CRITICAL CONCEPTS IN LINGUISTICS
Series editor: Carlos P. Otero

Selected titles in this series:

English for Academic Purposes
Edited and with a new introduction by Helen Basturkmen
4 volume set

Historical Linguistics
Edited and with a new introduction by Hope C. Dawson and Brian D. Joseph
6 volume set

History of the English Language
Edited and with a new introduction by Elly van Gelderen
4 volume set

Language and Migration
Edited and with a new introduction by Ingrid Piller
4 volume set

Language and the Media
Edited and with a new introduction by Helen Kelly-Holmes
4 volume set

Language Policy and Planning
Edited and with a new introduction by Thomas Ricento
4 volume set

Language Testing and Assessment
Edited and with a new introduction by Antony John Kunnan
4 volume set

Language Variation and Change
Edited and with a new introduction by Robert Bayley and Richard Cameron
5 volume set

Multimodality
Edited and with a new introduction by Sigrid Norris
4 volume set

Systemic Functional Linguistics
Edited and with a new introduction by James Martin and Yaegan Doran
5 volume set

Vocabulary
Edited and with a new introduction by Stuart Webb
4 volume set

SINO-TIBETAN LINGUISTICS

Critical Concepts in Linguistics

Edited by
Randy J. LaPolla

Volume I

Establishing the Relationships

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published 2019
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

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

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

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

Library of Congress Cataloging-in-Publication Data
Names: LaPolla, Randy J., editor.
Title: Sino-Tibetan linguistics : critical concepts in linguistics/edited by Randy J. LaPolla.
Description: London ; New York, NY : Routledge, 2018. | Includes bibliographical references and index. Contents: v. 1. Establishing the relationships — v. 2. Language Contact and Areal Features — v. 3. Sinitic — v. 4. Tibeto-Burman.
Identifiers: LCCN 2018012985 | ISBN 9780415577397 (set) | ISBN 9780415579766 (v. 1) | ISBN 9780415579759 (v. 2) | ISBN 9780415579742 (v. 3) | ISBN 9780415579735 (v. 4)
Subjects: LCSH: Sino-Tibetan languages.
Classification: LCC PL3521 .S565 2018 | DDC 495—dc23
LC record available at https://lccn.loc.gov/2018012985

ISBN: 978-0-415-57739-7 (set)
ISBN: 978-0-415-57976-6 (Volume I)

Typeset in Times New Roman
by Apex CoVantage, LLC

Publisher's Note
References within each chapter are as they appear in the original complete work

Dedicated to the great pioneers
in Sino-Tibetan linguistics

CONTENTS

VOLUME I Establishing the Relationships

VOLUME II Language Contact and Areal Features

ACKNOWLEDGEMENTS

The publishers would like to thank the following for permission to reprint their material:

Journal of Chinese Linguistics for permission to reprint Fang-Kuei Li, 'Languages and Dialects of China', *Journal of Chinese Linguistics* 1, 1, 1973, 1–13.

Linguistics of the Tibeto-Burman Area for permission to reprint Paul K. Benedict, 'Where it All Began: Memories of Robert Shafer and the "Sino-Tibetan Linguistics Project" Berkeley 1939–40', *Linguistics of the Tibeto-Burman Area* 2, 1, 1975, 81–92.

John Wiley & Sons for permission to reprint Paul K. Benedict, 'Thai, Kadai, and Indonesian: A New Alignment in Southeastern Asia', *American Anthropologist* 44, 1942, 576–601.

The International Linguistic Association for permission to reprint Robert Shafer, 'Classification of the Sino-Tibetan Languages', *Word* 11, 1, 1955, 94–111.

Journal of Chinese Linguistics for permission to reprint James A. Matisoff, 'Notes on Fang—Kuei Li's "Languages and Dialects of China"', *Journal of Chinese Linguistics* 1, 3, 1973, 471–474.

The Linguistic Society of America for permission to reprint James A. Matisoff, 'On Megalocomparison', *Language* 66, 1, 1990, 106–120.

Linguistics of the Tibeto-Burman Area for permission to reprint Paul K. Benedict, 'Comment on Matisoff's Comparison between Greenberg and Benedict', *Linguistics of the Tibeto-Burman Area* 14, 2, 1991, 169–170.

Annual Reviews, Inc., for permission to reprint James A. Matisoff, 'Sino-Tibetan Linguistics: Present State and Future Prospects', *Annual Review of Anthropology* 20, 1991, 469–504.

Linguistics of the Tibeto-Burman Area for permission to reprint Robbins Burling, 'The Sal Languages', *Linguistics of the Tibeto-Burman Area* 7, 2, 1983, 1–32.

Linguistics of the Tibeto-Burman Area for permission to reprint Karen H. Ebert, 'On the Evidence for the Relationship Kiranti-Rung', *Linguistics of the Tibeto-Burman Area* 13, 1, 1990, 57–78.

Linguistics of the Tibeto-Burman Area for permission to reprint Jackson T.-S Sun, 'The Linguistic Position of Tani (Mirish) in Tibeto-Burman: A Lexical Assessment', *Linguistics of the Tibeto-Burman Area* 16, 2, 1993, 143–188.

Cambridge University Press for permission to reprint Jean Przyluski and Gordon H. Luce, 'The Number "A Hundred" in Sino-Tibetan', *Bulletin of the School of Oriental Studies* 6, 3, 1931, 667–668.

Cambridge University Press for permission to reprint Stuart N. Wolfenden, 'Concerning the Variation of Final Consonants in the Word Families of Tibetan, Kachin, and Chinese', *Journal of the Royal Asiatic Society of Great Britain and Ireland* 4, 1937, 625–655.

The Institute of History and Philology for permission to reprint Randy J LaPolla, 'Variable Finals in Proto-Sino-Tibetan', *Bulletin of the Institute of History and Philology* 65, 1, 1994, 131–73.

The Institute of History and Philology for permission to reprint Gong Hwang-cherng, 'A Comparative Study of the Chinese, Tibetan, and Burmese Vowel Systems', *Bulletin of the Institute of History and Philology* 51, 3, 1980, 455–489.

Disclaimer

The publishers have made every effort to contact authors/copyright holders of works reprinted in *Sino-Tibetan Linguistics: Critical Concepts in Linguistics.* This has not been possible in every case, however, and we would welcome correspondence from those individuals/companies whom we have been unable to trace.

Chronological table of reprinted articles and chapters

Date	*Author*	*Article/Chapter*	*Source*	*Vol.*	*Chap.*
1930	Yuen-Ren Chao	A system of tone "letters"	*Le Maître Phonétique* 45, 24–7.	III	45
1931	J. Przyluski and G. H. Luce	The number "a hundred" in Sino-Tibetan	*Bulletin of the School of Oriental Studies* 6, 3, 667–8.	I	13
1933	Bernhard Karlgren	Word families in Chinese	*Bulletin of the Museum of Far Eastern Antiquities* 5, 9–120.	III	35
1933	Fang-Kuei Li	Certain phonetic influences of the Tibetan prefixes upon the root initials	*Bulletin of the Institute of History and Philology* 6, 2, 135–57.	IV	54
1934	Yuen-Ren Chao	The non-uniqueness of phonemic solutions of phonetic systems	*Bulletin of the Institute of History and Philology, Academia Sinica* IV, 4, 363–97.	III	46
1937	Stuart N. Wolfenden	Concerning the variation of final consonants in the word families of Tibetan, Kachin, and Chinese	*Journal of the Royal Asiatic Society of Great Britain and Ireland* 4, 625–55.	I	14
1942	Paul K. Benedict	Thai, Kadai, and Indonesian: a new alignment in southeastern Asia	*American Anthropologist* 44, 576–601.	I	3
1945	Li Fang-kuei	Some old Chinese loan words in the Tai languages	*Harvard-Yenching Journal of Asiatic Studies* 8, 333–42.	II	20
1947	Charles F. Hockett	Peiping phonology	*Journal of the American Oriental Society* 67, 4, 253–67.	III	47
1955	Robert Shafer	Classification of the Sino-Tibetan languages	*Word* 11, 1, 94–111.	I	4
1956	Bernhard Karlgren	Cognate words in the Chinese phonetic series	*Bulletin of the Museum of Far Eastern Antiquities* 28, 1–18.	III	36
1957	Eugénie J. A. Henderson	Colloquial Chin as a pronominalized language	*Bulletin of the School of Oriental and African Studies* 20, 323–7.	IV	58

Chronological table continued

Date	*Author*	*Article/Chapter*	*Source*	*Vol.*	*Chap.*
1959	G. B. Downer	Derivation by tone-change in classical Chinese	*Bulletin of the School of Oriental and African Studies, University of London* 22, 1/3, 258–90.	III	37
1962	R. A. D. Forrest	The linguistic position of Rong (Lepcha)	*Journal of the American Oriental Society* 82, 3, 331–5.	II	24
1962	Theodore Stern	Language contact between related languages: Burmese influences upon Plains Chin	*Anthropological Linguistics* 4, 4, 1–28.	II	26
1965	Eugénie J. A. Henderson	The topography of certain phonetic and morphological characteristics of South East Asian languages	*Lingua* 15, 400–34.	II	17
1965	John Okell	Nissaya Burmese: a case of systematic adaptation to a foreign grammar and syntax	*Lingua* 15, 186–227.	II	30
1966	Robbins Burling	The addition of final stops in the history of Maru (Tibeto-Burman)	*Language* 42, 3, 581–6.	IV	57
1966	Fang-Kuei Li	The zero initial and the zero syllabic	*Language* 42, 2, 300–2.	III	48
1967	Hla Pe	A tentative list of Mon loan words in Burmese	*Journal of the Burma Research Society* 50, 1, 71–94.	II	28
1970	Mei Tsu-lin	Tones and prosody in Middle Chinese and the origin of the rising tone	*Harvard Journal of Asiatic Studies* 30, 86–110.	III	38
1972	James A. Matisoff	Lahu nominalization, relativization, and genitivization	John P. Kimball (ed.), *Syntax and Semantics, Vol. 1* (New York: Academic Press), pp. 235–57.	IV	64

1973	Fang-Kuei Li	Languages and dialects of China	*Journal of Chinese Linguistics* 1, 1, 1–13; originally published in *The Chinese Year Book* (Shanghai 1937), pp. 121–8.	I	1
1973	James A. Matisoff	Notes on Fang–Kuei Li's 'Languages and Dialects of China'	*Journal of Chinese Linguistics* 1, 3, 471–4.	I	5
1973	James A. Matisoff	Tonogenesis in Southeast Asia	Larry M. Hyman (ed.), *Consonant Types and Tone* (Southern California Occasional Papers in Linguistics, No. 1) (Los Angeles, CA: UCLA), pp. 72–95.	IV	66
1973	Jerry Norman	Tonal development in Min	*Journal of Chinese Linguistics* 1, 2, 222–38.	III	50
1973	E. G. Pulleyblank	Some new hypotheses concerning word families in Chinese	*Journal of Chinese Linguistics* 1, 1, 111–25.	III	39
1973	E. G. Pulleyblank	Some further evidence regarding Old Chinese *-s* and its time of disappearance	*Bulletin of the School of Oriental and African Studies* 36, 2, 368–73.	III	40
1974	Jim Bauman	Pronominal verb morphology in Tibeto-Burman	*Linguistics of the Tibeto-Burman Area* 1, 1, 108–55.	IV	59
1974	Edward H. Bendix	Indo-Aryan and Tibeto-Burman contact: as seen through Nepali and Newari verb tenses	*International Journal of Dravidian Linguistics* 3, 42–59.	II	32
1975	Alton L. Becker	A linguistic image of nature: the Burmese numerative classifier system	*International Journal of the Sociology of Language: Sociolinguistics in Southeast Asia* 5, 109–21.	IV	65

Chronological table continued

Date	*Author*	*Article/Chapter*	*Source*	*Vol.*	*Chap.*
1975	Paul K. Benedict	Where it all began: memories of Robert Shafer and the "Sino-Tibetan Linguistics Project" Berkeley 1939–40	*Linguistics of the Tibeto-Burman Area* 2, 1, 81–92.	I	2
1975	Denise Bernot	Influence of Burmese language on some other languages of Burma (writings systems and vocabulary)	*Journal of the Siam Society* 63, 96–109.	II	27
1976	Paul K. Benedict	Sino-Tibetan: another look	*Journal of the American Oriental Society* 96, 2, 167–97.	I	6
1976	W. South Coblin	Notes on Tibetan verbal morphology	*T'oung Pao* 62, 1/3, 45–70.	IV	55
1976	Mantaro J. Hashimoto	Language diffusion on the Asian continent: problems of typological diversity in Sino-Tibetan	*Computational Analyses of Asian and African Languages* (National Inter-University Research Institute of Asian and African Languages and Cultures, Tokyo) 3, 49–66.	II	18
1976	Oi-kan Yue Hashimoto	Southern Chinese dialects: the Tai connection	*Computational Analyses of Asian and African Languages* (National Inter-University Research Institute of Asian and African Languages and Cultures, Tokyo) 6, 1–9.	II	22
1976	Fang-kuei Li	Sino-Tai	*Computational Analyses of Asian and African Languages* (National Inter-University Research Institute of Asian and African Languages and Cultures, Tokyo) 3, 39–48.	II	21
1976	Jerry Norman and Tsu-lin Mei	The Austroasiatics in ancient South China: some lexical evidence	*Monumenta Serica* 32, 274–301.	II	23
1977	Walter Simon	Alternation of final vowel with final dental nasal or plosive in Tibetan	*Bulletin of the School of Oriental and African Studies* 40, 1, 51–7.	IV	56
1980	David Bradley	Phonological convergence between languages in contact: Mon-Khmer structural borrowing in Burmese	*Proceedings of the Sixth Annual Meeting of the Berkeley Linguistic Society*, 259–67.	II	29
1980	Hwang-cherng Gong	A comparative study of the Chinese, Tibetan, and Burmese vowel systems	*The Bulletin of the Institute of History and Philology* 51, 3, 455–89.	I	16
1983	Robbins Burling	The Sal languages	*Linguistics of the Tibeto-Burman Area* 7, 2, 1–32.	I	10
1983	Charles N. Li	Languages in contact in Western China	*Papers on East Asian Languages* 1, 1, 31–51.	II	33
1984	Mantaro J. Hashimoto	Origin of the East Asian linguistic structure: latitudinal transitions and longitudinal developments of East and Southeast Asian languages	*Computational Analyses of Asian and African Languages* (National Inter-University Research Institute of Asian and African Languages and Cultures, Tokyo) 22, 35–41.	II	19
1986–1987	W. South Coblin	*Fangyan* gleanings	*Monumenta Serica* 37, 113–43.	III	41
1987	Derek D. Herforth	A case of radical ambiguity in Old Chinese: some notes toward a discourse-based grammar	*Suzugamine Joshi Tanki Daigaku Bulletin of Humanities and Social Science Research* 34, 31–40.	III	43
1988	Nicholas C. Bodman	On the place of Lepcha in Sino-Tibetan: a lexical comparison	*Linguistics of the Tibeto-Burman Area* 11, 1, 1–26.	II	25
1990	Karen H. Ebert	On the evidence for the relationship Kiranti-Rung	*Linguistics of the Tibeto-Burman Area* 13, 1, 57–78.	I	11
1990	James A. Matisoff	On megalocomparison	*Language* 66, 1, 106–20.	I	7

