Contact Languages

Why do groups of speakers in certain times and places come up with new varieties of languages? What are the social settings that determine whether a mixed language, a pidgin, or a Creole will develop, and how can we understand the ways in which different languages contribute to the new grammar? Through the study of Malay contact varieties such as Baba and Bazaar Malay, Cocos Malay, and Sri Lanka Malay, as well as the Asian Portuguese vernacular of Macau, and China Coast Pidgin, the book explores the social and structural dynamics that underlie the fascinating phenomenon of the creation of new, or restructured, grammars. It emphasizes the importance and interplay of historical documentation, socio-cultural observation, and linguistic analysis in the study of contact languages, offering an evolutionary framework for the study of contact language formation – including pidgins and Creoles – in which historical, socio-cultural, and typological observations come together.

UMBERTO ANSALDO is Associate Professor in Linguistics at the University of Hong Kong. He was formerly a senior researcher and lecturer with the Amsterdam Center for Language and Communication at the University of Amsterdam. He has also worked in Sweden and Singapore and conducted fieldwork in China, the Cocos (Keeling) Islands, Christmas Island, and Sri Lanka. He is the co-editor of the Creole Language Library Series and has co-edited various journals and books including Deconstructing Creole (2007).
Cambridge Approaches to Language Contact

General Editor
Salikoko S. Mufwene, University of Chicago

Editorial Board
Robert Chaudenson, Université d’Aix-en-Provence
Braj Kachru, University of Illinois at Urbana
Raj Mesthrie, University of Cape Town
Lesley Milroy, University of Michigan
Shana Poplack, University of Ottawa
Michael Silverstein, University of Chicago

Cambridge Approaches to Language Contact is an interdisciplinary series bringing together work on language contact from a diverse range of research areas. The series focuses on key topics in the study of contact between languages or dialects, including the development of pidgins and creoles, language evolution and change, world Englishes, code-switching and code-mixing, bilingualism and second language acquisition, borrowing, interference, and convergence phenomena.

Published titles
Salikoko Mufwene, The Ecology of Language Evolution
Michael Clyne, Dynamics of Language Contact
Bernd Heine and Tania Kuteva, Language Contact and Grammatical Change
Edgar W. Schneider, Postcolonial English
Virginia Yip and Stephen Matthews, The Bilingual Child
Bernd Heine and Derek Nurse (eds.), A Linguistic Geography of Africa
J. Clancy Clements, The Linguistic Legacy of Spanish and Portuguese
Umberto Ansaldo, Contact Languages

Further titles planned for the series
Guy Bailey and Patricia Cukor-Avila, African-American English
Maarten Mous, Controlling Language
Jan Blommaert, The Sociolinguistics of Globalization
Bridget Drinka, Language Contact in Diachronic Perspective
Contact Languages

Ecology and Evolution in Asia

Umberto Ansaldo

The University of Hong Kong
A i miei genitori,
zia Assunta,
ed alla memoria di zio Riccardo
Contents

List of maps ix
List of tables and figures x
Series editor’s foreword xi
Acknowledgements xiii
List of abbreviations xvi

1 Introduction 1
  1.1 Introduction to contact language formation (CLF) 1
  1.2 Research questions 3
  1.3 The role of ecology in Asian contexts 5
  1.4 Theory of language and CLF 9
  1.5 CLF beyond exceptional evolution 13
  1.6 Outlook 16

2 The ecology of Monsoon Asia 18
  2.1 Monsoon Asia 19
  2.2 Sino–Javanese trade 22
  2.3 The city-ports 29
  2.4 Manpower in early Southeast Asia 35
  2.5 The Western impact 42

3 Linguistic ecologies of Southeast Asia 52
  3.1 Southeast Asia and the role of Malay 53
  3.2 Malay contact varieties 60
  3.3 Introducing contexts of formation 65
  3.4 The role of Portuguese in Southeast Asia and Southern China 73
  3.5 Summary 78

4 Methodological issues in the study of contact languages 81
  4.1 The ideology of theory 83
  4.2 Multilingualism and transmission 87
  4.3 Conclusions 97
## Contents

5 Contact language formation in evolutionary theory  
5.1 Competence, performance, and socialization  
5.2 Language evolution and contact languages  
5.3 Functionalist assets for contact linguistics  
5.4 Conclusions  

