can be applied to 个样 [kɤ⁵³ ian⁵³] and 个调 [kɤ⁵³ tiao⁵³] in Hubei Mandarin which have similar meanings as the above items. (HYFYDCD: 378).

¹⁹ And synonyms such as *jièlián* 介廉, *jièbái* 介白, *jiézhèng* 介正, *jiéjué* 介决, etc.

²⁰ And other expressions such as *yí jiè cǎo mǐn* 一介草民 'an unimportant commoner', *yí jiè wǔ fū* 一介武夫 'an ordinary fighter', etc.

²² The indigenous default classifier in Shanghai Chinese is [tsə?⁵] 只 (cf. Xu and Tang 1988: 404–5) which has a wide semantic scope, as it can be used for human and animal nouns, as well as inanimate objects and abstract concepts.

Note that [tx⁰] 头 is a suffix that functions as a nominalizer in some Wu dialects. The noun phrase [Num- ka⁵⁵ 介- tx⁰ 头] emphasizes the exclusiveness of the human agents involved, and is often preceded by personal pronouns. This provides further semantic evidence of the strong semantic affinity between 介 [JIÈ] and 个 [GÈ] using synchronic data.

8.6.3 Pathway of development for the Shaowu numeral ‘one’ [kei²¹³]

Insofar as the Shaowu dialect is concerned, we contend that [*keajH] 介 (etymologically an adjective meaning ‘lone’, ‘unique’, ‘different’, which in turn is derived from the noun meaning ‘boundary’, ‘partition’) is the origin of the numeral ‘one’ meaning for [kei²¹³]. After collapsing with the classifier [*kaH] 个, which may also be related to the earlier OC meaning of ‘lone’, and becoming a free variant of the latter as a result of syncretism, [*kaH] 个/ 介 underwent further conversion into the numeral ‘one’ (Figure 8.1). As this needs further in-depth diachronic research, we will not elaborate here on the grammaticalization pathway regarding the classifier usage.

We believe that the Sinitic numeral ‘two’ [LIANG]²³ 两 has also undergone a similar pathway of grammaticalization: i.e. [LIANG] 两 started out as a noun meaning ‘two-horse chariot’ and by semantic extension came to denote the state of being ‘in a pair’, which then bifurcated into a classifier for nouns in pairs and the numeral ‘two’. Like [KA]/ [JIÈ] 介, it branched into the following paths.²⁴

²³ [LIANG] is an abstract phonetic representation for the pan-Sinitic numeral ‘two’, for which we drew inspiration from the different MC reconstructions made by linguists such as: Bernhard Karlgren’s [*liaN], Li Rong’s [*liaN], Dong Tonghe’s [*ljaN], Shao Rongfen’s [*liAN]. (See also note 12.)

²⁴ Examples from historical texts to support the bifurcation pathway:

- a) [liāng] 两 functioned as a classifier in Early Medieval Chinese (third–sixth centuries AD) (Yang-Drocourt 1993: 181) but this usage became obsolete in the course of time:

白 布 襪 一 两 (郭简575)
bái bù wà yī liǎng
white cloth sock one CL_{pair}
'a pair of white socks'

白 革 履 一 两 (郭简 206.23)
bái gé lǚ yī liǎng
white leather shoe one CL_{pair}
'a pair of white leather shoes'

(extracted from Guo Jian, inscriptions on wooden tablets from the Han dynasty)

- b) [liang] 两 functioned as the numeral ‘two’ as early as in AD 500 and up until our days:

两 个 月 秋 耕【贾思勰：齐民要术】
liǎng gè yuè qiū gēng
two CL month autumn cultivation
'two months of cultivation in Autumn'

(Jia Sixie: *Ways of earning a living for common people*, written during the Northern Wei Dynasty).

As for most Chinese dialects today, [LIANG] 两 is the numeral ‘two’ preceding the classifier and the head noun in [Num-CL-N] constructions, just as in the example (Figure 8.2).

222 Sing Sing Ngai

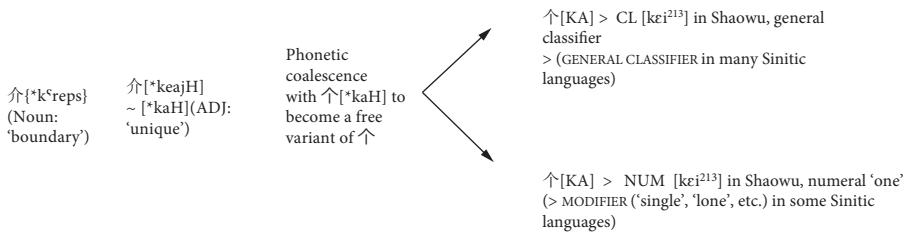


FIGURE 8.1 Grammaticalization pathways for 个 [KA]

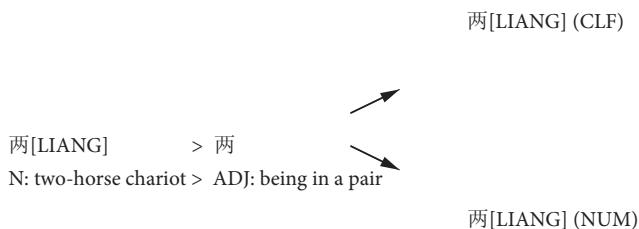


FIGURE 8.2 Grammaticalization pathways for 两 [LIANG]

8.6.2.3 Evaluation of the hypothesis

Given the arguments supported by diachronic and synchronic data, this account is plausible for the formation of the Shaowu numeral morpheme [kei²¹³] 个. Its origin may possibly be attributed to the lexical source [KA]/ [JIÈ] 介 which originally means 'single', 'unique'.

On this basis, we further the hypothesis of numeral 'one' as deriving from the notion of 'lone', 'singular', as shown by examples from other languages listed in Heine and Kuteva (2002: 219–20). This lexical conversion process would have involved both semantic change and decategorialization from being a member of the adjective class to becoming a member of a closed class, the set of basic numerals in Shaowu (Heine and Kuteva 2005). Independent typological support for the validity of our assumption comes from Amazonian languages of Jarawara and Paumari (Arawan family), where the verb meaning 'be alone' came to be used as a numeral 'one' and the verb meaning 'be a pair' came to be used as a numeral 'two' (Aikhenvald 2000: 100). Note that these two morphemes in question are both verbs of state, which are akin to adjectives (Heine and Kuteva 2007).

One might argue that [kei²¹³] 个 might have originally been a numeral, and the adjectival sense of 'lone' developed out of this lexeme. This is highly unlikely as this adjectival meaning made its first appearance in historical texts as early as 500 BC (Old Chinese), whereas in the long history of Chinese, there is no attested record

of [KA]/[GÈ] 个 being used as a numeral. Moreover, no other Sinitic dialect groups, except those in the vicinity of Shaowu, attest to a similar numeral usage of [KA]/[GÈ] 个.

The existence of a numeral ‘one’ in any language is hardly surprising, but the existence of three different numeral forms for ‘one’ in Shaowu is very unusual at least for Sinitic languages, and raises the question of whether they shared the same path of evolution before becoming counting and quantifying units. Our hypothesis, supported by typological and diachronic data, has shown that it is indeed plausible for a lexeme to have a concrete ‘pre-numerical’ semantic content such as ‘lone’ or ‘single’, before expressing the numerical notion of ‘one’. Moreover, since the classifier system only became mature in the period of Medieval Chinese during the Tang dynasty (Peyraube 1998), the numeral sense may well have preceded the classifier usage of KA 个. This topic awaits further research.

8.7 Typological implications

That a language has two (or more) sets of numeral paradigms is commonplace. In Russian, for example, there are at least two words for ‘one’: *raz* is the morpheme for numeral ‘one’ used in counting, whereas *odin* and its inflected forms (e.g. for number, case, and gender) are used as cardinal quantifiers having adjectival properties, while *pervy* is used for ordinal counting (see, for example, Grosberg 1957). Dialects located in northwestern Fujian and environs also possess two sets of numerals, albeit only for the basic ones, essentially ‘one’ and ‘two’: one set for pure counting, where the pan-Sinitic ‘one’ [IT] — is involved, and the other for pairing the indigenous numerals with classifiers in [Num-CL-N] constructions.

The indigenous numeral ‘one’ in [KA] 个 form is found mainly within the Shaojiang group and sporadically in some Wu dialects in the northern tip of Fujian. These two areas are however separated by the [TSI] form (the etymon is yet to be determined, but may well be the morpheme TE 特 meaning ‘special’ as suggested by Zhou 1982 for the Southern Min dialects), as identified in the dialects of Chong'an (崇安), Jianyang (建阳), and Jian'ou (建瓯). We posit that the [TSI] form came to northern Fujian at a later date, possibly brought by migration and commerce to the area, thus splitting the [KA]-occupied region, including the Yong'an area (永安), into two. If the opposite is true, we have to account for the parallel emergence of two [KA] regions and why they straddled the [TSI] area.

It is a well-established fact that synchronically, southern Sinitic languages have a larger stock of classifiers than the northern ones. Many of these dialects possess a general classifier distinct from [KA] 个, such as [TSEK] 隻²⁵ in the Eastern Min

²⁵ [TSEK] is an abstract phonetic representation for the general classifier found in Southern China, notably in Eastern Min, some Wu, Gan, and Xiang dialects. We drew inspiration from the different MC reconstructions such as Edwin Pulleyblank's [*cjk].

Group where the classifier can be used for human and animal nouns, as well as for inanimate objects and abstract concepts. The respective general classifiers [lei⁴⁴⁵] and [lE⁴⁴] in Guangfeng (广丰) and Jiangshan (江山) seem to be part of the native inventory. These local general classifiers constitute a resisting force against the encroachment of the pan-Sinitic general classifier [KA] 个 into their numeral classifier constructions. However, we cannot tell if it is also the case for Shaowu, as there is no diachronic or synchronic trace of another Shaowu general classifier being replaced by the pan-Sinitic general classifier [KA] 个, leading to our speculation that the classifier [kei²¹³] 个 has arisen from a grammaticalization process.

8.8 Cognitive implications

The earliest human numerical cognition involves the trichotomy of ‘one’ and ‘two’ and ‘many’. It is generally accepted that basic numerals in the world’s languages such as ‘one’ and ‘two’ are different from higher numerical orders in terms of semantic features and syntactic properties. The numeral ‘one’, for instance, can take gender in French; case and gender in German; case, gender, and number in Russian. It has been proposed that if a language has inflected numerals, it is invariably the lower numerals which have adjectival properties while the higher numerals appear more nominal in character (cf. Jespersen (1969), Corbett (1978), Hurford (1987)).

Although the lack of morphology in Sinitic languages makes it more difficult to distinguish the syntactic and morphological properties of basic numerals from adjectives, we nevertheless find that the indigenous numeral morphemes for ‘one’, as shown in §8.4, are all etymologically related to the notion of ‘single’, ‘special’, ‘lone’, and started out as adjectives before becoming numerals. And the basic numeral ‘two’ 两 could have stemmed from the notion of ‘pair’, although further research is needed to prove this. Support for our speculation comes from the fact that many Tai-Kadai languages possess the numeral ‘two’ morpheme [SOONG] which was likely acquired through language contact from Sinitic [SHUANG] 雙 which means ‘pair’. These include Southern Zhuang [tɔ:ŋ¹], Northern Yongbei Zhuang [so:ŋ¹], Tai Dam [sɔŋ¹], Lao [sɔ:ŋ¹], and Northeastern Thai [sɔ:ŋ¹].

On the cognitive side, it is not likely that basic numbers in a linguistic community stem from more abstract numerical notions, unless borrowing from other numerical systems is involved. It is more conceivable that a single entity is first recognized as an individual, ontological being, separate from others and is thus given the name ‘one’, meaning ‘singular’. The concept of ‘one’ might, we believe, have come from the notion of ‘singular’, which has been shown to be the case for a group of southeastern Sinitic languages, clustered in northwestern Fujian. More research into the origin of basic numbers in the world’s languages would be needed to support our claim.

8.9 Concluding remarks

In this analysis, we first discussed the singularity and singular nature of the Shaowu numeral ‘one’ [kei²¹³] by displaying various numeral paradigms and Num-CL constructions involving ‘one’ in Shaowu. We also presented the typological distribution of the numeral [kei²¹³] clustering around northwestern Fujian and other Sinitic morphemes for ‘one’ in Fujian province, whose source meaning could be ‘lone’, ‘singular’, and ‘unique’. We then posited a possible explanation for the formation of the numeral [kei²¹³] and argued that this morpheme most likely derived from the adjective [*keajH] 介 in Old Chinese meaning ‘unique’. However, its path of development was less straightforward than expected, as [*keajH] 介 later coalesced with the pan-Sinitic general classifier [*kaH] 个. This gave rise to the situation whereby the numeral [kei²¹³] existed side by side with its homonym, the classifier [kei²¹³], with both being represented by the same grapheme 个. Finally, we suggested that basic numerals such as ‘one’ and ‘two’ in many Sinitic languages have possibly derived from the notion of ‘unique’ and ‘pair’ respectively, although further research on this topic with more linguistic data from the world’s languages is needed to substantiate our claim.

Acknowledgements

My heartfelt gratitude goes to my thesis supervisors Prof. Hilary Chappell and Prof. Alain Peyraube, for their precious feedback and insightful comments on my analysis. I also wish to thank Prof. Jiang Shaoyu, Prof. Li Rulong, Prof. Hu Songbai, Prof. Laurent Sagart, Prof. Stephen Matthews, Dr Wang Hongzhong, Dr Zheng Wei, Dr Wang Jian, Dr Hilário de Sousa, and Jeremy Collins, for their helpful remarks and suggestions. I am also grateful to Prof. Hu Songbai for his Jiangshan information and native speaker’s judgement, and to Ms Xu Jie, Mr Feng Xiaoqing, and Mr Sha Yi, all natives of Shanghai, who provided and confirmed the Shanghai data.

I also extend my sincere thanks to my linguistic consultants Ms Li Jingxin and Mr Li Hougong, and the late Ms Wei Yixin, for their patience and passion in teaching me the Shaowu dialect, as well as providing useful linguistic data without which this analysis would not have been possible. I would like to thank as well the two anonymous reviewers for their insightful remarks and helpful comments on the chapter. All remaining errors are my own.

Research for this chapter has benefited from funding from the ERC SINOTYPE project.

Finally, I thank the Australian Academy of the Humanities for kindly giving its permission to reproduce Map B-12 from the *Language Atlas of China* (Wurm et al. (eds), 1987) for which it holds the copyright.

9

Complex pronouns in Wu Chinese: Focalization and topicalization

XUPING LI

9.1 Introduction

This chapter presents a case study of grammatical properties of complex pronouns in Wu Chinese, namely the Fuyang dialect, from both the synchronic and diachronic perspectives. We argue that the syntactic context for using complex pronouns is dependent on their information status as topic or focus. We make a distinction between stressed and unstressed complex pronouns: unstressed complex pronouns are restricted to topic positions, either primary or secondary, while stressed ones have no such restriction and can serve as contrastive topic or focus. We propose that complex pronouns in Fuyang Wu involve a ‘syntactic reanalysis’, being derived from the fusion of the copula and pronouns in cleft sentences and that the fused complex pronouns undergo a grammaticalization process from focalization to topicalization.

9.1.1 *Pronominal paradigms in Sinitic languages*

In most Sinitic languages, such as Standard Mandarin, there is only one paradigm of personal pronouns, which include singular and plural forms. The pronoun system in Mandarin Chinese is illustrated in Table 9.1.

Personal pronouns in Mandarin are not sensitive to case. There is no distinctive form for nominative, accusative, and genitive cases, etc. Putting it in a simpler way,

TABLE 9.1. Personal pronouns in Standard Mandarin (普通话)

	1 st person	2 nd person	3 rd person
Singular	wǒ 我	nǐ 你	tā 他
Plural	wǒ-men 我們	nǐ-men 你們	tā-men 他們

Complex pronouns in Wu Chinese: Focalization and topicalization 227

the same pronoun form may play different syntactic roles, such as subject, object, possessor, etc. These uses are illustrated in (1).

- (1) 他和我说，他见过你妹妹。

Tā hé wǒ shuō, tā jiàn-guò nǐ mèimei.¹
3SG to 1SG say 3SG see-EXP 2SG younger.sister
'He said to me, he saw your younger sister.'

However, in some central and northern Wu dialects, two different paradigms of personal pronouns exist, namely simple and complex pronouns (Z. Qian 1983; Z. Chen 1996; N. Qian 1999, among others). Specifically, complex pronouns are derived from simple pronouns by adding a prefix, such as the prefix /zəʔ/ in Shengxian Wu. The co-existence of dual paradigms of personal pronouns is mainly found in the Taihu Lake group of Northern Wu dialects.² Table 9.2 illustrates the singular forms for the complex pronouns in the Shengxian, Huzhou, and Shaoxing dialects of Zhejiang Province as well as for dialects spoken in certain districts of Shanghai. For ease of exposition, we only show the singular complex pronouns in these dialects while the simple counterparts are not set out separately (that is, the elements that these prefixes attach to are none other than corresponding simple personal pronouns).

TABLE 9.2. Complex pronouns in some Wu dialects

Dialects	1SG	2SG	3SG	Source
Shengxian Changle 嵊县长乐	zəʔ-ŋo	zəʔ-ŋ	zəʔ-i	Z. Qian 1983
Huzhou 湖州	zʔ-ŋ	zʔ-n	zʔ-dzi	Z. Chen 1996
Shaoxing 绍兴	zeʔ-ŋo	zeʔ-no?	zeʔ-fi	Chen and Pan 1999
Shanghai Nanhai Huinan 上海南汇惠南	zeʔ-βu	zeʔ-noŋ	zeʔ-fi	
Shanghai Baoshan Chengxiang 上海宝山城厢	zŋ-ŋ	zŋ-fi	zŋ-fi	
Shanghai Songjiang 上海松江		zəʔ-nu	zŋ-fi	
Shanghai Fengxian 上海奉贤		zeʔ-noŋ	zeʔ-fi	N. Qian 1999

¹ The romanizations used in this chapter are the IPA for Wu dialects, and *pīnyīn* for Standard Mandarin, unless examples are being cited from other references which follow the original in this case.

² According to the 'Inventory of Pronouns in Wu dialects' [Wuyu Rencheng Daici Biao 吴语人称代词表] provided on the website of the Wu Association [吴语论坛], this phenomenon is available in the dialects of the Linshao subgroup (临绍小片), the Tiaoxi subgroup (苕溪小片), and the Sujiahu subgroup (苏嘉沪小片), all belonging to the Taihu Lake group of Northern Wu.

TABLE 9.3. Three paradigms of personal pronouns in the Fuyang dialect

	Paradigm I		Paradigm II		Paradigm III	
	Simple pronouns		Complex pronouns		Complex pronouns	
	SG	PL	SG	PL	SG	PL
1 st Person	ŋy ³¹³	a? ¹ -la ³¹³	zə? ¹ -ŋy ³¹³		fɪə? ¹ -ŋy ³¹³	
2 nd Person	n ³¹³	na ³¹³	zə? ¹ -n ³¹³	zə? ¹ -na ³¹³	fɪə? ¹ -n ³¹³	fɪə? ¹ -na ³¹³
3 rd Person	i ³¹³	ia ³¹³	zə? ¹ -i ³¹³	zə? ¹ -ia ³¹³	fɪə? ¹ -i ³¹³	fɪə? ¹ -ia ³¹³

This analysis examines the syntactic and semantic properties of complex pronouns in Wu, based on the data from the Fuyang dialect of Wu Chinese.³ In the Fuyang dialect, there are three different paradigms of personal pronouns, as shown in Table 9.3.

The monosyllabic personal pronouns in the first paradigm in Table 9.3 are simple personal pronouns in Fuyang Wu. The second and the third paradigms are complex personal pronouns, and they are derived from the first paradigm by prefixing /zə?/ and /fɪə?/ respectively. In this research, we contrast complex personal pronouns with simple pronouns, which we label ‘common personal pronouns’.

The term ‘emphatic pronouns’ usually refers to a particular paradigm of pronouns in certain languages that has ‘some kind of discourse prominence, typically either contrast and/or intensification’ (Siewierska 2004: 67). As will be shown below, in many cases, complex pronouns in Wu can only be used with the accompaniment of stress, which expresses contrastive meaning. In other words, it is obligatory for complex pronouns to have such emphatic uses, determined by stress, in certain contrastive contexts, as will be shown. However, common personal pronouns have no such requirement.

We are interested in the following questions concerning the distribution and grammatical properties of complex pronouns in Fuyang Wu:

- (i) What are the factors affecting the distribution of stressed and non-stressed complex pronouns?

³ The Fuyang dialect (富阳方言) belongs to the Linshao group of Wu dialects [临绍小片]. It is spoken in the Fuyang county of Zhejiang Province and has about 620,000 native speakers. In the current research, we use the variant of Fuyang dialect spoken in the downtown area, as well as in Chunjiang Town and Lingqiao Town.

- (ii) What are the semantic interpretations of complex pronouns? How should we understand their functions in terms of Siewierska's 'discourse prominence' when they are interpreted emphatically?
- (iii) Z. Chen (1996) proposes a phonological reconstruction whereby the *z*-type prefix before pronouns in Wu is derived from the morpheme 是 /zɪ/, which can either be used as a copular verb 'be' or a demonstrative 'this' in early linguistic periods, such as in Middle Chinese (third–twelfth centuries). We are interested in the question of whether we are able to bring some syntactic/semantic evidence to bear on the synchronic perspective to decide upon the nature of this morpheme.

The remainder of the chapter will be organized as follows. Section 9.2 looks at when and how non-stressed complex pronouns are used. In particular, we will examine two different possibilities of generalizing the distribution of non-stressed complex pronouns. One concerns the subject and object asymmetry and the other is the correlation between complex pronouns and topic. In §9.3, we discuss the use of stressed complex pronouns. We will look at how phonetic stress helps to license complex pronouns in Fuyang Wu. Section 9.4 reexamines Z. Chen's (1996) phonological reconstruction and we argue that it is more appropriate to treat the prefix in complex pronouns as being derived from the copula 是 /zɪ/. In §9.5, we argue that complex pronouns in Fuyang Wu are reanalysed from the fusion of the copula and simple pronouns in cleft sentences. Section 9.6 concludes the chapter by summarizing the main claims we make.

9.2 Unstressed complex pronouns

In this section, we will investigate the question of how complex pronouns are different from common personal pronouns in terms of their syntactic distribution and function, without for the moment touching upon the issue of stress. We will investigate under what circumstances 'common' personal pronouns can or cannot be replaced by complex ones.

A preliminary study shows that the 2nd and the 3rd paradigms of personal pronouns in Fuyang Wu do not have any grammatical distinction and they are interchangeable with each other. We assume that the prefix /fiə?/ in the Fuyang dialect is a phonologically weakened form of the prefix /zə?/ via 'debuccalization'.⁴ In the rest of the analysis, we do not further distinguish between the *z*-type and the *h*-type of complex pronouns and we will merely contrast common personal pronouns with complex ones. For ease of exposition, when representing the examples in

⁴ The debuccalization from /z/ to /fi/ is a quite common phonological process in Wu dialects, such as Xuanzhou Wu (宣州片吴语) (Wang Jian: pers. comm.).

Chinese characters, we simply use /z/ and /h/ to stand for prefixes of the 2nd and the 3rd paradigms of personal pronouns respectively.

9.2.1 An illusory subject/object asymmetry

Common personal pronouns in the Fuyang dialect of Wu, that is, the first paradigm in Table 9.3, are able to appear in various syntactic positions, such as in the position of topic, subject and object.

- (2) 你, 娜妹子, 拨我去叫伊声。

n³¹³ na³¹³ mε³⁴tsi pəʔ⁵¹ iŋ³¹³ tɕ^hi³⁴ tɕio³⁴ i³¹³ ɕin⁵³.
2SG 2PL sister for 1SG go call 3SG call
'Your sister, could you please call her for me?'

For example, in (2), the pronoun *n*³¹³ 'you' acts as the subject of the clause, and the plural pronoun *na*³¹³ 'your' is a possessor of the head noun *mε*³⁴tsi 'sister' in the possessive phrase, which is co-referential with the pronoun in the object position, i.e. *i*³¹³ 'her'. Moreover, in this clause, the first person pronoun *iŋ*³¹³ 'I' is also the object of the preposition *pəʔ*⁵¹ 'for', which codes its beneficiary role.

Compared with common personal pronouns, unstressed complex pronouns in Fuyang Wu have a more restricted distribution. At first glance, unstressed complex pronouns can be used as subjects, as in (3), but they are not allowed in object positions, either as a canonical postverbal object (4) or as a preposed object marked by KE (5).⁵ On the contrary, common pronouns are allowed in all these three syntactic contexts from (3) to (5) without restriction.

- (3) Subject

早间, (z-)佢做礼拜去格。
tsɔʔ⁵¹³kan⁵³ (zəʔ¹-) ia³¹³ tsy³⁴ n̩i³¹³pa³⁴ tɕ^hi³⁴ 34go.
morning PREF-3PL do religious service go SFP
'They go to church in the morning.'

- (4) Canonical Object: S-V-O

勑去打(*z-)伊啊。
fio³⁴ tɕ^hi³⁴ tan⁵¹³ (*zəʔ¹-)i³¹³ a.
NEG go beat PREF-3SG SFP
'Don't beat him.'

⁵ Note that the Fuyang dialect of Wu has S-V-O word order, but when the object is definite, it is preferable to prepose it to a preverbal position, as marked by /kʰəʔ/. We call it a KE construction (Li and Bisang 2012).

(5) **Preposed KE object: S-KE-V-O**

我克(*z)你骂过芒?

ŋy³¹³ k^həʔ⁵¹ (*zəʔ¹-)n³¹³ mo³⁴ ku³⁴ man³⁵?
1SG OM PREF-2SG scold EXP Q-PRT

'Have I ever scolded you?'

From the stories and dialogues that we recorded in the Fuyang dialect, we found 25 sentences in which unstressed complex pronouns occurred.⁶ Among those 25 examples, 24 have complex pronouns occurring in the subject position and none of them is found in any object position. For the single exception, the complex pronoun is used as a possessor in a possessive phrase. Therefore, based on the representative examples from (3) to (5) and our statistics, it might initially appear that complex pronouns exhibit a subject/object asymmetry. However, the generalization about the correlation of unstressed complex pronouns with the subject role is immediately challenged by the following facts.

First, the generalization of subject/object asymmetry does not cover the use of complex pronouns as topic. As shown in (6), the complex pronoun zəʔ¹-ŋy³¹³ occupies two different syntactic positions in the sentence. In the first occurrence, zəʔ¹-ŋy³¹³ acts as the topic of the sentence, while in the second occurrence, it is the subject of the sentence. The subject constraint cannot explain the topic use of complex pronouns.

(6) **Topic**

z- 我哦, z-我小时光蹲勒富阳。

zəʔ¹- ŋy³¹³ o, zəʔ¹-ŋy³¹³ ciɔ⁵¹³zɪ³⁵kuan⁵³ təŋ⁵³-lə fu³⁴ian³⁵.
PREF-1SG TOP PREF-1SG childhood live-PFV Fuyang

'As for me, I lived in Fuyang in my childhood.'

Second, what is more problematic for the subject/object generalization is that while it is generally not acceptable to use complex pronouns in the postverbal object position or in the KE preverbal object position, as we saw in (4) and (5), it is nonetheless possible to use them in some other preposed object positions, such as in the sentence-initial position in (7a) or in the position between subject and predicate without an object marker, as in (7b).

(7) a. **Object preposed into sentence initial position: O-S-V**

z-你, 我骂过芒?

zəʔ¹-nŋ³¹³ ŋy³¹³ mo³⁴-ku³⁴ man³⁵?
PREF-2SG 1SG scold-EXP Q-PRT

'You, have I scolded you?'

⁶ We conducted a field trip to Fuyang City, Zhejiang Province in August 2011 and collected relevant data on emphatic pronouns in Fuyang Wu. We recorded three stories and dialogues in the Fuyang dialect. The collected data include a ten-minute dialogue between an eighty-nine-year-old woman and the author, and two five-minute stories told by two female speakers in their seventies.

b. Object preposed between Subject and VP without an object marker: S-O-V

我 z-你骂过芒?
ŋy³¹³ zə?¹-ŋ³¹³ mo³⁴-ku³⁴ man³⁵?
1SG PREF-2SG scold-EXP Q-PRT
'Have I scolded you?'

Taking into account these two problems, we do not think it appropriate to characterize the distribution of complex pronouns in the Fuyang dialect by way of a contrast between subject and object. The use of complex pronouns is therefore not dependent on their syntactic status as being subjects or objects, although this may be seen to play a minor role.

In §9.2.2, we will explore another possibility about the licensing of complex pronouns in Fuyang Wu, namely, the topic/focus asymmetry.

9.2.2 Complex pronouns as topics

In this section, we will examine whether or not the use of complex pronouns in Fuyang Wu is related to their information status, such as topics or foci.

Before proceeding, we adopt three basic assumptions concerning information structure in Wu Chinese:

- (i) As in Mandarin, Wu Chinese (represented here by the Fuyang dialect) is a topic-prominent language, where the sentence-initial position is the default topic position. Syntactically, it can be represented as [TopP [IP]], appearing on the left-periphery of a sentence; and
- (ii) Unlike in Mandarin, there is a secondary topic position between subject and the verb, i.e. between IP and VP, just as in other Wu dialects (Liu and Xu 1998 for Shanghai Wu; Hu, Pan, and Li 2003 for Ningbo Wu);
- (iii) Similar to Mandarin, the sentence-final position is the (information) focus position in Wu (for Mandarin, see Xu 2004).

The notion of topic is understood from two perspectives: it codes what the sentence is talking about, while it typically expresses 'old information' (see Chafe 1976). Focus refers to a constituent within a sentence that is highlighted or emphasized by grammatical means. When talking about focus, a distinction between information focus and contrastive focus has to be made. Information focus is understood to be relational: it concerns the information predicated about the topic. Contrastive focus is referential: it concerns material which the speaker calls to the addressee's attention, thereby often evoking a contrast with the other entities that might fill the same position (Kiss 1998, cf. Gundel and Fretheim 2003).

To start with, we make use of the question-answer pair as a diagnostic to find out what the information status of unstressed complex pronouns is when they are used

Complex pronouns in Wu Chinese: Focalization and topicalization 233

in those examples we collected from the recordings. Consider the examples in (8) and (9).

In example (8), the whole clause in B is the direct answer to the question, and the subject is part of the new information (focus) required by the question. In this case, it is infelicitous to use an unstressed complex pronoun in the subject position as a focused element (note: # marks an infelicitous sentence). Instead, a common pronoun can be used, such as $\eta\gamma^{313}$ 'I'.

- (8) Speaker A: 何尔行当?

Speaker A: koz^{35} $fian\eta^{35}tan^{513}$?
what matter

Speaker A: 'What happened?'

Speaker B: #z 我被蛇咬勒一口。

Speaker B: # $zəʔ^1-\eta\gamma^{313}$ $pəʔ^{51}$ dzo^{35} $\eta\gamma^{313}-lə$ $iʔ^1$ k^hiu^{513} .
PREF-1SG PASS snake bite-PFV one mouth

Speaker B: 'I was bitten by a snake.'