Chronological table continued

Date	*Author*	*Article/Chapter*	*Source*	*Vol.*	*Chap.*
1991	Paul K. Benedict	Comment on Matisoff's comparison between Greenberg and Benedict	*Linguistics of the Tibeto-Burman Area* 14, 2, 169–70.	I	8
1991	James A. Matisoff	Sino-Tibetan linguistics: present state and future prospects	*Annual Review of Anthropology* 20, 469–504.	I	9
1991	Chaofen Sun	The adposition *yi* and word order in Classical Chinese	*Journal of Chinese Linguistics* 19, 2, 202–18.	III	44
1992	Scott DeLancey	The historical status of the conjunct/disjunct pattern in Tibeto-Burman	*Acta Linguistica Hafniensia* 25, 1, 39–62.	IV	63
1992	Arienne M. Dwyer	Altaic elements in the Línxià dialect: contact-induced change on the Yellow River Plateau	*Journal of Chinese Linguistics* 20, 1, 160–79.	II	34
1992	M. A. K. Halliday	A systemic interpretation of Peking syllable finals	P. Tench (ed.), *Studies in Systemic Phonology* (London and New York: Pinter), pp. 98–121.	III	49
1992	Mantaro J. Hashimoto	Hakka in Wellentheorie perspective	*Journal of Chinese Linguistics* 20, 1, 1–48.	III	51
1992	Randy J. LaPolla	On the dating and nature of verb agreement in Tibeto-Burman	*Bulletin of the School of Oriental and African Studies* 55, 2, 298–315.	IV	60
1993	Anne Yue-Hashimoto	The lexicon in syntactic change: lexical diffusion in Chinese syntax	*Journal of Chinese Linguistics* 21, 2, 213–53.	III	52
1993	Randy J. LaPolla	Arguments against 'subject' and 'direct object' as viable concepts in Chinese	*Bulletin of the Institute of History and Philology* 63/4, 759–813.	III	53
1993	Jackson T.-S. Sun	The linguistic position of Tani (Mirish) in Tibeto-Burman: a lexical assessment	*Linguistics of the Tibeto-Burman Area* 16, 2, 143–88.	I	12
1994	Randy J. LaPolla	Variable finals in Proto-Sino-Tibetan	*The Bulletin of the Institute of History and Philology* 65, 1, 131–73.	I	15
1994	Randy J. LaPolla	Parallel grammaticalizations in Tibeto-Burman languages: evidence of Sapir's 'drift'	*Linguistics of the Tibeto-Burman Area* 17, 1, 61–80.	IV	61
1995	Jerry L. Norman and W. South Coblin	A new approach to Chinese historical linguistics	*Journal of the American Oriental Society* 115, 4, 576–84.	III	42
1997	Dipankar Moral	North-East India as a linguistic area	*Mon-Khmer Studies* 27, 43–53.	II	31
1999	Tej R. Kansakar	Verb agreement in Classical Newar and Modern Newar dialects	Yogendra P. Yadava and Warren W. Glover (eds), *Topics in Nepalese Linguistics* (Kathmandu: The Royal Nepal Academy), pp. 421–43.	IV	62

PREFACE

These four volumes grew out of the desire to gather together seminal journal articles in Sino-Tibetan linguistics to provide a useful source for those entering the field, or for those who want to understand more about the development of the field, or for those who are veterans in the field and want access to the classic works. I include or cite works that have been influential (or I think should have been), particularly from journals that were core to the field in the past but are no longer in existence or that younger scholars might not be familiar with, such as *Monumenta Serica*, *Acta Linguistica Hafniensia*, *The Bulletin of the Institute of History and Philology*,[1] and *Computational Analyses of Asian and African Languages*. I have also included a small number of articles from edited series that are hard to come by. Given space restrictions, we have limited the coverage to the twentieth century, when the scientific study of Sino-Tibetan languages began in earnest. There has been an explosion of work on Sino-Tibetan languages since then, due in part to the fact that key geographic areas, such as Northeast India and western China, which were earlier out of bounds politically, are more open to researchers, and so much more language data has become available. See Thurgood and LaPolla (2017) and the many references therein for some of that recent work. Another aspect is the new ways of thinking about language history stimulated by work in general linguistics in grammaticalization and contact linguistics.

The choice of articles is a personal one, reflecting my own understanding of the development of the field, based on more than 40 years of studying Sino-Tibetan languages and linguistics. I considered hundreds of articles and read and reread dozens of them in working on each volume to choose the few that were included in the volume and write up the introductions. I am sure there will be those who will think I omitted certain important articles or included some that might not be seen by all as important as the others, but I have tried to justify my selections in the introductions to the volumes, and have tried to mention many other articles that I was not able to include in the volume, to give pointers to people who want to go on to do further reading. I have also tried to weave the articles together to some extent in the introductions, to give readers a sense of the development of the field.

I sent a request for suggestions of articles to be included in the volumes to the Tibeto-Burman Linguistics Discussion Listserv (tibeto-burman-linguistics@listserv.linguistlist.org), but only heard back from a few people. I'd like to thank Zev Handel, Nathan Hill, Guillaume Jacques, James A. Matisoff, Graham Thurgood, William S-Y Wang and the *Journal of Chinese Linguistics* office, and Yap Fung Ha for suggestions. I'd also like to thank Anne Yue for help in sourcing her articles and also for information used in the introductions, and W. South Coblin for valuable advice.

As I mentioned in a recent article (LaPolla 2016a), I have tremendous respect for the pre-digital and pre-bibliometrics (pre-managerialism in the universities) era scholars, as they were allowed to and did take the time to immerse themselves in the primary sources, developing a deep personal understanding of the languages and the literature,[2] while staying up with the linguistics literature as well, even though both of these were difficult and time consuming. When they wrote of course it was also a tedious process, either handwritten or typed on a typewriter, with scraps of paper and notes everywhere, so they often didn't publish many papers per year, but what they did publish was often of amazing quality and erudition. So editing this set is something of a labour of love for me, a tribute to these great scholars, and that is why I have dedicated the volumes to them.

I would also like to mention to young scholars that they should be aware of the context in which the older scholars were working, and the particular constraints on their thinking or information at the time they were working, and not commit the "Historian's fallacy" of assuming that the scholars of that day had the same access, data, or understandings that we have today. I have seen some young scholars write things that are disrespectful of older scholars or miss the point of why the older scholar was writing what they were writing, which is inappropriate, as each of us is reacting to particular issues of our time, and constrained by the prevailing paradigm of the time. For example, even great scholars were constrained to think in terms of early generative or structuralist approaches in the late 1960s and early 1970s that we cannot imagine people going along with nowadays, but that was the paradigm of the time.

I'd like to thank Carly Jaques for her very efficient help in sourcing many of the articles, not only in the collection, but many of those I refer to in my introductions, and for her careful editing of the drafts of my introductions. I'd also like to thank Malcolm Campbell, Jillian Morrison, and Simon Alexander, of Routledge, for their help and their patience with me during the production of these volumes.

Notes

1 The nature of this journal changed somewhat when the Linguistics Section of the Institue of History and Philology split off and became the Institute of Linguistics at the Academia Sinica, and now has its own journal, *Language and Linguistics* (《語言與語言學》), which has become quite an important journal.

2 As an example of the sort of depth and breadth of knowledge I am talking about, back in the early 1980s, I was once discussing grammatical patterns that appeared in the 20-volume *Zu Tang Ji* (a Nan Tang dynasty [937–975 CE] Chinese vernacular text) with Prof. Tsu-lin Mei, and his knowledge of the text was so good that at one point, after I had suggested that a certain pattern might be found in the text, he said, "If I don't know about it, it doesn't exist". And he may actually have meant that about all of the early vernacular literature! See Georg (2017: 377–378) for criticism of the sort of thing that sometimes passes for scholarship now. See also LaPolla (2016b).

References

Georg, Stefan. 2017. 'The role of paradigmatic morphology in historical, areal and genealogical linguistics, thoughts and observations in the margin of paradigm change. *In the Transeurasian languages and beyond* (Robbeets and Bisang, eds)'. *Journal of Language Contact* 10: 353–381.

LaPolla, Randy J. 2016a. 'On scholarship in Sino-Tibetan linguistics: review article on *Studies in Chinese and Sino-Tibetan linguistics: Dialect, phonology, transcription and text* (Language and Linguistics Monograph Series 53)' (Richard VanNess Simmons and Newell Ann Van Auken, eds). *Journal of the American Oriental Society* 136.4: 867–876.

LaPolla, Randy J. 2016b. 'Once again on methodology and argumentation in linguistics: Problems with the arguments for recasting Sino-Tibetan as "Trans-Himalayan"'. *Linguistics of the Tibeto-Burman Area* 39.2 (December 2016), 282–297. doi 10.1075/ltba.39.2.03lap

Thurgood, Graham and Randy J. LaPolla (eds). 2017. *The Sino-Tibetan languages*, 2nd edn. Abingdon and New York: Routledge.

INTRODUCTION TO VOLUME I
Establishing the relationships

In this volume we present several of the articles that led to the understanding of the Sino-Tibetan family that we now have (Part 1), plus some of the reconstruction efforts based on that understanding (Part 2).

Part 1: Establishing the relationships

The first chapter is from Fang-Kuei Li, who laid out a basic outline of the languages of China in an article for *Journal of Chinese Linguistics*, originally published in *The Chinese Year Book* published by Commercial Press in Shanghai, 1937 (pp. 121–128). This became a very influential view of the family, and so when William S-Y. Wang established the *Journal of Chinese Linguistics* (*JCL*), he included a slightly revised version of this paper as the first article of the first issue of *JCL* (Li 1973). In the article, Prof. Li includes Hmong-Mien and Kam-Tai within what he calls the Indo-Chinese family together with Chinese and Tibeto-Burman as major branches. Within Chinese he recognizes the Northern Mandarin group, the Eastern Mandarin group, the Southwestern Mandarin group, the Wu group, the Gan-Hakka group, the Yue group, the Xiang group, and the Min group, and also mentions isolated unnamed groups in Anhui, Guangxi, and Hunan, the latter recently investigated by Hillary Chappell (e.g. 2015). Prof. Li divides the Tibeto-Burman (TB) branch into the Tibetan group, the Bodo-Naga-Katchin group (presaging comments in Benedict (1976) and Burling (1983)), the Burmese group (within which he includes the Kuki-Chin languages), and the Lolo group.

Also in the late 1930s there was a government sponsored make-work project created by Robert Shafer to collect and compare materials on Sino-Tibetan languages. Paul Benedict later led the project, and writes about the project in Chapter 2 in this volume (Benedict 1975). It is included because of the pioneering nature of the work that was done on that project, and the foundation it gave to the development of Sino-Tibetan historical linguistics. Shafer and Benedict had somewhat different methodologies and came to different conclusions about some of the relationships within the family, reflected in the following

two chapters, Benedict (1942), which argues that the Kam-Tai and Hmong-Mien languages should not be included in the Sino-Tibetan family, and Shafer (1955), which is something of a response to that proposal, arguing that what Benedict was saying was actually his idea. Shafer argues that Chinese and the Tai languages ("Daic") do not form a unit, and he expresses doubt that they are related at all, yet he still includes Tai within the Sino-Tibetan family, arguing that the main divisions of the family are "Sinitic (Chinese), Daic, Bodic, Burmic, Baric, and Karenic" (Shafer 1955: 99). He gives cognate counts to argue for a more or less close relationship between certain divisions of the family, for example claiming that Bodic is closest to Sinitic, but his high cognate counts for Sinitic and Tibetan are simply due to a bibliographic bias, as they are the two languages with the largest dictionaries and so it is easier to find cognates. Fang-Kuei Li and several others still had trouble accepting Benedict's exclusion of Tai from Sino-Tibetan, pointing to unanswered questions about correspondence sets. See for example Li (1976) (in Volume II—see the discussion in the introduction), Li (1978), as well as Luo (2008) for the history of the debate and new data that Prof. Luo brings to bear on the question.

The work Benedict did on the project led to the monograph *Sino-Tibetan: A Conspectus* (initially drafted in the 1940s but edited by James A. Matisoff and published in 1972), which in turn stimulated the whole field of Sino-Tibetan comparative studies, and led to the development of a similar project on a greater scale, the Sino-Tibetan Etymological Dictionary and Thesaurus project led by Prof. Matisoff (see Matisoff (2003) and the STEDT website: http://stedt.berkeley.edu).

Chapter 5 is a follow-up to the re-printing of Li (1937) by Prof. Matisoff, published in the third issue of *JCL* (Matisoff 1973a). It is a critique of Prof. Li's outline of the Indo-Chinese family, arguing that based on Benedict's work, the Tai and Hmong-Mien languages should not be included in the family, now called Sino-Tibetan. Matisoff also argues that the outline of Tibeto-Burman given by Prof. Li is out of date, and gives the view current at the time, with Burmese and Lolo being closely related, and Kuki-Chin being more closely related to the Naga languages than Burmese or Lolo, Bodo-Garo being a separate group, and Katchin (Kachin/Jinghpaw/Chinghpaw) being "in a class by itself" (p. 473). Since that time much new data has appeared, and Matisoff himself has revisited the question of the position of Jinghpaw, arguing it is a Luish language (Matisoff 2013).

Chapter 6, "Sino-Tibetan: another look" was published by Paul K. Benedict (1976) four years after the publication of his *Conspectus*, and is due to the feedback and criticisms he received in reviews of that book and also from discussions at the annual Sino-Tibetan conferences (see Matisoff (1973b) for a review of the first five conferences, now called the International Conference on Sino-Tibetan Languages and Linguistics (ICSTLL), and LaPolla and Lowe (1994) for a bibliography of the first 25 conferences—the 50th ICSTLL was held in Beijing in November 2017). In the article he responds to critics and presents a large amount

of data, with the article structured around the following key questions: a) Is Sino-Tibetan clearly a family?, b) If so, does it include the Tai and Hmong-Mien languages?, and c) Should Chinese be set off against Tibeto-Burman (Tibeto-Karen) and then Karen from Tibeto-Burman? Benedict uses comparisons of core vocabulary (Swadesh 100 list) to assert that Sino-Tibetan does in fact form a clear family, but Tai and Hmong-Mien should not be included, and Chinese should be set off as a major branch from Tibeto-Burman, though there was contact with Tibeto-Burman languages after the split, and Karen should not be set off from the rest of Tibeto-Burman. There is also a lengthy discussion updating the *Conspectus* with new findings related to the influence of prefixes in Chinese and Tibeto-Burman languages.[1]

Chapter 7, Matisoff (1990), is largely a criticism of Greenberg (1987), but it is included here because it is a discussion of the methodology used in determining genetic relatedness, and also compares and contrasts Greenberg's methodology with that of Paul K. Benedict, treating Benedict's methods quite critically. It also introduces the concepts of "Indosphere" and "Sinosphere" to identify the cultural spheres of the two major influences on the Tibeto-Burman languages. This paper is frequently cited in the Sino-Tibetan literature on methodology in historical and contact linguistics. Chapter 8 (Benedict 1991) is a two-page reply by Benedict to Matisoff's comments about his methodology.