6 Congruence and frequency in Sri Lanka Malay  
6.1 The SLM community  
6.2 Selection and replication in SLM  
6.3 Freeing SLM from the chains of exceptionalism  
6.4 Final remarks  

7 Identity alignment in Malay and Asian-Portuguese Diaspora  
7.1 The ecology of identity alignment  
7.2 Multiple alignments in contact settings  
7.3 Identity alignment and admixture  
7.4 Conclusions  

8 Pidgin ecologies of the China coast  
8.1 Socio-historical background of Europe–China relations  
8.2 The ecology of Macau and the Pearl River Delta  
8.3 Grammatical features of China Coast Pidgin  
8.4 The missing Makista link?  
8.5 Discussion  

9 Implications, conclusions, and new horizons  
9.1 Theoretical and methodological implications  
9.2 Conclusions and new horizons  

References  
Index
Maps

2.1 Monsoon Asia (with wind patterns)  page 20
2.2 Mongol Empire  
   ca. AD 1300  24
2.3 China: coastal regions and languages  25
3.1 Contact Malay varieties and their main locations in the
   Malay-Indonesian archipelago  58
3.2 Cocos (Keeling) Islands  67
6.1 Malay communities in Sri Lanka  126
8.1 Pearl River Delta  185
Tables and figures

Tables

3.1 Linguistic ages in Singapore (adapted from Lim 2007)  page 70
6.1 Sri Lanka Malay communities – sociolinguistic profile  127
6.2 Etymology of case markers in SLM  129
6.3 Functions of case in adstrates Sinhala and Tamil  130
6.4 Thematic roles in SLM and adstrates Sinhala and Tamil  130
6.5 Assignment of Dative in SLM and adstrates Sinhala and Tamil  131
6.6 Assignment of Accusative in SLM and adstrates Sinhala and Tamil  132
6.7 Morpheme sources for TA markers in SLM  135
6.8 Aspect in PDM varieties  136
7.1 Kinship terms in BM, Hokkien, and Malay  165
7.2 Cocos Malay terms of address  165
7.3 Pronominal systems of BM, Bazaar Malay, Cocos Malay, and Hokkien  166
7.4 Reduplication patterns in Singlish and adstrates  168
7.5 Polyfunctional ‘have’ in BM/BZM, Malay, and Sinitic  171
7.6 Morphological material in Cocos Malay  178
8.1 Contacts in the Canton region  195

Figures

5.1 Sources of learning in a homogeneous and normative linguistic environment  100
5.2 Sources of learning in a heterogeneous naturalistic linguistic environment  101
5.3 Internal and external ecology  112
6.1 The Control feature  131
9.1 Multilingual and monolingual ecologies  226
The series *Cambridge Approaches to Language Contact* (CALC) was set up to publish outstanding monographs on language contact, especially by authors who approach their specific subject matter from a diachronic or developmental perspective. Our goal is to integrate the ever-growing scholarship on language diversification (including the development of Creoles, pidgins, and indigenized varieties of colonial European languages), bilingual language development, code-switching, and language endangerment. We hope to provide a select forum to scholars who contribute insightfully to understanding language evolution from an interdisciplinary perspective. We favour approaches that highlight the role of ecology and draw inspiration both from the authors’ own fields of specialization and from related research areas in linguistics or other disciplines. Eclecticism is one of our mottoes, as we endeavour to comprehend the complexity of evolutionary processes associated with contact.

We are very proud to add to our list Umberto Ansaldo’s *Contact Languages: Ecology and Evolution in Asia*. This is, to my knowledge, the most comprehensive account to date of ‘contact language formation’ in Monsoon Asia, including Sri Lanka, the Malaysian-Indonesian region, and Southern China. Providing a rich socio-economic history of layers of colonization of various kinds in the region, by Asians and Europeans alike, the book sheds light on the spread of indigenous languages such as Malay and Chinese, which, incidentally, have not developed identical ethnographic functions in the new geographical spaces of their expansion. It articulates clearly the relevant ecological factors that favoured both the geographical expansion of some languages indigenous to Asia and the emergence of new varieties such as Bazaar, Baba, and Cocos Malay, especially the ‘city-ports’ in relation to which the actual agents of these particular evolutions are identified. It is likewise informative to find out where and how colonial European languages indigenized, producing ‘pidgins’, ‘Creoles’, or other less divergent local varieties, and even why the evolution of each language was not uniform from one colony to another.