In contrast, in (9), the individual under discussion in the context is first introduced in the preceding question by a simple pronoun in (9A).⁷ The use of the complex pronoun in (9B) is part of the answer to the question, but only the predicate in (9B) expresses new information. In other words, in (9B), the subject is a topic and the predicate makes a comment about it. Here, it is completely plausible to use an unstressed complex pronoun as the subject.

- (9) Speaker A: 伊做咯?

Speaker A: i^{313} tso^{34} lo ?
2SG do PRT

Speaker A: 'What happened to him?'

Speaker B: z 伊被蛇咬了一口。

Speaker B: $zəʔ^1-i^{313}$ $pəʔ^{51}$ dzo^{35} $\eta\gamma^{313}-lə$ $iʔ^1$ k^hiu^{513} .
PREF-1SG PASS snake bite-PFV one mouth

Speaker B: 'He was bitten by a snake.'

The contrast between (8) and (9) suggests that unstressed complex pronouns cannot act as a direct answer to a question that expresses new information, as in (8). Put otherwise, unstressed complex pronouns in Fuyang Wu can only be topics but not foci, as is the case in (9).

We thus propose that the use of unstressed complex pronouns in Fuyang Wu is not restricted to subjects, while it is clearly dependent on their information status as topic but never as focus.

⁷ A complex pronoun is also possible in this case.

Our proposal that unstressed complex pronouns are topics makes the correct prediction that unstressed and non-contrastive complex pronouns cannot be used as objects, unless they serve the discourse function of (secondary) topics, but not that of foci.

Recall the examples in (4), (5), and (7). As we noted, unstressed complex pronouns are ruled out in default postverbal object positions, as in (4) or in the KE object-marking construction, as in (5). In contrast, when the object is preposed to a sentence-initial position or between the subject and the verb without the object marker, it is possible to use complex pronouns, as in (7). This puzzle is now ready for explanation.

We first account for the grammaticality of (7). With regard to example (7a), as assumed at the beginning of this section, the sentence-initial position is the default topic position in Chinese languages in general, and the complex pronoun in (7a) happens to fall precisely into this topic position. According to Xu and Liu (1999) and Hu et al. (2003), Wu Chinese is a language which is more topic-prominent than Mandarin, in which there is also a secondary topic or subtopic position between the VP and the IP. We assume that Fuyang Wu is no exception to this generalization. Accordingly, the preposed complex pronoun in (7b) is located in a secondary topic position.

An independent piece of evidence about the topic status of complex pronouns in (7) comes from the use of resumptive pronouns in Fuyang Wu. See (10) for an illustration.

- (10) a. 我, z 你, 骂你过芒?
 $\eta\gamma^{313}$ zə?¹-n¹³ mo³⁴ n³¹³ ku³⁴ manj³⁵?
 1SG PREF-2SG scold 2SG EXP Q-PART
 ‘Me, have I scolded you?’
- b. (z)你, 我骂过你芒?
 zə?¹-n³¹³ $\eta\gamma^{313}$ mo³⁴ n³¹³ ku³⁴ manj³⁵?
 PREF-2SG 1SG scold 2SG EXP Q-PART
 ‘You, have I scolded you?’

When an object is preposed, it is possible to have a resumptive pronoun in the default object positions, as in (10a) and (10b). If we assume that topicalized elements are realized by movement from positions like subject and object, to the left-periphery of a sentence, then the resumptive pronoun can be seen as the overt realization of the trace left by the topicalized elements. Therefore, it further supports the claim that preposed elements in (7) are topics, even though there are no resumptive pronouns in either (7a) or (7b).

The ungrammaticality of the relevant examples in (4) and (5) is also explained by our generalization that unstressed complex pronouns can only be topics but not foci. Xu (2004), among others, claims that in (Mandarin) Chinese, the informational focus is restricted to a particular syntactic position, namely, the clause-final position. This observation is also applicable to the Fuyang dialect of Wu Chinese. Accordingly, the canonical postverbal object in the Fuyang dialect is a focus element. Therefore, in (4),

when an unstressed complex pronoun is used in a focus position, expectedly, it results in the ungrammaticality of the sentence.

In terms of (5), we suggest that the KE object-marking construction in the Fuyang dialect is a normal object position, parallel to the postverbal object position. This being the case, it is impossible to have a resumptive pronoun in the postverbal object position. As Li and Bisang (2012) have demonstrated, these two object positions are distinguished by the parameter of definiteness: the KE marked objects are realized by definite phrases, while postverbal objects can only be realized by indefinite phrases. However, no matter whether they are preposed and marked by KE or are located in a base-generated postverbal position, the object always acts as focus in Fuyang Wu.⁸ Given that an unstressed complex pronoun in (5) falls into a focus position, this explains why it is unacceptable.

To sum up, unstressed complex pronouns in the Fuyang dialect can only appear in topic positions, which crosscut the grammatical distinction of subject and object.

9.3 Stressed complex pronouns

In §9.2, we discussed the distributional constraints on the use of unstressed complex pronouns, with a brief comparison with common personal pronouns. In this section, we will turn our attention to stressed complex pronouns. What we mean by ‘stressed complex pronouns’ are those which bear a certain phonetic prominence. In Chinese, stress is realized via manipulation of phonetic primitives, such as pitch range and duration, which help to mark or realize contrastive topic or contrastive focus. In this respect, Xu (2004: 291) states that ‘(Mandarin) Chinese makes use of length and intensity rather than the rise and fall of pitch to indicate focus...’.

Stress is an important device in licensing complex pronouns in Fuyang Wu. As we will soon see, with the additional presence of the feature of stress, complex pronouns are not only able to appear in topic positions but also in focus positions. It is well known that stress can be used to express contrastiveness (Büring 2003). Consequently, stressed complex pronouns in Wu can function as contrastive topic and contrastive focus. We will discuss these two cases in turn.

Before proceeding, we would like to make two points that are worth a special mention with respect to the characteristics and function of stress in Chinese.

First, as reported in Xu (2004), ‘compared with European languages, it (Mandarin) uses more syntax and less phonology in focus realization’. Xu (2004: 291) states explicitly that in Mandarin, ‘even in a context where one of the constituents is clearly in focus, it need not always be stressed. In fact, when it takes the default focus position, stress is generally unnecessary. But it can be used for other purposes, for

⁸ This is different from the Mandarin BA construction, which is argued to be a secondary topic position (Li and Thompson 1981).

instance, for contrastive purposes.⁹ We propose that this generalization also applies to the Fuyang dialect of Wu Chinese.

Second, the phonetic characteristics of stress may be different from each other when they fall on topicalized and focalized elements. For example, in her study on Shanghai Wu, Y. Chen (2009) shows that contrastive focus is realized with a magnified Fo contour characteristic of the domain-initial lexical tone while contrastive topic is realized with the raising of the whole contour of the initial tone. In other words, ‘the effect of contrastive topic on F_o realization is rather local and restricted to the contrastively topicalized noun only’, while focus shows ‘a much more global effect on F_o ’ (Y. Chen 2009: 10). As will be shown soon, similar phonetic patterns may be found on stressed pronouns in Fuyang Wu.

9.3.1 Stressed complex pronouns as contrastive topics

We first consider the use of stressed complex pronouns as contrastive topics. A contrastive topic refers to an entity that the speaker wants to talk about and at the same time, due to its contrastiveness, it implies that there exist alternatives that the speaker may want to talk about too (Umbach 2001; see also Büring 1997).

One of the representative contexts in which contrastive topic is used is illustrated by the exchange in (11) (cf. Büring 1997, 2003).

- (11) A: Tell me about John. Did he watch Harry Potter?
B: I don't know about John, but Mary watched Titanic.

The question by A is about John, but B's reply is about Mary. Mary is therefore a topic that is interpreted contrastively. By uttering this sentence, B implicates that an alternative assertion regarding John cannot be made. A contrastive topic in English carries a certain accent, but a non-contrastive topic need not (Büring 2003).

We saw in §9.2 that unstressed complex pronouns are restricted to topic positions. However, complex pronouns in topic positions can also be given a certain phonetic prominence, which allows them to become stressed complex pronouns. Note that stressed complex pronouns in topic positions are characterized by the lengthening of the second syllable but not the first; that is, in the form of prefixed pronoun, only the pronoun is lengthened but not the prefix. As in Mandarin, stress in Wu expresses contrastiveness. Thus we claim that stressed complex pronouns in topic positions are contrastive topics.

⁹ According to Hartmann and Zimmermann (2009) (cf. Zimmermann 2008), the use of stress differs between intonational languages and tonal languages in general. They claim intonational languages obligatorily mark the existence of a contextually salient set of alternatives, i.e. focus, by using a pitch accent. As a result, every focus, contrastive or not, carries a pitch accent, often blurring the distinction between the two. In contrast, tonal languages, such as west Chadic languages, make use of pitch accent to express contrastiveness only.

Complex pronouns in Wu Chinese: Focalization and topicalization 237

Contrastive topics are possible both in complex sentences and in simple sentences, as will be illustrated in (13) and (14). Note that stressed elements are represented by bold and bracketed letters.

In complex sentences, such as coordinated sentences, the contrastive meaning is expressed by contrasting the individual referred to by the complex pronoun in one clause with some other individual(s), coded by nouns in another clause. Consider (12) and (13).

In (12), the speaker narrates the unfortunate events which happened to her and her husband, namely, that she was knocked down by a motorcycle, while her husband was robbed by a thief. These two events constitute a contrast.

- (12) 交交关年数喋, [z我]马达开撞勒一头, 娜外公么, 着贼。
qiɔ⁵³-qiɔ⁵³-kuan⁵³n.in³⁵sy³⁵ die, [zəʔ¹-ŋy³¹³] mo³¹³da³¹³k^hɛ⁵³ zan³¹³-lə
many-many years PRT PREF-1SG motorcycle knock-PFV
i?¹ dei³⁵, na¹³ ɳa³⁴kon⁵³ mə, dza?¹zə?¹.
one CL_{head} 2PL grandpa TOP get-robbed
'It has been many years since I was knocked down by a motorcycle and your grandpa got robbed.'

In (13), the speaker first talks about her grandson having travelled around a great deal, and then she brings herself into the limelight by saying 'I've also visited quite a few places'.

- (13) 你个侄孙子, 跑来跑去, [z我]也去过地方不少喋。
n³¹³ k^y³⁴ tsəʔ¹-sən⁵³tsɿ, bɔ⁵¹³-lɛ³⁵-bɔ⁵¹³-tç^hi³⁴
2SG-VOC CL grandson run-come-run-go
[zəʔ¹-ŋy³¹³] iə⁵¹³ tç^hi³⁴-ku³⁴ di³¹³faj⁵³ fə?⁵¹sɔ⁵¹³ die.
PREF-1SG also go-EXP place many SFP
'You, my grandson, have travelled around; I myself have also visited quite a few places.'

In simple sentences, the contrastive meaning is conveyed by contrasting the individual represented by a complex pronoun with some other individuals inferable from the context. Consider the example (14).

- (14) z伊, 格个, 春游去弗去格。
zəʔ¹-i³¹³ kəʔ⁵¹k^y³⁴ ts^hən⁵³iu³⁵ tç^hi fə?⁵¹ tç^hi³⁴ go³⁴.
PREF-3SG, then, spring.outing go NEG go PRT
'He, well, never went on spring outings.'

Example (14) means that he did not go on the travels organized by the Party Committee in the village, but his comrades and colleagues did, as inferred from the context. Thus, it is a contrast between his comrades and him.

These two groups of examples suggest that complex pronouns can be used contrastively when they are used as topics of the sentence. To achieve contrastiveness, phonetic stress is placed on the complex pronouns. Semantically, contrastive topics realized by complex pronouns usually relate the individual referred to by a complex pronoun to some other individual(s) in the discourse.

9.3.2 Stressed complex pronouns as contrastive focus

The conclusion we reached in §9.2 is that unstressed complex pronouns are not allowed in postverbal object positions nor in the KE object-marking construction, which we argued to both be focus positions. However, if we place stress onto this type of complex pronoun, it turns out that they become acceptable in these two cases. Note that when complex pronouns are emphasized in focus positions, both syllables of the complex pronouns are given expanded pitch range and lengthened. We will argue that stressed complex pronouns are found in contrastive focus. Compare (a) and (b) in examples (15) and (16).¹⁰

- (15) a. * 伊伽讲: “伊欢喜z 你”。
* i³¹³ ga³¹³kanj⁵¹³ “i¹³ fiuæn⁵³ci⁵¹³ zəʔ¹-n³¹³.”
3SG say 3SG like PREF-2SG
'She said: 'She likes you.'
- b. 伊伽讲: “伊欢喜[z 你]”。
i³¹³ ga³¹³kanj⁵¹³ “i¹³ fiuæn⁵³ci⁵¹³ [zəʔ¹-n³¹³].”
3SG say 3SG like PREF-2SG
'She said: 'She likes you.'
- (16) a. * 伊克z 我骂。
*i³¹³ k^həʔ⁵¹ zəʔ¹-ŋy³¹³ mo³⁴.
3SG OM PREF-1SG scold
'He rebuked me.'
- b. 伊克[z 我]骂。
i³¹³ k^həʔ⁵¹ [zəʔ¹-ŋy³¹³] mo³⁴.
3SG OM PREF-1SG scold
'He rebuked me.'

In (15a) and (16a), the object is realized by non-stressed complex pronouns. They are not acceptable. In contrast, in (15b) and (16b), when the complex pronouns are stressed, as indicated by bold letters in brackets, the sentences become grammatical. With stress, the pronouns express contrastive meaning. Example (15b) means that it is *you* that she likes and not somebody else. Example (16b) expresses that he rebuked even *me*.

¹⁰ The examples in (15) and (16) are not taken from our recordings but elicited during my field trip.

Complex pronouns in Wu Chinese: Focalization and topicalization 239

We argued earlier that postverbal objects and KE objects are information focus positions in Wu. Accordingly, we assume that with the accompaniment of stress, such stressed complex pronouns in these two positions are understood as contrastive focused elements. Moreover, as further supportive evidence, the examples in (15b) and (16b) can be treated in the same way as the canonical focus constructions in (17) and (18).

Examples in (17) represent two types of focus constructions, namely, the pseudo-cleft sentence (17a) and the cleft sentence, i.e. the *z1...ko* construction, in (17b). In these two constructions, it is the element following the copula 是 /z1/ that is focalized, granted that the copula is a focus marker. In both examples, stressed complex pronouns can be used.

- (17) a. 哭格是[z-你]，笑格也是[z-你]。

k^huo?⁵¹ ko z1³¹³ [zə?¹-n¹³], ciɔ³¹³ ko ie¹³ z1³¹³ [zə?¹-n¹³].
Cry MOD COP PREF-2SG laugh MOD also COP PREF-2SG
'The one who cried is you and the one who laughed is also you.'

- b. 只碗是[z-伊]敲破格。

tsə?⁵¹ uan⁵¹³ z1³¹³ [zə?¹-i³¹³] k^hɔ⁵³p^ha³⁴ ko.
CL bowl COP PREF-3SG break PRT
'It is he who broke the bowl.'

The examples in (18) show a different type of focus construction. The marker /lə/ 'even' is a focus-sensitive particle, or a marker that has an association with focus. In this case, it is also possible to use stressed complex pronouns. Compare (18a) and (18b).

- (18) a. *小王不认识 z-伊。

*ciɔ⁵¹³uan³⁵ fə?⁵¹ n̩in³⁴dza?¹ zə?³¹³-i¹³.
Xiao Wang NEG know PREF-3SG
'Xiao Wang did not know him.'

- b. 小王z-伊勒不认识。

ciɔ⁵¹³uan³⁵ za?¹-i³¹³ lə fə?⁵¹ n̩in³⁴dtsa?¹.
Xiao Wang PREF-3SG FOCUS NEG know
'Xiao Wang didn't even know him.'

The example in (18a) is a normal SVO sentence in Wu, which cannot take any unstressed complex pronouns as its object. In (18b), the object is preposed to a preverbal position between the subject and VP, and is marked by the focus-sensitive marker /lə/ 'even'. Example (18b) consequently uses the stressed complex pronoun, /za?-i/ 'he', to express the focus meaning that Xiao Wang did not even know him. Here, the complex pronoun is under the scope of the focus marker *le* and is thus a contrastive focus element.

In this subsection, we showed that stressed complex pronouns can be licensed in various focus positions to express contrastive meanings. These include the sentence-final position, such as objects in SVO sentences and in KE object-marking constructions, and significantly, special focus structures such as the cleft, pseudo-cleft, and /lə/ ‘even’ constructions. This shows that focus is indeed a syntacticized position in Fuyang Wu as well as being coded by stress.

In a combined conclusion to §9.2 and §9.3, the following recapitulation can be made:

- a) Unstressed complex pronouns are restricted to just preverbal positions, more precisely to topic positions, including primary and secondary topics.
- b) Stress is able to license complex pronouns in various syntactic positions, including even in focus positions. Given that stress in Wu expresses contrastiveness, stressed complex pronouns can serve as contrastive topics or contrastive foci respectively.

9.4 Decomposing complex pronouns

In this section, relying on Z. Chen’s (1996) assumption that the z-type prefix attached to pronouns in some Wu dialects is derived from the morpheme 是 /zɪ/, we look into the internal structure of complex pronouns in Wu.

We will address the question under what circumstances can /zɪ/ and pronouns form complex pronouns as a single lexical item. We argue that complex pronouns are the result of the fusion of the pronoun and the sentence-initial copula /zɪ/ in cleft sentences.

9.4.1 Reconstructing the prefix /zəʔ/ as the morpheme 是 /zɪ/

Z. Chen (1996) reconstructs the /z/ type of pronominal prefix in Wu dialects to be the morpheme 是 /zɪ/. Recall Table 9.1. He argues that the prefixes in different Wu dialects all start with the voiced consonant /z/, which is the same as the initial of the morpheme *shi* 是, which is phonetically /zɪ/ in the Fuyang dialect, and that the shape of the prefix, such as /zəʔ/ in the Fuyang dialect, is the result of glottalization of the vowel /ɪ/.¹¹ Chen also states that, in the Shaoxing dialect, the pronominal prefix and the verbal copula 是 have the same pronunciation, both of which are represented as /zeʔ/.

Additionally, Z. Chen provides us with some historical data on the occurrence of /zɪ/ with pronouns, which can be traced back to the (late) Tang dynasty or even earlier, such as in the *Dūnghuáng Biànwén Jí* 《敦煌变文集》, ‘Collection of Dun

¹¹ According to ZhengZhang (1995), glottalization by /ʔ/ is a common phenomenon in Wu dialects.

Huang Transformation Texts', (dated to the period from fifth to early eleventh century) and *Zǔtáng Jí*, 《祖堂集》'Collection From the Patriarchs' Hall', (written in the tenth century). Consider his examples in (19).

- (19) a. 是我今日莫逃得此难。

shì wǒ jīnrì mò táo dé cǐ nàn.
COP 1SG today NEG escape able DEM disaster
'It is me who cannot escape the disaster.'

From *Hán Qínhǔ Huàběn* of the *Dūnhuáng Biànwén Jí*
《敦煌变文集》之《韩擒虎话本》

- b. 是渠不得知东西……

shì qú bù dézhī dōngxī……
COP 3SG NEG know thing
'It is he who does not know the things.'

From *Cáo Shān Héshàng* of the *Zǔtáng Jí* Volume 8, p. 379
《祖堂集》第八卷《曹山和尚》

In terms of the grammatical function of the prefix on pronouns, Z. Chen simply mentions that the use of a prefix before pronouns is to 'highlight the deictic function' of the pronouns. Later, due to the weakening of this function, they behave like normal pronouns. We know that in Middle Chinese (third–twelfth centuries), the morpheme *是* /zɪ/ can either be a copula or a demonstrative (Wang 1937). Chen's argument that /zəʔ/ has a deictic function commits him to the standpoint that /zəʔ/ derives from the demonstrative /zɪ/ and not the copula /zɪ/.

We agree with Z. Chen's phonological reconstruction that the prefix /zəʔ/ is the morpheme *是* /zɪ/. However, we disagree with Chen's argument that the prefix /zəʔ/ has the same deictic function as a demonstrative. In the subsequent section, we will argue that the prefix /zəʔ/ develops from the copula /zɪ/ and not from the demonstrative /zɪ/.

9.4.2 Deriving the prefix /zəʔ/ from cleft sentences

In this subsection, we claim that complex pronouns in Fuyang Wu are reanalysed from the sentence-initial copula /zɪ/ combined with pronouns, when they are used in bare /zɪ/ cleft sentences.

We will make the assumption that cleft sentences in Wu work in a similar way to Mandarin Chinese. We now briefly introduce cleft sentences in Mandarin and their counterparts in Fuyang Wu.

Cheng (2008) and Paul and Whitman (2008) distinguish two types of cleft sentences in Mandarin: the *shì...de* cleft proper, as in (20a) and the sentence-initial bare *shi* cleft, as in (20b).

- (20) a. XiǎoWáng shì qù Shànghǎi chūchāi de.
Xiao Wang COP go Shanghai on:business DE
'It is to Shanghai that Xiao Wang goes on business.'
- b. shì XiǎoWáng qù Shànghǎi chūchāi.
COP Xiao Wang go Shanghai on:business
'It is Xiao Wang who goes to Shanghai on business.'

These two types of cleft sentences differ in which elements are focused. On the one hand, both Cheng (2008) and Paul and Whitman (2008) point out that in the *shì...de* cleft, the element immediately following the copula *shi* is the focus element. For example, in (21a), it is the subject *tā* that bears focus; in (21b), the focused element is the prepositional phrase *zài Běijīng* 'in Beijing'; in (21c), what is focused is the verb phrase, *xué yǔyánxué* 'study linguistics'.

- (21) a. shì [tā] zài Běijīng xué yǔyánxué de.
COP 3SG at Beijing learn linguistics DE
'It is he that studied linguistics in Beijing.'
- b. tā shì [zài Běijīng] xué yǔyánxué de.
3SG COP at Beijing learn linguistics DE
'It is in Beijing that he studied linguistics.'
- c. tā zài Běijīng shì [xué yǔyánxué] de.
3SG at Beijing COP learn linguistics DE
'It is linguistics that he studied in Beijing.'

On the other hand, Cheng (2008) claims that in the initial bare *shì* cleft, it is either the subject following *shì* that is focused, as in (22a) or the whole sentence that is focused, as in (22b).

- (22) a. shì [tā] zài Běijīng xué yǔyánxué. [Narrow focus]
COP 3SG at Beijing learn linguistics
'It is he that studied linguistics in Beijing.'
- b. shì [tā zài Běijīng xué yǔyánxué]. [Broad focus]
COP 3SG at Beijing learn linguistics
'It is the case that he studied linguistics in Beijing.'

When the subject is the focused element, it is called a narrow focus and when the whole sentence is focused, it is called a broad focus (Cheng 2008). Importantly, when the subject is focused, i.e. the narrow focus reading, it must have some intonational prominence. For broad focus, there is no stress needed.¹²

¹² However, according to Paul and Whitman (2008), only (22a) has a focus reading, i.e. the subject is focalized and (22b) is considered to be a 'propositional assertion'. Note that the difference between Cheng

Complex pronouns in Wu Chinese: Focalization and topicalization 243

If we go back to the examples of Middle Chinese discussed in Z. Chen (1996), such as those in (19), it is easy to see that all these examples start with the copula *shì*, which forms cleft sentences. However, there is no particle *de* at the end of the sentence. They, thus, behave like the initial bare *shì* cleft in (20b), and not the *shì...de* cleft as in (20a). Accordingly, the example in (19a) has the following two interpretations: a narrow focus reading in (23a) and a broad focus reading or propositional assertion in (23b).

- (23) a. 是[我]今日莫逃得此难。

shì [wǒ] jīnrì mò táo dé cǐ nàn.
COP 1SG today NEG escape able DEM disaster
'It is me who cannot escape the disaster.'

- b. 是[我今日莫逃得此难]。

shì [wǒ jīnrì mò táo dé cǐ nàn].
COP 1SG today NEG escape able DEM disaster
'It is the case that I cannot escape the disaster.'

In the Modern Fuyang dialect, it is possible to find the counterparts of both the bare *shì* clefts and the *shì...de* clefts. Given that the prefix to complex pronouns more likely develops from the bare copula clefts, we only introduce this type for the Fuyang dialect. In Modern Fuyang Wu, bare copula clefts are frequently used. They are formed by placing the copula /z/ at the sentence initial position while there is no particle, such as /ko/, at the end of the sentence. One such example is given in (24a).

- (24) a. 伽看来，是併勿好。

ga³¹³ k^han³⁴ l^e³⁵, zⁱ ia³¹³ fə?⁵¹ fiɔ⁵¹³.
this look come COP 3PL NEG good
'It seems that it's their fault.'

In the same vein, the example in (24a) can be interpreted with a narrow focus reading, as in (24b), and a broad focus reading or a propositional assertion, as in (24c). The element following the sentence-initial /z/ always carries stress, expressing contrastiveness.

- b. 伽看来，是[併]勿好，不是[你]勿好。 [Narrow focus]
ga³¹³ k^han³⁴ læ³⁵, zⁱ ia¹³ fə?⁵¹ fiɔ⁵¹³, fə?⁵¹ zⁱ ia³¹³ [n³¹³] fə?⁵¹ fiɔ⁵¹³.
This look come COP 3PL NEG good NEG COP 2SG NEG good
'It seems that it's their fault and not your fault.'

(2008) and Paul and Whitman (2008) in the details of their analyses of these two types of cleft sentences will not be of concern to us in this analysis.

- c. 伽看来，是[併勿好]. 你夔难过。 [Broad focus]
ga³¹³ k^han³⁴ læ³⁵, z1³¹³ [ia³¹³ fər⁵¹ fiɔ⁵¹³], n³¹³ fiɔ³⁴ nan³⁵ku³⁴.
This look come COP 3PL NEG good 2SG NEG sad
'It seems that it's their fault. Don't be sad.'

On the basis of the above arguments, we suggest that the prefix /zəʔ/ of the complex pronouns in the Fuyang dialect most likely comes from the sentence-initial copula /z1/ in cleft sentences.

9.5 Complex pronouns: From focalization to topicalization

In this section, we will address the two questions: (a) how is it syntactically possible that the copula followed by any of the pronouns allow them to be reanalysed into complex pronouns? (b) why are unstressed complex pronouns restricted to topic positions?

We assume that the frequent use of pronouns with the copula /z1/ in cleft sentences leads to the fusion of these two constituents into the single constituent of a complex complex pronoun. We hypothesize the following three-stage grammaticalization path which also involves morphologization whereby a copula verb develops into a prefix:

Stage I	Stage II	Stage III
z1+pronoun	> z1-pronoun	> zəʔ-pronoun.

In the first stage, /z1/ is a copula in cleft sentences and it is juxtaposed to the pronoun linearly. In the second stage, /z1/ is used as a focus marker before the pronoun, while it still has the same pronunciation as for its copular use. In this stage, the early forms of the complex pronouns emerge. In the third stage, the complete decategorialization of the prefix /z1/ is achieved, whereby it loses all features of its former verbal status to become a true pronominal prefix or clitic to the complex pronoun it forms, which is, itself, not focus-sensitive any more. Phonological attrition is well known to accompany grammaticalization (see Heine and Kuteva 2005), a process which specifically involves morphologization in this case. The weakening of pronunciation of the prefix from /z1/ to /zəʔ/ is one of the parameters of grammaticalization. Now we will show how this grammaticalization process is syntactically possible.

Kiss (1998, 2008) points out that identificational focus, i.e. contrastive focus, is a structural focus. Furthermore, this identificational focus is syntactically dependent. For example, in Hungarian, the immediate position before the VP is the focus element. In English, this type of focus is expressed by cleft sentences. According to Kiss, contrastive focus appears on the left-periphery of a sentence. It is an XP occupying the Specifier position of a FocP. In the case of Mandarin cleft sentences,

Complex pronouns in Wu Chinese: Focalization and topicalization 245

it is agreed that the focus element appears on the left periphery of a sentence, though there is no agreement about the status of the copula in clefts, such as *shì* in Mandarin. We simply assume that cleft sentences have the following structure: [FocP [IP]].

Topic elements may appear on the left-periphery of the sentence. As stated previously, Wu Chinese is a topic-prominent language, in which the sentence-initial position is the default topic position. We assume that there is an independent syntactic position for topic phrases. Namely, topic phrases occur in a position before the subject, i.e. [TopP [IP]].

- (25) z-伊_{Topic}哦，我叫伊过嚟。
zə?¹-i³¹³ Topic o ɿ³¹³ tɕio³⁴ i³¹³ ku³⁴ die.
PREF-3SG Top 1SG call 3SG EXP PRT
'As for him, I have asked him already.'

In addition, the topic element and the focus element can co-occur in the same sentence. The topic must precede the focus, not vice versa.

- (26) z-伊_{Topic}哦，是王老师_{Focus}克伊骂。
zə?¹-i³¹³ Topic o zi³¹³ uan³⁵ lɔ¹³ sɪ¹³ Focus kʰə²⁵¹ i³¹³ mo³⁴.
PREF-3SG Top COP Wang teacher OM 3SG scold
'As for him, it is Mr. Wang who scolded him.'

We assume that there are two extra projections above IP, namely, TopP and FocP. They are arranged hierarchically as: [TopP [FocP [IP]]].¹³

In the first stage, cleft sentences have simple pronouns as subject. They are preceded by the focus marker, namely, the verbal copula /zi/. As the subject of the sentence, simple pronouns are located at [Spec, IP]. In this case, either the Specifier is focused, which leads to the narrow focus reading, or the whole IP is focused, which leads to the wide focus reading.

In the second stage, we assume that due to the fact that the copula and simple pronouns are adjacent to each other, they are fused into a single lexical item. This is the early form of complex pronouns. This fused form is basically a pronoun in nature, but it also inherits the focus-marking property of the copula. In other words, at the early stage, the fused form of *copula + pronoun* is also focus-sensitive. Syntactically, it is highly possible that the focus-sensitive complex pronouns are located in the FocP domain, say, [Spec FocP] for instance.

In terms of its interpretation, complex pronouns in the old Fuyang dialect were able to express 'exhaustiveness' in cleft sentences à la Kiss (1998). Kiss (1998: 245) argues that identificational focus is characterized as exhaustiveness: it 'represents a

¹³ See Paul (2005) for the hierarchy of the elements occurring on the left periphery of sentences in Mandarin.

subset of the set of contextually or situationally given elements for which the predicate phrase can potentially hold; it is identified as the exhaustive subset of this set of which the predicate actually holds'. Along this line, we assume that complex pronouns in (old) Fuyang Wu always presuppose an existential set which is referential, out of which a certain member is picked out to be contrasted with the rest.

However, in Modern Fuyang Wu, i.e. in the third stage, the prefix in the paradigms of complex pronouns is no longer a focus marker and complex pronouns in Fuyang are used independently of stress and also non-contrastively. It is also evidenced by the phonetic reduction from the copula /zɪ/ to the pronominal clitic /zəʔ/. In other words, complex pronouns have evolved from focalized elements to non-focalized elements in Modern Fuyang Wu. However, as we observed before, this particular kind of use is still syntactically restricted. They are only possible in topic positions.