Chapter 9 is an overview of the state of the art of Sino-Tibetan linguistics around 1990, published by Prof. Matisoff in 1991. It covers the history of the development of the field and what had been done on all aspects of the family up to that point. Aside from genetic relations, it also includes discussion of areal and typological relations, including a discussion of the position of Tai and Hmong-Mien languages. It has copious (216!) references, giving a very good introduction to the literature up to that point. (See also the bibliography at http://tibeto-burman.net/bib/ for more references.)

Chapters 10–12 are influential articles that deal with establishing subgroups within Tibeto-Burman. The first of these is Robbins Burling's often-cited paper establishing the "Sal" languages as a separate grouping (1983). This grouping includes the Bodo-Garo and "(North)-Eastern Naga" (Konyak) languages, as well as Jinghpaw, and he compares forms from these languages with reconstructions in Benedict (1972) to show the uniqueness of some of the lexical items or their reflexes found in these languages. See Coupe (2012) for an up-to-date critical assessment of the different schemes for sub-grouping these languages.

Following that is Karen Ebert's (1990) paper showing the remarkable similarities in the person marking affix paradigms found in Kiranti languages, rGyalrong (Gyarong), and Rawang-T'rung (problematically referred to as "Nungish"). She is very conservative in her conclusions, simply stating that she found:

> striking similarities in the verbal paradigms of Gyarong and Eastern Kiranti on the one hand and Nungish and Khaling-Dumi on the other

> hand . . . an independent invention of the complex verbal paradigms of Gyarong and Eastern Kiranti is most unlikely.
>
> (p. 76)

She gives reasons why they are unlikely to be retentions from Proto-Tibeto-Burman, and argues that they must be a shared innovation. This hypothesis was later used as a test case for a more rigorous approach to establishing genetic relatedness in LaPolla (2013) (first presented at ICSTLL33 in 2000; see also LaPolla (2012)). The methodology is based on the fact that in Indo-European linguistics what was used to establish the family was not arbitrary word lists, but morphological paradigms (Nichols 1996). The results show strong statistical evidence that these three groups, plus the Western Himalayan languages, form one group within Tibeto-Burman due to the shared innovation of this paradigm, and the results also dovetail with work on the migrations of the Tibeto-Burman people (LaPolla 2001), where these groups are seen to have migrated from north-western China down along the river valleys skirting the eastern and southern sides of the Tibetan Plateau (see also LaPolla (2006, 2017) for the broader picture). The idea that this particular person marking paradigm[2] might be used to establish a subgroup within Tibeto-Burman was first mentioned by Paul K. Benedict in Chapter 6 (Benedict 1976), and LaPolla argues it defines what he calls the "Rung" group (distinct from the grouping by that name proposed in Thurgood 1984, 1985).

The last chapter in this part, Sun (1993a), is a summary of the findings of the UC Berkeley PhD dissertation of Jackson T-S. Sun (Sun 1993b). Through careful fieldwork and comparative work, it establishes the Tani languages (formerly known as Mirish or Abor-Miri-Dafla) as a branch within Tibeto-Burman. The dissertation also includes reconstructed Proto-Tani forms. Prof. Sun's work on this group has been accepted by all scholars in the field, and it has laid the foundations for work done since its publication by Mark Post and others in elaborating on this family (see Post and Sun (2017) for a summary, and Post and Burling (2017) for the larger picture of the languages of Northeast India).[3]

Part 2: Sino-Tibetan historical reconstruction

The first chapter in this part is a two-page note by Jean Przyluski and Gordon H. Luce (1931) on reconstructing the Proto-Sino-Tibetan (PST) form for "hundred". It is included here because it is one of the earliest published attempts at such reconstruction, and it is surprisingly modern in its approach, accepting the possibility of bisyllabic roots in the proto-language, when most scholars had assumed the roots to be monosyllabic.

Chapter 14, Wolfenden (1937), is a broader application of Karlgren's idea (1933—see Chapter 35 in Volume III) of comparing word families as a first step in identifying cognate forms. In this paper, Wolfenden compares word families in Kachin, Tibetan, and Chinese with a view to understanding variations in the

final consonants of the forms, and argues, for example, that Written Tibetan *-s* represents *-ds in a number of cases, based on comparisons with stop-final forms in the other languages. The idea of using word families as a way of identifying roots and cognates has become standard practice in the field, and as evidenced in many works assuming this methodology, it is the first step in identifying possible morphological forms (suffixes and prefixes and initial alternations), in that it gives us the forms we need to explain.

The particular question addressed by Wolfenden is returned to in LaPolla (1994), Chapter 15, though aside from giving many examples of word families, it discusses methodological issues related to the identification of word families and also to reconstruction methodology. Wolfenden, following Karlgren, limited his word families to only those where the final consonant had the same place of articulation, but LaPolla shows that this assumption is problematic, and would lead us to miss a large number of words that have the same meaning and form except for the final consonant. LaPolla also argues that instead of trying to account for all the variation found in the families by reconstructing overly complex systems or abstract symbols representing unknown variables, we should reconstruct a simple system (see the next paper for one example) and then try to explain the variation or accept it as is for the time being, if it cannot yet be explained. The paper then goes on to show the statistical preponderance of *-Ø ~ -k* variation, and attempts to explain this and other variations. The paper also argues for rigour in talking about variation, that is, variation can only be talked about in the context of regularity: if we want to say that two forms that differ in the final are cognate, we should make sure the other segments involved all correspond regularly, otherwise anything goes.

Our last chapter in this part, Gong (1980), is an independent comparison of just the three oldest written languages in Sino-Tibetan, Tibetan, Chinese, and Burmese, with a view towards reconstructing the vowels of Proto-Sino-Tibetan. Prof. Gong's methodology differs from many other scholars, in that a) he reconstructs on the basis of only three languages, and b) he does not treat the finals (rimes) as units in opposition to the initials, as is usually done in Sino-Tibetan studies, but in this paper looks only at the vowels, regardless of the final consonant (or not) of the word. His results are quite straightforward, with clear cognate sets and the proposal of a simple system of four vowels (i, a, u, ə) and two rising diphthongs (ia, ua). He argues that within Sino-Tibetan (or at least within these three languages) only Chinese maintained *ə, and the vowels e and o in Tibetan are secondary.

Notes

1 For other views of the structure of Tibeto-Burman or Sino-Tibetan not represented here, see DeLancey (1987), Matisoff (1991), Bradley (1997, 2002).

2 Person marking has been independently innovated many times in Tibeto-Burman (see LaPolla (2001)); the relevant paradigm used here to identify what is called the "Rung" group in LaPolla (2013) is only one of them, though is the oldest identified.

3 After his work on the Tani languages, Prof. Sun turned to fieldwork on and comparison of the rGyalrong languages of northern Sichuan and has similarly given us a much better understanding of that whole group of languages (e.g. Sun 2000, 2017), and in the process has also been investigating the Tibetan and Qiang dialects in the same area (e.g. Sun 2014), improving our understanding of the language situation of the whole of northern Sichuan.

References

Benedict, Paul K. 1942. 'Thai, Kadai, and Indonesian: a new alignment in southeastern Asia'. *American Anthropologist* 44: 576–601.

Benedict, Paul K. 1972. *Sino-Tibetan: A Conspectus (Princeton–Cambridge Studies in Chinese Linguistics II*; James A. Matisoff, contributing editor). Cambridge: Cambridge University Press.

Benedict, Paul K. 1975. 'Where it all began: memories of Robert Shafer and the "Sino-Tibetan Linguistics Project" Berkeley, 1939–40'. *Linguistics of the Tibeto-Burman Area* 2.1: 81–92.

Benedict, Paul K. 1976. 'Sino-Tibetan: another look'. *Journal of the American Oriental Society* 96.2: 167–197.

Benedict, Paul K. 1991. 'Comment on Matisoff's comparison between Greenberg and Benedict'. *Linguistics of the Tibeto-Burman Area* 14.2: 169–170.

Bradley, David. 1997. 'Tibeto-Burman languages and classification'. In D. Bradley (ed.), *Tibeto-Burman Languages of the Himalayas*. Canberra: Australian National University.

Bradley, David. 2002. 'The subgrouping of Tibeto-Burman'. In C. Beckwith (ed.), *Medieval Tibeto-Burman Languages*. Leiden: Brill, 73–112.

Burling, Robbins. 1983. 'The "Sal" languages'. *Linguistics of the Tibeto-Burman Area* 7.2: 1–32.

Chappell, Hilary. M. (ed.). 2015. *Diversity in Sinitic Languages*. Oxford: Oxford University Press.

Coupe, Alexander R. 2012. 'Overcounting numeral systems and their relevance to subgrouping in the Tibeto-Burman languages of Nagaland'. *Language and Linguistics* 13.1: 193–220.

DeLancey, Scott. 1987. 'The Sino-Tibetan languages'. In B. Comrie (ed.), *The World's Major Languages*. New York: Oxford University Press, 799–810.

Ebert, Karen H. 1990. 'On the evidence for the relationship Kiranti-Rung'. *Linguistics of the Tibeto-Burman Area* 13.1: 57–78.

Gong Hwang-cherng. 1980. 'A comparative study of the Chinese, Tibetan, and Burmese vowel systems'. *The Bulletin of the Institute of History and Philology* 51.3: 455–489.

Greenberg, Joseph H. 1987. *Language in the Americas*. Stanford, CA: Stanford University Press.

Karlgren, Bernhard. 1933. 'Word families in Chinese'. *Bulletin of the Museum of Far Eastern Antiquities* 5: 9–120.

LaPolla, Randy J. 1994. 'Variable finals in Proto-Sino-Tibetan'. *The Bulletin of the Institute of History and Philology* 65.1: 131–173.

LaPolla, Randy J. 2001. 'The role of migration and language contact in the development of the Sino-Tibetan language family'. In R. M. W. Dixon and A. Y. Aikhenvald (eds), *Areal Diffusion and Genetic Inheritance: Case Studies in Language Change*. Oxford: Oxford University Press, 225–254.

LaPolla, Randy J. 2006. 'Sino-Tibetan languages'. In Keith Brown (ed.), *Encyclopedia of Language and Linguistics, 2nd edn.* London: Elsevier, 393–397.

LaPolla, Randy J. 2012. 'Comments on methodology and evidence in Sino-Tibetan comparative linguistics'. *Language and Linguistics* 13(1) (*special issue on Sino-Tibetan Comparative Studies*), 117–132.

LaPolla, Randy J. 2013. 'Subgrouping in Tibeto-Burman: can an individual-identifying standard be developed? How do we factor in the history of migrations and language contact?' In Balthasar Bickel, Lenore A. Grenoble, David A. Peterson and Alan Timberlake (eds), *Language Typology and Historical Contingency*. Amsterdam/Philadelphia, PA: John Benjamins Publishing Company, 463–474.

LaPolla, Randy J. 2017. 'Overview of Sino-Tibetan morphosyntax'. In Graham Thurgood and Randy J. LaPolla (eds), *The Sino-Tibetan Languages, 2nd edn.* Abingdon/New York: Routledge, 40–69.

LaPolla, Randy J. and John B. Lowe. 1994. *Bibliography of the Papers from the International Conferences on Sino-Tibetan Languages and Linguistics*. STEDT Monograph Series, No. 1, 2nd edn, including Conferences I–XXV.

Li, Fang-Kuei. 1936–1937. 'Languages and dialects'. In *The Chinese Year Book*, 121–128; reprinted in *Journal of Chinese Linguistics* 1.1: 1–13, 1973.

Li, Fang-Kuei. 1976. 'Sino-Thai'. *Computational Analyses of Asian & African Languages* (National Inter-University Research Institute of Asian and African Languages and Cultures, Tokyo) 3: 39–48.

Li, Fang-Kuei. 1978. 'Siamese *khot*'. *Monumenta Serica* 33: 403–406.

Luo, Yongxian. 2008. 'Sino-Tai and Tai-Kadai: another look'. In Anthony V. N. Diller, Jerold A. Edmondson and Yongxian Luo (eds), *The Tai-Kadai Languages*. Abingdon/New York: Routledge, 9–28.

Matisoff, James A. 1973a. 'Notes on Fang—Kuei Li's "Languages and dialects of China".' *Journal of Chinese Linguistics* 1.3: 471–474.

Matisoff, James A. 1973b. 'The annual Sino-Tibetan Conferences: the first five years, 1968–1972'. *Journal of Chinese Linguistics* 1.2: 152–162.

Matisoff, James A. 1990. 'On megalocomparison'. *Language* 66.1: 106–120.

Matisoff, James A. 1991. 'Sino-Tibetan linguistics: present state and future prospects'. *Annual Review of Anthropology* 20: 469–504.

Matisoff, James A. 2003. *Handbook of Proto-Tibeto-Burman: System and Philosophy of Sino-Tibetan Reconstruction*. Berkeley, CA: University of California Press.

Matisoff, James A. 2013. 'Re-examining the genetic position of Jingpho: putting flesh on the bones of the Jingpho/Luish relationship'. *Linguistics of the Tibeto-Burman Area* 36.2: 15–95.

Nichols, Johanna. 1996. 'The comparative method as heuristic'. In M. Durie and M. Ross (eds), *The Comparative Method Reviewed*. New York/Oxford: Oxford University Press, 39–71.

Post, Mark W. and Jackson T.-S. Sun. 2017. 'Tani languages'. In Graham Thurgood and Randy J. LaPolla (eds), *The Sino-Tibetan Languages, 2nd edn.* Abingdon/New York: Routledge, 322–337.

Post, Mark W. and Burling, Robbins. 2017. 'The Tibeto-Burman languages of Northeast India'. In Graham Thurgood and Randy J. LaPolla (eds), *The Sino-Tibetan Languages, 2nd edn.* Abingdon/New York: Routledge, 213–242.

Przyluski, Jean and Gordon H. Luce. 1931. 'The number "a hundred" in Sino-Tibetan'. *Bulletin of the School of Oriental Studies* 6.3: 667–668.