The reader will certainly be curious to find out why the languages of the economically and/or politically more powerful did not always prevail. Some of the explanations are to be sought in the ways in which the European colonization...
of Monsoon Asia not only varied internally, such as between the Portuguese, the Dutch, and the British, but also differed from the layers of the Asian colonization that preceded it. It is certainly as important to ask why the Asian language varieties that spread are the non-standard ones as to query why neither Arabic nor Hokkien or Cantonese became major lingua francas, when both the Arabs and Southern Chinese had been trade superpowers in the region before the arrival of Europeans in the sixteenth century. The ecologies of the contact and competition between Chinese and Malay, both before and during the colonization of the region by the Europeans, are especially very informative.

No less significant is the question of how ‘contact language formation’ in Monsoon Asia differs from or resembles what we have learned to date about language contact in the Caribbean and Indian Ocean, regions which have informed our hypotheses the most about ‘language change under contact conditions’. Umberto Ansaldo contributes critically to the old debate about universalist, substratist, and superstratist hypotheses, highlighting the merits of the ecological and uniformitarian approach to the subject matter. *Contact Languages* is impressive in its eclecticism too. Not only does it fruitfully bridge views by various creolists, historical linguists, and language typologists, but it also shows how various theories of evolution outside linguistics enable us to better understand both the mechanisms and the ecological factors that drive language speciation. The theoretical bridge between the linguistic effects of colonization in Monsoon Asia and the hitherto better-documented cases of the emergence of Creoles both around the Atlantic and in the Indian Ocean is as commendable in this case as the comprehensive coverage of the relevant literature in various complementary research areas, which is quite in the spirit of CALC. The reader will find in this book several useful references to language contact outside creolistics, to language acquisition, to linguistic anthropology, to the philosophy of science, to language evolution, and to colonial history, among others. Experts and non-experts alike will appreciate especially the wealth of information that the book provides about its subject matter.

*SALIKOKO S. MUFWENE, University of Chicago*
The research and thoughts that led to this book started many years ago and I fear I may not be able to recall everyone who helped me get to this point. I first of all wish to acknowledge the lecturers who inspired me most in my undergraduate years at the University of Venice, in particular: Magda Abbiati and Maurizio Scarpari for Chinese and historical linguistics respectively, and Massimo Raveri for Asian philosophy. Next I wish to thank PhD supervisor and friend Maria Koptjevskaja-Tamm for her generous teachings in linguistic typology, research methods, and dissertation writing, and her continuous support in the past years.

Much gratitude goes to a large number of individuals in two communities where I have conducted fieldwork: the Cocos Malays and the Sri Lanka Malays. In the former, I wish to thank the following for their time and hospitality: Nek Aman (Cree Haig), Nek Harun (Capstan Benjamin), Wak James (James Benjamin), the late Nek Ketie lakilaki (Yapat Carey), Nek Ketie perumpuan (Ebetie Sebie), Nek Fatima (Pirus), Nek Ashari (Medous Bynie), Wak Pitah, Pak Adam (Rabuhu Anhoney), Pak Aini (Radal Feyrel), Mak Aini (Yean Yapat), Marias Blacket, Pak Shafi (Zaitol Wallie), Morin (Catalina Capstan), Matt Macrae, Mak Yazid (Nollia Zaibie), Mohd Noor Jada, Rosita Oemar, Aini Radal and Aindil Minkom, in particular Pak and Mak Aini and their family for their warm friendship. In relation to the history of the Cocos Islands, I also wish to thank Ron Grant, Pearl Tan, John Clunies-Ross, Jack Carlsen, John Hunt, and Kenneth McPherson for their input. In Sri Lanka, my most sincere gratitude goes to the following individuals for their interest, generous support, and warm hospitality and friendship: T. M. Muhaj Hamin, B. D. K. Saldin, T. K. Azoor, Salma and Upali Peiris and all their family, B. Jalaldeen, M. T. M. Rihan, B. J. Faizal Cassim; others whom I thank for their interest and time include: in Colombo, Stanley Oemar, K. Moulana, Fathima Mohamed, Darlene Burah, Kartini Mohamed, Eileen Dane, Mohamed Ismath Suhood, Romola Rassool, Fathima Shahdia Jamaldeen, Talisha Rozianthi Jamildeen, M. Naveed Jamaldeen, Dane Danish Suhood; in Hambantota, the late Proctor T. S. Doole, M. Doole, the late T. A. Amath, T. M. Farook, Barmatha Cassim; in Kirinda, J. M. Rafaidin, the late A. S. Damani (Nenek), the late A. D. Prenantha, N. P.
Alawudeen, R. M. Nahar, S. M. Nawas, Siti Fatima, T. F. Salma. I am indebted to Thiru Kandiah for providing the initial point of contact with the Sri Lanka Malays.