We suggest that when complex pronouns are fully grammaticalized and they lose their focus-sensitivity, they can no longer stay in the focus position and have to be raised to a higher position than FocP, such as [Spec TopP].

To summarize so far, diachronically speaking, according to our hypothesis, unstressed complex pronouns develop from stressed ones in cleft sentences, where they are interpreted as contrastive focus elements at this early stage of their development. However, due to the de-focalization of complex pronouns, they are forced to leave the focus position to move to a higher position, such as, a topic position, where stress is lost. In contradistinction to this, when complex pronouns receive prosodic stress, they may on the other hand be re-employed in these focus positions, thus possibly beginning a new cycle.

9.6 Conclusion

This chapter argues that the seeming subject/object asymmetry of complex pronouns in Fuyang Wu is better re-cast as a distinction between topic and focus. In so doing, we made a distinction between stressed and unstressed complex pronouns: unstressed complex pronouns are available only in topic positions, while stressed complex pronouns are possible in both topic and focus positions, where they function as contrastive topics and contrastive focus elements respectively.

We also claimed that complex pronouns are derived from the fusion of the sentence- initial copula and pronouns in cleft constructions, which are characterized by 'exhaustiveness' à la Kiss (1998). We explained that the reason why (unstressed) complex pronouns are restricted to topic positions is that the loss of contrastiveness for these complex pronouns, due to the loss of stress after fusion of the copula and pronoun has occurred, leads to a further syntactic change which shifts them from a focus position to a topic position.

Acknowledgements

The data used in this chapter were mainly collected during a field trip conducted in Fuyang, Zhejiang Province, in August 2011 (the author himself is a native speaker of Fuyang Wu). I would like to express thanks to Hilary Chappell, Yiya Chen, Katia Chirkova, Alain Peyraube, and the SINOTYPE team members for providing constructive feedback when presenting earlier versions of this analysis. Special thanks goes to Hilary Chappell and YiYa Chen, who provided me with helpful written comments. Victor Pan also deserves a ‘thank you’ for discussing topic/focus issues with me on several occasions. I am solely responsible for any errors.

10

Comparative constructions of inequality in the Southern Min dialect of Hui'an

WEIRONG CHEN

10.1 Introduction

This analysis aims to examine comparative constructions of inequality in the Southern Min dialect of Hui'an 惠安, a variety of Southern Min spoken in Hui'an County in Fujian 福建 province of China, which belongs to the Quan-Zhang, that is, Quanzhou 泉州 and Zhangzhou 漳州 subgroup of the Southern Min dialect family. The comparative construction of inequality (hereafter termed ‘comparative construction’) is defined as one which ‘has the semantic function of assigning a graded (i.e. non-identical) position on a predicative scale to two (possibly complex) objects’ (Stassen 1985: 24). Comparative constructions may involve four constituents: the comparee (COM), the standard of comparison (STA), the (comparative) predicate (PRED), and the comparative marker (CM) (Stassen 1985: 24–6; Ansaldi 1999: 39; Chappell and Peyraube this volume). The term ‘the (comparative) predicate’ here refers to the constituent that ‘states the dimension of the comparison’ (Ansaldi 1999: 39). This is different from ‘the predicate’ in traditional grammar, which functions as one of two main parts of a sentence, with the other part being the subject.

As stated in Chappell and Peyraube (this volume), the predicate in comparative constructions is predominantly expressed by an adjective, but not confined to an adjective in Sinitic languages. We follow Ansaldi (1999) and Chappell and Peyraube (this volume) in using the term ‘comparative marker’ to refer to markers for both the standard and the predicate: the former is typically used to introduce the standard of comparison, such as the comparative marker *bǐ* 比 in Mandarin Chinese; whereas the latter is generally used to modify the predicate, such as the comparative marker *gwo*³ 過 in Cantonese, which functions as a verbal complement of the predicate. These two

TABLE 10.1. Comparison of analytic approaches for comparative schemas of superiority

	STASSEN 1985, 2005, 2011	HEINE 1997
1:	Separative	Source
2:	Allative	Goal
3:	Locative	Location
4:	Exceed	Action
5:	Conjoined	Polarity
6:	Particle	Sequence
7:	–	Similarity
8:	–	Topic

types of comparative marker are referred to as a dependent-marker of comparison (DMC) and a head-marker of comparison (HMC) respectively in Ansaldi (1999: 39).

In his typological study on comparative constructions, Stassen (1985: 28–47) identified six basic types of comparative in a sample of 110 languages based on the encoding of the standard NP: the separative comparative; the allative comparative; the locative comparative; the exceed comparative; the conjoined comparative; and the particle comparative. Another important typological study on comparative constructions, Heine (1997), argued that most comparative constructions in the languages of the world are derived from a limited number of conceptual source structures (or event schemas), and identified eight main event schemas: Action; Location; Source; Goal; Polarity; Sequence; Similarity; and Topic. A comparison between these two proposals is given in Chappell and Peyraube (this volume), in which a table showing the comparison is reproduced in Table 10.1

In terms of Sinitic languages, much work has been done on comparative constructions during the past two decades, for example, Li and Lien (1995); Ansaldi (1999), (2010); Zhao (2002a); Li (2003); Chang and Kwok (2005); Wu (2010); Chappell and Peyraube this volume, among others. Ansaldi (1999: 105) identifies three main types of (target) comparative construction in Sinitic languages: (i) Head-marking comparatives based on the Surpass source construction; (ii) Dependent-marking based on the Similarity source construction; and (iii) Double-marking, as follows.

10.1.1 Three main types of comparative proposed in Ansaldi (1999)

(i) Head-marking/Surpass type

In the Surpass type of comparative, the standard of comparison is required to be constructed as the direct object of the main verb, observed to be a property of the source construction (Ansaldi 1999: 64). An example from Cantonese is given in (1),

where the standard of comparison *Soeng⁶ Hot²* 上海 ‘Shanghai’ functions as the direct object of the verb phrase *jit⁶ gwo³* 热过 ‘is warmer than’.

- (1) 香港 热 過 上海¹。
Hoeng¹ Gong² jit⁶ gwo³ Soeng⁶ Hot².
Hong Kong warm HMC Shanghai
'Hong Kong is warmer than Shanghai.' (Ansaldi 1999: 105)

This type of comparative construction is one of the two predominant comparative construction types in contemporary Sinitic languages, and widely attested in Southern and Southwestern China, in particular Yue, Hakka, and Southwestern Mandarin (cf. Li 2003). This type of comparative basically equates with the second structural type (Type II, Surpass type, NP_A – VERB – CM – NP_B), proposed by Chappell and Peyraube (this volume), the Exceed type in Stassen (1985), and is associated with the ‘Action schema’ in Heine (1997).

(ii) Dependent-marking/Similarity type

In this type of comparative, the standard of comparison in the source construction is marked by an element originally meaning ‘to look like’, ‘like’, ‘comparatively’, and ‘to be compared to’ (Ansaldi 1999: 64, 119). An example from Mandarin Chinese is given in (2), where the comparative marker *bì* 比 ‘compared to’ is used to introduce the standard of comparison *jiějie* 姐姐 ‘elder sister’.

- (2) 弟弟 比 姐姐 聰明。
didi bì jiějie cōngmíng.²
brother DMC elder.sister smart
'Little brother is smarter than elder sister.' (Ansaldi 1999: 105)

As mentioned in Ansaldi (1999: 64), the term ‘Similarity’ is used in a slightly different way from that in Heine (1997: 118), where, in the Similarity Schema, a relationship of similarity or equivalence is asserted between the comparee and the standard of comparison. The basic structure of the Similarity Schema in Heine (1997: 118) is [A_{COM} is PRED (like) B_{STA}], and can be paraphrased as [A_{COM} is PRED-er compared to B_{STA}].

Unlike Ansaldi (1999), comparative constructions such as example (2) are separately categorized as a new ‘Compare’ schema in Chappell and Peyraube (this volume), since they claim that the comparative marker *bì* ‘compared to’ does not have the meaning of ‘be similar to’ in the comparative construction, and thus cannot, strictly speaking, be treated as a similarity comparative. This type of comparative is the most common type in Sinitic languages: it is widely used in Northern China and has been adopted elsewhere in Sinitic (cf. Ansaldi 1999; Li 2003).

² The romanizations used in this chapter are the IPA for Hui'an Southern Min, and *pīnyīn* for Standard Mandarin, unless examples are being cited from other references which follow the original in this case.

(iii) Double-marking type

Double-marking refers to a comparative construction which has both a marker for the standard and a marker for the predicate. An example from the Southern Min dialect of Hui'an (hereafter termed ‘the Hui'an dialect') is given in (3), where the dependent-marker of comparison *pi*³ 比 ‘compared to' is used to introduce the standard of comparison *i*³ 椅 ‘chair', and the head-marker of comparison *kha?*⁷ 怡 modifies the comparative predicate *tay*⁵ 重 ‘heavy' to denote comparison.

- (3) 桌 比 椅 怡 重。
tɔ?7 pi³⁻² i³ kha?7-8 tay⁵.
desk DMC chair HMC heavy
'The desk is heavier than the chair.'

This type of comparative construction is also termed a hybridized construction and is so far attested only in Min and Hakka (Li 2003; Chappell this volume).

Besides these, in Ansaldi (1999: 43–4), comparative constructions with and without an overt standard of comparison are regarded as ‘long comparatives’ and ‘short comparatives’, respectively.

So far, it is quite clear that Types (i) and (ii) among the three types of comparative proposed in Ansaldi (1999) are the two main types of comparative in Sinitic languages, as mentioned in Chappell and Peyraube (this volume), since Type (iii) has a comparatively limited distribution in Sinitic languages. Significantly, Type (iii) is an important type of comparative construction in contemporary Southern Min (cf. Wang and Zhang 1994; Li and Lien 1995; Zhou and Ouyang 1998, among others). In addition, previous studies such as Li (2003) have shown that certain other types of comparative construction are also mainly attested in Min, while they are rarer in many other Sinitic languages. This analysis attempts to further examine these ‘special’ comparatives in Min which are quite limited in their distribution, with examples from the Hui'an dialect.

We have identified six main types of comparative construction in the Hui'an dialect, as shown in §10.1.2.

10.1.2 *Six main types of comparative in the Hui'an dialect in order of decreasing frequency in textual data:*

- a. Double-marking/hybridized comparatives
i.e ‘A_{COM} + *pi*³比+ B_{STA} + *kha?*⁷ 怡+ PRED (+ MW)’
- b. Short comparatives
i.e ‘A_{COM} + *kha?*⁷ + PRED (+ MW)’
- c. Head-marking comparatives with the marker *kha?*⁷
i.e ‘A_{COM} + *kha?*⁷ + PRED + B_{STA} (+ MW)’
- d. Zero-marked comparatives
i.e ‘A_{COM} + PRED + B_{STA} (+ MW)’

- e. Comparatives with the marker *khui*⁵
i.e. ' $A_{COM} + PRED + khui^5$ 去 + B_{STA} '
- f. Dependent-marking comparatives with the marker *pi*³
i.e. ' $A_{COM} + pi^3 + B_{STA} + PRED$ (+ MW)'

Note that type (a) and type (b), i.e. the hybridized or double-marked comparative ' $A_{COM} + pi^3 + B_{STA} + kha?^7 + PRED$ (+ MW)' and the short comparative ' $A_{COM} + kha?^7 + PRED$ (+ MW)' are the two dominant types, examples of which can be found in the spoken data we collected. In addition, these two types are commonly used not only by the younger generation, but also by the elder generation. Examples of type (c), i.e. the head-marking comparative ' $A_{COM} + kha?^7 + PRED + B_{STA}$ ' can also be found in the spoken data. However, type (c) is more common in the speech of the elder generation than in the speech of the younger generation. The other three types, i.e. types (d), (e), and (f), are attested mainly via elicitation. Among these three types, type (d) is more common than the other two types, whereas type (f), the most common one found in Sinitic languages, is the least common type of comparative in the Hui'an dialect.

I will now present each type of comparative construction in turn.

10.2 Double-marking/hybridized comparatives: $A_{COM} pi^3 B_{STA} kha?^7 PRED$ (MW)

This type of comparative construction is widely used not only in the Hui'an dialect, but also in other Southern Min varieties in Fujian, Taiwan, and Singapore (cf. Chen 1991a: 299; Chen 1991b: 458; Yang 1991: 260; Wang and Zhang 1994: 143; Li and Lien 1995: 72; Zhou and Ouyang 1998: 398; Li 2000: 131; Zhou and Zhou 2000: 130, among others). Ansaldi (1999) calls this type of comparative construction in Southern Min the double-marking type, since there exist two markers: one marker for the standard of comparison (i.e. *pi*³) and one marker for the predicate (i.e. *kha?*⁷). Chappell and Peyraube (this volume), however, use the term 'hybridized construction' to refer to this type, since it involves a blend of different types of comparative in Sinitic languages, as will be mentioned later.

In the Hui'an dialect, the comparee and the standard of comparison in this type of comparative construction can be encoded by substantive or predicative words/phrases. Two examples are given in (4) and (5).

- (4) 即蕊 比 迹蕊 怡 水 蠻 多。
*tsit*⁷⁻⁸*liu*³ *pi*³⁻² *hit*⁷⁻⁸*liu*³ *kha?*⁷⁻⁸ *sui*³ *ban*⁵⁻⁴ *tsue*⁵.
this-CL DMC that-CL HMC beautiful rather much
'This (flower) is much more beautiful than that (flower).'

Comparative constructions of inequality in Hui'an 253

- (5) 有 比 無 恰 好。
*u*⁴ *pi*³⁻² *bo*² *kha?*⁷⁻⁸ *ho*³.
have DMC not.have HMC good
'Something is better than nothing.' (Literally: 'to have' compared with 'to not have' comparatively good)

In (4), the comparee and the standard of comparison are encoded by the demonstrative phrase *tsit*⁷⁻⁸ *liu*³ 'this (flower)' and *hit*⁷⁻⁸ *liu*³ 'that (flower)', respectively. In (5), the comparee and the standard are expressed by the verbs *u*⁴ 'have' and *bo*² 'not have', respectively. In both example (4) and (5), *pi*³ functions as a dependent-marker of comparison introducing the standard of comparison, i.e. *hit*⁷⁻⁸ *liu*³ 'that (flower)' and *bo*² 'not have', while *kha?*⁷ is used as a head-marker of comparison modifying the predicate, which is encoded by the adjective *sui*³ 'beautiful' and *ho*³ 'good' respectively. Unlike (5), the adjective *sui*³ 'beautiful' in (4) is followed by the measure phrase *ban*²⁻⁴ *tsue*⁵ 'much' denoting the degree of the difference between the comparee and the standard, which suggests that the measure phrase is optional in this type of comparative construction.

The following example shows that the predicate can also be encoded by a non-monosyllabic adjective such as *u*⁴ *tshui*⁵⁻³ *tsi?*⁸ 有喙舌 'talkative'.

- (6) 伊 比 我 恰 有 喙舌。
*i*¹ *pi*³⁻² *ua*³ *kha?*⁷⁻⁸ *u*⁴ *tshui*⁵⁻³-*tsi?*⁸.
3SG DMC 1SG HMC have mouth-tongue
'She is more talkative than me.'

In (6), the comparative predicate *u*⁴ *tshui*⁵⁻³ *tsi?*⁸, consisting of the verb *u*⁴ 'have' and the noun *tshui*⁵⁻³ *tsi?*⁸ 'mouth and tongue', is a fixed expression functioning like an adjective expressing 'talkative'.

A topic may be involved in this type of comparative construction to denote the parameter based on which a comparison is made. The topic can be placed before the comparee, i.e. the comparative construction takes the form of 'Topic A_{COM} *pi*³ B_{STA} *kha?*⁷ PRED (MW)', as in (7).

- (7) 坐車 廣州 比 泉州 恰 加 時間。
*tsə*⁴*tshia*¹ *kŋ*³⁻²*tsiu*¹ *pi*³⁻² *tsuan*²⁻⁴*tsiu*¹ *kha?*⁷⁻⁸ *ke*¹ *si*²⁻⁴*kan*¹.
sit-bus PN DMC PN HMC more time
'It takes more time to take the bus to Guangzhou than to Quanzhou.'

In (7), the verb phrase *tsə*⁴*tshia*¹ 'take bus', preceding the comparee (*kŋ*³⁻²*tsiu*¹ 'Guangzhou'), functions as a topic indicating that the comparison between the comparee and the standard of comparison is based on 'taking the bus'.

The following example illustrates that the topic can also be inserted between the comparee and the dependent-marker of comparison, i.e. the comparative construction takes the form of ‘A_{COM} Topic pi³ B_{STA} kha?⁷ PRED (MW)’.

- (8) 我 讀 比 伊 怡 好。
*ua*³ *thak*⁸ *pi*³⁻² *i*¹ *kha?*⁷⁻⁸ *ho*³.
1SG study DMC 3SG HMC good
'I study better than him.'

In (8), the comparee (*ua*³ 'I') and the standard of comparison (*i*¹ 'he') are compared in terms of studying, which is encoded by the verb *thak*⁸ 'study'. In addition, *thak*⁸ functioning as a topic, is inserted between the comparee (*ua*³) and the dependent-marker of comparison (*pi*³).

The topic can also be placed between the standard of comparison and the head-marker of comparison, as in (9), where the topic *huay*¹ 風 'wind' is inserted between the standard of comparison *ku*⁴*lin*² 舊年 'last year' and the head-marker of comparison *kha?*⁷. In other words, the comparative construction takes the form of ‘A_{COM} pi³ B_{STA} Topic *kha?*⁷ PRED (MW)’.

- (9) 今 年 比 舊 年 風 怡 透。
*ken*¹*lin*² *pi*³⁻² *ku*⁵⁻⁴*lin*² *huay*¹ *kha?*⁷⁻⁸ *thau*⁵.
today-year DMC old-year wind HMC strong
'The wind of this year is stronger than that of last year.'

Examples (7)–(9) show that the parameter on which a comparison is based can be overt and encoded by a topic, unlike examples (4)–(6) where the comparative parameter is covert, but can be deduced from the predicate. For example, the predicate *sui*³ 'beautiful' in (5) indicates that the comparee *tsit*⁷⁻⁸ *liu*³ 'this (flower)' and the standard of comparison *hit*⁷⁻⁸ *liu*³ 'that (flower)' are compared in terms of their appearance.

It is generally known that the comparative construction in Standard Mandarin takes the form of [A_{COM} + *bī* 比 + B_{STA} + PRED], i.e. the standard of comparison is placed between the comparative marker (*bī* 比) and the comparative predicate, which is disharmonic with the basic SVO word order pattern of Standard Mandarin (cf. Dryer 1992; see also Chappell and Peyraube this volume), and not compatible with Greenberg's Universal 22 as shown here.

No. 22. If in comparisons of inequality the only order, or one of the alternative orders, is standard-marker-adjective, then the language is postpositional. With overwhelmingly more than chance frequency, if the only order is adjective-marker-standard, the language is prepositional.

(Greenberg 1966b: 89)

Similarly, in the double-marking/hybridized type of comparative construction in the Hui'an dialect as in examples (4)–(9), the standard of comparison is placed between the marker of the standard (dependent-marker of comparison) and the

predicate, i.e. 'DMC-STA-HMC-PRED', which also happens to be neither harmonic with its basic SVO word order pattern nor compatible with Greenberg's Universal 22.

Liu (2003b: 17–18), however, points out that this double-marking/hybridized type of comparative construction satisfies the principle concerning relators put forward by Dik (1997: 406): relators have their preferred position in between their two relata. According to Dik (1997: 398), 'The class of relators contains those grammatical elements which serve to link two constituents together, and/or to mark the function(s) of a constituent as specified in underlying clause structure'. In the case of comparatives, in order to be compatible with the principle concerning the relators, the comparative marker needs to be placed between the predicate and the standard of comparison. The comparative construction [A_{COM} + *bǐ* 比 + B_{STA} + PRED] in Standard Mandarin does not satisfy the principle concerning relators, since the marker *bǐ* 'compared to' is not placed between the standard of comparison and the predicate, whereas the marker of the predicate (head-marker of comparison) such as *kha?*⁷ 怪 in the double-marking/hybridized type of comparative in the Hui'an dialect fills in the 'gap' between the standard of comparison and the predicate (Liu 2003b: 17–18).

As with other Southern Min varieties, the head-marker of comparison *kha?*⁷ can be substituted by *ko?*⁷⁻⁸ *kha?*⁷ 傍怪 'even more', which is similar to the adverb *gèng* 更 '(even) more' in Mandarin Chinese, as in (10).

- (10) 我 比 伊 佧 怪 懸。
*ua*³ *pi*³⁻² *i*¹ *ko?*⁷⁻⁸ *kha?*⁷⁻⁸ *kuin*².
1SG DMC 3SG even DMC tall
'I am even taller than her.'

Example (10) implies that both the comparee (*ua*³ 'I') and the standard of comparison (*i*¹ 'she') are tall, based on which, the comparee is taller than the standard of comparison. However, without *ko?*⁷, i.e. *ua*³ *pi*³⁻² *i*¹ *kha?*⁷⁻⁸ *kuin*² 我比伊怪懸 'I am taller than her' does not imply that both the comparee and the standard of comparison are tall. In other words, it could be the case that both the comparee and the standard of comparison are short, while the comparee is nonetheless taller than the standard of comparison.³

The double-marking type of comparative is probably not uncommon crosslinguistically. At least, it is attested in many European languages such as German, English, and Italian. An example from English is given in (11), where the comparative predicate *old* is marked by the suffix *-er*, and the standard of comparison *me* is marked by the conjunction *than*. This is known as the Particle comparative in Stassen's framework and the Sequence comparative in Heine's.

³ Note also that the English translation provided is not semantically equivalent to the comparative construction of the Hui'an dialect in example (10), since the English sentence 'I am even taller than her' does not necessarily imply that both the comparee and the standard are tall.

(11) *He is older than me.*

However, unlike the double-marking type of comparative in European languages as in (11), the double-marking type of comparative in Sinitic languages (e.g. Min and Hakka) tends to be a result of hybridization due to language contact (see also Li and Lien 1995; Ansaldi 1999; Chappell and Peyraube this volume). This type of comparative is not attested in the *Li Jing Jì* 荔鏡記, a play with a mixture of both the Chaozhou and Quanzhou dialects of Southern Min written in Ming dynasty (AD 1368–1644), and may have arisen due to language contact (Li and Lien 1995: 74, 77). Ansaldi (1999: 107, 140) suggests that the double-marking type of comparative construction may arise in Sinitic as a consequence of a hybrid of two different source constructions, i.e. the Surpass type (e.g. with *kha?*⁷ as a head-marker of comparison) and the Similarity type (with *pi*³ as a dependent-marker of comparison), due to northern influence. Following this proposal, the double-marking/hybridized comparative in Southern Min, including the Hui'an dialect, may be the result of interaction of the Similarity/Compare type (i.e. A_{COM} *pi*³ B_{STA} PRED) and head-marking comparatives with the marker *kha?*⁷ (i.e. A_{COM} *kha?*⁷ PRED B_{STA} (MW), see §10.4). This is also what Chappell (this volume) suggests. However, we are not able to rule out another possibility: the double-marking/hybridized comparative ‘A_{COM} *pi*³ B_{STA} *kha?*⁷ PRED (MW)’ may be a blend of the Similarity/Compare type (i.e. A_{COM} *pi*³ B_{STA} PRED) and short comparatives (i.e. A_{COM} *kha?*⁷ PRED (MW), see §10.3), since, as will be shown in the following discussion, short comparatives are one of the predominant comparative construction types in both earlier Southern Min and contemporary Southern Min.

Note that unlike *gwo*³ in Cantonese which functions as a verbal complement in postverbal position as in example (1), *kha?*⁷ functions as a comparative marker in preverbal position. As suggested in §10.1, *kha?*⁷ can be used in three types of comparative construction, which is the same as the distribution of *k'ah* in Southern Min, pointed out by Ansaldi (1999: 114). In the Hui'an dialect, however, *kha?*⁷ can also be used as a degree adverb meaning ‘a bit; fairly’ in non-comparative constructions, as in (12).

(12) Non-comparative constructions with *kha?*⁷ as a degree adverb

- a 湯 怡 無 甜。
*thy*¹ *kha?*⁷⁻⁸ *bo*²⁻⁴ *tin*¹.
soup a.bit not.have sweet
'The soup is not tasty enough.'

- b 汝 著 怡 主動。
*lu*³ *tio?*⁸⁻⁴ *kha?*⁷⁻⁸ *tsu*³⁻²-*tɔŋ*⁵.
2SG should fairly active
'You need to be fairly active.'

In (12a), the subject *thy*¹ ‘soup’ is not explicitly compared to anything else, and the adjective phrase *kha?*⁷⁻⁸ *bo*²⁻⁴ *tin*¹ ‘not tasty enough’ is used to denote the degree of tastiness of the soup, i.e. somewhere between tasty and non-tasty, but closer to tasty, which suggests that adding a bit of salt, but not too much, will be enough. The context for (12b) is that the speaker is giving some advice on how to find a job to the hearer (i.e. the subject *lu*³ ‘you’) who is about to graduate from university and look for a job. The speaker neither makes a comparison between the hearer and anyone else, nor compares the present behaviour of the hearer and his future behaviour. The use of *kha?*⁷ here means that the active approach of the subject in taking the initiative to look for a job should achieve a certain degree, otherwise the subject will not get what he wants.

The two examples in (12) take the form of [NP *kha?*⁷ VP] with *kha?*⁷ functioning as a degree adverb meaning ‘a bit; fairly’, rather than a comparative marker. This adverbial use of *kha?*⁷ may suggest that the short comparatives (i.e. A_{COM} *kha?*⁷ PRED (MW), see §10.3) are not highly grammaticalized comparative constructions, since similar syntactic structures may be regarded as a non-comparative construction where *kha?*⁷ is a degree adverb without denoting comparison. However, in examples of double-marking/hybridized comparatives above and head-marking comparatives with *kha?*⁷ (i.e. A_{COM} *kha?*⁷ PRED B_{STA} (MW)) described in §10.4, *kha?*⁷ indicates comparison. This may suggest that *kha?*⁷ in the Hui'an dialect undergoes grammaticalization from a degree adverb to a marker of comparison, although the relevant pathway remains unclear.

The counterparts of *kha?*⁷ in other Min dialects have also been reported to function as a degree adverb, or, if not, also have a meaning of ‘a bit, a little’. For example, in a dictionary of cognates between Southern Min and Ancient Chinese, Lin (1998: 374) points out that *kah* in Southern Min is a degree adverb meaning ‘a bit, a little’. The counterpart of *kha?*⁷ in the Singapore variety of Southern Min has also a meaning of ‘a bit, a little’ (Zhou and Zhou 2002: 291). Zhou and Ouyang (1998: 379, 398–400) and Zhou (1999: 220, 222) suggest that *k'a?* in the Xiamen 廈門 variety of Southern Min is originally a degree verb meaning ‘fairly’, but has been gradually grammaticalized to be a marker of comparative constructions. *K'a?* in the Chaozhou 潮州 variety of Southern Min and in the Fuzhou 福州 variety of Eastern Min functions as a degree adverb meaning ‘too’ (cf. Cai 1991: 235; Feng 1998: 395), which may represent an earlier meaning of the word which subsequently became bleached.

Similar extensions from a degree adverb to a comparative marker are also attested in other Sinitic languages, although a post-predicate degree adverb, rather than a pre-predicate degree adverb such as *kha?*⁷ in the Hui'an dialect, may be involved. For example, in Xianghua 乡话, an unclassified Sinitic language spoken in Xiangxi 湘西 in Hunan 湖南 province of China, the morpheme *sa*⁵⁵些 is originally a degree adverb meaning ‘a little’, but functions as a comparative marker in polarity comparatives (cf. Chappell this volume), as in example (13), where *sa*⁵⁵ follows the predicate *fi*²¹³ 肥 ‘fat’ and *ua*⁵⁵ 瘦 ‘thin’.

- (13) 你 肥 些， 我 瘦 些。
n̩i²⁵ *f²¹³* *s^{a⁵⁵},* *wu²⁵* *ua⁵⁵* *s^{a⁵⁵}.*
2SG fat little 1SG thin little
'You're fatter and I'm thinner.' = 'You're fatter than me.'

In Cantonese, the morpheme *dī* 咩 can be used to simply denote comparison in the short comparative [A_{COM} PRED *dī*], i.e. where the standard of comparison is not explicitly mentioned but implied by the context, as in (14), although it may also mean 'rather' or 'a bit' in a similar structure (Matthews and Yip 1994: 166).

- (14) 今日 热 咩。
gāmyaht *yih^t* *dī.*
today hot a.bit
'It's hotter today.' (Matthews and Yip 1994: 166)

10.3 Short comparatives: A_{COM} *kha?*PRED (MW)

As stated in Ansaldi (1999: 43), comparatives often appear in a shortened, or truncated form without an overt standard of comparison in conversation and real language use. This is also true for comparatives in the Hui'an dialect. The term 'short comparative' is used by Ansaldi (1999) to refer to the comparative construction with a shortened form, i.e. the comparative without an overt standard of comparison. The short comparative in the Hui'an dialect typically takes the form of [A_{COM} *kha?*PRED], in which *kha?* serves as a head-marker of comparison, like the comparative marker *kha?* in the double-marking type of comparative examined in §10.2.

The short comparative is also widely used in other modern Southern Min varieties such as Taiwanese Southern Min (Lu 2003: 92–3) and the Jieyang 揭陽 variety of Southern Min (Xu 2007: 278–9). In Taiwanese Southern Min, the comparative marker in the short comparative is the counterpart of *kha?* in the Hui'an dialect, whereas the short comparative in the Jieyang variety is syntactically indicated by markers such as *iau⁵³*, *lau⁵³*, and *zu⁵³* (cf. Lu 2003: 92; Xu 2007: 278). Xu (2007: 278) also mentions that short comparatives in the Jieyang variety are 'only possible if the Sta [standard of comparison] has been mentioned before in the immediate discourse'. In the Hui'an dialect, however, the standard of comparison is not expressed and this is usually because it can be inferred from the context. An example is given in (15).

- (15) 口 喉 怡 有聲 咯。
tse⁵ *au²* *kha?*⁷⁻⁸ *u⁴siā¹* *lɔ⁰.*
now throat HMC have-voice SFP
'(Your) throat is better now (than before).'

In (15), the comparee *tse⁵* 'now' is compared to 'before' (the standard of comparison) which is not expressed but can be inferred from the context. Example (15) also

Comparative constructions of inequality in Hui'an 259

suggests that a topic such as *au*² 'throat' can be overt in short comparatives. In addition, the topic *au*² is placed between the comparee *tse*⁵ 'now' and the head-marker of comparison *kha?*⁷. The comparee *tse*⁵ 'now' is a temporal word which is substantive, and the comparative predicate *u*⁴ *siā*¹ is a verb phrase with *siā*¹ as the object of the existential verb *u*⁴.

In the following, we give an example in which the comparee is encoded by a verb phrase such as *tshij*⁵⁻⁴ *tsit*⁷⁻⁸ *liā*³ 穿即領 'wear this piece (of clothing)', rather than a substantive word.