Shafer, Robert. 1955. 'Classification of the Sino-Tibetan languages'. *Word* 11.1: 94–111.
Sun, Jackson T.-S. 1993a. 'The linguistic position of Tani (Mirish) in Tibeto-Burman: a lexical assessment'. *Linguistics of the Tibeto-Burman Area* 16.2: 143–188.
Sun, Jackson T.-S. 1993b. 'A historical comparative study of the Tani (Mirish) branch of Tibeto-Burman'. PhD dissertation, Berkeley, CA: University of California.
Sun, Jackson T.-S. 2000. 'Parallelisms in the verb morphology of Sidaba rGyalrong and Guanyinqiao in rGyalrongic'. *Language & Linguistics* 1.1: 161–190.
Sun, Jackson T.-S. (ed.). 2014. *Phonological Profiles of Little-Studied Tibetic Varieties. Language and Linguistics Monograph* Series 55. Taipei: Institute of Linguistics, Academia Sinica.
Sun, Jackson T.-S. 2017. 'Tshobdun Rgyalrong'. In Graham Thurgood and Randy J. LaPolla (eds), *The Sino-Tibetan Languages, 2nd edn.* Abingdon/New York: Routledge, 557–571.
Thurgood, Graham. 1984. 'The "Rung" languages: a major new Tibeto-Burman subgroup'. *Proceedings of the 10th Annual Meeting of the Berkeley Linguistics Society*. Berkeley, CA: Berkeley Linguistics Society, 338–349.
Thurgood, Graham. 1985. 'Pronouns, verb agreement systems, and the subgrouping of Tibeto-Burman'. In G. Thurgood, J. A. Matisoff and D. Bradley (eds), *Linguistics of the Sino-Tibetan Area: The State of the Art. Papers Presented to Paul K. Benedict for his 71st Birthday (Pacific Linguistics C, 87)*. Canberra: Australian National University, 376–400.
Wolfenden, Stuart N. 1937. 'Concerning the variation of final consonants in the word families of Tibetan, Kachin, and Chinese'. *Journal of the Royal Asiatic Society of Great Britain and Ireland* 4: 625–655.

Part 1

ESTABLISHING THE RELATIONSHIPS

1

LANGUAGES AND DIALECTS OF CHINA*

Fang-Kuei Li

Source: *Journal of Chinese Linguistics* 1, 1, 1973, 1–13; originally published in *The Chinese Year Book* (Shanghai, 1937), pp. 121–8.

The languages and dialects of China present a complicated linguistic picture. In this article these languages are systematically classified into families and branches according to their characteristics. The simplified picture looks thus:

1. Indo-Chinese: a) Chinese, b) Kam-Tai, c) Miao-Yao, d) Tibeto-Burman
2. Austro-Asiatic: Mon-Khmer
3. Altaic: a) Turkish, b) Mongolian, c) Tungus
4. Indo-European: Tokharian (extinct)

0. Introduction

The linguistic situation in China is a very complicated one. Aside from the Chinese language with its numerous dialects, there are many other languages, our knowledge of which is incomplete. Some of them have not been adequately studied, some of them are scarcely known to us, and many of them have not been sufficiently recorded. The material of these languages is therefore scanty, their history unknown, and their relation with other groups very vaguely understood.

In the following description the languages in China are grouped into branches and families, and under each group some of the characteristies which distinguish it from the other groups are called to the reader's attention. We need a few words of introduction about the classification of these languages. Languages are classified into families, with the assumption that they are historically related. Among a group of languages, say, the Indo-European languages, we find some similarities or correspondence in word forms, in grammatical elements such as prefixes, suffixes, vocalic and consonantal alternations, etc., and in general syntactic structure. These

similarities and correspondences can sometimes be formulated precisely in the form of sound laws. With the relations among the different languages thus formulated, it is apparent that these similarities cannot be due to chance or mere borrowing, but are due to the fact that these languages are the descendents of a common parent speech. Over a period of time, this parent speech splits into various dialects and through successive evolutions will develop into such languages as English, Russian, and Bengali. Sometimes the changes that these languages undergo are far reaching. The farther they have become differentiated from each other, the more difficult it is to trace their relations unless ancient documents which describe the older stage are available. With languages which have no records and which are known only in their modern forms, we encounter great difficulty in establishing their relationships. Many languages in China not only lack records of their own, but are known to us very fragmentarily. Their classification is therefore tentative.

1. Indo-Chinese family

One of the largest language families in China is known as Indo-Chinese (Tibeto-Chinese, Sino-Tibetan, or Sinitic in a wider sense). Languages of this family are spoken throughout China Proper and Tibet, and extend into northeastern China and Xinjiang (Chinese Turkestan), as well as into Southeast Asia. One of the characteristics of this family is the tendency towards monosyllabism. By monosyllabism we do not mean that all words in these languages consist of single syllables, but that a single syllable is an important phonologic unit and often is a morphemic unit, the structure of which is rigidly determined by the phonologic rules of the language, and serves as the basis for the formation of words, phrases, and sentences. The Tibeto-Burman branch of this family still possesses some of the prefixes, sometimes syllabic and sometimes asyllabic, but the Chinese and the Kam-Tai group have lost all active use of the prefixes quite early. It is probable that prefixes, suffixes, vocalic changes, and consonantal alternations have been in use in primitive Indo-Chinese to form causatives, denominatives, and to modify the meaning of the stem in general as in classical Tibetan: *hgeṅs-pa* (present), *bkaṅ* (perfect), *dgaṅ* (future), *khoṅ* (imperative) 'to fill'. Such derivative processes have long been dead in most languages of this family, so that we have only inflexible monosyllabic stems such as are found in Chinese now.

The tendency to develop a system of tones is another characteristic of this family. We do not know whether tones existed in early Indo-Chinese speech, and it is doubtful whether tones existed in classical Tibetan, However, modern Chinese, the Kam-Tai languages, the Miao-Yao languages as well as varieties of modern Tibetan all possess tones. These tones are influenced by the nature of the initial consonant. Such has been the case with Chinese, Tibetan, Burmese, the Kam-Tai, and the Miao-Yao languages, and has been considered the most powerful argument for the common origin of these languages.

Another phonetic tendency that this family of languages shares in common is the unvoicing of the original voiced initial consonants, but this has not taken place in all dialects. It occurs in most Chinese dialects except the Wu 吳 and the Xiang 湘,

in practically all the Kam-Tai languages, and in many Tibeto-Burman languages, so that what was originally a voiced initial is now retained in the nature of the tone.

Aside from these there are, of course, many items in the lexicon which are shared in common in this family. Exact correspondences of sounds, however, have not been worked out. Four main branches are known in the family: Chinese, Kam-Tai, Miao-Yao, and Tibeto-Burman.

1.1. Chinese

Chinese is the most important member of this family. The earliest records consist of numerous bone and tortoise shell inscriptions dating around 1400 B.C., the excavation of which has been systematically carried out in Henan. The reading of these inscriptions is in progress but still presents many difficulties. Our knowledge of the archaic phonological system is largely derived from a study of the rimes in archaic texts, principally the *Shi Jing* 詩經, and of the phonetic compounds which determine a class of written characters 諧聲字. It has been shown that initial consonant clusters such as *gl-, bl-, ml-,* etc., and many final consonants such as *-b, -d, -g, -p, -t, -k,* etc., existed in Archaic Chinese.[1] But up to about 600 A.D., when we have the system of Ancient Chinese well represented by rime books such as *Qie Yun* 切韻, the initial consonant clusters were already simplified and final *-b, -g, -d* dropped. From that time on the Chinese language underwent a series of evolutions such as the unvoicing of the initial consonants, the dropping final *-p, -t, -k* and the simplification of rimes. The modern Peking dialect which is the National Language 國語 now has only 400 some possible syllables, and each syllable may have theoretically four tones. This phonetic simplification, which gives rise to many homophones, is counterbalanced by a great increase in the use of compounds, so that what was formerly expressed by one syllable must now be expressed in the colloquial by two syllables.

We may divide the Chinese dialects into the following groups:[2]

(i) The Northern Mandarin group occupies a large area in North China, in the provinces of Hebei 河北, Shanxi 山西, Shanxi 陝西, Gansu 甘肅, Henan 河南, and Shandong 山東, and extends into Xinjiang 新疆, Inner Mongolia 內蒙古, and Manchuria 滿洲 in the north, and into Hubei 湖北, Anhui 安徽, and in Jiangsu 江蘇 in the south. It is characterized by the unvoicing of the ancient voiced stops, affricates, and fricatives, and by the disappearance of the 'entering tone' 入聲. There are as a rule only four tones: 'ying-ping' 陰平, 'yang-ping' 陽平, 'shang' 上, 'qu' 去. Further division into subgroups is possible.

(ii) The Eastern Mandarin group is spoken along the lower Yangtze in the provinces of Anhui 安徽 and Jiangsu 江蘇. It is differentiated from the Northern group by the existence of the 'entering tone' as a short tone, but the original final consonants *-p, -t,* and *-k,* which accompanied the 'entering tone', are substituted by the glottal stop. It has, therefore, five tones.

(iii) The Southwestern Mandarin group is a fairly uniform type of speech spoken in Sichuan 四川, Yunnan 雲南, Guizhou 貴州, and parts of Hubei 湖北 and Guangxi 廣西. It has as a rule no 'entering tone'. but in the central part of Sichuah 四川 along the Yangtze, the 'entering tone' is preserved as a special tone, but the final consonants have completely disappeared. Further division into subgroups is possible.

(iv) The Wu 吳語 group of dialects is spoken south of the Yangtze in Jiangsu 江蘇, Zhejiang 浙江, and in a few districts in the eastern part of Jiangxi 江西. It is characterized by the preservation of the ancient voiced stops, etc., as aspirated voiced consonants and by the preservation of the 'entering tone' as a short tone with the loss, however, of the final *-p, -t, -k* (or rather, substituted by the glottal stop). It often has six or seven tones.

(v) The Gan-Hakka group 贛客家 is spoken principally in the provinces of Jiangxi 江西 and Guangdong 廣東. It is characterized by the change of the ancient voiced stops, etc., into aspirated surds in all four original tone classes (aspirated in ping-sheng only in the three Mandarin groups.) The 'entering tone' is preserved and the final *-p, -t, -k* are more or less preserved according to different dialects and there are often six or seven tones. The Northern or Gan贛 group, particularly around Boyang Lake 鄱陽湖, has the tendency to voice all aspirated surds in connected speech. The Hakka 客家 group preserved the final consonants such as *-m, -p, -t, -k* much better. Settlements of Hakka people can be found in various districts in Guangdong 廣東 and Guangxi 廣西, and in Southeast Asia, and the South Seas.

(vi) The Min 閩語 can be further divided into two subgroups. The Northern group is spoken in the northern part of Fujian 福建 and the Southern group is spoken in the southern part of Fujian, in the eastern part of Guangdong, in Hainan Island 海南島, and in parts of the Leizhou Peninsula 雷州半島. It is characterized by the change of the original voiced stops, etc., into unaspirated surds, even in ping-sheng where the aspirated pronounciation is the prevalent one; by the preservation of the Ancient Chinese prepalatal plosives *t̂, t̂', d̂'* (知, 徹, 澄) as dental plosives which were the archaic forms from which the ancient prepalatals were derived; and by the preservation of final *-p, -t, -k* (sometimes in modified and simplified forms). It has as a rule seven tones. The Hainan dialects of the Southern group possess many phonetic peculiarities, possibly under the influence of an aboriginal speech, presumably a Tai language. Settlements of speakers of the Southern group (Amoy, Swatow, Hainan, etc.) may be found in large number in Formosa, Indo-China, Burma, Siam, Malay Peninsula, and the South Seas.

(vii) The Cantonese or the Yue group 粵語 is spoken in the provinces of Guangdong and Guangxi. It is characterized by the preservation of the final consonants *-m, -n, -ŋ, -p, -t, -k.* It has a system of eight, nine, or more tones. This distinction of long and short vowels, as in Cantonese, is also a special

feature. Certain distinctions of tone depend on the length of the vowel. Settlements of speakers of this group can be found in large numbers in Southeast Asia and the South Seas.

(viii) The Xiang group 湘語 is spoken principally in Hunan 湖南. The Ancient voiced stops, etc., are as a rule kept as truly voiced consonants (except the Changsha 長沙 dialect). The final *-p, -t, -k* are usually lost, but the 'entering tone' is preserved as distinct tone classes.

(ix) Certain isolated groups such as dialects spoken in the southern part of Anhui 安徽, certain dialects in Hunan 湖南 and in the northeastern part of Guangxi 廣西 may be mentioned here.

Aside from the phonological features specific to the groups mentioned above, there are also lexical items more or less peculiar to each of these groups, but these will not be discussed here. Among the various groups some are mutually intelligible.

1.2. Kam-Tai

The Kam-Tai branch is proposed by the author to include the Tai languages on the one hand and the Kam-Sui languages on the other.[3] The term Tai is used here in a narrow sense and does not include languages whose kinship has not been sufficiently clarified, such as Annamite (studied by Maspero) and Miao-Yao (studied by Schmidt). The Kam-Sui languages, on the other hand, can be shown to be definitely related to the Tai, but must have separated from Primitive Tai sufficiently early to develop their particular features. while the Tai languages have developed fairly uniformly. This is illustrated by words such as Kam *q'wa:u, k'wa:u,* Sui *q'a:u, k'a:u, ha:u,* Mak *la:u,* T'en *la:u* against *lau* 'wine' in all Tai languages, or Kam *pa:u,* Sui *pa:u, qa:u,* Mak *ka:u,* T'en *pa:u* against *k'au, kau, xau* 'horn' in the Tai languages. It seems proper, therefore, to include the Kam-Sui and Tai languages under one general group Kam-Tai, keeping the other related languages, Siamese, Lao, Shan, Lu, Nung, Tho. Zhuang, etc., under Tai. It may be noted that the name Tai with its various dialectical pronounciations is only used by a portion of the Tai speakers, and are not known to Tho 土, Zhuang 僮, Zhong-jia 仲家, and Dioi.

This branch is closely related to Chinese and possesses four tone classes analogous to the 'ping', 'shang', 'qu', and 'ru' of the Chinese. These four tone classes are each further divided into two according to whether the initial consonant was originally voiced or voiceless, so that the modern Kam Tai languages often possess eight tones. There are sometimes nine or more tones as a further development according to vocalic length (I have counted those tones with final *-p, -t,* or *-k* separately according to the customary treatment of Chinese tones).[4] It has a series of preglottalized consonants, limited to *'b, 'd,* and *'j* in Primitive Tai but far more extensive as by evidence in Kam Sui. Several Sui dialects possess *'b, 'd, 'm, 'n, 'ŋ, 'y,* etc., beside the usual *b, d, m, n, ŋ, y,* etc. Initial consonant clusters

such as *kl-, pl-,* etc., are preserved by some dialects to this day, i.e. Siamese and Tai languages of Wuming 武鳴 and Long-an 隆安 in Guangxi 廣西, but the original voiced stops, etc., have practically all become voiceless in the modern dialects. Word order in Kam-Tai is also slightly different from Chinese. For instance, 'good man' in Chinese becomes 'man good' in Kam-Tai. The earliest monument is a Siamese inscription in the thirteenth century. In China most of the Kam-Tai languages have no writing of their own, except those in Yunnan which use either the Shan alphabet (derived from Burmese) or one closely related to the Southern Tai alphabet, both derived from Hindu sources.