The following colleagues and friends have offered comments to various parts of this book: I thank Christopher Stroud for his suggestions on sociolinguistic theory, Wolfram Hinzen for his view on linguistic theory and the philosophy behind it, Perry Johansson for his comments on the historical section, Hugo Cardoso for his feedback on Indo-Portuguese, and especially Stephen Matthews for comments on the data and many inspirational discussions on language contact and typology over the years.

Other colleagues have helped with materials, suggestions and advice. I want to thank Sander Adelaar and Jim Collins who have made their immense knowledge on Malay linguistics available to me by sharing their works and views. I thank Alan Baxter for generously providing me with materials on Makista, and Clancy Clements for making available some early drafts of his book on the linguistic legacy of Spanish and Portuguese. I also want to thank Geoff Smith and Stephen Matthews for sharing with me valuable materials and views on China Coast Pidgin. More thanks for their ideas, suggestions and materials also go to: Aone van Engelenhoven, John Holm, Brian Joseph, Yaron Matras, Miriam Meyerhoff, Pieter Muysken, Sebastian Nordhoff, Scott Paauw, Rachel Selbach, and Uri Tadmor.

There are colleagues in many places and universities around the world who were kind enough to listen to the ideas developed in this book at seminars, workshops, conference presentations, and invited talks over the past few years. I remember in particular the colleagues in Stockholm, Nijmegen, Leiden, Giessen, Edinburgh, Coimbra, Cape Town, Western Cape, Colombo, Kelaniya, Hong Kong, and Seoul, for interesting discussions.

At the University of Amsterdam, I would like to thank Kees Hengeveld for his constant support throughout the project, Enoch Aboh for many constructive discussions on Creole languages, as well as Norval Smith and Hans den Besten for being always available with expert advice.

I also acknowledge valuable assistants who helped with transcribing and organizing the Malay data and metadata: Diane Chang, Christine Angelique Kang, and Konrad Rybka.

For hospitality during my sabbatical year, which I devoted to this project, my gratitude goes to my good friends Charles Goddard, Tom Mitchell, and Martin Webster.

There are those I cannot mention: they know why.

The role of Salikoko Mufwene cannot be overestimated, having been most supportive of this project from the earliest phases. His numerous comments, suggestions, and ideas, and his constructive criticism on content, style, and
Acknowledgements

organization have been fundamental in arriving at what is hopefully a clear and complete product.

I also wish to thank Helen Barton and the whole production team at CUP for their efficiency.

My deepest and most emphatic gratitude must, however, go to the one person in particular who, more than anyone else, has supported this book project in too many ways to mention. Lisa Lim has contributed phonetic and sociolinguistic expertise, as well as professional advice as linguistic and copy editor, and has been a valuable fieldwork and life partner throughout. She has also afforded me unique, insider’s perspectives on the ecology and evolution of new languages derived from her Peranakan heritage and Singlish competence. ‘Thank you’, Lisa, is barely enough, I know.

The following grants should be acknowledged: ‘Contact languages of Southeast Asia: The role of Malay’ (R-103–000–020–112), Academic Research Grant, National University of Singapore, for fieldwork on Cocos Malay and (partially) Sri Lanka Malay; ‘The documentation of Sri Lanka Malay: Linguistic and cultural creolization endangered’, Volkswagen Stiftung initiative for the Documentation of Endangered Languages (DoBeS), for fieldwork on Sri Lanka Malay; and, last but not least, ‘Contact languages of Asia’, The Netherlands Organisation for Scientific Research (NWO), for precious time to write this book.