- (16) 穿 即領 恰 爽。
*tshij*⁵⁻⁴ *tsit*⁷⁻⁸ *liā*³ *kha?*⁷⁻⁸ *sɔŋ*³.
wear this-CL HMC comfortable
'Wearing this piece (of clothing) is more comfortable than (wearing that one).'

In (16), the comparee *tshij*⁵⁻⁴ *tsit*⁷⁻⁸ *liā*³ 'wear this piece of (clothing)' is compared to *tshij*⁵⁻⁴ *hit*⁷⁻⁸ *liā*³ 'wear that piece of (clothing)', which is not expressed but can be inferred from the context. The comparative predicate in (16) is encoded by the monosyllabic adjective *sɔŋ*³ 'comfortable'.

As in the double-marking/hybridized type of comparative, the predicate in the short comparative in the Hui'an dialect may also be followed by a measure word, as exemplified by (17), where the comparative predicate *kiau*³ 巧 'clever' precedes the measure word *tiəm*³⁻² *a*³ 點仔 'a bit, a little'.

- (17) 伊 恰 巧 點仔。
*i*¹ *kha?*⁷⁻⁸ *kiau*³ *tiəm*³⁻² *a*³.
3SG HMC clever a.bit
'He is a bit cleverer (than me).'

According to Li and Lien (1995: 74), the short comparative [A_{COM} *kha?*⁷ PRED] can be found in the *Li Jing Ji*. They also mentioned that the predicate in the short comparative in the *Li Jing Ji* is more complex than that in other comparative constructions attested in the *Li Jing Ji*, such as the zero-marked comparative [A_{COM} PRED B_{STA}] and the head-marking comparative with *kha?*⁷ [A_{COM} *kha?*⁷ PRED B_{STA} (MW)]. This may suggest that the short comparative [A_{COM} *kha?*⁷ PRED] was already in wide use in the period corresponding to the composition of the *Li Jing Ji*, that is, the Ming dynasty (1368–1644) and can be regarded as a comparative construction native to Southern Min, unlike the double-marking/hybridized type. However, as mentioned in §10.2, the short comparative [A_{COM} *kha?*⁷ PRED] in the Hui'an dialect may not yet have become a highly grammaticalized comparative construction, since similar syntactic structures may be regarded as non-comparative constructions where *kha?*⁷ is a degree adverb without denoting comparison.

Although we adopt the term 'short comparative' for the comparative construction [A_{COM} *kha?*⁷ PRED] in the Hui'an dialect, it does not mean that we are suggesting

that the comparative [A_{COM} *kha?*⁷ PRED] is a shortened form that has developed from a long comparative with an overt standard of comparison such as the double-marking/hybridized type [A_{COM} + *pi*³ + B_{STA} + *kha?*⁷ + PRED (+ MW)] examined in §10.2, or from the long comparative [A_{COM} *kha?*⁷ PRED B_{STA} (MW)], which will be examined in the following section.

First, it is impossible for the short comparative [A_{COM} *kha?*⁷ PRED] to be derived from the double-marking/hybridized comparative [A_{COM} + *pi*³ + B_{STA} + *kha?*⁷ + PRED (+ MW)]. According to previous studies, such as Li and Lien (1995: 74–5) and Yue-Hashimoto (1999: 63), the comparative construction with the comparative marker *pi*³ 比, which can be found in different modern Southern Min varieties such as the Hui'an dialect and Taiwanese Southern Min, is not attested in earlier Southern Min documents such as the *Li Jing Ji* and the *Doctrina Christiana*, which dates from the early seventeenth century. The comparative [A_{COM} + *pi*³ + B_{STA} + *kha?*⁷ + PRED (+ MW)] is most likely a late development due to the influence of the comparative construction in Northern Chinese, as mentioned in §10.2. It is thus clear that the short comparative [A_{COM} *kha?*⁷ PRED] is not a shortened form of the double-marking comparative [A_{COM} + *pi*³ + B_{STA} + *kha?*⁷ + PRED (+ MW)].

Second, it is also superfluous to claim that the short comparative developed from the long comparative [A_{COM} *kha?*⁷ PRED B_{STA} (MW)]. The reason is that the short comparative [A_{COM} *kha?*⁷ PRED] can directly undergo syntactic reanalysis from the non-comparative construction [NP *kha?*⁷ VP] with *kha?*⁷ functioning as a degree adverb meaning ‘a bit; fairly’, as in example (12). Another piece of evidence comes from the short comparative in other Sinitic languages, such as the comparative [A_{COM} PRED *di*] in Cantonese, as in example (14). The long comparative in Cantonese typically takes the form of [A_{COM} PRED *kwo*³ B_{STA}]. It is obvious, then, that the short comparative [A_{COM} PRED *di*] is not a shortened form of the long comparative [A_{COM} PRED *kwo*³ B_{STA}].

10.4 Head-marking comparative with *kha?*⁷: A_{COM} *kha?*⁷PRED B_{STA}(MW)

This type of comparative construction is regarded as an adverbial strategy in Chappell (this volume). As mentioned in §10.1, it is not widely used in the modern Hui'an dialect compared to the double-marking type and the short comparative, and is more commonly used in the speech of the elder generation than in that of the younger generation, even although its examples are indeed attested in modern daily conversation. The following is an example collected from the spoken data.

- (18) 阮 即厝 怡 有嘅食 懈 口面。
*un*³ *tsit*⁷⁻⁸*tshu*⁵ *kha?*⁷⁻⁸ *u*⁴*ay*¹*tsia?*⁸ *len*³ *khau*³⁻²-*ben*².
1PL this-house HMC have-can-eat 2PL outside
'We here have more things to eat than you, outside the home.'

In (18), the comparee and the standard of comparison are realized as ‘pronoun + locative word’, i.e. *un*³ *tsit*⁷⁻⁸ *tshu*⁵ ‘we here’ and *len*³ *khau*³⁻² *ben*² ‘you outside’, respectively. The comparative predicate is encoded by the verb phrase *u*⁴ *ay*¹ *tsia?*⁸ ‘have things to eat’. Not only the comparee NP and the standard NP, but also the comparative predicate, are expressed by a comparatively complex noun or verb phrase, rather than a simple monosyllabic noun or adjective respectively. This kind of evidence may indicate that the comparative construction [A_{COM} *kha?*⁷ PRED B_{STA} (MW)] is, or at least once was, a very highly grammaticalized and mature comparative construction in the Hui'an dialect, even though it is not common nowadays. The morpheme *kha?*⁷ in (18) simply denotes comparison, rather than functioning as a degree adverb meaning ‘a bit; fairly’. The standard of comparison *len*³ *khau*³⁻² *ben*² syntactically functions as an object of the verb phrase *kha?*⁷⁻⁸ *u*⁴ *ay*¹ *tsia?*⁸ ‘have more things to eat’, i.e. example (18) is a transitive construction. This shows that the comparative construction [A_{COM} *kha?*⁷ PRED B_{STA} (MW)] is associated with the ‘Action schema’ in Heine (1997) or the Exceed type in Stassen (1985) in terms of the transitivity of the construction. However, unlike the typical Exceed/Surpass type of comparative based on the ‘Action schema’ such as the comparative construction [A_{COM} PRED *kwo*³ B_{STA}] in Cantonese, there is no verb expressing the notion ‘defeat’, ‘exceed’, or ‘surpass’ in the Hui'an comparative construction [A_{COM} *kha?*⁷ PRED B_{STA} (MW)].

The following are more examples of this type of comparative in the Hui'an dialect.

- (19) 口 恰 大 口。
*tsat*⁸ *kha?*⁷⁻⁸ *tua*⁵ *hat*⁸.
this HMC big that
'This one is bigger than that one.'
- (20) 今年 恰 寒 舊年。
*ken*¹*lin*² *kha?*⁷⁻⁸ *kuā*² *ku*⁵⁻⁴*lin*².
today-year HMC cold old-year
'This year is colder than last year.'
- (21) 阿母 恰 愛食 嬰仔。
*a*¹*bu*³ *kha?*⁷⁻⁸ *ai*⁵⁻⁴*tsia?*⁸ *in*¹*ia*³.
Prefix-Mum HMC love-eat child
'Mother loves eating more than the child.'

Examples (19)–(21) show that the comparee and the standard of comparison in the comparative [A_{COM} *kha?*⁷ PRED B_{STA} (MW)] can be encoded by demonstratives (e.g. *tsat*⁸ ‘this’ and *hat*⁸ ‘that’ in (19)), temporal words (e.g. *ken*¹*lin*² ‘this year’ and *ku*⁵⁻⁴*lin*² ‘last year’ in (20)), or common nouns (e.g. *a*¹*bu*³ ‘mother’ and *in*¹*ia*³ ‘child’ in (21)). The comparative predicate can be a monosyllabic adjective such as *tua*⁵ ‘big’ in (19) and *kuā*² ‘cold’ in (20), or a disyllabic adjective such as *ai*⁵⁻⁴*tsia?*⁸ ‘love eating’ in (21).

A topic can also be overt in this type of comparative, as in example (22).

- (22) 我 手 恰 長 汝。
ua³ tshiu³ kha?⁷⁻⁸ tʃ² lu³.
1SG arm HMC long 2SG
'My arm is longer than yours.'

In (22), the noun *tshiu³* 手 'arm, hand' is placed between the comparee *ua³* 我 'I' and the head-marker of comparison *kha?⁷* 恰, functioning as a secondary topic. Semantically speaking, the comparee *ua³* and the secondary topic *tshiu³* function as a possessor and possessee, respectively.

As illustrated by example (23), a measure word can be further added after the standard of comparison, rather than after the comparative predicate, as in the double-marking/hybridized type and the short comparative, examined in §10.2 and §10.3 respectively.

- (23) 大車 恰 繫 細車 淡薄。
tua⁵⁻⁴tshia¹ kha?⁷⁻⁸ ken³ sue⁵⁻⁴tshia¹ tam⁴po?⁸.
big.car HMC fast small.car a.little
'The big car runs a bit faster than the small car.'

In (23), the measure word *tam⁴po?⁸* 'a little' follows the standard of comparison *sue⁵⁻⁴tshia¹* 'small car' to modify the degree of difference between the comparee *tua⁵⁻⁴ tshia¹* 'big car' and the standard *sue⁵⁻⁴tshia¹* 'small car' in terms of speed. The comparee and the standard of comparison in this example are both encoded by a noun phrase, and the predicate is expressed by the monosyllabic adjective *ken³* 'fast'.

This type of comparative construction has also been reported in other Southern Min varieties in Fujian and Taiwan (cf. Chen 1982: 62; Wang and Zhang 1994: 143; Li and Lien 1995: 72; Zhou and Ouyang 1998: 198). In modern Taiwanese Southern Min, for example, this type of comparative is attested in both the younger and older generation, even though it has an extremely low frequency (Li and Lien 1995: 81). In the modern Xiamen variety of Southern Min, it is typically attested in the speech of older and lower educated informants (see Wu 1958: 82–102; Crosland 1995; also cited in Li and Lien 1995: 79; Ansaldo 1999: 116).

Ansaldo (1999: 116) also mentioned that the existence of this type of comparative construction in Southern Min can be traced back to at least the beginning of the century in, for example, the grammar of Amoy by Piñol (1928), where Piñol points out that this type of comparative is the primary option for comparison of superiority. According to Li and Lien (1995: 73), this head-marking comparative construction [A_{COM} *kha?* PRED B_{STA}] is nonetheless attested even as early as the *Li Jing Ji* and its comparative predicate can be both monosyllabic and disyllabic. In addition, this type of comparative is the most common and mature comparative construction in the *Li Jing Ji*, without taking the short comparative [A_{COM} *kha?* PRED] into account.

It is thus clear that this type has served as one of the basic comparative constructions in early Southern Min, and can be regarded as one of the ‘native’ comparative constructions in this group of dialects, as also suggested by Li and Lien (1995: 74) and Ansaldi (1999: 116).

10.5 Zero-marked comparatives: A_{COM} PRED B_{STA} (MW)

Previous works on Southern Min such as Chen (1982), Li and Lien (1995), Zhou and Ouyang (1998), and Xu (2007) have all reported that there exists a type of comparative construction with the structure of ‘A_{COM} PRED B_{STA}’. This type is regarded as the absent-marking type by Ansaldi (1999), and as the zero-marked type by Chappell (this volume).

In the modern Hui'an dialect, this type of comparative construction is more commonly used by the elder generation than by the younger generation. The comparee and the standard of comparison are predominantly encoded by a simple noun or noun phrase, and the comparative predicate is typically a monosyllabic adjective, as illustrated by examples (24)–(26).

- (24) 我 肥 汝。
*ua*³ *put*² *lu*⁰.
1SG fat 2SG
'I am fatter than you.'
- (25) 大車 緊 細車。
*tua*⁵⁻⁴*tshia*¹ *ken*³⁻² *sue*⁵⁻⁴*tshia*¹.
big.car fast small.car
'The big car runs faster than the small car.'
- (26) 即張 鋪 閑 迄張 鋪。
*tsit*⁷⁻⁸*tiū*¹ *phɔ*¹ *khua?*⁷ *hit*⁷⁻⁸*tiū*¹ *phɔ*¹.
this.CL bed wide that.CL bed
'This bed is wider than that bed.'

These three examples show that the comparee and the standard of comparison in this type of comparative construction can be encoded by a personal pronoun such as the first person singular pronoun *ua*³ and the second person singular pronoun *lu*⁰ in (24), a noun such as *tua*⁵⁻⁴*tshia*¹ 'big car' and *sue*⁵⁻⁴*tshia*¹ 'small car' in (25), or a noun phrase such as *tsit*⁷⁻⁸*tiū*¹*phɔ*¹ 'this bed' and *hit*⁷⁻⁸*tiū*¹*phɔ*¹ 'that bed' in (26). The comparative predicate in all these three examples is encoded by a monosyllabic adjective, i.e. *put*² 'fat', *ken*³ 'fast', and *khua?*⁷ 'wide' in (24)–(26), respectively. In these examples, the standard of comparison such as *lu*⁰ 'you' in example (24) serves as the object of the comparative predicate *put*² 'fat'. In other words, like

the comparative [A_{COM} *kha?* PRED B_{STA}] examined in §10.4, the comparative [A_{COM} PRED B_{STA}] is also a transitive construction and associated with the ‘Action schema’ in Heine (1997).

However, the comparee in this type of comparative in the Hui'an dialect can also be a verb phrase, as in (27).

- (27) 穿 即領 爽 迹領。
*tshin*⁵⁻⁴ *tsit*⁷⁻⁸*liā*³ *sɔŋ*³ *hit*⁷⁻⁸*liā*³.
wear this.CL comfortable that.CL
'Wearing this piece of (clothing) is more comfortable than wearing that one.'

In (27), the comparee is encoded by the verb phrase *tshin*⁵⁻⁴ *tsit*⁷⁻⁸*liā*³ ‘wear this piece (of clothing)’, whereas the standard of comparison is simply expressed by the demonstrative phrase *hit*⁷⁻⁸*liā*³ ‘that piece (of clothing)’, rather than the verb phrase *tshin*⁵⁻⁴ *hit*⁷⁻⁸*liā*³ ‘wear that piece (of clothing)’. In other words, the repetition of the verb *tshin*⁵ is not necessary for coding the standard of comparison. A similar phenomenon is also attested in the Surpass type of comparative in Cantonese [A_{COM} PRED *kwo*³ B_{STA}], but not in the Compare type of comparative in Standard Mandarin and Mandarin varieties (see Chappell and Peyraube, this volume). Moreover, for example (27), the construction without a second *tshin*⁵ is more natural than the one with it, i.e. *tshin*⁵⁻⁴ *tsit*⁷⁻⁸*liā*³ *sɔŋ*³ *hit*⁷⁻⁸*liā*³ 穿即領爽迹領 is more natural than *tshin*⁵⁻⁴ *tsit*⁷⁻⁸*liā*³ *sɔŋ*³ *tshin*⁵⁻⁴ *hit*⁷⁻⁸*liā*³ 穿即領爽穿迹領, for native speakers of the Hui'an dialect.

Besides being a monosyllabic adjective as in examples (24)–(27), the comparative predicate in this type of comparative can also be encoded by a disyllabic adjective such as *tua*⁵⁻⁴*tsia?*⁷大隻 ‘big’ in (28), or a trisyllabic adjective such as *u⁴tshui⁵⁻³tsi?*⁸有喙舌 ‘talkative’ in (28).

- (28) 鵝 大隻 雞。
*gia*² *tua*⁵⁻⁴*tsia?*⁷⁻⁸ *kue*¹.
goose big.CL chick
'The goose is bigger than the chicken.'

- (29) 我 有喙舌 汝。
*ua*³ *u⁴tshui⁵⁻³tsi?*⁸⁻⁴ *lu*³.
1SG have.mouth.tongue 2SG
'I am more talkative than you.'

In (28), the predicate is encoded by the adjective *tua*⁵⁻⁴*tsia?*⁷ ‘big’ which consists of the adjective *tua*⁵ ‘big’ and the classifier *tsia?*⁷. In (29), the predicate *u⁴tshui⁵⁻³tsi?*⁸ is a fixed expression functioning like an adjective expressing ‘talkative’ (see also example (6) in §10.2).

Comparative constructions of inequality in Hui'an 265

Following are two examples with an overt topic.

- (30) 我 骸 短 汝。
ua³ kha¹ tɔ³ lu⁰.
1SG foot short 2 SG
'My feet are shorter than yours.'
- (31) 坐車 廣州 加時間 泉州。
tsə⁴tshia¹ kŋ³⁻²tsiu¹ ke¹si²⁻⁴kan¹ tsuan²⁻⁴tsiu¹.
1 SG PN more.time PN
'It takes more time to take the bus to Guangzhou than to Quanzhou.'

The topic *kha¹* 'foot' in (30) is placed between the comparee *ua³* 'I' and the predicate *tɔ³* 'short', whereas the topic *tsə⁴tshia¹* 'take the bus' in (31) is placed before the comparee *kŋ³⁻²tsiu¹* 'Guangzhou'.

A measure word following the standard of comparison such as *ban²⁻⁴ tsue⁵* 蠻多 'much' in example (32), is optional in the Hui'an dialect.

- (32) 我 勇 伊 (蠻 多)。
ua³ iŋy³ i⁰ (ban²⁻⁴ tsue⁵).
1SG strong 3SG (rather many)
'I am (much) stronger than her.'

The comparative construction [A_{COM} PRED B_{STA} (MW)] *without* the requirement of a measure word is so far attested in the Min group, including the Hui'an, Quanzhou, Xiamen, and Jieyang varieties of Southern Min, and in Southwestern Mandarin such as the Jishou dialect in Hunan province of China, whereas the comparative [A_{COM} PRED B_{STA} MW] *with* an obligatory overt measure word is widely used in various Sinitic languages such as Standard Mandarin, Wu, Hakka, Cantonese, and Xiang dialects (cf. Li 2003: 218–9; Ansaldi 1999: 105–6; Xu 2007: 272). It should be mentioned, however, that the zero-marked comparative with a measure word, [A_{COM} PRED B_{STA} MW], rather than the same form without, [A_{COM} PRED B_{STA}], is also attested in modern Taiwanese Southern Min (Li and Lien 1995: 77).

As with the comparative [A_{COM} *kha?* PRED B_{STA}], the comparative [A_{COM} PRED B_{STA}] can be attested in the Ming dynasty *Li Jing Ji*, where only a monosyllabic comparative predicate is allowed in this type (see Li and Lien 1995: 73–4). In addition, as mentioned above, the zero-marked comparative [A_{COM} PRED B_{STA}] is less common and mature than the head-marked comparative [A_{COM} *kha?* PRED B_{STA}] in the period of the *Li Jing Ji*. Nonetheless, in terms of its 'maturity', the zero-marked comparative construction [A_{COM} PRED B_{STA}] can also be regarded as one of the more 'native' comparative constructions in Southern Min, like the head-marked comparative [A_{COM} *kha?* PRED B_{STA}].

10.6 Comparatives with the marker *khui*⁵: A_{COM} PRED *khui*⁵ B_{STA}

This type of comparative construction is also not commonly used in the modern Hui'an dialect and is more common in the speech of the elder generation than in the speech of the younger generation. The use of this type of comparative is more limited than the zero-marked comparative [A_{COM} PRED B_{STA}] exemplified in (24)–(32), not to mention than the four other types of comparative we have already examined.

The comparee and the standard of comparison of this type of comparative are typically nouns or pronouns, personal or demonstrative, rather than a VP, and the comparative predicate is typically a monosyllabic adjective, as exemplified by (33)–(35).

- (33) 口 大 去 口。
*tsat*⁸ *tua*⁵⁻⁴ *khui*⁵⁻⁴ *hat*⁸.
this big go that
'This one is bigger than that one.'
- (34) 我 凶 去 汝。
*ua*³ *hiɔŋ*¹ *khui*⁵⁻⁴ *lu*³.
1SG ugly go 2SG
'I am uglier than you.'
- (35) 張三 重 去 李四。
*tiū*¹*sā*¹ *tay*⁵⁻⁴ *khui*⁵⁻⁴ *li*³⁻²*si*⁴.
PN heavy go PN
'Zhangsan is heavier than Lisi.'

Examples (33)–(35) show that the comparee and the standard of comparison in this type of comparative can be encoded by a demonstrative such as *tsat*⁸ 'this' or *hat*⁸ 'that' in (33), a personal pronoun such as *ua*³ 'I' or *lu*³ 'you' in (34), if not a proper noun such as *tiū*¹*sā*¹ 'Zhangsan' or *li*³⁻²*si*⁴ 'Lisi' in (35).

Not all monosyllabic adjectives can be used in this type of comparative, as shown by (36b).

- (36) a 我 巧 汝。
*ua*³ *khiau*³ *lu*⁰.
1SG clever 2SG
'I am cleverer than you.'
- b *我 巧 去 汝。
*ua*³ *khiau*³ *khui*⁵⁻⁴ *lu*³.
1SG clever go 2SG
('I am cleverer than you.')

Comparative constructions of inequality in Hui'an 267

In (36), the monosyllabic adjective *khiau*³ 'clever' can function as a comparative predicate in the zero-marked comparative construction [A_{COM} PRED B_{STA}], but not in the *khui*⁵-comparative [A_{COM} PRED *khui*⁵ B_{STA}]. This may be associated with the fact that the verbal complement *khui*⁵ in the Hui'an dialect tends to be used with a negative, rather than a positive attribute.

However, not all positive monosyllabic adjectives are excluded from this type of comparative, as seen in (37).

- (37) a 我 好 去 汝。
*ua*³ *ho*³ *khui*⁵⁻⁴ *lu*³.
1SG good go 2SG
'I am better than you.'
- b 我 否 去 汝。
*ua*³ *phai*³ *khui*⁵⁻⁴ *lu*³.
1SG bad go 2SG
'I am worse than you.'

As shown in (37), both the positive monosyllabic adjective *ho*³ 'good' and its negative form *phai*³ 'bad' can serve as a comparative predicate followed by the comparative marker *khui*⁵. One of the reasons may be that the adjective *ho*³ 'good' is more basic and common than the adjective *khiau*³ 'clever'.

In addition, not all non-positive monosyllabic adjectives can be naturally used with the comparative marker *khui*⁵, as in (38).

- (38) a 即搭 热 迂搭。
*tsit*⁷⁻⁸*ta*⁷ *lua*⁷ *hit*⁷⁻⁸*ta*⁷.
this.LOC hot that.LOC
'It is hotter here than there.'
- b? 即搭 热 去 迂搭。
*tsit*⁷⁻⁸*ta*⁷ *lua*⁷ *khui*⁵⁻⁴ *hit*⁷⁻⁸*ta*⁷.
this. LOC hot go that. LOC
'It is hotter here than there.'

The monosyllabic adjective *lua*⁷ 'hot' used with the comparative marker *khui*⁵ as in example (38b) is acceptable for native speakers of the Hui'an dialect. However, comparatively speaking, the zero-marked form [A_{COM} PRED B_{STA}], without the marker *khui*⁵ as in (38a), is preferred in use, rather than the form with the marker *khui*⁵ as in (38b). One of the reasons may be that the sequence *khui*⁵⁻⁴ *hit*⁷⁻⁸*ta*⁷ is ambiguous in the Hui'an dialect, having another possible, and probably more common interpretation, i.e. the locative word *hit*⁷⁻⁸*ta*⁷ 'there' functions as a locative object of the motion verb *khui*⁵ 'go'.

As demonstrated by examples (39) and (40), it seems that the comparative predicate in this type of *khui*⁴-comparative cannot be other than monosyllabic.

- (39) a 今日 嘴喉 有聲 昨日。
kiā¹let⁸ *lau²⁻⁴au²* *u⁴siā¹* *tsa⁴let⁸.*
today throat have.voice yesterday
'(Your) throat is better today (than yesterday).'
- b? 今日 嘴喉 有聲 去 昨日。
kiā¹let⁸ *lau²⁻⁴au²* *u⁴siā¹* *khui⁵⁻⁴* *tsa⁴let⁸.*
today throat have.voice go yesterday
'(Your) throat is better today (than yesterday).'
- (40) *我 有喙舌 去 汝。
ua³ *u⁴tshui⁵⁻³tsi?*⁸ *khui⁵⁻⁴* *lu³.*
1SG have.mouth.tongue go 2SG
'I am more talkative than you.'

In (39), *u⁴siā¹* 'have voice' serving as a comparative predicate and followed by the comparative marker *khui*⁵ is acceptable. However, the form without the marker *khui*⁵ as in (39a) is better than the form with the marker *khui*⁵ as in (39b). In (40), *u⁴tshui⁵⁻³tsi?*⁸ 'talkative' cannot occur in the comparative with the comparative marker *khui*⁵.

Based on the analysis above, it may be safe to say that the comparative predicate in this type of comparative in the modern Hui'an dialect tends to be encoded by a negative monosyllabic adjective, though positive monosyllabic adjectives with high frequency such as *ho*³ 好 'good' are also possible.

Compared to the comparative predicate, the comparee is less restricted, as seen in (41), where the comparee is encoded by the verb phrase *tshiy*⁵⁻⁴ *tsit*⁷⁻⁸*liā*³ 'wear this piece of (clothing)'.

- (41) 穿 即領 爽 去 迄領。
*tshiy*⁵⁻⁴ *tsit*⁷⁻⁸*liā*³ *sɔŋ*³ *khui*⁵⁻⁴ *hit*⁷⁻⁸*liā*³.
wear this.CL comfortable go that. CL
'Wearing this piece of (clothing) is more comfortable than wearing that one.'

This example also shows that the positive monosyllabic adjective *sɔŋ*³ 'comfortable' can be used with the comparative marker *khui*⁵. In other words, the positive negative distinction may not be, after all, the most important factor in explaining the limited use of the comparative construction with the marker *khui*⁵ but rather the requirement of monosyllabicity for the predicate.

Unlike the previous four types of comparative construction as examined in §10.2–10.5, this type of comparative construction has not been elsewhere reported in the literature on Southern Min, but can be found in the Fuzhou variety of Eastern Min (Yuan 1989: 305; Zhao 2002a: 49). Wu Fuxiang (pers. comm.) suggests that the

comparative marker *khui*⁵ here functions as a verbal complement of the predicate. As mentioned above, the head-marker of comparison *gwo*³ in Cantonese is also regarded as a verbal complement in the literature, including Li (1994), Mok (1998), Ansaldi (1999: 121), and Chappell (this volume). In other words, as with *gwo*³ in Cantonese, *khui*⁵ in the Hui'an dialect can also serve as a head-marker of comparison, similarly observed by Liu (2003b: 14) where he suggests that *qù* 去 in Northern Min and *qǐ* 起 in the Shandong dialect are markers for the predicate in comparative constructions and that this function can be derived from their functions as directional complements.

Similar kinds of head-markers of comparison and comparative constructions are attested in many other Sinitic languages, including Wu, Hakka, Gan, Xiang, and Mandarin varieties (Li 2003: 219). We agree that the comparative marker *khui*⁵ in the Hui'an dialect may also be a verbal complement of the predicate. In this case, the comparative with the marker *khui*⁵ is a transitive construction with the standard of comparison as the object of the complex [PRED *khui*⁵], while it is associated with the ‘Action schema’ in Heine (1997). However, we also wonder whether there is another possible interpretation. The counterpart of *khui*⁵ in Mandarin Chinese, i.e. *qù*, originally means ‘leave, depart from’ in the Pre-Qin period (before 221 BC) and was also used as a verb meaning ‘be apart from, be at a distance of’ in the same period, which may imply comparison (cf. He 2011: 49).

Thus, is there any possible connection between the comparative marker *khui*⁵ and the Separative comparatives in Stassen (1985)? However, as pointed out by Wu Fuxiang (pers. comm.), the comparative markers associated with Separative comparatives are typically morphemes indicating cases such as adpositions or affixes, while *khui*⁵ in the Hui'an dialect is not used as an adposition or a coverb. While this suggests that comparatives with the marker *khui*⁵ may not be a typical example of the Separative type, the possibility cannot be ruled out that the comparative with the marker *khui*⁵ belongs to the Separative type, not the Surpass type. Wu (2011: 9) also suggests that the comparative construction with the marker *qù* 去 is an innovation in Eastern Min based on the zero-marked comparative construction ‘A_{COM} PRED B_{STA}’.

10.7 Dependent-marking comparatives with *pi*³: A_{COM} *pi*³ B_{STA}PRED (MW)

It is generally known that the basic comparative construction in Mandarin Chinese is A_{COM} *bì* 比 B_{STA} PRED, which is called the Similarity type by Ansaldi (1999) and the Compare type by Chappell and Peyraube (this volume), respectively, as mentioned in §10.1. We have also mentioned that this type of comparative is the most common one found in Sinitic languages. Not surprisingly, it is also attested in

different Southern Min varieties. In the Jieyang variety of Southern Min, for example, this type of comparative is the most versatile and plays an important role in Jieyang syntax (see Xu 2007: 277). In Taiwanese Southern Min, this type of comparative is found in the speech of different age groups ranging from 15 years old to 74 years old (Li and Lien 1995: 80–1). In the Hui'an dialect, however, it is not widely used and is quite limited. The double-marking/hybridized type of comparative is preferred when the comparative marker *pi*³ is involved. In addition, this type of comparative construction usually requires a measure phrase following the comparative predicate, as in (42), where the measure phrase *sā¹ hō⁵* 三歲 ‘three years’ is preceded by the predicate *tua⁵* 大 ‘big’.

- (42) 我 比 伊 大 三 歲。
ua³ pi³⁻² i¹ tua⁴ sā¹ hō⁵.
1SG DMC 3SG big three year
'I am three years older than him.'

10.8 Conclusion

This analysis has examined comparative constructions of inequality (termed ‘comparative construction’) in the Southern Min dialect of Hui'an, a Southern Min variety spoken in Hui'an County in Fujian province of China. Six main types of comparative construction are attested in the Hui'an dialect:

- (a) Double-marking/hybridized comparatives:
[A_{COM} *pi*³ B_{STA} *kha?*⁷ PRED (MW)]
- (b) Short comparatives: [A_{COM} *kha?*⁷ PRED (MW)]
- (c) Head-marking comparatives with the marker *kha?*⁷:
[A_{COM} *kha?*⁷ PRED B_{STA} (MW)]
- (d) Zero-marked comparatives:
[A_{COM} PRED B_{STA} (MW)]
- (e) Comparatives with the marker *khui*⁵: [A_{COM} PRED *khui*⁵ B_{STA}]
- (f) Dependent-marking comparatives with the marker *pi*³ ‘compare’:
[A_{COM} *pi*³ B_{STA} PRED (MW)]

Type (a) and Type (b), i.e. [A_{COM} *pi*³ B_{STA} *kha?*⁷ PRED] and [A_{COM} *kha?*⁷ PRED] are the dominant types. Type (c)–(e) are more commonly used in the speech of the elder generation than that in the speech of the younger generation. In addition, the use of Type (e) is much more restricted than that of Types (c) and (d). Type (f), the most common one in Sinitic languages, however, is quite limited in the Hui'an dialect and usually requires a measure word following the comparative predicate.