(i) The Kam-Sui group is spoken in southeastern Kweichow and in a few districts in northern Guangxi, and may be divided into four subgroups, Kam, 洞話, Sui 水話, Mak 莫話, and T'en 羊黃話. Initial consonant clusters like *kl, pl,* etc., are not allowed but must have existed. There is a series of voiceless nasals in Kam and Sui, but it disappears in Mak and T'en There is also a distinction of velar and palatal plosives in Kam and Sui, prepalatal and palatal in Mak, but confused in T'en. Furthermore, there is a series of preglottalized consonants, extensive in Sui, limited in Mak and T'en. but none in Kam. The lengthening of the corresponding short vowels in the Tai languages is apparent in this group in many words common to them both.

(ii) The Tai group may be divided into two subgroups.[5] (a) The Zhuang group consists of many dialects spoken in a great part of Guangxi 廣西 (known as Zhuang 僮 or Tho 土) and in the southern part of Guizhou 貴州 (known as Zhong-jia 仲家), Man 蠻, Bendi 本地, or Dioi, and also in the southeastern part of Yunnan 雲南 (known as Sha 沙 or Tho 土). The language of Shu Li 熟黎, spoken in the northern part of Hainan Island, in Ling-gao 臨高, Chengmai 澄邁, and Qiongshan 瓊山, belongs here also, but the Li 黎 languages in the center and in the south of the island seem to show great divergence from the rest of the Tai languages. Their relation to this group is therefore doubtful. The languages of this group are characterized by the lack of aspirated surds such as *p', t', k';* the preservation of the distinction between original *k'-* and *x-, g-,* and *ɣ-*; and by the preservation of an original *hr-, thr-,* etc. as *r-* (in Wuming 武鳴), as *l-* (in Tianzhou 田州), as *ð-* (in Dioi of Guizhou 貴州), or *ɣ-* (in Qianjiang 黔江), corresponding to the *h-* of Shan, Siamese, Lao, Nung, etc. The development of vowels also shows many peculiar features.

(b) The Southwestern group consists of some of the best known of the Tai languages and lies mostly outside of China. We may divide this group into several subdivisions: (1) Ahom, once spoken in Assam but now extinct, (2) Kamti and Shan, spoken in Burma and western Yunnan, (3) Siamese and Lao spoken in Thailand and Laos, (4) Lü spoken in southern Yunnan, (5) Thai Blanc, Nung, Tho, etc. spoken in Laos and Cambodia, in the southwestern part of Guangxi, and in southern Yunnan. This group is characterized by the preservation of aspirated consonants, *p'-, t'-, k'-,* by the change of the

original guttural spirants *x-* and *ɣ-* into stops, by the appearance of *hr-* as *h-* (except Ahom where *r-* is preserved), and by a very uniform system of vocalic correspondences shared between them.

1.3. Miao–Yao 苗傜

The Miao-Yao branch of the Indo-Chinese family is monosyllabic like Chinese and Kam-Tai, and is known to possess tones.[6] The relationship between Miao 苗 and Yao 傜 seems to be definitely established with a study of the Yao languages in southern Guizhou where they are not so strongly influenced by Chinese or Tai as in Guangdong, Guangxi, Laos, and Cambodia. Word order resembles Kam-Tai. It is spoken by groups of mountaineers throughout the southwest. Aside from the occasional use of Chinese characters, there is no writing system of their own.

(i) The Miao 苗 group is spoken under various tribal names in the western mountain regions of Hunan, in a large part of Guizhou, and is found scattered here and there in northern Guangxi, southern Sichuan, Yunnan, Indo-China, and Thailand. It is characterized by the dropping of the final consonants, so that only *-ŋ* and rarely *-n* are allowed to occur in final position. There is a distinction between palatal and velar consonants, *k-, q-,* etc.; a series of prenasalized consonants, *mp-, mp'-, nt-, nt'-, ŋk-, ŋk'-,* etc.; and consonant clusters, *pl-, pr-, mpl-, mpr-, tl-, kl-,* etc., are still preserved by some dialects. The number of tones is usually eight or more. The Hei-Miao 黑苗 chiefly spoken in southeastern Guizhou seems to form a special subgroup. It allows no consonant clusters and no prenasalized consonants, but possesses a bewildering number of aspirated consonants, *p'-, t'-, k'-, tɕ'-, q'-, f'-, s'-, ɕ'-, t'-, ɱ'-, ɳ'-,* etc.

(ii) The Yao 傜 group is also spoken under various tribal names in the northwestern mountain regions of Guangdong, in southern Guizhou, and is scattered here and there among the various mountain regions of Guangxi, Yunnan, Indo-China, and Thailand. It preserves the final consonant better than the Miao; final *-m, -n, -ŋ, -p, -t, -k,* all exist. The number of tones varies from five to eight or more. It is greatly influenced by Tai and Chinese, and some have entirely adopted either the Chinese or Tai language.

There are certain features in common among the Chinese, Kam-Tai, and Miao-Yao groups. Notably the word order, subject-verb-object, stands in contrast to the Tibeto-Burman branch where we have subject-object-verb. The system of tones in Chinese and Kam-Tai consists of originally four tone classes, and this may ultimately prove to be the case with Miaŏ-Yao. It seems possible to group them together under one branch, and it does not seem improper to give it the name Sinitic, since all the Kam-Tai and Miao-Yao languages have deep relationships and close contacts with China historically, geographically, and culturally. Annamite

(Vietnamese) may possibly be included in this group, although it shows strong affinities with the Mon-Khmer languages.

1.4. Tibeto-Burman[7]

This branch of the Indo-Chinese family is one which presents most clearly the use of prefixes, alternations of voiced and voiceless consonants, and the use of suffixes such as in classical Tibetan. Tones depend upon whether the initial is voiced or voiceless and are further influenced by the prefixes, but the system of tones seems to be much simpler than that of Chinese, Kam-Tai, or Miao-Yao. As a rule word order is subject-object-verb. Four known divisions comprise the Tibeto-Burman branch:

(i) The Tibetan group is spoken principally in Tibet and Xikang 西康 and extends into Qinghai 青海 and the western part of Sichuan 四川. The earliest record of this group dates from the ninth century. The alphabet was derived from the devanagari form of the Hindu alphabet. A great amount of literature, largely Buddhistic, exists. The main groups of dialects may be distinguished. The Western group, Balti, Ladak, etc. preserves to a great extent the prefixes, initial consonant clusters, and final stops, generally transcribed as *-b, -d, -g*. The Central dialects, including that of Lhasa, are characterized by the loss of prefixes, simplification of consonant clusters, and the dropping of final consonants. The eastern dialects, the Khams and the Jarong, preserve the prefixes and final consonants very faithfully. There are some Tibeto-Himalayan dialects and some north Assam dialects spoken along the southern border of Tibet, and some Xifan dialects spoken in Xikang and Qinghai which belong to the Tibetan group. Of interest are Trung and Nung, called by Chinese Quzi 曲子 and Nuzi 怒子, in the northwestern corner of Yunnan 雲南. Like some Nepalese dialects, the reduced forms of the personal pronouns are used as prefixes and suffixes of the verb to form verbal conjugations as in Trung.

ŋa	ɣaŋ	dza	k'ai	tɕia-ŋ	'I can eat' < I food eat can
na	ɣaŋ	dza	k'ai	nə-tɕia	'you can eat'
ɣaŋ	ɣaŋ	dza	k'ai	tɕia	'he can eat'
ʔiŋ	ɣaŋ	dza	k'ai	tɕia-i	'we can eat' (pl.)
ne niŋ	ɣaŋ	dza	k'ai	nə-tɕia-n	'you can eat' (pl.)
ɣaŋ niŋ	ɣaŋ	dza	k'ai	tɕia	'they can eat' (pl.)

(The preceding examples were kindly furnished by Kun Chang.)

(ii) Katchin of the Bodo-Naga-Katchin group is spoken in the northwestern border of Yunnan.

(iii) Speakers of the various languages of the Burmese group, Burmese Kuki-Chin, 'Old Kuki', etc. are found mostly in Burma and Assam.

(iv) Among the Lolo group, Lolo and its dialects is spoken in a large portion of Yunnan, in northwestern Guizhou, and in southern Sichuan and and Xikang. It extends into Indo-China and Thailand. Lolo has an independent syllabic writing of its own, used largely in religious texts. Moso is spoken in the northwest of Yunnan and extends into Xikang. It possesses two systems of writing, one hieroglyphic and the other syllabic, like Lolo. This group is characterized by the simplification of the phonetic system, such as the complete disappearance of the final consonants and the rarity of dipthongs. Tones are usually five or six in number, and word order resembles Tibetan. Minkia may possibly belong to this group, but it shows strong Chinese influence in its vocabulary and word order, and its relationship remains doubtful.

2. Austro–Asiatic family

Of this large family, proposed by P. W. Schmidt, to which Munda, Mon-Khmer, and according to some authors, Annamite belong, we only mention the Mon-Khmer group of which there are representatives in China.[8] The earliest records of this group are some Khmer inscriptions of the seventh century and a Mon inscription of the eleventh century, the alphabets being derived from Hindu sources. This group of languages has no tones, and makes use of prefixes and infixes for the derivation of words. The stem is generally monosyllabic; the word order is subject-verb-object.

Dialects of this group spoken in China are Palaung, Wa, and some others along the Yunnan-Burmese border. We know very little about Wa, but Palaung is known to have no tones, and has a number of prefixes, both syllabic and asyllabic, *p-, pan-, ra-, kar-,* as in *yam* 'to die', *p-yam* 'to kill', *pan-p-yam* 'the killing, one who is killed'. A special series of initials: *hl-, hr-, hm-,* and *hn-,* exists. The language shows close resemblance to the Tai languages.

3. Altaic family[9]

This family of languages is spoken along the northern territory of China from Chinese Turkestan through Mongolia to Northeastern China. It extends further southwest to Asia Minor and northeast to Siberia up to the Arctic coast. It consists of three main branches of languages, Turkish, Mongolian, and Tungus. The relationship among the three groups has not been decisively established, although their phonetic structure, syntax, and vocabulary show great similarities. The exclusive use of suffixes, either derivative or syntactical, is one of the characteristics, so that the stem or root always remains at the beginning of a word. The suffixes are loosely joined one after another according to a fixed order, for example, Turkish, *baba* (father) *-lar* (plural) *-um* (our) *-dan* (from) 'from our fathers'.

A specific phonetic feature known as vocalic or vowel harmony operates on the Altaic languages. Briefly, all vowels within a word must be either front vowels,

i, y, e, ö, or back vowels, *ï, u, a, o,* as in Turkish *sev-il-dir-eme-mek* 'not to be able to cause to be loved' and *jas-il-dir-ama-mak* 'not to be able to cause to be written'. Different dialects, of course, possess slightly different rules for vocalic harmony. Vocalic harmony may influence the consonants, i.e., palatals in the neighborhood of front vowels and velars in the neighborhood of back vowels.

Word order is subject-object-verb, the verb always occurring at the end of the sentence. Modifying words are placed before the modified.

3.1. Turkish

This branch of the Altaic family is spoken in China in Chinese Turkestan, in the northwest corner of Mongolia, and in certain parts of Gansu. The Turkish dialects are divided into several groups, but the differences among them are slight. The eastern dialects are characterized by the wide application of the rules of vocalic harmony and by the existence of only surds in initial and final positions, and only voiced consonants in medial position. Dialects spoken in the northwest corner of Mongolia, (Tangnu Uriankhai) belong to this group. The central dialects possess voiced initials and have an indifferent /i/ with respect to vocalic harmony. They are chiefly spoken in Chinese Turkestan, and include Tarantchi, the dialects of Hami, Aksu, Kashgar, and Yarkand, among others. The dialects spoken in the northern part of Xinjiang 新疆, such as Kirghis, belong to the western group. The southern group of the Turkish dialects is not represented in China. The oldest texts are of some Siberian inscriptions from the eighth century, and a Latin-Persian-Turkish vocabulary from the fourteenth century. Several forms of writing have been known to be in use, including the Runiform, Uigur, Brahmi, Tibetan, etc., but most dialects have adopted the Arabic alphabet under the influence of the spread of Mohammedanism.

3.2. Mongolian 蒙文 [10]

The Mongolian language is centered in Mongolia and extends to central Asia in the west, to Siberia in the north, and to the northern provinces of China in the south. We find here, as in Turkish, the use of suffixes; vocalic harmony, although distorted in certain ways, is still observable. There are several divisions of Mongolian, although the differences among them are very slight. The Khalkha group occupies a vast area in Outer Mongolia 外蒙古; the Buriat group is spoken chiefly in Siberia, but also in certain parts in northern Mongolia and in the western part of Heilongjiang 黑龍江; the Kalmuck group is spoken in western Mongolia and in the northern part of Xinjiang 新疆; the Southern (or Eastern group is spoken in Chahar 察哈爾, Suiyuan 綏遠, Rehe 熱河, Ningxia 寧夏, and in some parts of Manchuria. Some Mongolian dialects are spoken in Qinghai 青海 and Gansu 甘肅; another one is spoken outside of China in Afghanistan.

The differences among these groups are slight. The most important one deals with the treatment of the palatal affricates *dz-* and *ts-* of literary Mongolian. The Southern group preserves the palatal position, Khalkha changes them to *dž-* and

tš except before *i.* Buriat changes *tš-* and *ts-* into *š-* and *s-* respectively, Kalmuck changes *dz-* into *z-*.

The Mongols adopted the Uigur alphabet in the thirteenth century. It is still in use with slight modifications. It writes from top to bottom like Chinese, but begins from the left side of the page. During the thirteenth and the fourteenth centuries another alphabet known as the hphags-pa, derived from Tibetan, was in use, but was soon discarded.

3.3. Tungus

The Tungus branch is spoken in eastern Siberia and northern Manchuria. An exact classification of the dialect is impossible on account of the lack of material. It is generally known to consist of two groups; the Northern group and the Southern group. Manchu, Gold, Oroch, and Solon, which form the Southern groups, are spoken in the provinces of Heilongjiang 黑龍江 and Kirin 吉林. Manegir and Birar, of the Northern group, are spoken in Heilongjiang also. Most of the languages of the Northern group are spoken in Siberia A small group of Manchu speakers, the descendents of earlier Manchu garrisons, is found in Ili in Xinjiang.