Amsterdam, November 2008

Much of the content of this book has not appeared in press before but I would like to acknowledge the following: parts of chapter 4 have appeared in Ansaldo and Matthews (2007); a version of chapter 6 was the basis for Ansaldo (2008); for some sections of chapter 7, I am indebted to earlier work appearing in Ansaldo and Matthews (2004), Lim and Ansaldo (2007), and Ansaldo, Lim and Mufwene (2007); an earlier version of chapter 8 constitutes the basis of and is informed by later developments of Ansaldo, Matthews and Smith (to appear). I thank all these co-authors for the inspiration they have provided.
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABL</td>
<td>Ablative</td>
</tr>
<tr>
<td>ACC</td>
<td>Accusative</td>
</tr>
<tr>
<td>ANIM</td>
<td>animacy</td>
</tr>
<tr>
<td>BM</td>
<td>Baba Malay</td>
</tr>
<tr>
<td>BZM</td>
<td>Bazaar Malay</td>
</tr>
<tr>
<td>CCP</td>
<td>China Coast Pidgin</td>
</tr>
<tr>
<td>CL</td>
<td>classifier</td>
</tr>
<tr>
<td>CLF</td>
<td>contact language formation</td>
</tr>
<tr>
<td>CM</td>
<td>Cocos Malay</td>
</tr>
<tr>
<td>CONT</td>
<td>continuous</td>
</tr>
<tr>
<td>DAT</td>
<td>Dative</td>
</tr>
<tr>
<td>DEM</td>
<td>demonstrative</td>
</tr>
<tr>
<td>DEF</td>
<td>definiteness</td>
</tr>
<tr>
<td>DOM</td>
<td>definite object marker</td>
</tr>
<tr>
<td>FT</td>
<td>foreigner talk</td>
</tr>
<tr>
<td>LC</td>
<td>linguistic citizenship</td>
</tr>
<tr>
<td>LOC</td>
<td>Locative</td>
</tr>
<tr>
<td>MM</td>
<td><em>Macau Monograph</em></td>
</tr>
<tr>
<td>N</td>
<td>noun</td>
</tr>
<tr>
<td>NEG</td>
<td>negation</td>
</tr>
<tr>
<td>NP</td>
<td>noun phrase</td>
</tr>
<tr>
<td>NOM</td>
<td>nominative</td>
</tr>
<tr>
<td>NUM</td>
<td>number</td>
</tr>
<tr>
<td>O</td>
<td>object</td>
</tr>
<tr>
<td>PASS</td>
<td>passive</td>
</tr>
<tr>
<td>P.c.</td>
<td>personal communication</td>
</tr>
<tr>
<td>PDM</td>
<td>Pidgin-derived Malay</td>
</tr>
<tr>
<td>PFV</td>
<td>perfective</td>
</tr>
<tr>
<td>PK</td>
<td>Papia Kristang</td>
</tr>
<tr>
<td>PL</td>
<td>plural</td>
</tr>
<tr>
<td>POSS</td>
<td>Possessive</td>
</tr>
<tr>
<td>PP</td>
<td>prepositional phrase</td>
</tr>
</tbody>
</table>

xvi
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PROG</td>
<td>progressive</td>
</tr>
<tr>
<td>PST</td>
<td>past</td>
</tr>
<tr>
<td>Psych V</td>
<td>psychological verb</td>
</tr>
<tr>
<td>RG</td>
<td>Redhaired Glossary</td>
</tr>
<tr>
<td>SVC</td>
<td>serial verb constructions</td>
</tr>
<tr>
<td>SLM</td>
<td>Sri Lanka Malay</td>
</tr>
<tr>
<td>S</td>
<td>subject</td>
</tr>
<tr>
<td>TA</td>
<td>tense and aspect</td>
</tr>
<tr>
<td>TM</td>
<td>typological matrix</td>
</tr>
<tr>
<td>TMA</td>
<td>tense, mood, and aspect</td>
</tr>
<tr>
<td>TRANS</td>
<td>transitive</td>
</tr>
<tr>
<td>V</td>
<td>verb</td>
</tr>
<tr>
<td>VP</td>
<td>verb phrase</td>
</tr>
</tbody>
</table>