A topic may be involved in these comparative constructions to denote the parameter based on which a comparison is made. In addition, the topic may be placed before the comparee, be inserted between the comparee and the dependent-marker of

comparison, or even be placed between the standard of comparison and the head-marker of comparison, especially in the comparative construction with the presence of all these relevant constituents such as the double-marking/hybridized type.

The morpheme *kha?*⁷ in the Hui'an dialect can also be used as a degree adverb meaning 'a bit; fairly' in non-comparative constructions, which suggests a possible grammaticalization path from a degree adverb to a marker of comparison. In fact, a similar extension from a degree adverb to a comparative marker is also attested in other Min varieties and other Sinitic languages such as Xianghua and Cantonese.

The non-comparative constructions, in which the degree adverb *kha?*⁷ 'a bit, fairly' occurs, share the same structure with comparatives of Type (b), which suggests that the Type (b) comparative, i.e. [A_{COM} *kha?*⁷ PRED (MW)], is not highly grammaticalized. In Type (b) short comparatives, the standard of comparison is usually not expressed because it can be inferred from the context. This is different from the Southern Min dialect of Jieyang in which the standard of comparison needs to have been mentioned in the prior discourse (see §10.3).

The head-marking Type (c), [A_{COM} *kha?*⁷ PRED B_{STA} (MW)], was the primary option for comparison at the beginning of the twentieth century and was also already widely used in the Ming dynasty *Lì Jīng Jí*. Similarly, the zero-marked Type (d) is also attested in the *Lì Jīng Jí*. These may show that these two types are both 'native' comparative constructions in Southern Min. However, neither is widely used in the modern Hui'an dialect. In addition, these two types are both associated with the 'Action schema' in Heine (1997) or the Exceed type in Stassen (1985) in terms of the transitivity of their forms. However, unlike the typical Exceed/Surpass type of comparative based on the 'Action schema' such as the comparative construction [A_{COM} PRED *kwo*³ B_{STA}] in Cantonese, there is no verb expressing the notion 'defeat', 'exceed', or 'surpass' in these two types.

Type (e) comparatives have not been reported in the literature on other varieties of Southern Min, but are found in the Fuzhou variety of Eastern Min. There are two possible sources of this type of comparative construction with the marker *khui*⁵: (a) similar to the comparative marker *gwo*³ in Cantonese, the marker *khui*⁵ functions as a verbal complement of the predicate and this type of comparative construction can be classified as the Surpass type; and (b) the counterpart of *khui*⁵ in Mandarin Chinese, i.e. *qù*, originally means 'leave, depart from' in the Pre-Qin period (before 221 BC) and was also used as a verb meaning 'be apart from, be at a distance of' in the same period, which may involve comparison. This suggests that the marker *khui*⁵ may be derived from its use as a verb meaning 'leave, depart from'. In the first case, it may be regarded as being associated with the 'Action schema' in Heine (1997), whereas, in the second case, it may also be associated with the Separative type in Stassen (1985), if not the Source schema in Heine.

Acknowledgements

Preliminary versions of this chapter appeared in *Kansas Working Papers in Linguistics* and as part of the author's doctoral dissertation. This research has been partially supported by a postgraduate studentship from the University of Hong Kong, by the European Research Council under the European Community's Seventh Framework Programme (FP7/2007–2013) for the ERC project SINOTYPE under grant agreement n° 230388 (awarded to Prof. Hilary Chappell), by the Fundamental Research Funds for the Central Universities in UIBE (13QD21), and by 211 PROJECT of UIBE (XK2014303). We are deeply grateful to Dr. Stephen Matthews, Prof. Hilary Chappell, Dr. Umberto Ansaldi, Dr. Picus Sizhi Ding, Dr. Xuping Li, and Sing Sing Ngai for their valuable comments on this study.

References

- Ahrens, Kathleen and Huang Chu-Ren (1996). ‘Classifiers and semantic type coercion: Motivating a new classification of classifiers’, in B.-S. Park and J.B. Kim (eds), *Proceedings of the 11th Pacific Asia Conference on Language, Information and Computation*, Seoul: Kyung Hee University, 1–10.
- Aikhenvald, Alexandra (1999). ‘Areal diffusion and language contact in the Içana-Vaupés basin, North-West Amazonia’, in R.M.W. Dixon and Alexandra Aikhenvald (eds), *The Amazonian languages*. Cambridge: Cambridge University Press, 385–415.
- Aikhenvald, Alexandra (2000). *Classifiers. A Typology of Noun Categorization Devices* (Oxford Studies in Typology and Linguistic Theory). Oxford: Oxford University Press.
- Aikhenvald, Alexandra and R.M.W. Dixon (eds) (2001). *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*. Oxford: Oxford University Press.
- Aissen Judith (2003). ‘Differential object marking: Iconicity vs.economy’, *Natural Language and Linguistic Theory* 21: 435–83.
- Akitani, Hiroyuki 秋谷裕幸 (2001). Wuyu Jiangshan Guangfeng fangyan yanjiu 吴语江山广丰方言研究 [A study of the Jiangshan and Guangfeng Wu dialects]. Matsuyama 松山市: Ehime Daigaku Hö bun Gakubu Sōgō Seisaku Gakka 愛媛大学法文学部総合政策学科.
- Akitani, Hiroyuki 秋谷裕幸 (2008). *Mínbeiqū Sān Xiànsì Fāngyán Yānjiū 閩北區三縣市方言研究 [Studies of the Dialects in Three Counties/Cities in the Northern Min Area]*. Taipei: Institute of Linguistics, Academia Sinica 中央研究院語言學研究所.
- Anderson, Stephen R. and Edward L. Keenan (1985). ‘Deixis’, in T. Shopen (ed.), *Language Typology and Syntactic Description: Grammatical Categories and the Lexicon* Vol. 3. Cambridge: Cambridge University Press, 259–308.
- Ansaldo, Umberto (1999). ‘Comparative constructions in Sinitic: Areal typology and patterns of grammaticalization’, PhD dissertation, Department of Linguistics, Universitet Stockholm.
- Ansaldo, Umberto (2010). ‘Surpass comparatives in Sinitic and beyond: Typology and Grammaticalization’, *Linguistics* 48.4: 919–50.
- Bauer, Robert S. (1996). ‘Identifying the Tai Substratum in Cantonese’, in *Pan-Asiatic Linguistics. Proceedings of the Fourth International Symposium on Languages and Linguistics* V. Bangkok: Institute of Language and Culture for Rural Development, Mahidol University, 1806–44.
- Baxter, William H. (1992). *A Handbook of Old Chinese Phonology*. Berlin: Mouton de Gruyter.
- Baxter, William H. and Laurent Sagart (2014). *Old Chinese: A New Reconstruction*. New York: Oxford University Press.
- Bennet, Paul A. (1979). ‘A critique of the Altaïcisation Hypothesis’, *Cahiers de Linguistique—Asie Orientale* 6: 91–104.
- Bisang, Walter (1992). *Das Verb im Chinesischen, Hmong, Vietnamesischen, Thai und Khmer*. Tübingen: Gunter Narr Verlag.

- Bisang, Walter (1996). ‘Areal typology and grammaticalization: Processes of grammaticalization based on nouns and verbs in East and mainland South East Asian languages’, *Studies in Language* 20.3: 519–97.
- Bisang, Walter (1999). ‘Classifiers in East and Southeast Asian languages: Counting and beyond’, in Jadranka Gvozdanovic (ed). *Numerical Types and Changes Worldwide*. Berlin: Mouton de Gruyter, 113–85.
- Bossong, Georg (1985). *Differentielle Objektmarkierung in den Neuiranischen Sprachen*. Tübingen: Gunter Narr Verlag.
- Bu Lianzeng 步连增 (2011). ‘Nánfāng Fāngyán zhōng Liàngcí Zhǐdìng Xiànxìàng Láiyuán Chūtàn—Yí Guinán Pínghuà wéi Lí 南方汉语中量词定指现象来源初探—以桂南平话为例 [A study about the source of the phenomenon referring to the classifiers’ identifiability: With the example of Guinan Pinghua]’, *Studies in Language and Linguistics 语言研究* 31(3): 95–102.
- Büring, Daniel (1997). *The Meaning of Topic and Focus—The 59th Street Bridge Accent*. London: Routledge.
- Büring, Daniel (2003). ‘On D-Trees, Beans, and B-Accents’, *Linguistics and Philosophy* 26(5): 511–45.
- Bybee, Joan L., William Pagliuca, and Revere D. Perkins (eds) (1990). ‘On the asymmetries in the affixation of grammatical material’, in William Croft, Keith Denning, and Suzanne Kemmer (eds), *Studies in Typology and Diachrony*. Papers Presented to Joseph H. Greenberg on his 75th Birthday. Amsterdam: John Benjamins, 1–42.
- Cai, Lunming 蔡伦明 (ed.) (1991). *Chaozhou Fangyan Cihui: Putonghua duizhao 潮州方言詞彙：普通話對照* [Lexicon of the Chaozhou Dialect: A Comparison with Mandarin Chinese]. Hong Kong: T.T. Ng Chinese Language Research Centre Institute of Chinese Studies, The Chinese University of Hong Kong.
- Cao, Guangshun 曹广顺 (1994). ‘Shuo zhuci ge说助词‘个’ [A discussion on the grammatical particle ‘ge’]’, 古汉语研究 *Gu Hanyu Yanjiu* 4.
- Cáo Zhiyún 曹志耘 and Rixin Zhao 赵日新 (eds) (2008). *Hànyǔ Fāngyán Ditújí 汉语方言地图集* [Linguistic Atlas of Chinese Dialects]. Beijing: The Commercial Press 商务印书馆.
- Chafe, Wallace (1976). ‘Givenness, contrastiveness, definiteness, subjects, topics, and point of view’, in Charles N. Li (ed.), *Subject and Topic*. New York: Academic Press, 25–55.
- Chang Song-hing 张双庆 and Bit-chee Kwok 郭必之 (2005). ‘Xianggang Yueyu liangzhong chabiju de jiaoti 香港粤语两种差比句的交替 [Alternation between two types of comparative sentences in Hong Kong Cantonese]’, *Zhongguo Yuwen 中國語文* 3: 232–8.
- Chao Yuen Ren (1968). *A Grammar of Spoken Chinese*. Berkeley and Los Angeles: University of California Press.
- Chappell, Hilary (1986). ‘Formal and colloquial adversity passives in standard Chinese’, *Linguistics* 24.6: 1025–52.
- Chappell, Hilary (1992). ‘Causativity and the bā construction in Chinese’, in Hansjakob Seiler and Walfried Prempel (eds.), *Partizipation: Das sprachliche Erfassen von Sachverhalten*. Tübingen: Gunter Narr Verlag, 509–30.
- Chappell, Hilary (2001) (ed.), *Sinitic Grammar: Synchronic and Diachronic Perspectives*. Oxford: Oxford University Press.

- Chappell, Hilary (2001a). ‘Synchrony and diachrony of Sinitic languages: A brief history of Chinese dialects’, in H. Chappell (ed.), *Sinitic Grammar: Synchronic and Diachronic Perspectives*. Oxford: Oxford University Press, 3–28.
- Chappell, Hilary (2001b). ‘Language contact and areal diffusion in Sinitic languages: Problems for typology and genetic affiliation’, in Alexandra Aikhenvald and R.M.W. Dixon (eds), 328–57.
- Chappell, Hilary (2001c). ‘A typology of evidential markers in Sinitic languages’, in H. Chappell (ed.), *Sinitic Grammar: Synchronic and Diachronic Perspectives*. Oxford: Oxford University Press, 56–84.
- Chappell, Hilary. (2006). ‘From Eurocentrism to Sinocentrism: The case of disposal constructions in Sinitic languages’, in Felix Ameka, Alan Dench, and Nicholas Evans (eds.), *Catching Language: The Standing Challenge of Grammar Writing*. Berlin: Mouton de Gruyter, 441–86.
- Chappell, Hilary (曹茜蕾) (2007a). Hanyu fangyan de chuzhibiaoj de leixing 《汉语方言的处置标记的类型 [A typology of Object-Marking constructions: A Pan-Sinitic view]. *Yuyanxue Luncong* 《语言学论丛》36.1 : 184–209.
- Chappell, Hilary (2007b). ‘Grammaticalization zones for the identical marking of agents and patients in Sinitic languages’, paper presented at EACL5 (*Fifth International Conference of the European Association for Chinese Linguistics*) held at the Max Planck Institute, Leipzig, 4–7 September 2007.
- Chappell, Hilary (2008). ‘Variation in the grammaticalization of complementizers from *verba dicendi* in Sinitic languages’, *Linguistic Typology* 12-1: 45–98.
- Chappell, Hilary (2013a). ‘Pan-Sinitic Object Markers: Morphology and syntax’, in Cao Guangshun, H. Chappell, R. Djemouri, and Thekla Wiebusch (eds.), *Breaking Down the Barriers: Interdisciplinary Studies in Chinese Linguistics and Beyond*. 2 vols. Taipei: Academia Sinica, 785–816.
- Chappell, Hilary (2013b). ‘Ditransitives and the grammaticalization of *Give* verbs in early Southern Min’, paper presented at the French ANR-Taiwanese NSC workshop on *Typologization of Synchronic and Diachronic Processes in Southern Min, a Sinitic Language (TYSOMIN)*. Barcelona (Spain), 19–22 December.
- Chappell, Hilary (this volume). ‘Linguistic areas in China for differential object marking, passive, and comparative constructions.’
- Chappell, Hilary (forthcoming (a)). ‘Southern Min’, in Alice Vittrant and Justin Watkins (eds.), *Mainland Southeast Asian Languages*. Berlin: Mouton de Gruyter.
- Chappell, Hilary (forthcoming (b)). ‘Major non-canonical clause types in Mandarin Chinese: *bă, bēi*, and Ditransitives’, Chapter 15 in Chu-Ren Huang and Dingxu Shi (eds.), *Cambridge Reference Grammar of the Chinese Language*. Cambridge: Cambridge University Press.
- Chappell, Hilary (forthcoming (c)). ‘Languages of China in their East Asian context’, in Raymond Hickey (ed.), *The Cambridge Handbook of Areal Linguistics*. Cambridge: Cambridge University Press.
- Chappell, Hilary and Huali Liu (forthcoming). ‘Wait Causatives and Passives in Sinitic Languages: Typology and Diachrony’.
- Chappell, Hilary and Christine Lamarre (2005). *A Grammar and Lexicon of Hakka: Historical Materials from the Basel Mission Library* (Collection des Cahiers de Linguistique—Asie Orientale 8). Paris: EHESS.

276 *References*

- Chappell, Hilary, Li Ming, and Alain Peyraube (2007). ‘Chinese linguistics and typology: The state of the art’, *Linguistic Typology* 11.1: 187–211.
- Chappell, Hilary and William B. McGregor (1996). ‘Prolegomena to a theory of inalienability’, in H. Chappell and W. McGregor (eds), *The Grammar of Inalienability. A Typological Perspective on Body Part Terms and the Part-Whole Relation*. Berlin: Mouton de Gruyter, 3–30.
- Chappell, Hilary and Alain Peyraube (2006a). ‘The analytic causatives of Early Modern Southern Min in diachronic perspective’, in Dah-An Ho, Hung-nin S. Cheung, Wuyun Pan, and Fuxiang Wu (eds.), *Linguistic Studies in Chinese and Neighboring Languages*. Nankang, Taiwan: Academia Sinica, Institute of Linguistics, 973–1011.
- Chappell, Hilary and Alain Peyraube (2006b). ‘Polygrammaticalization of *say* verbs in Sinitic languages’, paper presented at the 20^{ème} *Journées de linguistique d’Asie Orientale*. Paris, 22–24 June.
- Chappell, Hilary and Alain Peyraube (2006c). ‘The diachronic syntax of causative structures in Early Modern Southern Min’, in Dah-an Ho (ed.), *Festschrift for Ting Pang-Hsin*. Taipei: Academia Sinica, 973–1011.
- Chappell, Hilary and Alain Peyraube (2007). ‘The diachronic syntax of ditransitive constructions from Archaic Chinese to early Southern Min (Sinitic)’, paper presented at the *Conference on Ditransitive Constructions*. Leipzig: Max Planck Institute for Evolutionary Anthropology, 23–25 November.
- Chappell, Hilary and Alain Peyraube (2013). ‘Lun hanyu yanshuo dongci de lishi fazhan 论汉语言说动词的历史发展 [On the historical evolution of the verbs of saying]’, in Wu, Fuxiang and Xing, Xiangdong 吴福祥, 邢向东 (eds.), *Yufahua Yu Yufa Yanjiu 语法化与语法研究* 6: 1–12.
- Chappell, Hilary and Alain Peyraube, (this volume). ‘The comparative construction in Sinitic: Synchronic and diachronic variation’.
- Chappell, Hilary, Alain Peyraube and Yunji Wu (2011). ‘A comitative source for object markers in Sinitic languages: 跟 *kai*⁵⁵ in Waxiang and 共 *kang*⁷ in Southern Min’, *Journal of East Asian Linguistics* 20.4: 291–338.
- Chen Fajin 陳法今 (1982). ‘Minnan fangyan de liangzhong bijiaoju’, 閩南方言的兩種比較句 [Two kinds of comparative sentences in Southern Min], *Zhongguo Yuwen* 中國語文 [*Studies of the Chinese Language*] 1: 62–5.
- Chen Fanfan 陈凡凡 and Lin, Lunlun 林伦伦 (2003). ‘Guangdong Chenghai Min fangyan liangci de yufa tedian 广东澄海闽方言量词的语法特点 [The characteristics of the grammar of the classifiers in the Chenghai dialect in Guangdong province]’, *Shantou Daxue Xuebao (Supplement)*: 70–6.
- Chen Hui 陈晖 (1999). *Lianyuan fangyan yanjiu* 涟源方言研究 [Research on the Lianyuan Dialect]. Changsha: Hunan Education Publishing House.
- Chen I-Hsuan and Chin-fa Lien (2008). ‘The interaction between the construction *kong* + Topic and its thematic markers in Taiwanese Southern Min’, in Redouane Djamousi and Rint Sybesma (eds), *Chinese Linguistics in Europe* 2. Paris: EHESS, 175–90.
- Chen Jianchu 陈建初 (1995). ‘Hunan Lengshuijiang fangyan zhong de daici 湖南冷水江方言中的代词 [The pronouns in the Lengshuijiang dialect of Hunan]’, *Gu Hanyu Yanjiu* 5:16–21.

- Chen Minyan 陈敏燕, Sun Yizhi 孙宜志, and Chen Changyi 陈昌仪 (2003). 'Jiangxi jingnei Gan fangyan zhishi daici de jinzhi he yuanzhi 江西境内赣方言指示代词的近指和远指 [The distance categories for the demonstratives of the Gan dialects in Jiangxi Province]', *Zhongguo Yuwen* 中國語文 6: 496–504.
- Chen Ping 陈平 (1987). 'Shi hanyu zhong yu mingcixing chengfen xiangguan de sizu gainian 释汉语中与名词性成分相关的四组概念 [The definitions for four groups of basic terms for noun phrases]', *Zhongguo yuwen* 中國語文 2: 81–92.
- Chen Ping (2004). 'Identifiability and definiteness in Chinese', *Linguistics* 6: 1129–84.
- Chen Xiaojin 陈晓锦 and Chén Tāo 陈滔 (2005). *Guāngxī Běihǎishi Yuéfāngyán Diàochá Yánjiū* 广西北海市粤方言调查研究 [Investigative Studies of the Yue dialects in Bēihǎi City, Guāngxī]. Beijing: China Social Science Press 中国社会科学出版社.
- Chen Xiaoyan 陈小燕 (2007). *Duōzúqún Yúyán de Jiéchū yǔ Jiāoróng—Hèzhōu Bēndihuà Yánjiū* 多族群语言的接触与交融—贺州本地话研究 [The Contact and Merger of Multi-Group Languages—Studies of the Hèzhōu Local Dialect]. Beijing: The Ethnic Publishing House 民族出版社.
- Chen Xingwei 陈兴伟 (1992). 'Yiwu fangyan liangci qian zhishi daici he shuci de shenglue 义乌方言量词前指示代词和数词的省略 [The omission of demonstratives and numerals preceding the classifiers in the Yiwu dialect]', *Zhongguo Yuwen* 中國語文 2: 206.
- Chen Yiya (2009). 'Prosody and information structure mapping: Evidence from Shanghai Chinese', *Chinese Journal of Phonetics* 2: 123–33.
- Chen Yujie 陈玉洁 (2010). *Hanyu zhishici de leixingxue yanjiu* 汉语指示词的类型学研究 [Typological Research on Chinese Demonstratives] Beijing : Zhongguo Shehui Kexue (CASS) Publishers.
- Chen Yujie (this volume). 'The Semantic differentiation of demonstratives in Sinitic languages'.
- Chen Zeping 陈泽平(1998). 福州方言研究 *Fuzhou fangyan yanjiu* [A Study of the Fuzhou Dialect]. Fuzhou: Fujian renmin chubanshe.
- Chen Zhangtai 陈章太(1991a). 'Datianxiannei de fangyan 大田縣內的方言 [Dialects of the Datian County]', in Zhangtai Chen and Rulong Li 陈章太, 李如龙 (eds), *Minyu Yanjiu* 閩語研究 [Studies on the Min Dialects]. Beijing: Language and Culture Press, 266–303.
- Chen Zhangtai 陈章太(1991b). 'Shunchangxian pushang minnan fangyandao 順昌縣埔上閩南方言島 [Pushang, a Southern Min Island in Shunchang County]', in Zhangtai Chen and Rulong Li 陈章太, 李如龙 (eds), *Minyu Yanjiu* 閩語研究 [Studies on the Min Dialects]. Beijing: Language and Culture Press, 421–491.
- Chen Zhangtai and Li Rulong 陈章太, 李汝龙 (1991). 閩語研究 *Minyu yanjiu* [A study of Min Languages]. Beijing: Yuwen Chubanshe.
- Chen Zhangtai 陈章太 and Li Xingjian 李行健 (1996). *Putonghua Jichu Fangyan Jiben Cihui* 普通话基础方言基本词汇 [Basic Vocabulary of Mandarin]. Beijing: Yuwen Publishers.
- Chen Zhongmin 陈忠民 (1996). 'Lun beibu wuyu yizhong daici qianzhu 'shi' 论北部吴语一种代词前缀‘是’ [On the PREF ‘shi’ in complex pronouns in Northern Wu]', *Yuyan Yanjiu* 语言研究 2: 62–4.
- Chen Zhongmin and Wuyun Pan 陈忠民, 潘悟云 (1999). 'Lun Wuyu de rencheng daici 伦吴语的人称代词 [On personal pronouns in Wu]', in Li, Rulong, and Zhang Shuangqing (eds), 李如龙, 张双庆 主编 *Daici* 代词 [Pronouns]. Guangzhou: Jinan University Press, 1–24.

- Cheng, Lisa (2008). ‘Deconstructing the *shi...de* Construction,’ *The Linguistic Review*. 25 (3–4): 235–66.
- Chin, Andy C. 錢志安 (2010). ‘Two types of indirect object markers in Chinese: Their typological significance and development’, *Journal of Chinese Linguistics* 38(1): 1–25.
- Chu Zexiang 储泽祥 and Deng Yunhua 邓云华 (2003). ‘Zhishi daici de leixing he gongxing 指示代词的类型和共性 [The type and general characteristics of demonstratives]’, *Dangdai Yuyanxue [Contemporary Linguistics]* 4: 299–306.
- Clark, Marybeth (1974). ‘Submissive verbs as adversatives in some Asian languages’, in Nguyen Dang Liem (ed.), *South-East Asian Linguistic Studies*, Vol. 1. Canberra: Pacific Linguistics, Australian National University, 89–110.
- Clark, Marybeth (1989). ‘Hmong and areal Southeast Asia’, in David Bradley (ed.), *South-East Asian Syntax* (Pacific Linguistics. Series A, no. 77. *Papers in South-East Asian Linguistics*, no. 11). Canberra: The Australian National University, Research School of Pacific Studies, 177–230.
- Comrie, Bernard (1989). *Language Universals and Linguistic Typology*. Oxford: Blackwell.
- Comrie, Bernard and Maria Polinsky (eds.). (1993). *Causatives and Transitivity*. Amsterdam: John Benjamins.
- Corbett, Greville G. (1978). ‘Universals in the syntax of cardinal numerals’, *Lingua* 46: 355–68.
- Craig, Colette (ed.) (1986). *Noun Classes and Categorization* (Proceedings of a Symposium on Categorization and Noun Classification, Eugene, Oregon, October 1983) (Typological Studies in Language 7). Amsterdam: John Benjamins.
- Croft, William (1994). ‘Semantic universals in classifier systems’, *Word* 45, (2): 145–71.
- Croft, William (2000). *Explaining Language Change: An Evolutionary Approach*. Harlow, UK: Pearson Education.
- Croft, William (2003). *Typology and Universals* (2nd edition). Cambridge: Cambridge University Press.
- Crosland, Jeff (1995). ‘Xiamen dialect discourse data’. *Yuen Ren Society Treasury of Chinese Dialect Data* Vol I: 183–241.
- Cui Zhenhua 催振华 (1998). *Yiyang fangyan yanjiu 益阳方言研究* [Research on the Yiyang Dialect]. Changsha: Hunan Jiaoyu.
- Dahl, Osten (2001). ‘Principles of areal typology’, in M. Haspelmath et al. (eds.), *Language Typology and Language Universals: An International Handbook*. Berlin: De Gruyter Mouton, 1456–70.
- Dai Qingxia 戴庆厦 (ed.) (1992). *Hanyu Yu Shaoshu Minzu Yuyan Guanxi Gailun 汉语与少数民族语言关系概论* [General Outline of the Relationship between Chinese and Minority Languages]. Beijing: Zhongyuan minzu xueyuan chubanshe [Press of the Central University for the Nationalities].
- Dai Qingxia 戴庆厦 (2005). *Langsuyu Yanjiu 浪速语研究* [Research on Langsu]. Beijing: Minzu Chubanshe.
- Dai Yaojing 戴耀晶 (1999). ‘Ganyu Taihe fangyan zhong de daici 赣语泰和方言中的代词 [The pronouns in the Taihe dialect (Gan)]’, in Li and Zhang (eds), 167–75.
- de Sousa, Hilário 蘇沙 (2013). ‘Nánníng Shàngyáo Pínhuà de Yīxiē Míngcí Duǎnyǔ Xiànxìàng Duìbǐ Yánjiū 南宁上尧平话的一些名词短语现象对比研究 [Comparative studies of some noun phrase phenomena in Nánníng Shàngyáo Pínhuà]’, in Dānqīng Liú 劉丹青, Léi Zhōu

- 周磊, and Caidé Xuē 薛才德 (eds), *Hànyǔ Fāngyán Yǔfǎ Yánjiū de Xīnshǐjiǎo—Dì Wǔ Jiè Hànyǔ Fāngyán Yǔfǎ Guójí Xuéshù Yántáohuì Lùnwénjí* 汉语方言语法研究的新视角—第五届汉语方言语法国际学术研讨会论文集 [New Viewpoints in the Studies of Grammar of the Chinese Dialects—Proceedings of the Fifth International Conference on Syntax of Chinese Dialects]. Shanghai: Shanghai Educational Publishing House 上海教育出版社, 141–60.
- de Sousa, Hilário (2015). ‘The Far Southern Sinitic languages as part of Mainland Southeast Asia’. In N.J. Enfield and Bernard Comrie (eds.), *The Languages of Mainland Southeast Asia—The State of the Art*, Berlin/Boston: De Gruyter Mouton, 356–439.
- de Sousa, Hilário (this volume). ‘Language contact in Nanning: Nanning Pinghua and Nanning Cantonese’.
- Diessel, Holger (1999). *Demonstratives: Form, Function and Grammaticalization*. Amsterdam: John Benjamins.
- de Sousa, Hilário, François Langella, and N.J. Enfield (2015). ‘Temperature terms in Lao, Southern Zhuang, Southern Pinghua and Cantonese’. In Maria Koptjevskaja-Tamm (ed.), *The Linguistics of Temperature*, Amsterdam/Philadelphia: John Benjamins Publishing Company, 594–638.
- Diessel, Holger (2005). ‘Distance contrasts in demonstratives’, in Martin Haspelmath, David Gil, and Bernard Comrie (eds), *World Atlas of Language Structures*. Oxford: Oxford University Press, 170–3.
- Dik, Simon C. (1997). (ed. Kees Hengeveld). *The Theory of Functional Grammar. Part I: The Structure of the Clause*, Second, revised version. Berlin and New York: Mouton de Gruyter.
- Djamouri, Redouane (1994). ‘L’emploi des signes numériques dans les inscriptions Shang’, Extrême-Orient, Extrême-Occident. 16: 12–42.
- Dong Tonghe 董同龢 (1967). *Shanggu yinyun biao gao* 上古音韵表稿 [Draft Tables for Old Chinese Phonology]. Reprint. Taipei: Institute of History and Philology, Academia Sinica.
- Dryer, Matthew (1992). ‘The Greenbergian word order correlations’, *Language* 68: 81–138.
- Dryer, Matthew (2003). ‘Word order in Sino-Tibetan languages from a typological and geographical perspective’, in Graham Thurgood and Randy LaPolla (eds.) *The Sino-Tibetan Languages*. London/New York: Routledge, 43–55.
- Dryer, Matthew (2008). ‘Word order in Tibeto-Burman languages’, *Linguistics of the Tibeto-Burman Area* 31.1: 1–88.
- Duanmu San 端木三 (2011). ‘Chinese syllable structure’, in Marc Van Oostendorp, Colin Ewen, Elizabeth Hume, and Keren Rice (eds), *The Blackwell Companion to Phonology* volume 5. Malden, MA/Oxford: Wiley-Blackwell, 2151–77.
- Emeneau, Murray (1956). ‘India as a linguistic area’, *Language* 32: 3–16.
- Endo, Masahiro (遠藤雅裕) (2004). ‘Hanyu fangyan chuzhishi biaoji de dili fenbu yu gezhong chuzhiji 汉语方言处置式标记的地理分布与各种处置式 [The geographical distribution of disposal markers in Chinese dialects and several types of disposal structures]’, *Chuugoku Gogaku* 251: 1–15.
- Enfield, N.J. (2003). *Linguistic Epidemiology: Semantics and Grammar of Language Contact in Mainland Southeast Asia*. London: Routledge Curzon.
- Enfield, N.J. (2005). ‘Areal linguistics and mainland Southeast Asia’, *Annual Review of Anthropology* 34: 181–206.
- Enfield, N.J. (2007). *A Grammar of Lao*. (Mouton Grammar Library) Berlin: Mouton de Gruyter.