The best known language of this group is Manchu 滿文. The writing is derived from Mongolian in the sixteenth century with slight modifications. Vocalic harmony in Manchu does not follow a uniform set of rules; we may however, summarize it briefly thus:

Back vowels: a, o, o
Front vowels: ä, ü
Indifferent vowel: i

In the above description of various languages in China I have purposely avoided many of the tribal names, particularly of the southwest, which are numerous and confusing. Our present knowledge of the many languages and dialects in China is still very limited. Field work and systematic studies of the various language groups should be done before we can have a more complete picture of the linguistic situation in China.

4. Extinct languages

Aside from the living languages in China, which were presented in this paper, there are several languages which have left records but are no longer in existence. Among them is Tokharian 吐火羅, an Indo-European language which was once spoken in Chinese Turkestan.[11] It seems to form an independent group among the Indo-European languages. It was written with a kind of Hindu alphabet, and there are a number of texts, mostly fragmentary. Some speakers of an Iranian dialect may still be found in western Chinese Turkestan. Another important language was Xixia 西夏, once spoken in Gansu.[12] It was written with

a system of characters evidently modeled after Chinese, but much more complicated. The deciphering of these texts, largely Buddhistic, is still in the early stages, but the language seems to belong to the Lolo group. Others such as the language of the Kitan 契丹 are still less known, although some inscriptions resembling those of the Xixia have been found.

Notes

* [A condensed version of this article first appeared in the Chinese Year Book, Shanghai, 1937. Since that time, it has been regarded as the standard reference on the subject. Although research over the past three decades has brought us a more refined understanding of the individual dialects, Li's broad outline remains essentially accurate and useful. The present version was contributed by the author with minor modifications of the original article, change of Wade-Giles romanization to the Pinyin system, and adapted for the Journal of Chinese Linguistics by Teresa M. Cheng - Editor]

1 For further details see Karlgren 1954 and 1957, Pulleyblank 1962 and 1963, and Li 1971.

2 For recent works on Chinese dialects, see Yuan et all 1960.

3 See Li 1964.

4 For a more detailed treatment of initials and tones see Zide 1966, Pp. 82–8.

5 More recently the author has divided the Tai group into three subgroups. See Diamond 1960, Pp. 951–9.

6 For a thorough study of Miao-Yao tones, see Chang 1973 (in print) where an extensive biliography of recent works on Miao-Yao may be found.

7 A recent general treatment of the Tibeto-Burman languages may be found in Benedict 1972.

8 See Schmidt 1905, and Zide 1966.

9 See the recent work of Poppe 1960.

10 See Poppe 1955.

11 See Pedersen 1941.

12 Xixia studies have made great progress in recent years. See Nishida 1964–66, and Sofronov 1968.

References

BENEDICT, P. K. 1972. Sino-Tibetan, a Conspectus. 1–152.

CHANG, KUN. 1973. The Reconstruction of Proto-Miao-Yao Tones. BIHP. Academia Sinica. (in press)

DIAMOND, S. 1960. Culture in History. 951–959.

KARLGREN, B. 1954. Compendium of Phonetics in Ancient and Archaic Chinese. BMFEA 26.211–267.

———. 1957. Grammata Serica Recensa. BMFEA 29.1–332.

LI, F. K. 1964. The Tai and the Kam-Sui Languages. Lingua 14.1.148–179.

———. 1971. Shang-gu Yin Yen-jiu (Studies on Archaic Chinese). Tsing Hua Journal of Chinese Studies, N.S. 9.1–2.1–61.

NISHIDA, TATSUO. 1964–66. A Study of the Hsi-hsia Language, 2 vols.

PEDERSEN, H. 1941. Tocharisch.

POPPE, N. N. 1955. Introduction to Mongolian Comparative Studies.

———. 1960. Verleichende Grammatik der Altaischen Sprachen, Teil 1, Vergleichende Lautlehre.

PULLEYBLANK, E. G. 1962. The Consonantal System of Old Chinese. Part 1, AM 9.59–144.

———. 1963. The Consonantal System of Old Chinese. Part 2, AM 9.206–265.

SCHMIDT, W. 1905. Grundzuge einer Lautlehreder der Mon-Khmer Sprachen.

SOFRONOV, M. V. 1968. Grammatika Tangutskogo Yazyka. 2 vols.

YUAN, JIA-HUA et al. 1960. Hanyu Fangyan Gaiyao.

ZIDE, N. H. 1966. Studies in Comparative Austroasiatic Linguistics.

2

WHERE IT ALL BEGAN

Memories of Robert Shafer and the "Sino-Tibetan Linguistics Project" Berkeley 1939–40

Paul K. Benedict

Source: *Linguistics of the Tibeto-Burman Area* 2, 1, 1975, 81–92.

I first met Robert Shafer in the winter of 1938–39, after I had gone to Berkeley to assume supervision of the "Sino-Tibetan Philology Project." I had been invited by Professor A. L. Kroeber, at the time the dominant figure in American anthropology, to take over where Shafer had left off, completing the work of the project. Kroeber was under the erroneous impression that I was a master of Oriental languages, hence his apparent delight when I readily consented to leaving my doctoral (anthropology) program at Harvard for the Berkeley scene. I had other — largely extraneous — reasons for wanting to be on the Berkeley campus, so I eagerly accepted the offer and set off for Berkeley in the late fall of 1938.

Shafer was not in town when I arrived and I had to fend for myself. Kroeber greeted me with much warmth and even enthusiasm, and naturally took me around to the Project offices at the first opportunity. Being an impostor of sorts, I felt some trepidation on this first visit. Apparently I carried it off well enough, though, since I have reason to believe that Kroeber never became aware of the deception — nor Shafer later.

The Project was located in a small frame house (see photo) near the campus, with work being carried out in two or three large rooms, as I recall. Kroeber had assured me that everything was shipshape, and that in view of my extensive knowledge of the field there should be no problem about completing the Project. Knowing that the second part of this proposition was fallacious, I was hoping that the first part would hit the mark, but my first visit to the Project dispelled that notion also. I was confronted by what appeared to be mountains of linguistic data in all shapes and forms: dictionaries, articles (some torn out of journals, others offprinted or xeroxed), and a vast array of papers, much of it covered with what I quickly came to recognize as the Shafer scrawl — firm enough and fairly legible.

(JAM informs me that much of it is still to be found in the margins of books in the Berkeley library.) There were piles of papers everywhere, on the floor as well as on tables, but hardly any appeared to be other than the roughest kind of "working paper." Shipshape it was not. Organized, maybe, but only in the sense that any of Shafer's productions were ever organized. Shafer himself, I was later to learn, took a certain pride in his ability to "organize" materials, but certainly for me, at that small crisis in my academic life, all was a monstrous mess that left me utterly dismayed. I strongly considered simply telling Kroeber to go find himself another whipping boy — but then, as I have said, there were compelling reasons for me to stay on campus. So I smiled as brightly as possible under the circumstances, assured Kroeber that everything was under control, and set about seeing what I might be able to salvage from the wreckage (as I saw it).

It wasn't easy — especially for an impostor! At that time, despite my spurious reputation, I had only a mediocre knowledge of Chinese, an even poorer acquaintance with Japanese and a smattering of Vietnamese, having worked on the language earlier that year (1938) in Hanoi. But I hadn't the slightest scrap of knowledge about Thai, Tibetan, Burmese or the like. I remember Kroeber escorting me about the Project offices, waving at various mountains of material here and there, and ending up with expressions such as, "But of course you know all that, Benedict." A worker came up to us with a question about Burmese transcription. Having never gazed upon Burmese before, I was rather taken aback at this, but I did come up with a quick, arbitrary response (luckily I had been given a choice, say, between k- and kh-), and we continued on our journey of inspection. I simply should have packed my bags and left Berkeley. But I decided to stay, to go along with the unwitting deception. I managed to bluff my newly acquired staff, who had been "set up" for it by the reputation that had preceded me to Berkeley, by working frenetically, often long into the night after the staff had left, acquiring sufficient mastery of these "other" languages for carrying on comparative linguistic work. By the time that Shafer had arrived on the scene, a matter of a few months, I had become a complete "expert" on Oriental languages, willing to discuss anything from Tibetan a-chung to Burmese auk-myit.

Shafer never really filled me in on the details concerning his leaving the Project, nor his subsequent return to Berkeley, nor do I recall any comments from Kroeber on the subject. One of the stories had him inheriting a large sum of money from a deceased uncle, but he never appeared well off, even by the modest standards of a scholar. He spent a good deal of time around Berkeley, hovering about the Project in the manner of a midwife about an expectant mother. The Himalayish volumes had been completed or were in the final stages of typing, and the Tibetan materials were also in fairly good shape, so that Shafer was able to give these volumes his last, loving touches. We spent a considerable amount of time together, often along with Don Walters, the senior (and only) "linguist" on the Project staff, mostly in the Project offices or in nearby lunchrooms. The talk was largely "shop talk" with an admixture of politics.

I never really got acquainted with Shafer on any personal basis, and his private life hardly ever entered into our conversations. I never met anyone whom he

described as a friend, nor remember his talking about friends, and in general he impressed me as an isolated person in many ways. He was amiable enough, however, and highly verbal, talking in rather staccato fashion, very intent in manner, with flashing brown eyes darting ceaselessly about. He walked in a brisk manner, much like the late President Truman on his morning strolls, even bouncing at times. The general picture was that of a tireless, machine-like man, constantly ferreting about, forever asking questions and looking for answers. As I write this, I realize that I am also drawing the portrait of a newspaperman, or perhaps the caricature of one, as played by a Cary Grant. Shafer was a newspaperman — or should we say, he had been one — this much I learned, but little more. At heart, however, he was a linguist, and in later years he turned to teaching, not to newspaper work, when his fortune (if he ever had one) became depleted. He retained the cigars, along with a certain carelessness in dress, from his newspaperman days, and talked about politics and politicians (he was against them both, as aren't we all) when not holding forth on his favorite subject of languages and linguistics. In the photograph he is shown sitting on the right, socks about his ankles. I remember him as sitting in strange postures, more or less folded up and "holding on," possibly another reflection of his sitting about newspaper offices.

Shafer's attitude toward me was very much that of a father yielding up his child to a foster parent. He wanted to see the Project completed, but at the same time he appeared to be clinging to it, as if reluctant to entrust this prized possession to a stranger. I do not recall his actually interfering in the operation of the Project, but he did, as I have said, hover about much of the time. I remember him as patient in dealing with me, as he had always been with his staff, whether in explaining his system of transcription (see below) or in advising me about some detail of the operation. He appeared to see me as something of a young upstart (I was nearly a generation younger and already the "perpetual graduate student"), and, very clearly on occasion, as a member of the Eastern Establishment. At times he would become almost irrational on the subject of Harvard and the Establishment élite, to the point where I would feel compelled to defend a social institution about which I harbored my own misgivings. He was also anti-academic to a fault, given to long discourses on the smugness of the professional scholars and their incestuous relationships whereby all honors — and opportunities for publication — were kept within the academic family. I would listen to these expostulations with only mild demurrals from time to time. Over and over the theme of publication would come up, with Shafer insisting that the Project volumes would never be published (here you can still get odds) and that none of his articles would ever be accepted. Although I half suspected that he was right, I kept urging him to send in articles for publication. As we now know, Shafer did send in articles — and articles — and books — and they got published here, there and everywhere, with the ironic twist that Shafer finally succeeded in proving that he was wrong about the Establishment. Or was he, altogether?

Shafer's work habits were remarkable. I myself had acquired something of a reputation as a hard worker, and I do admit to having been afflicted with bouts of exhausting work — a condition which has continued to cast a shadow over what

might otherwise have been a happy, carefree life — but I found myself feeling positively otiose when confronted with Shafer, who appeared to live one long life of work interrupted only by hasty eating, scant sleep, and probably at times other of the more rudimentary physiological functions — in other words, the nightmare of the graduate student struggling with his Ph.D. dissertation. But the man seemed to thrive on this diet of work, and certainly never complained about it. He habitually carried about with him an assortment of papers of one sort or another, seemingly in complete disarray, along with reprints, pamphlets or the like, and when not otherwise engaged (usually in talking) he would work on these papers, scribbling notes — he was a great marginal note writer – or roughing in the outlines of a table. I am not sure just what his record-size chart was, but I recall having seen tables composed of several standard size (8 1/2 × 11) sheets of paper stuck together somehow, at times trailing behind him to the floor as he moved about or spread across a desk — or two! Shafer thought primarily in terms of tables, the longer the better — and damn the white spaces! As told to me by Walters, who was much closer to him, Shafer would work on these tables for hours at a time, into the small hours of the morning, then return to the task after a brief sleep. He would also spend countless hours scrounging up sources — and I use the term advisedly. Shafer must certainly be rated as one of the great scroungers of all time, determined and tireless, a master of the art. The unrivalled sources at the disposal of the Project were a monument to this aspect of Shafer's genius, just as the huge amount of research turned out on the Project was a testimony to his prodigious capacity for work, at a level hardly ever reached by an amateur, perhaps never by a professional.

Shafer and his Project staff were something else again. I write "his" staff rather than "my" staff, because all these people had been hired by Shafer and a kind of family bond had developed, still very much in evidence when I arrived on the scene. All were on a first-name basis, and there was a certain amount of socializing outside the office, which at times had also included Shafer in the days before he left the Project. In many ways I became a kind of younger sibling (only Walters, perhaps, was younger than I), despite my role as director of the Project. There was very little disciplining to be done, since this was a "make work" project, as were all WPA operations, and the general idea was that any work that could be extracted from the staff was better than the dole system. I had anticipated finding a group of complete non-workers, as one might encounter around a political clubhouse, and I was pleasantly surprised to discover that the staff actually did work, mostly to please Shafer, the father figure (or was he only an older brother figure?). They had been trained to do specific things, and took a special pride in their "specialties." I do not recall all the details, or all the names, but I do remember the people, as I look at them now in the photograph.

The two Chinese, at least one of whom was a Lee, formed their own little Chinese clique within the extended family, much as our Sinologist brethren constitute a distinct grouping within our great Sino-Tibetan family (pace everyone). They were well trained, but unfortunately as laundrymen rather than Sinologists — as

in the Army, WPA assignments did not necessarily match the individuals' backgrounds! They did know Chinese characters — at least some of them — and this gave them a certain allure for Shafer, who never really got into this arcane field of study. I never quite understood just what their function on the Project was. Shafer seemed to be groping toward a comparative study of Chinese dialects, anticipating later developments in this area, and I believe that these two workers might have been developing Cantonese materials for this aspect of the Project. My knowledge of characters, rather better than theirs, gave me a special relationship to them, but I was never able to satisfy myself as to just what they were up to on the Project — and Shafer, if he knew, never revealed the secret to me. They did keep busy enough, mostly producing what passed for Chinese calligraphy, and I quickly learned that it was better to be discreet about pressing my workers for details.