280 *References*

- Erbaugh, Mary S. and Yang Bei (2006). 'Two general classifiers in the Shanghai Wu dialect: A comparison with Mandarin and Cantonese', *Cahiers de linguistique—Asie orientale*, 35: 169–207.
- Faarlund, Jan (1990). *Syntactic Change: Toward a Theory of Historical Syntax*. Berlin: Mouton de Gruyter.
- Fang Mei 方梅 (2006). 'Beijinghua li shuo de yufahua, cong yanshuo dongci dao congju biaoji 北京话里说的语法化，从言说动词到从句标记 [The grammaticalization of the saying verbs into complementizers in Pekingese]', *Zhongguofangyanxuebao 中国方言学报* 1: 107–21.
- Feng Aizhen 馮愛珍 (1998). *Fuzhou Fangyan Cidian 福州方言詞典* [A Dictionary of the Fuzhou Dialect]. Nanjing: Jiangsu Education Press.
- Fillmore, Charles J. (1982). 'Towards a descriptive framework for spatial deixis', in Remo Jarvella and Wolfgang Klein (eds), *Speech, Place, and Action: Studies in Deixis and Related Topics*. Chichester: John Wiley and Sons.
- Fought, Carmen. (2010). 'Ethnic identity and linguistic contact', in Raymond Hickey (ed.), *The Handbook of Language Contact*. West Sussex: Wiley-Blackwell, 282–98.
- Gan, Rui-Jing, Shang-Ling Pan, Laura F. Mustavich, Zhen-Dong Qin, Xiao-Yun Cai, Ji Qian, Cheng-Wu Liu, Jun-Hua Peng, Shi-Lin Li, Jie-Shun Xu, Li Jin, Hui Li, and The Genographic Consortium (2008). 'Pinghua population as an exception of Han Chinese's coherent genetic structure', *Journal of Human Genetics* 53: 303–13.
- Gan Yuen 甘于恩 (2010). *Guangdong Siyi fangyan yufa Yanjiu 广东四邑方言语法研究* [Research on the Grammar of the Siyi Dialect in Guangdong Province]. Guangzhou: Jinan University Press.
- Giacalone-Ramat, Anna (1998). 'Testing the boundaries of grammaticalization', in Anna Giacalone-Ramat and Paul Hopper (eds), *The Limits of Grammaticalization*. Amsterdam: John Benjamins, 197–227.
- Glaser, Elvira (2006). 'Zur Syntax des Lützebuergeschen: Skizze und Forschungsprogramm', in Claudin Moulin and Damaris Nübling (eds.), *Perspektiven einer linguistischen Luxemburgistik: Studien zur Diachronie und Synchronie*. Heidelberg: Universitäts Verlag Winter, 225–44.
- Goméz Rendón, Jorge (2008). 'Typological and social constraints on language contact: Amerindian languages in contact with Spanish', PhD thesis, Universiteit van Amsterdam.
- Gould, Stephen (1983). *Hen's Teeth and Horse's Toe: Further Reflections in Natural History*. New York: W.W. Norton and Company.
- Gould, Stephen and Elizabeth Vrba (1982). 'Exaptation—a missing term in the science of form', *Paleobiology* 8: 4–15.
- Greenberg, Joseph (1963). 'Some universals of grammar with particular reference to the order of meaningful elements', in Joseph Greenberg (ed.), *Universals of Language*. Cambridge, MA: MIT Press, 58–90.
- Greenberg, Joseph (1966a). *Language Universals, with Special Reference to Feature Hierarchies*. (Janua Linguarum, Series Minor 59.) The Hague: Mouton.
- Greenberg, Joseph (1966b). 'Some universals of grammar with particular reference to the order of meaningful elements', in Joseph Greenberg (ed.), *Universals of Language*. Cambridge, MA: MIT Press, 73–113.
- Greenberg, Joseph (1977). 'Numeral classifiers and substantival number: problems in the genesis of a linguistic type', in A. Makkai, V.B. Makkai, and Luigi Heilmann (eds), *Linguistics at the Crossroads*. Lake Bluff, Illinois: Jupiter Press, 276–300.
- Greenberg, Joseph (1978). 'Generalizations about numeral systems', in Joseph Greenberg (ed.), *Universals of Human Language*, Volume 3. Stanford, CA: Stanford University Press, 249–97.

- Greenberg, Joseph (1991). ‘The last stages of grammatical elements: Contractive and expansive desemanticization’, in Elisabeth Traugott and Bernd Heine (eds), *Approaches to Grammaticalization*. Amsterdam: John Benjamins, 301–14.
- Greenberg, Joseph (2001). ‘The methods and purposes of linguistic genetic classification’, *Language and Linguistics* 2.2: 111–35.
- Grosberg, John (1957). ‘The numeral in idiomatic Russian’, in *Études Slaves et Est-Européennes / Slavic and East-European Studies* 2.3: 177–9.
- Güldemann, Tom (2008). *Quotative Indexes in African Languages: A Synchronic and Diachronic Survey*. Berlin: Mouton de Gruyter.
- Gundel, K. Jeanette and Thorstein Fretheim (2003). ‘Topic and focus’, in G. Ward and L. Horn, (eds), *Handbook of Pragmatic Theory*. Oxford: Blackwell, 175–96.
- Guo Xiliang 郭锡良 (1997). ‘Jieci ‘yu’ de qiyuan he fazhan 介词 « 』的起源和发展 [On the origins and development of the preposition *yù*]’, *Zhongguo Yuwen* 中國語文 2: 131–8.
- Hagège, Claude (1982). *La Structure des Langues*. Paris : Presses Universitaires de France (*Que sais-je ?*).
- Hagège, Claude (1993). *The Language Builder*. Amsterdam: John Benjamins.
- Haiman, John (2011). *Cambodian Khmer*. (London Oriental and African Language Library 16). Amsterdam/Philadelphia: John Benjamins.
- Harris, Alice and Lyle Campbell (1995). *Historical Syntax in Cross-Linguistic Perspective*. Cambridge: Cambridge University Press.
- Hartmann, Katharina, and Malta Zimmermann (2009). ‘Morphological focus marking in Gürültüm (West Chadic)’, *Lingua* 119 (9): 1340–65.
- Hashimoto, Mantaro (1976a). ‘Language diffusion on the Asian continent: Problems of typological diversity in Sino-Tibetan’, *Computational Analyses of Asian and African Languages* 3: 49–65.
- Hashimoto, Mantaro (1976b). ‘The double object construction in Chinese’, *Computational Analyses in Asian and African Languages* 6: 33–42.
- Hashimoto, Mantaro (1978). *Gengo ruikei chiriron* 语言类型地理论 [*Linguistic Typogeography*]. 东京: 宏文堂. Tokyo: Kobundo.
- Hashimoto, Mantaro (1986). ‘The Altaicization of Northern Chinese’, in J. McCoy and T. Light (eds.), *Contribution to Sino-Tibetan Studies*. Leiden: E. J. Brill, 76–97.
- Hashimoto, Mantaro 桥本万太郎 (1987). ‘Hànyǔ bēidòngshì de lishi 『汉语被动式的历史』 [The history of the Chinese passive construction]’, *Zhōngguó Yǔwén* 《中国语文》 1: 36–49.
- Hashimoto, Mantaro (1988). ‘The structure and the typology of the Chinese passive construction’, in M. Shibatani (ed.), 329–54. = Translation of Hashimoto (1987).
- Haspelmath, Martin (1990). ‘Grammaticalization of passive morphology’, *Studies in Language* 14.1: 25–72.
- Haspelmath, Martin (1999). ‘Why is grammaticalization irreversible?’, *Linguistics* 37: 1043–68.
- Haspelmath, Martin (2004). ‘On directionality in language change with particular reference to grammaticalization’, in Olga Fischer, Muriel Norde, and Harry Perridon (eds), *Up and Down the Cline—the Nature of Grammaticalization*. Amsterdam/Philadelphia: John Benjamins, 17–44.
- Haspelmath, Martin (2005). ‘Ditransitive constructions: The verb *give*’, in Martin Haspelmath, Matthew Dryer, David Gil, and Bernard Comrie (eds.) *The World Atlas of Language Structures*. Oxford: Oxford University Press, Chapter 105.

282 *References*

- Haspelmath, Martin and the Comparative Constructions Consortium (2013). ‘The typology of comparative constructions revisited,’ paper presented at the *Tenth Biennial Conference of the Association for Linguistic Typology* (AL10), August 2013, at the University of Leipzig.
- Haugen, Einar (1950). ‘The analysis of linguistic borrowing’, *Language* 57: 183–91.
- He Hongfeng 何洪峰 (2011). ‘Dongci “qù” xiāng chusuo jiècì yufahua de zhōngzhí yǔ huiguī’ 動詞“去”向處所介詞語法化的終止與回歸 [The termination and retroversion of grammaticalization of qù in the process of becoming a locative preposition], *Yuyan Yanjiu* 語言研究 [*Studies in Language and Linguistics*] 2: 48–55.
- He Wenping 贺文平 (2006). ‘Xiangxiang (Hutian) hua-zhong ‘zhi+ming’ jiégou yanjiu 湘乡(壶天)话中“隻+名”结构研究 [A study of ‘隻+noun’ phrases in the Xiangxiang (Hutian) dialect]’, MA thesis, Yanbian University, Yanji.
- Heine, Bernd (1997). *Cognitive Foundations of Grammar*. Oxford: Oxford University Press.
- Heine, Bernd (2003). ‘On degrammaticalization’, in Kate Burridge, Barry Blake, and Jo Taylor (eds), *Historical linguistics 2001: Selected Papers from the 15th International Conference on Historical Linguistics*. Melbourne, Australia, 13–17 August 2001. (*Amsterdam Studies in the Theory and History of Linguistic Science. Series IV: Current Issues in Linguistic Theory*, 237.) Amsterdam and Philadelphia: John Benjamins, 163–79.
- Heine, Bernd and Tania Kuteva (2002). *Word Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.
- Heine, Bernd and Tania Kuteva (2005). *Language Contact and Grammatical Change*. Cambridge: Cambridge University Press.
- Heine, Bernd and Tania Kuteva (2006). *The Changing Languages of Europe*. Cambridge: Cambridge University Press.
- Heine, Bernd and Tania Kuteva (2007). *The Genesis of Grammar*. Oxford/New York: Oxford University Press.
- Himmelmann, Nikolaus P. (1996). ‘Demonstratives in narrative discourse: A taxonomy of universal uses’, in Barbara A. Fox (ed.), *Studies in Anaphora*. Amsterdam: John Benjamins, 201–54.
- Hirata, Shoji (平田昌司) (1998). *Huizhou Fangyan Yanjiu* 徽州方言研究. Tokyo: Kobun Shuppen.
- Hong Bo 洪波 (1998). ‘Lun Hanyu shici xuhua de jizhi 论汉语实词虚化的机制 [On the mechanisms of full words becoming empty words in Chinese]’, in Hong Bo (ed.), *Hanyu lishi yufa yanjiu* 汉语历史语法研究 [*Research on Chinese Historical Grammar*]. Beijing: Shangwu Yinshuguan, 299–310.
- Hong Cheng 洪诚 (1963). *Lüelun liangci ge de yuyuan ji qi zai Tang yiqian de fazhan qingkuang* 略论量词‘个’的语源及其在唐以前的发展情况 [A brief discussion on the origin of GE and its development before the Tang Dynasty] 南京大学学报 *Nanjing daxue xuebao* 2: 40–5.
- Hong Ji 洪济 (2009). *Taining Xianzhi* 泰宁县志 [A Survey of Taining County]. Xiamen: Xiamen Daxue Chubanshe.
- Hopper, Paul (1991). ‘On some principles of grammaticalization’, in Elizabeth Traugott and Bernd Heine (eds), *Approaches to Grammaticalization*. Amsterdam: John Benjamins, 17–36.
- Hopper, Paul J. and Sandra A. Thompson (1980). ‘Transitivity in grammar and discourse’, *Language* 56 (1): 251–99.

- Hopper, Paul and Elizabeth Traugott (1993). *Grammaticalization*. Cambridge: Cambridge University Press. 2nd edition in 2003.
- Hu Guangbin 胡光斌 (2010). *Zunyi Fangyan Yufa Yanjiu* 遵义方言语法研究 [Research on the Grammar of the Zunyi Dialect]. Chengdu: Bashu Press.
- Hu Jianhua, Pan Haihua, and Li Baolun 胡建华, 潘海华, 李宝伦 (2003). Ningbohua yu putonghua-zhong huati he cihuati de jufa weizhi 宁波话与普通话中话题和次话题的句法位置 [The syntactic positions of topic and secondary topic in Ningbo Wu and Mandarin], in Xu, Liejiong, and Liu Danqing (eds) 徐烈炯, 刘丹青主编, *Huati Yu Jiaodian Xinlun* 话题与焦点新论 [New Arguments on Topic and Focus]. Shanghai: Shanghai Education Press, 164–75.
- Hu Songbo 胡松柏 and Ge Xin 葛新 (2003). ‘Gan dongbei Gan, Wu, Huiyu jieyuan didai fangyan de chuzhi jieci he beidong jieci 赣, 东北赣, 吴, 徽语接缘地带方言的处置介词 和被动介词 [The disposal and passive prepositions in northeastern Jiangxi where Gan, Wu and Hui dialects co-exist]’, in *Hanyu fangyan yufa yanjiu he tansuo—shoujie guoji Hanyu fangyan yufa xueshu yantaohui lunwenji* 汉语方言语法研究和探索—首届国际汉语方言语法学术研讨会论文集 [Research on grammar of Chinese dialects—anthology of the First International Conference on the Grammar of Chinese], Dai Zhaoming (ed.), 戴昭铭. Harbin: Heilongjiang Renmin Chubanshe, 241–5.
- Huang Jianyun 黄剑云 (1990). *Táishān Fāngyán* 台山方言 [Táishān Dialect]. Guangzhou: Sun Yat-sen University Press 中山大学出版社.
- Huang Shuang-fang (2003). ‘Doubts about complementation: A functionalistic analysis’, *Language and Linguistics* 4–2: 429–55.
- Huang Xianfan 黄现璠 (2004). *Gushu jiedu chutan—Huang Xianfan xueshu lunwenji* 古书解读初探—黄现璠学术论文选 [Exploring and explaining old scripts—Selected academic papers by Huang Xianfan], 广西师范大学出版社 Guilin: Guangxi Shifan Daxue Chubanshe.
- Huang Xiaohui 黄晓惠 (1992). ‘Xiandai Hanyu chabi geshi de laiyuan ji yanbian 现代汉语差比格式的来源及演变 [The origin and development of the Modern Chinese comparative of inequality]’, *Zhongguo Yuwen* 中國語文 3: 213–24.
- Huang Yingfu 黄英富 (2005). ‘Binyángxiàn Dàqíáo Pínghuà Yīnxì 宾阳县大桥平话音系 [The phonology of Dàqíáo Pínghuà in Binyáng County]’, *Journal of Guangxi College of Education* 广西教育学院学报 78(4): 112–15.
- Huang Yuanwei (1997). ‘The interaction between Zhuang and the Yue (Cantonese) dialects’, in Jerold A. Edmondson and David B. Solnit (eds), *Comparative Kadai: The Tai Branch*. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington, 57–76.
- Huddleston, Rodney and Geoffrey K. Pullum (2002). *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- Hurford, James (1987). *Language and Number: The Emergence of a Cognitive System*. Oxford: Blackwell.
- Jaisser, Annie and Martha Ratliff (1995). *Hmong for Beginners. Part 7*. Berkeley: Center for Southeast Asia Studies, UC Berkeley.
- Jenny, Mathias Jenny, Mathias (2015). *The Far West of Southeast Asia: ‘Give’ and ‘get’ in the Languages of Myanmar*, in Nick J. Enfield and Bernard Comrie, *Languages of Mainland Southeast Asia*. Berlin, 155–208.

284 *References*

- Jespersen, Otto (1969). *Analytic Syntax*. New York: Holt, Rinehart, and Winston.
- Jiang Lansheng 江蓝生 (1999). ‘Hanyu shiyi yu beidong jianyong tan yuan 汉语使役与被动兼用探源 [On the common uses of causatives and passives]’, in Alain Peyraube and Chaofen Sun (eds.), *In Honor of Mei Tsu-lin—Studies on Chinese Historical Syntax and Morphology*. Paris: École des Hautes Études en Sciences Sociales, 57–72.
- Jiang Shaoyu 蒋绍愚 (2002). ‘Gei ziju, jiao ziju biao beidong de laiyuan 给字句，叫字句表被动的来源 [Origin of the passive markers *gei* and *jiao*]’, *Yuyanxue luncong 语言学论丛* 26: 159–77.
- Jin Youjing 金有景 (1962). ‘Suzhou fangyan de fangwei zhishici 苏州方言的方位指示词 [The demonstrative pronouns for location in the Suzhou dialect]’, *Zhongguo Yuwen 中國語文* 4: 183.
- Karlgren, Bernhard (1964). *Grammata Serica Recensa*. Göteborg: Elanders Boktryckeri Aktiebolag.
- Keenan, Edward L. and Matthew S. Dryer (2007). ‘Passives in the world’s languages’, in Timothy Shopen (ed.), *Language Typology and Syntactic Description*. Vol.1. 2nd edition. Cambridge: Cambridge University Press, 325–61.
- Kiparsky, Paul (2005). ‘Grammaticalization as optimization’, paper delivered at Santigao de Compostela, July 17.
- Kiparsky, Paul (2012). ‘Grammaticalization as optimization’, in Dianne Jonas, John Whitman, and Andrew Garrett (eds), *Grammatical Change: Origins, Nature, Outcomes*. Oxford: Oxford University Press, 15–51.
- Kiss, E. Katalin (1998). ‘Identificational focus and information focus’, *Language* 74: 245–73.
- Kiss, E. Katalin (2008). ‘Contrastive focus and emphasis’, *Acta Linguistica Hungarica* 55: 347–60.
- Kuryłowicz, Jerzy (1965). ‘The evolution of grammatical categories’, reprinted in Jerzy Kuryłowicz (1976), *Esquisses linguistiques* 2: 38–54. Munich: Fink.
- Kwok Bit-Chee 郭必之 (2010). ‘Yüyán Jiēchù zhōng de Yǔfǎ Biànshù: Lùn Nánníng Yüeyǔ “Shýyǔ + Bínnyǔ + Bǔyǔ” Jiégòu de Láiuyuán 語言接觸中的語法變化：論南寧粵語「述語 + 賓語 + 補語」結構的來源 [Grammatical change in language contact: On the origin of the “verb + object + complement” construction in Nanning Yue]’, in Hung-nin Samuel Cheung 張洪年 and Song Hing Chang 張雙慶 (eds), *Diachronic Change and Language Contact—Dialects in South East China 歷時演變與語言接觸—中國東南方言*, *Journal of Chinese Linguistics Monograph* 24, 201–16.
- Langacker, Ronald (1977). ‘Syntactic reanalysis’, in Charles Li (ed.), *Mechanisms of Syntactic Change*. Austin: University of Texas Press, 57–139.
- LaPolla, Randy (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), 225–54.
- Lass, Roger (1990). ‘How to do things with junk: Exaptation in language evolution’, *Journal of Linguistics* 26: 79–102.
- Lau Chun-fat 刘镇发 (2007). ‘Xiàndài Fāngyán de [j] Zhuócháhuà—Fùlùn Zhōnggǔ Xiémǔ de Xíngchéng 现代方言的[j]浊擦化—附论中古邪母的形成 [The voicing and fricativization of [j] in the modern dialects—and the formation of the Xié initial in Middle Chinese]’, *Zhōngguó Yǔwén 中国语文* 317(2): 169–74.
- Lazard, Gilbert (1994). ‘Le *râ* persan et le *ba* chinois’, *Cahiers de Linguistique—Asie Orientale* 23: 169–76.

- Leblanc, Jacques (1910). *Cours de langue chinoise parlée: dialecte cantonais*. Hanoi/Haiphong : Imprimerie d'Extrême-Orient.
- Lenz, Alexandra (2008). ‘Zur Grammatikalisierung von *geben* im Deutschen und Lëtzebuer-geschen’, *Zeitschrift für Germanistische Linguistik* 35.1–2: 52–82.
- Lewis, M. Paul, Gary F. Simons, and Charles D. Fennig (eds) (2014). *Ethnologue: Languages of the World*. 17th edition. Dallas, Texas: SIL International.
- Li, Charles N. and Sandra A. Thompson (1981). *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press.
- Li Chia-chun and Lien Chinfa (李家純, 連金发) (1995). ‘Lun Minnanyu bijiaoshi leixing ji lishide tantao 論閩南語比較式類型及歷史的探討 (An exploration of the typology and history of the comparative in Southern Min)’, in Tsao Fengfu and Meihui Tsai (eds), *Papers from the 1994 Conference on Language Teaching and Linguistics in Taiwan*, Vol I: Southern Min. Taipei: The Crane Publishing Co., 71–87.
- Li Fang-Kuei (1960). ‘A tentative classification of Tai dialects’, in S. Diamond (ed.), *Culture in History: Essays in Honor of Paul Radin*. New York: Columbia University Press, 951–9.
- Li Fang-Kuei (1971). ‘Shan-ku In Yen-chiu [Research on Old Chinese Phonology]’, *Tsinghua Journal of Chinese Studies* New Series IX : 1–2.
- Li Fang-Kuei (1977). *A Handbook of Comparative Tai*. Honolulu: The University of Hawaii Press.
- Li Jiachun and Chinfa Lien 李佳純, 連金發 (1995). ‘Lun Minnanyu bijiaoshi—leixing ji lishi de tantao 論閩南語比較式—類型及歷時的探討 [A discussion on comparatives in South-ern Min—types and history]’, in Fengfu Tsao and Meihui Tsai 曹逢甫, 蔡美慧 (eds), *Taiwan Minnanyu Lunwenji* 臺灣閩南語論文集 [*Collections of Theses on Taiwan Southern Min*]. Taipei: Crane Publishing Company, 71–85.
- Li Jinfang 李锦芳 (2002). *Dòng Tái Yǔyán yǔ Wénhuà* 侗台语言与文化 [*Tai-Kadai: Lan-guage and Culture*]. Beijing: The Ethnic Publishing House 民族出版社.
- Li Jingzhong 李敬忠 (1990). ‘Yueyu shi hanyu zuqun zhong de duli yuyan 粤语是汉语族群中 的独立语言[Cantonese is different from other Sinitic languages]’, *Yuyan Jianshe Tongxun* 语言建设通讯 27: 28–48.
- Li Jingzhong 李敬忠 (1994). *Yǔyán Yǎnbìan Lùn* 语言演变论 [*Theory of Language Change*]. Guangzhou: Guangzhou Publishing House 广州出版社.
- Li Lan 李藍 (2003). ‘Xiandai hanyu fangyan chabiju de yuxu leixing 現代漢語方言差比句的 語序類型 [Word Order Typology of Comparative Constructions in Modern Chinese Dialects]’, *Fangyan 方言* [Dialect] 3: 214–32.
- Li Lan (李藍) (2006). ‘Zhuo zi shi de beidong ju de gongshi fenbu yu leixing chayi. 着字式的 被动句的共时分布与类型差异 [Synchronic distribution and typological differences for passive constructions formed with *zhuo*]’, *Zhongguo Fangyan Xuebao* 中国方言学报, 194–205.
- Li Lan (李藍), Cao Xilei (曹茜蕾) (= H. Chappell) (2013a). ‘Hanyu fangyan-zhong de chuzhishi he ‘bā’ ziju (shang) 汉语方言中的处置式和把字句(上) [Disposal constructions and bā constructions in Chinese dialects, (Part I)]’, *Fangyan* (Dialects) 1: 11–30.
- Li Lan (李藍), Cao Xilei (曹茜蕾) (= H. Chappell) (2013b). ‘Hanyu fangyan-zhong de chuzhishi he ‘bā’ ziju (xia) 汉语方言中的处置式和把字句(下) [Disposal constructions and bā constructions in Chinese dialects, (Part II)]’, *Fangyan* (Dialects) 2: 97–110.

- Li Lianjin 李连进 (2003). ‘Pínghuà Shì Dúlì Fāngyán Hái shì Shúyú Yuè Fāngyán 平话是独立方言还是属于粤方言 [Is Pinghua an independent dialect or affiliated to the Yue dialects?]', in Bóhùi Zhān 詹伯慧 (ed.), *Dì Bā Jiè Guójì Yuè Fāngyán Yántǎohùi Lùnwénjí* 第八届国际粤方言研讨会论文集 [*Proceedings of The Eighth International Conference on Yuè Dialects*]. Beijing: China Social Science Press 中国社会科学出版社, 70–94.
- Li Lianjin 李连进 and Yan'e Zhu 朱艳娥 (2009). *Guǎngxī Chóngzuò Jiāngzhōu Zhèyuánhuà Bǐjiào Yánjiū* 广西崇左江州蔗园话比较研究 [*Comparative Studies of the Sugarcane Field Dialect in Jiāngzhōu, Chóngzhō, Guangxi*]. Guilin: Guangxi Normal University Press 广西师范大学出版社.
- Li Ming 李明 (2003). ‘Shi tan yanshuo dongci xiang renzhi dongci de yinshen 试谈演说动词向认知动词的引申 [On the verbs of saying becoming cognitive verbs]', in Wu Fuxiang and Hong Bo (eds) 吴福祥, 洪波 (主编), *Yufahua Yu Yufa Yanjiu 1* 语法化与语法研究(一). Beijing; Shangwu yinshuguan. 350–70.
- Li Rong 李榮 (1952). 切韵音系 *Qieyun Yinxi* [*The Phonological System of the Qieyun*]. Beijing: Kexue Chubanshe.
- Li Rulong 李如龍 (2000). ‘Minnan fangyan de jieci 閩南方言的介詞 [Adpositions in Southern Min]', in Rulong Li and Song-hing Chang 李如龍, 張雙慶 (eds.), *Jieci 介詞 [Adpositions]*. Guangzhou: Jǐ'nan University Press, 126–43.
- Li Rulong (李如龍) and Zhang Shuangqing (張雙慶) (eds) (1992). *Ke-Gan Fangyan Diaocha Baogao* 客贛方言調查報告 [*A Report on a Survey of Kejia and Gan Dialects*]. Xiamen: Xiamen University Publishing House.
- Li Rulong 李如龍 and Zhang Shuangqing 張双庆 (eds) (1997). *Dongci weiyuju* 动词谓语句. [Verbal Predicate Sentences]. Guangzhou: Jinan Daxue Chubanshe.
- Li Rulong 李汝龙 and Zhang Shuangqing 张双庆 (eds) (1999). *Daici 代词 [Pronouns]*. Guangzhou: Jinan Daxue Chubanshe.
- Li Xiaofan 李小凡 (1998). *Suzhou Fangyan Yufa Yanjiu* 苏州方言語法研究 [*On the Grammar of the Suzhou Dialect*]. Beijing: Beijing University Press.
- Li Xinkui 李 (1994). *Guangdong de Fangyan* 廣東的方言 [*Dialects of Guangdong*]. Guangzhou: Guangdong Renmin Chubanshe.
- Li Xinkui and Huang Jiajiao 李新魁, 黄家教 (1995). *Guangzhou fangyan yanjiu* 广州方言研究 [*Studies on the Guangzhou Dialect*]. Guangzhou: Guangdong Renmin Chubanshe.
- Li XuPing (2013). *Numerical Classifiers in Chinese: The Syntax-Semantics Interface*. Berlin: De Gruyter Mouton.
- Li XuPing and Walter Bisang (2012). ‘Classifiers in Sinitic languages: From individuation to definiteness-marking’, *Lingua* 122(4): 335–55.
- Li Yunbing 李雲兵 (2003). ‘Miáo-Yáoyǔ Shēngdiào Wèntí 苗瑤語聲調問題 [Some issues regarding tones in the Miao-Yao languages]', *Language and Linguistics 語言暨語言學* 4(4): 683–712.
- Liang Jinrong 梁金荣 (1997). ‘Guibei Pínghuà Yüyín Yánjiū 桂北平话语音研究 [Phonological studies of Northern Pinghua]', PhD thesis, Jinan University.
- Liang Jinrong 梁金荣 (2000). ‘Cóng Yüyín Tèzhēng Kàn Guibei Pínghuà yǔ Yuè Fāngyán de Guānxì 从语音特徵看桂北平话与粤方言的关系 [Viewing the relationship between Northern Pinghua and Yue dialects from their phonological characteristics]', in Chow Yiu Sin 单周尧 and Kang-kwong Luke 陆镜光 (eds), *Yuèyǔ Yánjiū: Dì Qī Jiè Guójì*

- Yuè Fāngyán Yántāohuì Lùnwénjí* 粤语研究：第七届国际粤方言研讨会论文集 [Yue Research: Proceedings of The Eighth International Conference on Yuè Dialects]. Beijing: The Commercial Press 商务印书馆.
- Liang Min 梁敏 (2009a). ‘Dòngyǔ Jiānzhì 哆语简志 [Sketch of Kam]’, in Xiǎoyù Zēng 曾晓渝 and Xùliàn Lǐ 李旭练 (eds), *Zhōngguó Shǎoshù Mínzú Yúyán Jiānzhì Cóngshū* 中国少数民族语言简志丛书. Volume 3. Beijing: The Ethnic Publishing House 民族出版社, 165–238.
- Liang Min 梁敏 (2009b). ‘Máonànyǔ Jiānzhì 毛南语简志 [Sketch of Maonan]’, in Xiǎoyù Zēng 曾晓渝 and Xùliàn Lǐ 李旭练 (eds), *Zhōngguó Shǎoshù Mínzú Yúyán Jiānzhì Cóngshū* 中国少数民族语言简志丛书 Volume 3. Beijing: The Ethnic Publishing House 民族出版社, 623–90.
- Liang Weihua 梁伟华 and Yi Lin 林亦 (2009). *Guāngxī Chóngzuò Xīnhé Zhèyuánhuà Yánjiū* 广西崇左新和蔗园话研究 [Studies of the Sugarcane Field Dialect in Xīnhé, Chóngzhō, Guangxi]. Guilin: Guangxi Normal University Press 广西师范大学出版社.
- Lien Chinfa (1994). ‘Lexical diffusion’, in R. E. Asher (ed.), *Encyclopedia of Language and Linguistics*. (Vol. 4). Oxford: Pergamon Press, 2141–4.
- Lien Chinfa (2001). ‘Competing Morphological Changes in Taiwan Southern Min’, in H. Chappell (ed.), *Sinitic Grammar: Synchronic and Diachronic Perspectives*. Oxford: Oxford University Press, 309–39.
- Lien Chinfa (2005). ‘Families of ditransitive constructions in *Li Jing Ji*’, *Language and Linguistics* 6.4: 707–37.
- Lien Chinfa (2011). ‘Verbs of saying *serh4*, *tann3* and *kong2* in the *Li Jing Ji*: A constructional approach’, *Cahiers de linguistique Asie Orientale* 40–1: 33–72.
- Lin Baoqing 林宝卿 (1998). *Minnan Fangyan yu Gu Hanyu Tongyuan Cidian* 閩南方言與古漢語同源詞典 [A Dictionary of Cognates in Southern Min and Ancient Chinese]. Xiamen: Xiamen University Press.
- Lin Huadong 林华东 (2008) *Quanzhou fangyan yanjiu* 泉州方言研究 [Research on the Quanzhou Dialect]. Xiamen: Xiamen University Press.
- Lin Qinjuan 林钦娟 (2008). ‘Qīnzhōuhuà Tónghǐnzhì 欽州话同音字汇 [The homophony of the Qinzhou dialect]’, *Journal of Guilin Normal College* 桂林师范高等专科学校学报 22(1): 1–7.
- Lin Yi 林亦 (2003). ‘Guānyú Pínghuà děng de Zhuàngyǔ Jiècí 关于平话等的壮语借词 [On the Zhuàng loanwords in Pínghuà and others]’, *Minority Languages of China* 民族语文 2003(2): 46–50.
- Lin Yi (林亦) (2005). *Guibeī Pínghuà yu tuiguāng Putonghua yanjiu—Xīng’ān Gāoshāng Ruàntuhuà yanjiu* 桂北平话与推广普通话研究—兴安高尚软土话 [Research on Northern Pinghua and Promoting Standard Mandarin—The Soft Patois of Gaoshang in Xing'an County]. Nanning: Guangxi Minzu.
- Lin Yi 林亦 (2010). ‘Hàn-Zhuàngyǔ Jīchù xià de Guāngxī Yuè Fāngyán Gěiyūyì Dòngcí de Biānyì 漢壯語接觸下的廣西粵方言給與動詞的變異 [Development of the Give verb in the Yue dialects of Guangxi: A contact linguistic approach]’, in Hung-nin Samuel Cheung 張洪年 and Song Hing Chang 張雙慶 (eds), *Diachronic Change and Language Contact—Dialects in South East China* 歷時演變與語言接觸—中國東南方言, *Journal of Chinese Linguistics Monograph* 24 (2010), 217–26.

- Lin Yi 林亦 and Fèngyǔ Qín 覃凤余 (2008). *Guāngxī Nánníng Bái huà Yánjiū* 广西南宁白话研究 [Studies on the Nánníng Cantonese of Guangxi]. Guilin: Guangxi Normal University Press 广西师范大学出版社.
- Liu Baoyuan 刘保元 (2009). ‘Yáoosú Lājìayú Jiānzhi 瑶族拉珈语简志 [Sketch of Lakkja of the Yao nationality]’, in Xiāoyù Zēng 曾晓渝 and Xùliàn Lǐ 李旭练 (eds), *Zhōngguó Shǎoshù Mínzú Yǔyán Jiānzhi Cóngshū* 中国少数民族语言简志从书 Volume 3. Beijing: The Ethnic Publishing House 民族出版社, 239–85.
- Liu Cunhan 劉村漢 (1995). *Liúzhōu Fāngyán Cídiǎn* 柳州方言詞典 [Dictionary of the Liúzhōu Dialect]. Nanjing: Jiangsu Education Publishing House 江蘇教育出版社.
- Liu Danqing 刘丹青 (1995). ‘Wujiang fangyan de zhishi fanchou 吴江方言的指示范畴 [The category of demonstrative in the Wujiang dialect]’, presented at *Quanguo Hanyu Fangyan-xue Nianhui* [The annual conference on Chinese dialects]. Wuhan.
- Liu Danqing 刘丹青(1999). ‘Wujiang Fangyan de daici xitong jí neibu chayi 吴江方言的代词系统及内部差异 [The demonstrative systems in the Wujiang dialect and their internal differences]’, in Li and Zhang (eds), 102–25.
- Liu Danqing (劉丹青) (2003a). *Yǔxù Lèixíngxué yǔ Jiècí Lìlùn* 語序類型與介詞理論 [Word Order Typology and Theory of Prepositions]. Beijing: Shangwu Yinshuguan.
- Liu Danqing 劉丹青(2003b). ‘Chabiju de diaocha kuangjia yu yanjiu silu’ 差比句的調查框架與研究思路 [The investigation framework and research ideas on comparative sentences], in Qingxia Dai and Yang Gu 戴慶夏, 顧陽 (eds), *Xiandai Yuyanxue Lilun yu Zhongguo Shaoshu Minzu Yuyan Yanjiu* 現代語言學理論與中國少數民族語言研究 [Modern Linguistic Theory and the Study of Chinese Minority Languages]. Beijing: The Ethnic Publishing House, 1–21.
- Liu Danqing 刘丹青(2004). ‘Hanyu li de yige neirong binyu biaojudi—cong shuo-dao de dao shuoqi 汉语里的一个内容宾语标句词—从‘说道’的‘道’说起 [A substantive complementizer in Chinese—on the *dao* of *shuodao*]’, Qingzhu «Zhongguo Yuwen» chuangkan 50 zhounian xueshu lunwenji 《庆祝〈中国语文〉创刊50周年学术论文集》 [Anthology celebrating the 50th anniversary of *Zhongguo Yuwen* 中國語文]. Academy of Social Sciences, Beijing: Shangwu Yinshuguan.
- Liu Danqing 刘丹青 and Xu Liejiong 徐烈炯 (1998). ‘Putonghua yu Shanghaihua-zhong de kaobeishi huati jiegou 普通话与上海话中的拷贝式话题结构’ [Topic-copying in Standard Mandarin and the Shanghai dialect], *Yuyan Jiaoxue yu Yanjiu* 《语言教学与研究》 1998: 1.
- Liu Danqing 刘丹青 and Chen Yujie 陈玉洁 (2008). ‘Hanyu zhishici xiāngsìxing de kuafangyan kaocha (Shang) 汉语指示词像似性的跨方言考察 (上) [Studies in the phonetic iconicity of Chinese demonstratives across dialects (Part 1)]’, *Dangdai Yuyanxue* 4: 289–97.
- Liu Danqing 刘丹青 and Chen Yujie 陈玉洁 (2009). ‘Hanyu zhishici xiāngsìxing de kuafangyan kaocha (Xia) 汉语指示词像似性的跨方言考察 (下) [Studies in the phonetic iconicity of Chinese demonstratives across dialects (Part 2)]’, *Dangdai Yuyanxue* 1:1–9.
- Liu Danqing 刘丹青 and Liu Haiyan 刘海燕 (2005). ‘Chongming fangyan de zhishici—fanfu de xitong jiqi beihou de yuyan gongxing 崇明方言的指示词·繁复的系统及其背后的语言共性 [Demonstratives in the Chongming Dialect: A complicated system and the language universals Behind it]’, *Fangyan*, 2: 97–108.
- Liu Fuzhu 刘福铸 (2007). ‘Púxiān Fāngyán Biānchāyīn [4] Shēngmǔ Tànyuán 莆仙方言边擦音 [4] 声母探源 [Origin of the initial [4] of the lateral fricative in the Puxian dialect]’, *Journal of Putian University* 莆田学院学报 14(3): 93–8.

- Liu Jian and Alain Peyraube (1994). ‘History of some coordinative conjunctions in Chinese’, *Journal of Chinese Linguistics* 22–2: 179–201.
- Liu Lihua 刘丽华 (2001). *Loudi fangyan yanjiu* 娄底方言研究 [Research on the Loudi dialect]. Changsha: Central South University Press.
- Liu Shiru 刘世儒 (1965). *Lun Weijin Nanbeichao de liangci* 论魏晋南北朝的量词 [On Classifiers of the Wei Jin Nanbei Dynasties]. Beijing: Zhonghua Shuju.
- Liu Tanzhou 刘探宙 and Shi, Dingxu 石定栩 (2012). ‘Yantaihua-zhong budai zhishici huo shuci de liangci jiegou 烟台话中不带指示词或数词的量词结构 [Bare classifier phrases in the Yantai dialect]’, *Zhongguo Yuwen* 中國語文 1: 38–49.
- Liu Xinzhong 刘新中 (2006). *Hainán Mínyǔ de Yúyīn Yánjiū* 海南闽语的语音研究 [Studies of the Phonology of Hainan Min]. Beijing: China Social Science Press 中国社会科学出版社.
- Long, Yaohung 龙耀宏 (2003). *Dòngyǔ Yánjiū* 侗语研究 [Studies on Kam]. Guiyang: Guizhou Nationalities Publishing House 贵州民族出版社.
- Lord, Carol (1993). *Historical Change in Serial Verb Constructions*. Amsterdam/Philadelphia: John Benjamins.
- Lord, Carol, Foong Ha Yap, and Shoichi Iwasaki (2002). ‘Grammaticalization of give: African and Asian perspectives’, in Ilse Wischer and Gabrielle Diewald (eds), *New Reflections on Grammaticalization*. Amsterdam: John Benjamins, 217–35.
- Lu Guangcheng 卢广诚 (2003). *Taiwan Minnanyu Gaiyao* 台湾闽南语概要 [Outline of Taiwanese Southern Min]. Taipei: Nantian Shuju (Southern Materials Center).
- Luo Liming 罗黎明, Yaowu Qin 覃耀武, Zhenyu Lu 陆振宇, and Fulung Chen 陈福隆 et al. (eds) (2005). *Zhuàng Hán Yǐng Cídiǎn* 壮汉英词典 *Cuengh Gun Yingh Swzdenj Zhuang-Chinese–English Dictionary*. Beijing: The Ethnic Publishing House 民族出版社.
- Luo Xinru 罗昕如 (1998). *Xinhua fangyan yanjiu*. 新化方言研究 [Research on the Xinhua Dialect]. Changsha: Hunan Education Publishing House.
- Lyons, John (1977). *Semantics*. Cambridge: Cambridge University Press.
- Mai Yun 麦耘 (2010). ‘Yuèyǔ de Xíngchéng, fāzhǎn yǔ Yuèyǔ hé Pínghuà de Guānxì 粤语的形成、发展与粤语和平话的关系 [The formation and development of Yue, and the relationship between Yue and Pinghua]’, in Wùyún Pán 潘悟云 and Shěn Zhōngwěi 沈钟伟 (eds), *Yánjiū zhī Lè—Qīngzhú Wáng Shiyuán Xiānshēng Qīshíwǔ Shòuchén Xuéshù Lùnwénjí* 研究之乐庆祝王士元先生 七十五寿辰学术论文集 [The Joy of Research II—A Festschrift in Honor of Professor William S.-Y. Wang on His Seventy-fifth Birthday]. Shanghai: Shanghai Education Publishing House 上海教育出版社, 227–43.
- Malchukov, Andrej, Martin Haspelmath, and Bernard Comrie (eds) (2010). *Studies in Ditransitive Constructions*. Berlin: De Gruyter.
- Mao Zongwu 毛宗武 (2004). *Yáozú Miěn Fāngyán Yánjiū* 瑶族勉语方言研究 [Studies of the Mien Dialects of Yao Ethnicity]. Beijing: The Ethnic Publishing House 民族出版社.
- Mao Zongwu 毛宗武 et al. (1982). *Yaozu yuyan jianzhi* 瑶族语言简志 [Sketch of the Yao language]. Beijing: Minzu chubanshe.
- Matisoff, James. (1991). ‘Areal and universal dimensions of grammaticalization in Lahu’, in E.C. Traugott and B. Heine (eds). *Approaches to Grammaticalization. Volume 2: Focus on Types of Grammatical Markers*. Amsterdam: John Benjamins, 383–454.
- Matisoff, James A. (2001). ‘Genetic versus contact relationship: Prosodic diffusibility in South-East Asian languages’, in Alexandra Y. Aikhenvald and R.M.W. Dixon (eds), 291–327.

290 *References*

- Matthews, Stephen and Virginia Yip (2001). ‘Aspects of contemporary Cantonese grammar: The structure and stratification of relative clauses’, in H. Chappell (ed.), *Sinitic Grammar: Synchronic and Diachronic Perspectives*. Oxford: Oxford University Press, 266–81.
- Matthews, Stephen and Virginia Yip (1994). *Cantonese: A Comprehensive Grammar*. London: Routledge.
- McGregor, William B. (2002). *Verb Classification in Australian Languages*. Berlin and New York: Mouton de Gruyter.
- McMahon, April (1994). *Understanding Language Change*. Cambridge: Cambridge University Press.
- Meecham, William (1993). ‘On the ethno-linguistic identity of the Ancient Yue’, paper presented at the Symposium ‘Evolution of the Yue population and languages: Qualitative and quantitative approaches’. City Polytechnic of Hong Kong, December 16.
- Mei Tsu-lin 梅祖麟 (1988). ‘Beifang fangyan zhong diyi rencheng daici fushu baokuoshi he paichushi duili de laiyuan 北方言中第一人称代词复数包括式和排除式对立的来源 [Origin of the opposition between inclusive and exclusive forms of the 1st plural personal pronouns in Northern Chinese]’, *Yuyanxue luncong 语言学论丛* 15: 141–5.
- Meillet, Antoine (1912). ‘L’évolution des formes grammaticales’, in Antoine Meillet, *Linguistique historique et linguistique générale*. Paris: Champion, 1948: 130–48.
- Meng Qinghui 孟庆惠 (1981). ‘Huángshānhuà de t₁, t₂, ɿ jí Tànyuán 黄山话的 t₁, t₂, ɿ 及探源 [The t₁, t₂, ɿ in the Huángshān Dialect and Investigating its Source]’, *Zhōngguó Yǔwén 中国语文* 160 (1): 46–9.
- Mok Sui-Sang (1998). ‘Cantonese exceed comparatives’, PhD dissertation, University of California, San Diego.
- Moravcsik, Edith (1978). ‘Language contact’, in Joseph Greenberg, Charles Ferguson, and Edith Moravcsik (eds), *Universals of Human Language*. Stanford: Stanford University Press, 93–123.
- Mufwene, Salikoko S. (2001). *The Ecology of Language Evolution* (Cambridge Approaches to Language Contact). Cambridge: Cambridge University Press.
- Nedjalkov, Igor V. (1993). ‘Causative-Passive polysemy of the Manchu-Tungusic -bul-V(U)’, *Linguistica Antverpiensia* 27: 193–202.
- Newman, John (1996). *Give: A Cognitive Linguistic Study*. Berlin: Mouton de Gruyter.
- Newmeyer, Fritz (1998). *Language Form and Function*. Cambridge, MA: MIT Press.
- Nguyễn, Đinh-Hoà (1997). *Vietnamese—Tiếng Việt Không Son Phấn*. Amsterdam/ Philadelphia: John Benjamins.
- Norde, Muriel (2002). ‘The final stages of grammaticalization: Affixhood and beyond’, in Ilse Wischer and Gabriele Diewald (eds), *New Reflections on Grammaticalization*. Amsterdam: John Benjamins, 45–66.
- Norman, Jerry (1982). ‘Four Notes on Chinese–Altaic Linguistic Contacts’, *Tsinghua Journal of Chinese Studies* 14 (1–2): 243–7.
- Norman, Jerry (1988). *Chinese*. Cambridge: Cambridge University Press.
- Ostapirat, Weera (2000). ‘Proto-Kra’, *Linguistics of the Tibeto-Burman Area* 23(1): 1–251.
- Ostapirat, Weera (2004). ‘Proto-Hlai sound system and lexicons’, *Language and Linguistics Monograph Series W-4: 121–75.*

- Ostapirat, Weera (2008). ‘The Hlai language’, in Anthony V. N. Diller, Jerold A. Edmondson, and Yongxian Luo (eds), *The Tai-Kadai Languages*. Abingdon, Oxfordshire/New York: Routledge, 623–52.
- Ōta Tatsuo 太田辰夫 (1958). *Chūngokugo Rekishi Bumpō* 中国语历史文法 [An historical grammar of the Chinese language]. Tokyo: 江南书院 (Translated by Jiang Shaoyu and Xu Changhua into Chinese, 2003: 蒋绍愚, 徐昌华 Beijing : Beijing University Publishers.)
- Ouyang Jueya 欧阳觉亚 (1989). ‘Hànyǔ Yuèfāngyán lì de Gǔ Yuèyǔ Chéngfèn 汉语粤方言里的古越语成分 [Elements of the ancient Yuè language in the Yuè dialect of Chinese]’, in Shǔxiāng L 吕叔湘 (ed.), *Yǔyán Wénzì Xuéshù Lùnwénjí—Qīngzhú Wáng Lì Xiānshēng Xuéshù Huódòng Wǔshí Zhōunián* 语言文字学术论文集—庆祝王力先生学术活动五十周年 [Academic Proceedings of Language and Writing—Celebrating the Fiftieth Anniversary of Professor Wang Li's Academic Activities]. Beijing: Knowledge Publishing House 知识出版社.
- Pan Wuyun 潘伍云(2000). *Hanyu Lishi Yinyunxue* 汉语历史音韵学 Shanghai: Shanghai Jiaoyu.
- Pan Wuyun 潘悟云 and Tao Huan 陶寰 (1999). ‘Wuyu de zhidaici 吴语的指代词 [Pronouns in Wu dialects]’, in Li Rulong and Zhang Shuangqing 李如龙, 张双庆 *Daici* 代词 [Pronouns]. Guangzhou: Jinan University Press, 25–67.
- Paul, Waltraud (2005). ‘Low IP area and left periphery in Mandarin Chinese’, *Recherches Linguistiques de Vincennes* 33: 111–34.
- Paul, Waltraud and John Whitman (2008). ‘*Shi... de* focus clefts in Mandarin Chinese’, *The Linguistic Review* 25(3-4): 413–51.
- Peyraube, Alain (1981). ‘The dative construction in Cantonese’, *Computational Analysis of Asian and African Languages* 16: 29–65.
- Peyraube, Alain 贝罗贝(1986). ‘Shuang binyu jiegou—cong Han dai zhi Tang dai de lishi fazhan 双宾语结构 从汉代至唐代的历史发展 [Double-object constructions—from the Han dynasty to the Tang dynasty]’, *Zhongguo Yuwen* 中国语文 5, 204–16.
- Peyraube, Alain (1987). ‘The double-object construction in *Lunyu* and *Mengzi*’, in *Wang Li Memorial Volumes*. Hong Kong: Joint Publishing, 331–58.
- Peyraube, Alain (1988). *Syntaxe diachronique du chinois: évolution des constructions datives du 14^{ème} siècle av. J.-C. au 18^{ème} siècle*. Paris: Collège de France.
- Peyraube, Alain (1989a). ‘History of the passive construction in Chinese until the 10th century’, *Journal of Chinese Linguistics* 17.2: 335–72.
- Peyraube, Alain (1989b). ‘History of the comparative constructions in Chinese from the 5th century BC to the 14th century AD’, in *Proceedings of the 2nd International Conference on Sinology*. Taiwan, Academia Sinica, 589–612.
- Peyraube, Alain (1991). ‘Some remarks on the history of Chinese classifiers’, in Patricia M. Clancy and Sandra A. Thompson (eds), *Asian Discourse and Grammar. Santa Barbara Papers in Linguistics* 3: 106–26.
- Peyraube, Alain (1997). ‘Cantonese post-verbal adverbs’, in Anne Yue and Mitsuaki Endo (eds), *In Memory of Mantaro J. Hashimoto*. Tokyo: Neishan shudian, 303–13.
- Peyraube, Alain 贝罗贝 (1998). ‘Shanggu zhonggu hanyu liangci de lishi fazhan 上古汉语量詞的歷史發展 [Historical development of classifiers in Archaic and Medieval Chinese]’, *Yuyanxue Luncong* 21: 99–122.

292 *References*

- Peyraube, Alain 贝罗贝(2005). ‘Leitui, yufahua, qu yufahua yu gongneng gengxin 类推，语法化，去语法化与功能更新 [Analogy, grammaticalization, degrammaticalization, and exaptation]’, paper presented at the third International Conference on Grammaticalization, Luoyang, China. October 27–28 [in Chinese]
- Peyraube, Alain (2007). ‘Languages and genes in China and in East Asia’, *Bulletin of Chinese Linguistics* 2.1: 1–15.
- Peyraube, Alain (2013). ‘Syntactico-semantic change in Chinese: Processes of analogy, reanalysis, external borrowing’, paper presented at the 5th International Conference in Evolutionary Linguistics, Hong Kong, Chinese University of Hong Kong. August 17–19.
- Peyraube, Alain (forthcoming). ‘Can the Northwestern Sinitic languages be analyzed as having cases?’, in Cao Guangshun, Redouane Djamouri, and Alain Peyraube (eds), *Collection des Cahiers de linguistique Asie orientale*. Paris: EHESS.
- Peyraube, Alain and Liu Huali (2013). ‘Hanyu dengdai yi dongci de lishi kaocha 汉语等待义动词的历史考察 [Historical analysis of waiting verbs in Chinese]’, *Lishi yuyanxue yanjiu 历史语言学研究 [Historical Linguistics]* 6: 23–39.
- Peyraube, Alain and Liejiong Xu (1997). ‘On the double-object construction and the oblique construction in Cantonese’, *Studies in Language* 21(1): 105–27.
- Phua, Chiew Pheng 潘秋平 (2007). ‘Yuèfāngyán Géiyūyì Shuāngbīnyǔ Jiégòu de Láiyuán 粤方言给予义双宾语结构的来源 [The origin of giving type ditransitive constructions in the Yuè dialect]’, in Hung-nin Samuel Cheung 张洪年, Song Hing Chang 张双庆, and Hung Kan Chan 陈雄根 (eds), *Dì Shí Jiè Guójì Yuè Fāngyán Yántǎohuì Lùnwénjí 第十届国际粤方言研讨会论文集 [Proceedings of The Tenth International Conference on Yue Dialects]*. Beijing: China Social Science Press 中国社会科学出版社, 214–29.
- Piñol y Andreu, Francisco (1928). *Gramática China del dialecto de Amoy*. Hong Kong: Imprenta de Nazaret.
- Pittayaporn, Pittayawat (2009). ‘The phonology of Proto-Tai’, PhD dissertation, Cornell University.
- Pulleyblank, Edwin G. (1984). *Middle Chinese: A Study in Historical Phonology*. Vancouver: UBC Press.
- Qian Dianxiang 钱奠香 (2002). *Hainan Tunchang Minyu yufa yanjiu*. 海南屯昌闽语语法研究 [Research on the Grammar of the Tunchang Dialect in Hainan Province]. Haikou: Hainan People's Press.
- Qian Nairong 钱乃荣 (1997). *Shanghaihua yufa* 上海话语法 [The Grammar of the Shanghai dialect]. Shanghai: Shanghai People's Publishing House.
- Qian Nairong 钱乃荣 (1999). ‘Beibu wuyu de daici xitong 北部吴语的代词系统 [The pronoun system in Northern Wu]’, in Li, Rulong, and Zhang Shuangqing (eds), 李如龙, 张双庆 主编 *Daici* 代词 [Pronouns]. Guangzhou: Jinan University Press, 68–84.
- Qian Zengyi 钱曾怡 (1983). *Shengxian Changlehua san ze* 嵊县长乐话语法三则 [Three notes on the grammar of the Changle dialect in Shengxian], in Fudan Daxue Zhongguo Yuyan Wenxue Yanjiusuo Wuyu Yanjiushi (eds) 复旦大学中国语言文学研究所吴语研究室, *Wuyu Luncong* 吴语论丛 [Anthology on the Wu Dialects]. Shanghai: Shanghai Jiaoyu Chubanshe, 297–303.
- Qian Zengyi 钱曾怡 and Luo Futeng 罗福腾 (1992). *Weifang Fangyanzhi* 潍坊方言志 [Record of the Weifang Dialect]. Weifang: Weifangshi Xinwen Chubanju.

- Qian Zengyi 钱曾怡等, Zhang Shuzheng 张树铮, and Luo Futeng 罗福腾 (2001). *Shandong fangyan yanjiu* 山东方言研究 [Studies on the Shandong Dialects]. Jinan: Jilu Shushe.
- Qin Dongsheng 覃东生 (2012). 'Duì Guǎngxī Sān'gè Qūyùxìng Yǔfǎ Xiànxìàng de Kǎochá 对广西三个区域性语法现象的考察 [On three areal grammatical features in Guangxi]', PhD thesis, Hebei Normal University.
- Qin Fengyu 覃凤余 and Fúxiáng Wáng 吴福祥 (2009). 'Nánníng Bái hùa "Guò" de Liǎng-hǒng Tèshū Yòngfǎ 南宁白话“过”的两种特殊用法 [Two special uses of "Guò" in Nánníng Cantonese]', *Minzu Yuwen* 民族语文 2009 (3): 16–29.
- Qin Xiaohang 覃晓航 (1995). *Zhuàngyǔ Tèshū Yǔfǎ Xiànxìàng Yánjiū* 壮语特殊语法现象研究 [Studies of Special Syntactic Phenomena in Zhuang]. Beijing: The Ethnic Publishing House 民族出版社.
- Qin Xiaohang (2004). *Zhuang Lexicology: A Study on Zhuang Words and an Attached Zhuang-English Vocabulary*. Beijing: The Ethnic Publishing House.
- Qin, Yuanxiong 覃遠雄, Shùguān Wěi 韋樹關, and Chéglín Biàn 卞成林 (1997). *Nánníng Pínghuà Cídǎn* 南寧平話詞典 [Dictionary of Nánníng Pínghuà]. Nanjing: Jiangsu Education Publishing House 江蘇教育出版社.
- Ramat, Paolo (2001). 'Degrammaticalization or transcategorialization?', in Chris Schaner-Wolles, John Rennison, and Friedrich Neubarth (eds), *Naturally! Linguistic Studies in Honour of Wolfgang Ulrich Dressler Presented on the Occasion of his 60th Birthday*. Torino: Rosenbach and Sellier, 393–401.
- Saxena, Anju (1988). 'The case of the verb *say* in Tibeto-Burman', *Proceedings of the Berkeley Linguistic Society* 14: 375–88.
- Shao Jingmin 邵敬敏 (1993). 'Liangci de yuyi fenxi ji qi yu mingci de shuangxiang xuanze 量词的语义分析及其与名词的双项选择 [The semantic analysis of classifiers and the bilateral selection between the classifier and the noun]', *Zhongguo Yuwen* 中国语文 3:181–8.
- Shen Pei 沈培 (1992). *Yinxu jiagu buci yuxu yanjiu* 殷墟甲骨卜辭語序研究 [Word order in the oracle bone inscriptions of the Yin]. Taipei: Wenjin Chubanshe.
- Shi Ruijie 石汝杰 (1999). 'Suzhou fangyan de daici xitong 苏州方言的代词系统 [The demonstrative system in the Suzhou dialect]', in Li and Zhang (eds), 85–101.
- Shi Ruijie 石汝杰 and Liu Danqing 刘丹青 (1985). 'Suzhou fangyan de liangci dingzhi yongfa ji qi biandiao 苏州方言的量词定指用法及其变调 [Definite classifiers and their tone sandhi in the Suzhou dialect]', *Yuyan Yanjiu* 1: 160–6.
- Siewierska, Anna (1984). *The Passive: A Comparative Linguistic Analysis*. London: Croom Helm.
- Siewierska, Anna (2004). *Person*. Cambridge: Cambridge University Press.
- Siewierska, Anna and Dik Bakker (2013). 'Passive agents: Prototypical vs. canonical passives', in Dunstan Brown, Marina Chumakina, and Greville G. Corbett (eds), *Canonical Morphology and Syntax*. Oxford: Oxford University Press, 151–89.
- Sio, Joanna Ut-seong 蕭月娟 and Rint Sybesma 司馬翎 (2008). 'The nominal phrase in Northern Zhuang—a descriptive study', *Bulletin of Chinese Linguistics* 中國語言學集刊 3(1): 175–225.
- Song, Jae Jung (2001). *Linguistic Typology: Morphology and Syntax*. Harlow, UK: Pearson Education.
- Stassen, Leon (1985). *Comparison and Universal Grammar*. Oxford: Blackwell.

294 *References*

- Stassen, Leon (2005). ‘Comparative constructions’, in Martin Haspelmath, Matthew S. Dryer, David Gil, and Bernard Comrie (eds.), *The World Atlas of Language Structures*. Oxford: Oxford University Press (Map and Description 121).
- Stassen, Leon (2011). ‘Comparative constructions’, in Matthew S. Dryer and Martin Haspelmath (eds), *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library, feature 121A.
- Stolz, Thomas (1999). ‘Comitatives vs. instrumentals vs. agents’, in Walter Bisang (ed.), *Aspects of Typology and Universals*, (Studia Typologica 1). Berlin: Akademie Verlag, 153–74.
- Sun, Chaofen (1989). ‘The discourse function of numeral classifiers in Mandarin Chinese’, *Journal of Chinese Linguistics* 16(2): 298–322.
- Sun, Chaofen (1996). *Word Order Change and Grammaticalization in the History of Chinese*. Stanford: Stanford University Press.
- Sybesma, Rint (2008). ‘Zhuang: A Tai language with some Sinitic characteristics—Post-Verbal ‘can’ in Zhuang, Cantonese, Vietnamese and Lao’, in Pieter Muysken (ed.), *From Linguistic Areas to Areal Linguistics*. Amsterdam/Philadelphia: John Benjamins, 221–74.
- Tai, James H-Y. (1994). ‘Chinese classifier systems and human categorization’, in Matthew Chen and Ovid Tseng (eds), *In Honor of Professor William S-Y. Wang: Interdisciplinary Studies on Language and Language Change*. Taipei: Pyramid Publishing Company, 479–94.
- Tang, Aihua 唐爱华 (2000). ‘Anhui Susong fangyan de zhishi daici 安徽宿松方言的指示代词 [The demonstrative system of the Susong dialect in Anhui province]’, *Journal of Anqing Teachers College (Social Science Edition)* 19.2: 53–6.
- Tang, Aihua 唐爱华 (2005). *Susong fangyan yanjiu 宿松方言研究 [Research on the Susong Dialect]*. Beijing: China Social Sciences Publishing House.
- Thepkajana, Kingkarn (2010), ‘Ditransitive construction in Thai’, in Andrej Malchukov, Martin Haspelmath, and Bernard Comrie (eds), *Studies in Ditransitive Constructions: A Comparative Handbook*. Berlin; New York: de Gruyter Mouton, 409–26.
- Thomason, Sarah G. (2001a). *Language Contact: An Introduction*. Edinburgh: Edinburgh University Press.
- Thomason, Sarah G. (2001b). ‘Speakers’ attitudes in language change, contact-language genesis and language preservation’, *Estudios de Sociolinguística* 2(2): 13–26.
- Thomason, Sarah G. (2010). ‘Contact explanations in linguistics’, in Raymond Hickey (ed.), *The Handbook of Language Contact*. West Sussex: Wiley-Blackwell, 31–47.
- Thomason, Sarah G. and Terrence Kaufman (1988). *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.
- Traugott, Elizabeth C. (1989). ‘On the rise of epistemic meanings in English: An example of subjectification in semantic change’, *Language* 65–1: 31–55.
- Traugott, Elizabeth C. (2001). ‘Legitimate counter-examples to unidirectionality’, paper presented at Freiburg University, October 17.
- Traugott, Elizabeth C. (2004). ‘Exaptation and grammaticalization’, in Minoji Akimoto (ed.), *Linguistic Studies Based on Corpora*. Tokyo: Hituzi Syobo Publishing Co. 133–56.
- Ultan, Russell (1972). ‘Some features of basic comparative constructions’, *Working Papers on Language Universals* 9: 117–62.