The second person from the left, a tall and rather taciturn man by the name of Frank or Henry (I think), also remained something of an enigma for me. He talked in a grave manner about various aspects of the Project, and perhaps was better educated than the average worker (that would be high school graduate), but I'm not sure that he was ever really trained for any specific function. As I remember, I had him "looking for" various things for me, from among the mountainous piles of papers littering the offices, and I believe that Shafer had used him a good deal as someone to talk to — Shafer tended to have unidirectional conversations — so it can be seen that the man served some valuable functions.

The third person from the left is Marie, a pleasant French housewife. When I arrived on the Project she was busy erasing paper. Shafer had a tendency to mark up piles of paper with headings, e.g. "East Himalayish Final *-ak," intended for tables but frequently left with just headings. The goal of any WPA project was to hire people, and if anything useful (a road leading from Podunk to Podunk Corners, park benches for the elderly in Paducah, Sino-Tibetan Philology) came of it, well so much the better. I quickly learned, as Shafer had before me, that WPA funds were for people, not things. So we scrounged for everything — and Marie erased paper — as Shafer had performed prodigious feats in obtaining the necessary source materials for the Project. Finally the day came when all the paper had been erased, and I then put Marie to work at a job in which her language facility would be of some help, probably copying entries from some French source, thus affording her proper status in the extended family.

Next to her in the photograph, fourth from the left, is "Smiddie," whose lack of formal education was especially in evidence. As I recall, he worked mainly with the typists (none shown in the photograph), serving as a proofreader of sorts, a function which he carried out much better than one would have anticipated.

Finally, the sixth from the left in the photo is "Doc," a man of sufficient bearing to have gained that title, and possibly college-educated to some extent. (I believe that none of us pressed him for details of his education.) He had been ensconced as the reigning Tibetan expert on the Project, and actually knew how to transcribe Tibetan. Shafer had used him primarily in working with Tibetan dialects, and he

reacted to questions in other areas like a true expert - "Please, ask! I do Tibetan." "Doc" also had attained a kind of minimal status as a linguist, although well below that reached by Walters. Doc had a penchant for turning up cognate pairs involving English words like "hare" and "hair" — i.e. if a word meant "hare" in Language A, and "hair" in Language B, they were potential cognates, if they bore some similarity in phonetic appearance. Walters had been assigned by Shafer to comb such nits as this out of the system — and there were lots and lots of nits and nuts and bolts and the like, many entered carefully onto "official" Project papers, to be expunged finally by Shafer himself, or later by me.

As I have said, not everyone connected with the Project was in the photograph, which probably dates from 1938 or 1939. "Rosie," our Burmese expert, somehow didn't make it for the picture, although he was the unofficial "PR man" of the Project as well as its preeminent Burmese expert. He had no visible qualifications for either function, but was inordinately proud of his ability to transcribe Burmese, which he would carry out with a grand flourish, letter by letter. He also fancied himself as a comparativist, but was never officially entrusted with this function since he was given to straightforward comparisons by broad categories, matching "rat" and "horse" (animals), "toe" and "esophagus" (body-parts) with nary a question. (No, this never affected me, in case any of you are wondering!) Only once did I see Rosie at a loss for words. He had, without our advice or consent, committed our Project "spelling team" to a challenge on radio to the champion spelling team of the University, the prize being ten silver dollars. The sponsors accepted the challenge on the theory, no doubt, that any group representing the "Sino-Tibetan Philology Project" ought to be able to spell well, at the very least. The big night finally came, and we went to the radio station across the bay in Frisco — Walters, myself, and three non-spellers (Rosie, Smiddie, and Doc). We were introduced to the radio audience in grandiose terms, something like "one of the most gifted spelling teams of all time." Rosie, leading off, was asked "an easy one — just to warm you up — FRIEND." Walters and I looked at each other as Rosie spelled out the expected "F-R-E-I-N-D," then at the face of the MC as a look of utter astonishment spread over it, to be followed by one of disgust. Rosie couldn't think of a word to say at that moment, but later explained he had had "an off night — like the best of us." Walters and I upheld the honor of the Project, however, by spelling down the other team and gaining the ten silver dollars, which sufficed to provide for an evening of debauchery for the five of us. (Debauchery was much cheaper in those days.)

There were others, all Shafer people, although none so colorful as Rosie. Walters, who was taking a Master's in history at the University, was a key part of the Project. Shafer had trained him to carry out comparative work at a fairly high level, and I continued in this tradition, using him especially for the Bodo-Garo (Barish) volume preparation. This volume, which is (for me) the most satisfactory in the whole STL series, owes much to Walters, who actually served as a second linguist, although he was always very modest about his ability and probably would have refused to accept the designation. He remained

primarily an historian in his interests, and turned down suggestions made by me — and by Shafer before me — that he take formal linguistic training and go into the field.

The Project was also fortunate in having the services of two first-rate typists, able to set up on the special Project typewriters (I wonder how Shafer ever came by them!) the complicated tables fed to them by Shafer — and in the mad transcription that was "official" on the Project (see below). One of these typists, a man of about 30, had spent some time traveling about the Far East with a companion, a man of about the same age, and they both had been hired on the Project some time before I got there. They both had got into TM, or the equivalent of the time, and the typist would prepare himself each morning by sitting cross-legged upon a stool for an hour or so, in deep meditation, before doing any typing at all. He would more than make up for this loss of time, however, with a fantastic display of typing each day. I have never seen his like since, and I wish he would interrupt his meditations, wherever he is, and come back to do some more typing for me! His friend, the "intellectual" member of the twosome, carried out his own research on the Project, and I never quite knew what he was up to. He left in a jealous rage one day after his friend had got himself married.

I have spent some time on the Project staff because they constituted, as it were, extensions of the Master himself. They were all beautiful people (you can ignore my attempts at description if there is any discrepancy) and they made the Project go. But Shafer was the heart of it. I have described the manic pace at which he worked — until his death, I daresay. I should now write something about how he worked, his methods, his ways of handling material. Some may wonder about this, saying, "Whaddya mean, methods?" Shafer has been called "disorganized" or words to that effect (in reviews here and there). Actually, he was too organized, it has often seemed to me, in looking back on it. He had adapted a mad (I believe originally French) transcription system precisely because it was completely logical and "organized." Little did it matter to Shafer that so many of the common sound units with which he was working came out festooned with otiose diacritics, like so much phonetic bric-a-brac! The most common vowel symbol in the sources (mostly older ones by non-linguists), for example, was a, often probably for ə, especially in unstressed or prefixial syllables. This comes out ê, and the pages of the early (pre-Benedict) volumes of STL are literally covered with ê, and the like. Shafer was too organized here to be practical. For me, the product of an (ultimately) Yankee background in which "resourceful" was the highest word of praise that one could utter about a man, this transcription was almost criminally stupid, and I probably told Shafer so in about those words. (I was given to plain talk in those early days.) But he would hear none of it, countering to the effect that I was the stupid one, unable to see the beauties of a logical transcription. But Shafer was a great scrounger, as I have said, and that is a good Yankee trait, so that I was able to overlook his sins of transcription — but I did throw the whole system out in the later STL volumes, a crime for which Shafer probably never really forgave me.

In the area of classification, also, Shafer was very fond of setting up all sorts of hierarchical systems, with much attention to arbitrary suffixes and to historical names, hence terms such as "Daic" (the root is reconstructed with initial *d-, whence "Tai" or "Thai"). I felt (and still feel) that it is preferable (i.e., more practical) to operate with "nuclei" or the like (as in the Conspectus), but Shafer saw this as a chaotic approach and we would often argue about the point.

Shafer was also methodical (and traditional) in approaching the task of reconstructing PTB or PST, working from the local groupings, even dialects, to broader and broader supergroupings. He made a stab (JAOS articles) at setting up a vocalic system for PST, but it is clear that he regarded this as very provisional, and he contented himself with trying to work out certain "correspondences," with suggestions that the ultimate system would be much more complex, with three-vowel clusters, etc. Like all Tibeto-Burman scholars who had preceded him (and many since), he was markedly Tibetocentric, spending endless time mulling over Tibetan dialects. I soon got the impression that little was to be gained from this Tibetobsessionalism, hence I set about working out a framework of some sort, for PST as well as PTB, with a consideration also of Karen (largely ignored by Shafer). Shafer never really approved of this change in emphasis, preferring to continue along conventional lines. He would never have approved of "teleo-reconstruction," having been extremely critical of Simon's attempt at a direct comparison of Tibetan with Chinese.

Actually, Shafer was basically conservative at heart, as shown by his papers on such distinct groups as Nahali, Li, and Miao-Yao. True, he did turn out a few "wild" things indicating possible connections of Sino-Tibetan with one or another group of Amerindian languages. But even here he was following in the path of a fairly well-known linguist (Sapir!). He continued to plug along with Tai as one of the ST groups, despite an apparently growing uncertainty about the matter, never accepting the suggestion that Tai is basically related to Austronesian, this again reflecting a fundamental spirit of conservatism on his part. As for details of his comparative work, here too he was extremely conservative in approach, admitting very few semantic shifts and expressing a negative view of my published opinion that more leeway should be allowed in semantics than in phonology when making these broader comparisons. I point this out because of what seems to have become a widespread view that Shafer was a "wild" linguist in some sense, possibly because he was a pioneer in the field.

His conservatism perhaps was mainly responsible for his not accepting a fully phonemic approach, along with an insistence upon literal transcriptions, e.g. the writing of Thai (Siamese) medial oo as "ɔɔə", the idea that Written Burmese ui was somehow phonetically [ui], and that á (with the glottalized auk-myit accent) was a "short a." Shafer also refused even to consider Karlgren's reconstructions for Archaic Chinese — again an indication of his basic conservatism — but in all fairness I should add that when the Grammata Serica first came out, and I had George Trager (then at Yale) take a look at it, he quickly pronounced it a "non-language."

Haudricourt, in his review of the Conspectus, has described Shafer's approach as "analytical," as contrasted with mine as "synthesizing," but I feel that this misses the mark, the difference lying primarily in the way we approach the problem — the conservative vs. the radical, if you will: the working up to a framework, dialect by dialect and language by language, as opposed to the setting up of a series of provisional frameworks, then working within these frameworks to modify them as need be. (This accounts for many of the "contra Benedict's" in Benedict.) Shafer and I often discussed this basic difference in approach, and even the irony, as we both saw it, in the circumstance that he, the outsider and anti-academician, was actually the conservative one while I, a member of the academic élite (Shafer's point of view!), was the radical one. Shafer simply thought I was unsound, I suppose, an opinion that has clung to me in certain circles over the years.

What more can I write about Shafer? He was the explorer, the pioneer, venturing ever further and further into virgin fields. Wolfenden had come by that way some years before, it is true, but he had been looking for rather different things. They were the first whom we can fairly label "Tibeto-Burmanists," each having left the Tibetan "nest" and flown to distant ranges. In a still broader context, Shafer surely is to be regarded as the first "Sino-Tibetanist" ever, without the need for qualification. He had an idea — a seemingly preposterous idea — that if given the opportunity he would be able to train a group of WPA workers sufficiently to enable them to collect meaningful data on languages as esoteric as Tibetan, Burmese, and Chinese.

It is very probable that only one scholar in the world in a position to sponsor such a project would have done so. By some quirk of fate Shafer got to this scholar, Professor Kroeber, and convinced him that the project was practicable. Kroeber never really discussed this aspect of the Project with me in any detail, but I soon learned that he was a fervent believer in "mass" research. One of his dicta, which I have never forgotten, went like this: Never do anything yourself which you can get someone else to do for you. He was not being Machiavellian here, but was simply pointing out that there are levels of research, and that one shouldn't work at lower than his true level (an argument which generations of graduate students have been using with their wives). Kroeber did admit one day, however, when I was complaining about the "impossibility" of completing the Project, that he had been pushing the method to the limits when he accepted Shafer's proposal. He also remarked that a thing isn't "fun" unless it is difficult.

I think that Shafer also felt that way about his work. He gave his life to it, and I can't imagine his ever having any regrets. He made it possible for me to go even deeper into untrodden areas, and for all of us to go our separate ways in this still primeval wonderland of Tibeto-Birmanica and Sino-Tibetica. His goals were monumental, but they fitted the man, and few have lived to find so large a measure of achievement. May we all fare so well!

3

THAI, KADAI, AND INDONESIAN

A new alignment in southeastern Asia

Paul K. Benedict

Source: *American Anthropologist* 44, 1942, 576–601.

In the present paper the writer presents a general solution to the complex problem of the affinities of the Indonesian languages. The two following premises are basic to the thesis developed here:

1. The true Indonesian substratum on the Asiatic mainland is represented by four scattered languages in southern China, northern Tonkin, and Hainan, all of which constitute a single linguistic stock (Kadai).
2. The recognition of the Kadai stock, which shows numerous points of contact with Thai, opens the way to a new interpretation of the latter as a more distant member of an archaic Thai-Kadai-Indonesian linguistic complex.

Although these suggestions are new and perhaps unexpected, it can be said that they accord with the general picture as reconstructed from historical and cultural data. It is generally agreed that the Indonesian migrations have proceeded from the Asiatic mainland, but the evidence brought forward has been of a generic rather than specific nature, and the area of departure has not been delimited. The linguistic speculation has been notable for range rather than relevancy,[1] and the cultural treatment has in some instances been equally unsound.[2] It is hoped that the argument developed below will provide a number of solid *points d'appui* from which further ramifications can be anticipated.

The newly recognized Kadai stock comprises the Li dialects of the island of Hainan, the Kelao language of southcentral China, and the Laqua and Lati languages of the China-Tonkin border region. The term "Kadai" has been compounded by the writer from "Dai," one of the forms of the Li term for themselves,[3] and the *kă-* prefix found in Laqua *kădăŭ,* Kelao *kătsü* "man (homo)." These languages, with the exception of Li, are not generally known to the scientific world, and our available sources are rather meager. The Li dialects have been described by a number of European observers, the most thorough of whom have been

Savina and Stübel-Meriggi.[4] Bonifacy has published word-lists of Kelao, Laqua, and Lati,[5] while additional material on all three languages has been supplied by his compatriot, Lunet de Lajonquière.[6] A third Kelao source has been furnished by Samuel R. Clarke, the author of a popular account of the little-known tribal groups in southern China.[7]

The Kedai languages have received scant attention from anthropologists and linguists. Li has evoked occasional comment, yet no real analysis has been attempted. The obvious Thai element in the language was noted by Parker over half a century ago,[8] and this theme was further developed by Strzoda.[9] The less apparent Indonesian affinities were first pointed out by Terrien de Lacouperie, who suggested a relationship with the Indonesian languages of Formosa.[10] In more recent times P. Mus[11] and H. Maspero[12] have further extended this line of thought and have supplied the first concrete bits of evidence. Maspero, a sound and generally conservative scholar, concludes that the Li numerals "certainly" belong to the Indonesian family (*op. cit.*, p. 230).