- Umbach, Carla (2001). ‘Contrast and contrastive topic’, in Ivana Kruijff-Korbayová and Mark Steedman (eds), *Proceedings of the ESSLLI 2001 Workshop on Information Structure, Discourse Structure and Discourse Semantics*, Helsinki, 175–88.
- Van der Loon, Piet (1967). ‘The Manila incunabula and early Hokkien Studies’, *Asia Major-New Series*, Part II: 95–186.
- Wang Fushi 王辅世 (1985). *Miaoyu jianzhi* 苗语简志 [Sketch of the Miao Language]. Beijing: Minzu Chubanshe.
- Wang Huayun 汪化云 (2005). *E'dong fangyan yanjiu*. 鄂东方言研究 [Research on the E'dong Dialect]. Chengdu: Bashu Press.
- Wang Huayun 汪化云 (2008). *Hanyu Fangyan Daici Lunlüe* 汉语方言代词论略 [An Introduction to Pronouns in Sinitic languages]. Sichuan: Bashu Press.
- Wang Jian (this volume). ‘Bare classifier phrases in Sinitic languages: A typological perspective’.
- Wang Jian and Hilary Chappell (2012). ‘A Typological Study on Interrogative Words for Person in Sinitic languages’, paper presented at the 45th International Conference on Sino-Tibetan Languages and Linguistics, Singapore: Nanyang Technological University, October 2012.
- Wang Jianshe and Ganli Zhang 王建设, 张甘荔 (1994). *Quanzhou Fangyan yu Wenhua* 泉州方言與文化 [The Dialect and Culture of Quanzhou]. Xiamen: Lujiang Press.
- Wang Jun 王均 (1962). ‘Zhuàngyǔ zhōng de Hán yǔ Jiècí 壮语中的汉语介词 [Chinese loanwords in Zhuang]’, *Zhōngguó Yǔwén* 中国语文 1962 (6): 251–64.
- Wang Jun 王均, Guoqiao Zheng 郑国乔, and Jing Yuè 岳静 (2009). ‘Mùlǎoyǔ Jiǎnzhì 仫佬语简志 [Sketch of Mulam]’, in Xiāoyù Zēng 曾晓渝 and Xùliàn Lǐ 李旭练 (eds), *Zhōngguó Shǎoshù Mínzú Yǔyán Jiǎnzhì Cóngshū* 中国少数民族语言简志从书 Volume 3. Beijing: The Ethnic Publishing House 民族出版社, 551–621.
- Wang Li 王力 (1937). ‘Zhongguo wenfa zhong de xici 中国语法的系词 [The copula in Chinese grammar]’, *Journal of Tsinghua University*. Vol 12(1).
- Wang Li 王力 (1958). *Hanyu shigao* 汉语史稿 [A History of the Chinese Language]. Beijing: Kexue Chubanshe [New edition, 2009. Beijing: Zhonghua Shuju 中华书局].
- Wang Lianqing (1994). ‘Origin and development of classifiers in Chinese’, PhD dissertation, The Ohio State University, Columbus.
- Wang Rudong 汪如东 (2012). ‘Hai'an hua zhong biao dingzhi yi liangming jiegou de biandiao ji biaoqing gongneng 海安话中表定指义量名结构的变调及表情功能 [Tone sandhi and the pragmatic functions of definite bare classifier phrases in the Hai'an dialect]’, *Dongfang Yuyanxue* 11: 93–105.
- Wang Sen 王森 (1996). ‘Linxia fangyan qianzhi juzi chengfen de biaoshifa 临夏方言前置句子成分的表示法 [Ways to express the preposed sentence elements in the Linxia dialect]’, in Huang Borong (ed.), *Hanyu fangyan yufa leibian* 汉语方言语法类编 [A Concordance on the Grammar of Sinitic Languages]. Qingdao: Qingdao Chubanshe, 472–3.
- Wang S.-Y. William (1969). ‘Competing Changes as a Cause of Residue’, *Language* 45: 9–25.
- Wang Weihui 汪维辉 (2003). ‘Hanyu shuo lei ci de lishi yanbian yu gongshi fenbu 汉语说类词的历史演变与共时分布 [Diachronic evolution and synchronic distribution of Chinese say verbs]’, *Zhongguo Yuwen* 中国语文 4: 329–42.

296 *References*

- Wang Wenqing (王文卿) (2002). *Jinyuan Fangyan Yanjiu* 晋源方言研究 [Studies on the Jinyuan Dialect]. Beijing: Yuwen Chubanshe.
- Wang Yinzhī 王引之 (1797/2000). *Jingyi shuwen* 经义述闻 [An Account of Old Scripts]. Nanjing: Jiangsu Guji Chubanshe.
- Wei, Jingyun 韦景云 and Xiaohang Qin 覃晓航 (2006). *Zhuangyu Tonglun* 壮语通论 [General Study of Zhuang]. Beijing: Central University for Nationalities Press 中央民族大学出版社.
- Wei Pei-chuan (2007). ‘Guanyu chabiju fazhan guocheng de ji dian kanfa 关于差比句发展过程的几点看法 [Some views on the developmental process of comparatives of inequality]’, *Language and Linguistics* 8.2 : 603–37.
- Weinreich, Uriel (1953). *Languages in Contact: Findings and Problems*. The Hague: Mouton.
- Wen Bo, Hui Li, Daru Lu, Xiufeng Song, Feng Zhang, Yungang He, Feng Li, Yang Gao, Xianyun Mao, Liang Zhang, Ji Qian, Jingze Tan, Jianzhong Jin, Wei Huang, Ranjan Deka, Bing Su, Ranajit Chakraborty, and Li Jin (2004). ‘Genetic evidence supports demic diffusion of Han culture’. *Nature* 431: 302–5.
- Wiemer, Björn (2011). ‘The grammaticalization of passives’, in Bernd Heine and Heiko Narrog (eds), *A Handbook of Grammaticalization*. Oxford: Oxford University Press, 535–46.
- Wu Fuxiang 吴福祥 (2010). ‘Yueyu chabishi “X + A + guo + Y” leixingxue diwei 粤语差比式‘X+A+过+Y’类型学地位 [The typological position of the comparative “X + A + guo + Y” in the Yue dialects]’. *Zhongguo Yuwen* 中國語文 3: 238–55.
- Wu Fuxiang 吴福祥 (2012). ‘Minyu Chabishi de Lishi Cengci’ 閩語差比式的歷史層次 [The historical stratum of comparative constructions in the Min dialects], *Liyun Xuekan* 励耘學刊 1.
- Wu Fuxiang 吴福祥 (2013). ‘Nanfang minzu yuyan bijiaoju yuxu de yanbian he bianyi 南方民族语言的语法特点 [Special features of southern minority languages in China]’, in Cao Guangshun, Hilary Chappell, Redouane Djemouri, and Thekla Wiebusch (eds), *Breaking Down the Barriers: Interdisciplinary Studies in Chinese Linguistics and Beyond*. 2 vols. Taipei: Academia Sinica, 831–64.
- Wu Hezhong 伍和忠 (1998). ‘Lipu fangyan de yufa tedian. 荔浦方言的语法特点 [The grammatical characteristics of the Lipu dialect]’, *Guangxi shiyuan xuebao (Zhe-shi ban)* 广西师院学报(哲史版) 1.69–75.
- Wu Yunji 伍云姬 (1995). ‘谈雌雄动物名称的演变 [The development of gender markers for animal terms]’ in Shi Feng 石峰(ed.), *Hanyu yanjiu zai hawai* 汉语研究在海外 [Chinese Linguistics Overseas]. Beijing: Beijing Yuyan Xueyuan Publishers, 111–22.
- Wu Yunji (2005). *A Synchronic and Diachronic Study of the Grammar of the Chinese Xiang Dialects* (Trends in Linguistics 162). Berlin: Mouton de Gruyter.
- Wu Yunji (2013). ‘Hunan Nanbu chuzhi he beidong tong biaoji xianxiang de zai sikao 湖南南部處置和被動同標記現象的再思考 [Further reflections on the phenomenon of identical marking for the disposal and the passive in Southern Hunan]’, in Cao Guangshun, H. Chappell, R. Djemouri, and Thekla Wiebusch (eds), *Breaking Down the Barriers: Interdisciplinary Studies in Chinese Linguistics and Beyond*. 2 vols. Taipei: Academia Sinica, 995–1010.
- Wu Zhongping 吴仲平 (1958). ‘Xiamenhua de yufa tedian 厦門話的語法特點 [Grammatical characteristics of the Xiamen dialect]’, *Fangyan yu Putonghua Jikan* 方言與普通話集刊 [A

- Collection of Papers on Dialects and Putonghua]* Vol. 1. Peking: Language Reform Publishing House, 84–103.
- Wurm, Stephen, Rong Li, Theo Baumann, and Mei W. Lee (eds) (1987). *Language Atlas of China / Zhongguo Yuyan Dituji* 中国语言地图集. Hong Kong: Longman.
- Xiang Mengbing 项梦冰 (1997). *Liancheng Kejiahua yufa yanjiu* 连城客家语法研究 [Research on the Grammar of Liancheng Hakka]. Beijing: Language and Culture Press.
- Xiang Mengbing 项梦冰 (1999). ‘Qingliu fangyan de daici xitong 清流方言的代词系统 [The pronominal system in the Qingliu dialect]’, in Li and Zhang (eds), 201–32.
- Xiao, Richard, Tony McEnery, and Yufang Qian (2006). ‘Passive constructions in English and Chinese: A corpus-based contrastive study’, *Languages in Contrast* 6.1: 109–49.
- Xiaochuan Huanshu 小川环树 (1981). ‘Suzhou fangyan de zhishi daici 苏州方言的指示代词 [Demonstratives in the Suzhou dialect]’, *Fangyan* 4: 287–8.
- Xie Jianyou 谢建猷 (2007). *Guǎngxī Hánnyǔ Fāngyán Yánjiū* 广西汉语方言研究 [Studies of Chinese Dialects in Guǎngxī]. Nanning: Guangxi People’s Publishing House 广西人民出版社.
- Xing Xiangdong 邢向东 (2002). *Shenmu fangyan yanjiu*. 神木方言研究 [Research on the Shenmu Dialect]. Beijing: Zhonghua Book Company.
- Xiong Zhenghui 熊证辉 and Zhang Zhenxing 张振兴 (2008). 汉语方言的分区. *Hanyu fangyan de fenqu* [The distribution of Chinese dialects] 方言 *Fangyan* 2: 97–108.
- Xu Baohua (許寶華) and Tao Huan (陶寰) (1999). ‘Wuyu de Chuzhiju 吳語的處置句 [Disposal constructions in the Wu dialects]’, in Yunji Wu (伍云姬) (ed.), *Hanyu Fangyan Gongshi yu Lishi Yufa Yantao Lunwenji* 汉语方言共时与历时语法研讨论文集 [Anthology on the Synchronic and Diachronic Grammar of Chinese Dialects]. Guangzhou: Jinan Daxue Chubanshe, 135–67.
- Xu Baohua and Miyata Ichiroo (eds) 许宝华, 宫田一郎 (1999). *Hanyu fangyan da cidian* 汉语方言大词典 [A Comprehensive Dictionary of Chinese Dialects]. Beijing: Zhonghua Shuzhu.
- Xu Baohua and Zhenzhu Tang 许宝华, 汤珍珠 (1988). *Shanghai Shiqu Fangyanzhi* 上海市 区方言志 [A Record of the Dialect of Shanghai City]. Hong Kong: Joint Publishers.
- Xu Huiling (2007). *Aspects of Chaoshan Grammar: A Synchronic Description of the Jieyang Dialect*. Monograph series of *Journal of Chinese Linguistics*. Berkeley: University of California.
- Xu Huiling and Stephen Matthews 许惠玲, 马诗帆 (2007). ‘Cong dongci dao ziju jiegou biaoji: Chaozhou fangyan he Taiwan Minnan hua dongci ‘shuo’ he ‘kan’ de xuhua guocheng 從動詞到子句標記: 潮州方言和台灣閩南話動詞‘說’和‘看’的虛化過程 [From verb to complementizer: The grammaticalization process for the verbs ‘say’ and ‘see’ in the Chaozhou dialect and Taiwanese Southern Min]’. *Zhongguo Yuwen* 中國語文 *Yanjiu* 中國語文研究 [Studies in Chinese Linguistics] 23: 61–72.
- Xu Huiling and Stephen Matthews (2011). ‘On the polyfunctionality and grammaticalization of the morpheme *kai* in the Chaozhou dialect’, in Foong Ha Yap, Karen Grunow-Härsta, and Janick Wrona (eds), *Nominalization in Asian Languages: Diachronic and Typological Perspectives*. Amsterdam: John Benjamins, 109–24.
- Xu Jieshun 徐杰瞬 (1999). *Xuéqíú—Hàn Mínzú de Rénlèixué Fēnxī* 雪球—汉民族的人类学分析 [Snowball—Anthropological Analysis of the Han Ethnicity]. Shanghai: Shanghai People’s Publishing House 上海人民出版社.

- Xu Liejiong (2004). ‘The manifestation of informational focus’, *Lingua* 114: 277–99.
- Xu Liejiong and Liu Danqing 徐烈炯，刘丹青(1998). *Huati de jiegou yu gongneng* 话题的结构与功能 [*The Structure and Function of Topics*] Shanghai: Shanghai Education Press.
- Xu Liejiong and Alain Peyraube (1997). ‘On the double-object construction and the oblique construction in Cantonese’, *Studies in Language* 21-1: 105–27.
- Xu Qulian 许秋莲 (2007). ‘Hengdong Xintang fangyan liangming jiegou yanjiu. 衡东新塘方言量名结构研究 [Research on bare classifier phrases in the Hengdong Xintang dialect]’, MA thesis, Changsha: Hunan Normal University.
- Yang Jianqiao 杨剑桥 (1998). ‘Wuyu zhishici+liangci de shenglüe shi 吴语“指示词+量词”的省略式 [The elided form of the phrase ‘demonstrative+classifier’ in Wu dialects]’, *Zhongguo Yuwen* 中國語文 4: 286.
- Yang Xiufang 楊秀芳 (1991). *Taiwan Minnanyu Yufagao* 臺灣閩南語語法稿 [*Taiwanese Southern Min Grammar*]. Taipei: Da'an Press.
- Yang-Drocourt, Zhitang (1993). ‘Analyse syntaxique des expressions quantitatives en chinois archaïque’, *Cahiers de Linguistique—Asie Orientale* 22.2: 217–37.
- Yang-Drocourt, Zhitang (2004). *Évolution syntaxique du classificateur en chinois du XIIIe siècle av. J.-C. au XVIIe siècle* [*The Syntactic Evolution of the Chinese Classifier from the Thirteen Century BC to the-Seventeenth Century*]. Paris : École des Hautes Études en Sciences Sociales, Centres de Recherches Linguistiques sur l'Asie Orientale (Collections des Cahiers de Linguistique Asie Orientale, no. 7).
- Yap, Foong Ha and Shoichi Iwasaki (2003). ‘From causative to passive: A passage in some East and Southeast Asian Languages’, in Eugene Casad and Gary Palmer (eds), *Cognitive Linguistics and Non-Indo-European Languages* [Cognitive Linguistics Research 18]. Berlin: Mouton de Gruyter, 419–46.
- Yap, Foong Ha and Shoichi Iwasaki (2007). ‘The emergence of ‘give’ passives in East and Southeast Asian languages’, in Mark Alves, Paul Sidwell, and David Gil (eds), *SEALS VIII. Papers from the Eighth Annual Meeting of the Southeast Asian Linguistic Society*, Canberra: Pacific Linguistics, 193–208.
- Yeung, Kawai (2006). ‘On the status of the complementizer *waa6* in Cantonese’, *Taiwan Journal of Linguistics* 4-1: 1–43.
- Yiu, Carine Yuk-man 姚玉敏 (2010). ‘Zǎoqī Yuèyǔ zhōng de “Bi” zìjù 早期粵語中的「畀」字句 [*Bei-sentences in Early Cantonese*]’, in Hung-nin Samuel Cheung 張洪年 and Song Hing Chang 張雙慶 (eds), *Diachronic Change and Language Contact—Dialects in South East China 歷時演變與語言接觸—中國東南方言*, *Journal of Chinese Linguistics Monograph* 24, 162–85.
- You Rujie 游如杰 (1985). 汉语量词‘个’语源辨析 (*Hanyu liangci ge yuyuan bianxi* [An analysis of the etymological origin of the classifier ‘ge’]) 语文学研究 [*Yuwen Yanjiu*] (4): 23–4.
- You Rujie 游如杰 (1992). *Hanyu fangyanxue daolun* 汉语方言学导论 [*Chinese Dialectology*]. Shanghai: Shanghai Jiaoyu.
- You Rujie 游汝杰 (2003). ‘Wenzhou fangyan yufa gangyao 温州方言语法纲要 [Outline of Wenzhou dialect grammar]’, in *You Rujie Zixuanji* 游汝杰自选集 [*Selected Writings of You Rujie*]. Hefei: Anhui Jiaoyu Chubanshe, 164–235.
- Yu Cuirong 喻翠容 (2009). ‘Bùiyǔ Jiānzì 布依语简志 [Sketch of Bouyei]’, in Xiāoyù Zēng 曾晓渝 and Xuliàn Lǐ 李旭练 (eds), *Zhōngguó Shǎoshù Míngzú Yǔyán Jiānzì Cóngshū* 中

- 国少数民族语言简志从书 Volume 3. Beijing: The Ethnic Publishing House 民族出版社, 89–164.
- Yu Jin 余瑾 (2007). ‘Pínghuà Wèntí Yánjiū zhi Sikǎo 平话问题研究之思考 [Reflections on the research into the Pínghuà problem]’, in Yì Lín 林亦 and Jīn Yú 余瑾 (eds), *Dì Shíyī Jiè Guójì Yuè Fāngyán Yántǎohui Lùnwénjí 第 11 届国际粤方言研讨会论文集 [Proceedings of The Eleventh International Conference on Yuè Dialects]*. Nanning: Guangxi People's Publishing House 广西人民出版社, 1–5.
- Yuan Jiahua 袁家骅 (1989). *Hanyu Fangyan Gaiyao 漢語方言概要 [An Introduction to Chinese Dialects]*. Beijing: Language Reform Press.
- Yue, Anne O. (2003). ‘Chinese dialects: Grammar’, in Graham Thurgood and Randy LaPolla (eds), *The Sino-Tibetan Languages*. London: Routledge, 84–125.
- Yue-Hashimoto, Anne (1991). ‘Stratification in comparative dialectal grammar: A case in Southern Min’, *Journal of Chinese Linguistics* 19.2: 139–71.
- Yue-Hashimoto, Anne (1993). *Comparative Chinese Dialectal Grammar: Handbook for Investigators*. (Collection des Cahiers de Linguistique d'Asie Orientale 1). Paris: École des Hautes Études en Sciences Sociales.
- Yue-Hashimoto, Anne (1997). ‘Syntactic change in progress—Part I: the comparative construction in Hong Kong Cantonese’, in Anne Oi-kan Yue and Mitsuaki Endo (eds), *In Memory of Mantaro Hashimoto*. Tokyo: Uchiyama Shoten, 329–75.
- Yue-Hashimoto, Anne (1999). ‘The Min translation of the *Doctrina Christiana*’, in Pang-Hsin Ting (ed.), *Contemporary Studies on the Min Dialects*, *Journal of Chinese Linguistics* Monograph 13, 42–76.
- Zhang Cheng 张赪 (2004). ‘Mingdai de chabiju 明代的差比句 [Comparative constructions in the Ming dynasty]’, *Language and Linguistics* 5.3: 705–25.
- Zhang Cheng 张赪 (2005). ‘Cóng Hanyu bijiaoju kan lishi yanbian yu gongshi dili fenbu de guanxi 从汉语比较句看历史演变与共时地理分布的关系 [The relationship between historical evolution and synchronic distribution seen from the perspective of the Chinese comparative construction]’, *Yuwen Yanjiu 语文学研究* 1: 43–8.
- Zhang Junru 张均如 (1980). *Shuiyu Jianzhi 水语简志 [Sketch of the Sui Language]*. Beijing: Minzu chubanshe.
- Zhang Junru 张均如 (1987). ‘Jì Nánníng Xīnxū Pínghuà 记南宁心圩平话 [Notes on Pínghuà of Xīnxū, Nánníng]’, *Fāngyán 方言* 1987(4): 241–50.
- Zhang Junru 张均如 (2009). ‘Suīyǔ Jiǎnzhi 水语简志 [Sketch of Sui]’, in Xiǎoyù Zēng 曾晓渝 and Xùliàn Lǐ 李旭练 (eds), *Zhōngguó Shǎoshù Mínlú Yúyán Jiǎnzhi Cóngshū 中国少数民族语言简志丛书 Volume 3*. Beijing: The Ethnic Publishing House 民族出版社, 481–549.
- Zhang Junru 张均如, Min Liang 梁敏, Jueya Ouyang 欧阳觉亚, Yíqing Zheng 郑贻青, Xulian Li 李旭练, and Jianyou Xie 谢建猷 (1999). *Zhuàngyǔ Fāngyán Yánjiū 壮语方言研究 [Studies on Zhuang dialects]*. Chengdu: Sichuan Nationalities Publishing House 四川民族出版社.
- Zhang Min (2000). ‘Syntactic change in Southeastern Mandarin: How does geographical distribution reveal a history of diffusion?’, in Ting Pang-Hsin and Anne O. Yue (eds), *In Memory of Professor Li Fang-Kuei: Essays of Linguistic Change and the Chinese Dialects*. Taipei and Seattle: Academia Sinica and University of Washington, 197–242.

- Zhang Min 张敏 (2011). ‘漢語方言雙及物動詞結構南北差異的成因：類型學研究引發的新問題。Hanyu fangyan shuang jiwcj jiegou nan-bei chayi de chengyin: Leixingxue yanjiu de yinfa de xin wenti [The reasons for the North–South difference for ditransitives in Chinese dialects: New issues arising in typological research]’, *Bulletin of Chinese Linguistics* 4.2: 87–270.
- Zhang Min 张敏 and Lieting Zhou 周烈婷 (2003). ‘Gōulòupiàn Yuèyǔ hé Pínghuà de Jiūgé: Cóng Yùlínhuà de Guīshú Shuōqǐ 勾漏片粤语和平话的纠葛：从玉林话的归属说起 [The controversy about Gōulòu Yuè and Pínghuà: On the affiliation of the Yùlín dialect]’, in Bóhuì Zhān 詹伯慧 (ed.), *Dì Bā Jiè Guójì Yuè Fāngyán Yántáohuì Lùnwénjí* 第八届国际粤方言研讨会论文集 [*Proceedings of The Eighth International Conference on Yue Dialects*]. Beijing: China Social Science Press 中国社会科学出版社, 42–69.
- Zhang Peng 张彭 (2001). *Guangze Xianzhi* 光泽县志 [*A Survey of Guangze County*]. Beijing: Xianzhuang Shuju.
- Zhang Qingwen and Tang Sze-Wing (2010). ‘CL-NP phrases in five Chinese dialects’, paper presented at IACL-18 and NACCL-22, Harvard University.
- Zhang Xianmin 张贤敏 (2011). ‘Guangshan fangyan zhishi daici sanfen ji xianguan wenti 光山方言指示代词三分及相关问题 [The three-way system and other relevant issues in the Guangshan dialect]’, *Journal of Xinyang Normal University (Philosophy and Social Sciences Edition)* 31: 69–72.
- Zhang Xingliang 张兴良 (2008). ‘Hunan Ningxiang fangyan geti liangci yanjiu 湖南宁乡方言个体量词研究 [Research on sortal classifiers of the Ningxiang dialect]’, MA thesis, Hunan Normal University, Changsha.
- Zhang Xinxin and Sang Ye (1987). (eds) *Beijing Ren: Yibai ge putongren de zishu* 北京人：一百個普通人的自述 [*Beijing People: Narratives of 100 Ordinary People*]. Shanghai: Shanghai Wenyi Chubanshe.
- Zhang Yajun 张亚军 (2008). ‘Jiangsu Hai’an hua de liangci duyong biandiao xianxiang 江苏海安话的量词独用变调现象 [The tone sandhi of bare classifier phrases in the Hai’an dialect in Jiangsu province]’, *Zhongguo Yuwen* 中國語文 1: 61–4.
- Zhang Zhenxing 院语究 2013. 中国语言地图集 *Zhongguo Yuyan Dituji* [*Language Atlas of China*], Second edition, 2 vols. Beijing: Commercial Press.
- Zhao Jing 赵晶 (2008). ‘Hàn-Zhuàng Míngcízǔ Yǔxù de Bǐjiào Yánjiū 汉壮名词组语序的比较研究 [Comparative study of Chinese and Zhuang noun phrase word order]’, MA thesis, Guangxi University.
- Zhao Jinming 赵金铭 (2002a). ‘Hanyu chabiju de Nan-Bei chayi jí qí lishi shanbian 汉语差比句的南北差异及其历史嬗变 [The differences between North and South for Chinese comparative constructions and their historical transmutation]’, *Yuyan Yanjiu* 语言研究 3: 49–55.
- Zhao Jinming 赵金铭 (2002b). ‘Chabiju yuyi zhixiang leixing bijiao yanjiu 差比句语义指向类型比较研究 [Typological comparative research on the semantic orientation of comparative sentences]’, *Zhongguo Yuwen* 中国语文 5: 452–8.
- Zhao Xianzhi 赵先志 (1996). ‘Weifang de zhishi daici 潍坊方言的指示代词 [Demonstratives in the Weifang dialect]’, in Huang Borong (ed.), *Hanyu fangyan yufa leibian* 汉语方言语法类编 [*A Concordance on the Grammar of Sinitic Languages*]. Qingdao: Qingdao Chubanshe, 472–3.

- Zheng Wei 鄭偉 (2007). ‘吳語早期文獻所見的“等”字句 Wuyu zaoqi wenxian suojian de ‘deng’ziju [Sentences with *dēng* in early literature of the Wu dialects]’, *Zhongguo Yuwen* 中國語文 24.2: 33–42.
- Zheng Zuoguang and Lin Yi (eds) (2005). *Guibei Pinghua yu tuiguang Putonghua yanjiu*. 桂北平话与推广普通话研究 [Research on Northern Pinghua and the Popularization of Standard Mandarin]. Nanning: Guangxi Minzu Chubanshe.
- ZhengZhang, Shangfang 郑张尚芳 (1987). Shanggu yunmu xitong he sideng, jieyin shengdiao de laiyuan zenti 上古韵母系统和四等介音声调的来源问题 [The Old Chinese system of finals and the origins of the four divisions, the medials and the tones], *Wenzhou Shifan Xueyuan Xuebao* 温州师范学院学报 4.
- ZhengZhang, Shangfang 郑张尚芳 (1985). ‘Pucheng fangyan de nanbei qufen 浦城方言的南北区分 [The North-South division of the dialect of Pucheng]’, *Fangyan* 方言 1: 39–45.
- ZhengZhang, Shangfang 郑张尚芳(1995). ‘Fangyan zhong de shushengcuhua xianxiang 方言中的舒声促化现象 [On glottalization in Chinese dialects]’, *Zhongguo Yuyan Xuebao* 中国语言学报 1995 (5): 172–83.
- ZhengZhang, Shangfang 郑张尚芳 (2008). *Wenzhou fangyanzhi* 温州方言志 [A Study of the Wenzhou Dialect]. 中华书局 Zhonghua Shuju.
- Zhou Benliang 周本良, Xianghe Shen 沈祥和, Ping Li 黎平, Yujuan Wei 韦玉娟 (2006). ‘Nánníngshí Xiàguójīe Guānhuà Tóngyīn Zihùi 南宁下郭街官话同音字汇 [A list of homonyms of Xiaguo street dialect in Nanning]’, *Journal of Guilin Normal College* 桂林师范高等专科学校学报 20(2): 1–8.
- Zhou Changji 周长楫 (1982). ‘Shuo yi, zhi, he shu 说一, 植, 和蜀 [On the Minnan numerals for ‘one’: *yi*, *zhi* and *shu*]’, *Yuyan Yanjiu* 语言研究 2: 198–202.
- Zhou Changji 周長楫 (1999). ‘Xiamenhua de chabiju 廈門話的差比句 [Comparative sentences in the Xiamen dialect]’, in Yunji Wu 伍雲姬 (ed.), *Hanyu Fangyan Gongshi yu Lishi Yufa Yantao Lunwenji* 漢語方言共時與歷時語法研討論文集 [Synchronic and Diachronic Perspectives on the Grammar of Sinitic Languages]. Guangzhou: Jil'an University Press, 217–23.
- Zhou Changji and Ouyang Yiyun 周长楫, 欧阳忆 (1998). *Xiamen Fangyan Yanjiu* 厦门方言研究 [A Study of the Xiamen Dialect]. Fuzhou: Fujian Renmin Chubanshe.
- Zhou Changji and Qinghai Zhou 周長楫, 周清海 (2000). *Xingjiapo Minnanhua Gaishuo* 新加坡閩南話概說 [An Introduction to Southern Min in Singapore]. Xiamen: Xiamen University Press.
- Zhou Changji and Qinghai Zhou 周長楫, 周清海 (eds) (2002). *Xingjiapo Minnanhua Cidian* 新加坡閩南話詞典 [A Dictionary of Southern Min in Singapore]. Beijing: China Social Sciences Press.
- Zhou Fagao 周发高 (1968). *Zhongguo Yuyanxue Lunwenji* 中国语言学论文集 [Collected Writings on Chinese Linguistics]. Hong Kong: Chongji Shudian.
- Zhu Dexi 朱德熙 (1979). ‘Yu dongci gei xiangguan de jufa wenti 与动词‘给’相关的句法问题 [On the verb *gei* ‘to give’ and related syntactic problems]’, *Fangyan* 方言 2: 81–7.
- Zhu Dexi 朱德熙 (1982). *Yufa Jiangyi* 语法讲义 [Lectures on Grammar]. Beijing: The Commercial Press.
- Zhu Dexi 朱德熙 (1985). ‘Hànyǔ fāngyán-li de liǎng-zhōng fǎnfù wènjù’ 汉语方言里的两种反复问句 [Two kinds of V-NEG-V polar questions in the Chinese dialects]’, *Zhōngguó Yǔwén* 184: 10–20.

302 *References*

- Zhu Jiansong 朱建颂 (1992). *Wuhan Fangyan Yanjiu* 武汉方言研究 [Research on the Wuhan dialect]. Wuhan: Wuhan Press.
- Zhu Zhiyan 朱志燕 (2004). ‘Zhèyuán Rén de Rénwén Tèzhēng 蔗园人的人文特征 [Humanistic features of sugarcane field people]’, *Journal of Guangxi University for Nationalities (Philosophy and Social Science Edition)* 广西民族学院学报(哲学社会科学版) 2004: 23–4.
- Zimmermann, Malte (2008). ‘Contrastive Focus and Emphasis’, *Acta Linguistica Hungarica* 55: 347–60.
- Zubin, David and Mitsuaki Shimojo (1993). ‘How “general” are the general classifiers? with special reference to *ko* and *TSO* in Japanese’, in J.S. Guenter, B.A. Kaiser, and C.C. Zoll (eds), *Proceedings of the Nineteenth Annual Meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society. 490–502.