The three mainland languages (Kelao, Laqua, Lati) have attracted still less attention. Bonifacy, who recorded Laqua, noted the analogy between the Laqua and Cham numerals,[13] but this observation seems to have been overlooked by Maspero and other scholars. Kelao and Lati have gone almost entirely unnoticed, although W. Schmidt has seen fit to classify the latter as an independent linguistic entity.[14] It was Bonifacy's observation on Laqua that led to the writer's discovery of the relationship between Laqua and Li, and thus ultimately to the concept of a single unified Kadai stock.

The Kadai-speaking groups are all of marginal type, as should be expected on the basis of our substratum theory. The Li, who inhabit the mountainous central and south-central parts of Hainan, are under economic pressure from their powerful Ong-Be (Thai-speaking) and Hoklo (Chinese-speaking) neighbors. The Kadai groups on the mainland rank even below the Miao and Lolo, and generally regard themselves as autochthonous. The Laqua, who call themselves Ka Beo, in the upper Rivière Claire valley of northern Tonkin, are described by Lunet de Lajonquière as follows:

> *Ils se considèrent comme aborigènes et il est certain qu'ils sont venus dans la contrée avant toutes les autres tribus montagnardes* (*cit. supra,* p. 339).
>
> *C'est une variété* [of economic life] *en complete décadence. La plus grande partie des terres qu'ils cultivaient ont été déjà cédées aux Meo, qui paraissent devoir les absorber.*
>
> (*ibid.*, p. 341)

Of the Lati, also in the upper Rivière Claire valley, the same writer states simply that "*Ils se prétendent aborigènes*" (*cit. supra,* p. 358). Bonifacy places his estimate of the number of Lati at only 450 (76 families).

The Kelao or Lao, who call themselves Thü, range over an extensive area in south-central China and northern Tonkin, but their true home appears to be

Kueichou province, whence they have migrated into the northern Tonkin border region (cf. Lunet de Lajonquière, *cit. supra,* p. 356). Clarke, who has given us the fullest available account of the Kelao, stresses the aboriginal nature of the group:

> The Keh-lao, however, are now nearly extinct; many of them have married into Chung-chia [Thai] and Old Chinese families. Some writers have spoken of them as extinct. As far as we know, there are now only several hamlets of them in the Anshun prefecture [west-central

> Kueichou], which altogether do not number more than two or three hundred families. These people claim, and rightly, we believe, to be the real aborigines of that region. In some parts of the province the Miao claim to be the aborigines, but where the Miao and Keh-lao occupy the same district, the Miao allow that the Keh-lao were there before themselves (*cit. supra,* p. 13).

Another missionary writer, Aloys Schotter, also attributes a low rank to the Kelao:

> *Le plus bas dans l'échelle sociale c'est peut-être le groupe des Blancs* [White Miao]. *La tribu des Kë-lao est peut-être plus dégradée encore surtout quant aux moeurs.*[15]

The languages spoken by these primitive groups fall into two major divisions, viz. Li-Laqua and Lati-Kelao, which together constitute the Kadai stock. Dialectical differences can be established both for Li and Kelao, and are of such magnitude that they must fully be taken into account. The numerous Li dialects can be classified under the headings of "Southern Li" and "Northern Li" on the basis of their treatment of original nasal initials. In Northern Li these initials tend to be transformed into the homorganic stops, whereas in Southern Li they are uniformly retained; cf. N. Li *ba~pa,* S. Li *ma* "dog" (Thai **hma*); N. Li *dau~tau,* S. Li *nau* "long" (Thai **ńau*); N. Li *ka,* S. Li *nga* "horse" (Annamite *ngüa*). The "Central Dai" dialect recorded by Savina and most of the dialects recorded by Jeremiassen and Stübel belong in the Northern Li group, while the "Southern Dai" dialect of Savina, the Yulinkau dialect of Swinhoe and Calder, and the K'iung-Shan dialect of Parker belong in the Southern Li group. Kelao similarly shows a dialectical cleavage between "Northern Kelao" (dialect recorded by Clarke) and "Southern Kelao" (dialect recorded by Bonifacy and Lunet de Lajonquière). The distinctions here, both lexical and phonetic, are even more marked than those that obtain in Li, but conform to no easily recognizable pattern. It is apparent that a full treatment of the linguistic problems of Kadai would require detailed phonetic information on a wide range of dialects for at least four languages, and it is not unlikely that further exploration in the Tonkin-China border area will reveal still other members of this stock. Unfortunately, we lack the materials necessary to implement a complete study of the whole stock,[16] and must content ourselves with a survey of the more salient points.

All four Kadai languages are of monosyllabic, isolating type, with full tonal systems as in Thai. The Kadai word-order, like that of Thai and Indonesian, shows object following verb, and modifying elements (including genitive constructs) following modified elements; thus, Malay *mata hari,* Li *sa ven,* Thai **ta wăn* "sun," lit. "eye (of the) day." Kadai, like Thai, lacks the affixation apparatus of Indonesian, yet prefixed forms abound in the Lati-Kelao branch of the stock, e.g. Lati prefixed *m-* in *m-tšua* "moon," *m-bo* "sky," *m-ti* "earth," *m-ni* "ox," *m-go* "cat," *m-so* "elephant," *m-si* "beak, mouth," *m-tšu* "eye," *m-ngă* "oil"; Lati prefixed *a-* in *a-ńa* "rain," *a-lia* "rat," *a-k'o* "monkey," *a-ti* "tiger," *a-kŭ* "bird," *a-li*

"fish," *a-k'e* "frog," *a-k'u* "man," *a-sa* "hair," *a-ńu* "salt." Kelao has prefixed *bu-* occasionally corresponding to Lati prefixed *m-,* as in Kelao *bu-to* "earth," *bu-tsüe* "beak" (*bu-tsü-lüa* "mouth"). Laqua has prefixed *kă-* in *kă-dăŭ* "man," *kă-zio kă-pă* "boy," *kă-zio ka-măi* "girl," where *zio* stands for "child" and *pă* and *măi* are the sex modifiers.

On the phonetic side, the Kadai languages present a fairly uniform pattern of relatively simple type, though mixed (indeterminate) and front-rounded vowels are uncommonly abundant. Li exhibits the greatest range of initials and finals, with Laqua not far behind, while Lati and Kelao have undergone a process of extreme modification, in the course of which almost all final nasals and stops have been eliminated.[17] The phonetic attrition shown by Lati and Kelao has proved to be one of the chief stumbling blocks in our analysis of Kadai phonology. When it is realized that Li and Laqua, the better preserved pair of languages, stand in much the same relationship to Indonesian, some inkling of our difficulties can be gained. The investigation of the phonetic shifts exhibited by the Thai roots in Li, the best recorded of the Kadai languages, has brought to light a number of significant variations, especially as regards initials, which are useful in the study of Thai itself. Thus, the writer has reconstructed a separate phoneme **hr* (surd *r*) for archaic Thai on the basis of the equation Ahom *r-* = Siamese *h-* = Tho *t'-*; in this series Li significantly has *s-*, suggesting an original **sr-*:

	Ahom	Simese	Tho	Li
stone	*rin*	*hĭn*	*t'in*	*sien*
louse	*rau*	*hău*	*t'au*	*său*
break	*rak*	*hăk*	—	*săk*
carry	*rap*	*hap*	*t'ap*	*sap*

The variations in initials between S. Li and N. Li are often of unusual type, e.g. S. Li *d-* = N. Li *f-*, corresponding to Thai *đ-* (Siamese *đ-* = Shan *l-* = Khamti *n-*):[18]

	Siamese	Shan	Khamti	S. Li	N. Li
earth	*đĭn*	*lin*	*nin*	*dăn*, *den*	*fan*
bone	*kăđuk*	*luk*	*nuk*	*drü*	*füök*
raw	*đĭp*	*lip*	*nip*	*diep*	*fiep*

Aspiration of initial stops is characteristic of Li; cf. Li *hăn,* Thai **guăn* "smoke"; Li *ha,* Thai **ga* "thatching grass"; Li *hang,* Thai **gang* "jaw"; Li *hăn,* Thai **k'ăn* "crow of a cock"; Li *hău,* Thai **k'ău* "horn," also "mountain"; and Li *k'ău,* Thai **kău* "old"; Li *k'ăi,* Thai **kăi* "fowl." Li often simplifies the complicated diphthongs and triphthongs of Thai, but note Li medial *-ie-* = Thai *-ĭ-, -ĕ-,* as in Li *dien~tlien,* Thai **lĭn* "tongue"; Li *diep~fiep,* Thai **đĭp* "raw"; Li *liep,* Thai **lĕp* "fingernail," and Li *-öü* = Thai *-aü,* as in Li *böü,* Thai **ɓaü* "leaf"; Li *tšöü,*

Thai **tšaü* "heart"; Li *t'öü,* Thai **taü* "low." S. Li retains final *-k* after short vowels but substitutes a glottal stop[19] after long vowels, while N. Li uniformly retains the final velar stop; cf. Li *p'ăk,* Thai **văk* "hatch"; Li *t'ŏk,* Thai **tŏk* "fall"; Li *fi'*, Thai **pik* "wing"; Li *t'o'*, Thai **t'ok* "pour"; Li *sa'*, Thai **sak* "pestle"; S. Li *drü'*, N. Li *füök,* Thai **đuk* "bone."

The morphological and phonological points developed above point to Thai rather than to Indonesian, yet the lexical elements of Kadai bear the unmistakable imprint of the latter stock, along with an equally deep imprint of the former. In brief, the numerals and a scattering of nouns, pronouns, and adjectives show Indonesian affinities, while many of the remaining elements show Thai affinities. On the basis of this distribution, the writer at first regarded Kadai as a composite of Indonesian and Thai, with the former as the more likely substratum. Further analysis of Thai, however, has led to the view presented below; to wit, that Thai, Kadai, and Indonesian together constitute a single linguistic complex. Kadai is the "transitional" member of this triune, though in the main it approaches Thai rather than Indonesian. Both Thai and Kadai have reduced a number of disyllabic roots to monosyllables, have developed complete tonal systems, and have discarded the original morphological apparatus of affixes.[20] Throughout this elaborate linguistic metamorphosis, however, a number of basic lexical landmarks have persisted and it is to these that we shall direct our attention.

The Kadai numerals are of fundamental importance in the present connection, since the Indonesian affinities of the stock are more apparent there than elsewhere. The following table of Kadai numerals, in which reconstructed Indonesian (IN) roots taken from O. Dempwolff's recent work[21] have been incorporated, serves to illustrate this point.

	IN	Laqua	S. Li	N. Li	S. Kelao	N. Kelao	Lati
one	**'it'a'*	*tiă*	*kü*	*ü*	*tsi*	*si*	*tšăm*
two	**ḍuwa'*	*đe*	*dau*	*trau*	*đü*	*so*	*fu*
three	**təlu'*	*tău*	*su*	*su*	*tŏ*	*da*	*si*
four	**'ə(m)pat*	*pe*	*sau*	*so*	*pu*	*bu*	*pu*
five	**lima'*	*mö*	*ma*	*pa*	*mlĕn*	*mbu*	*ng*
six	**'ənam*	*nam*	*nom*	*tom*	*tšö*	*nang*	*nă*
seven	**pitu'*	*mö tău*	*t'u*	*t'au*	*đi*	*ši*	*ti*
eight	**walu'*	*mö dü*	*du*	*au*	*šiă*	*vleu*	*be*
nine	**t'iwa'*	*mö điă*	*pöü*	*föü*	*ku*	*su*	*lu*
ten	**puluh*	*păt*	*p'uot*	*fuot*	*tsü*	*beu*	*pa*

The following variants are worthy of comment: S. Kelao *mlĕn* "5" but *tsü mu* "15" (cf. N. Kelao *mbu* "5"); Lati *tšăm* "1" but *pa tšă* "11"; Lati *pa* "10" but *fu pe* "20", *sie pe* "30" (*si* "3").

Some of the leading features of Kadai phonology are illustrated in the above table of numerals. N. Li *pa* < *ma* "5," *tom* < *nom* "6", and *föü* < *pöü* "9," *fuot* < *p'uot* "10" are all regular developments (see the discussion above). Li *kü~ü* "1" are

probably independent of the IN root, and the analysis of Li *su* "3," *sau~so* "4" is not certain. For the latter, Maspero suggests a development comparable with that found in Tarema (Formosa), which has *suatto* < **suat* < **səwat* < **səbat* < **s-pat* "4." Li *du~au* "8" belong to a puzzling series in which S. Li initial *d-* corresponds to N. Li initial *h-* or vocalic anlaut, e.g. *dai~hai* "iron," *dai~hiai* "a Li," *duoi~ui* "fat" (n.), *döü~öü* "thin." These forms seem to have been derived from roots with labial+liquid initial cluster; cf. the variant form *b'lai* "a Li," and the frequent correspondences with Thai initial *r-*, as in S. Li *da,* Thai **rüa* "boat"; S. Li *dăt,* Thai **răt* "squeeze"; S. Li *düön,* Thai **rüan* "house" (N. Li *plong*). We can infer a bifurcate development of the type **walu'* > **wlu'* or **blu'* > *du* (S. Li), and **walu'* > **wau'* > *au* (N. Li). Li *p'uot~fuot* "10" attest to a pair of shifts, viz. final *-h* > *-t,* as in IN **darah,* Li *dat~tlat* "blood," and medial *-u-* > *-uo-,* as in Thai **ńŭng,* Li *ńuong* "mosquito." The development here has been of the type **puluh* > **p'luh* > *p'lut* > **p'ut* > *p'uot.*

Laqua parallels Li in the developments *mö du* > **walu'* "8" and *păt* > **puluh* "10." Laqua *tiă* < **'it'a'* "1," *tău* < **təlu'* "3," and *mö đia* < **t'iwa'* "9" reveal IN affinities not apparent in Li. The Laqua vocalic shift *a* < *e~ö* is found in the forms *đe* "2," *pe* "4," and *mö* "5"; cf. Laqua *pö,* IN **batu'* "stone"; Laqua *pe,* IN **bapa'* "father"; Laqua *te,* IN **mata'* "eye"; Laqua *ne* Li *na~ta,* Thai **na* "rice-field"; and, medially, Laqua *dön,* Li *dan* "100"; Laqua *nen,* Li *ńan* "moon." The appearance of *mö* "5" in the Laqua numerals "6" to "8" is suggestive of a quinary system; cf. the S. Kelao numerals cited by Lunet de Lajonquière: *sü-u* "2," *to-u* "3," *pu-u* "4," *nlĕ-u* "5," but *tšĕ-ni* "6," *dž-ni* "7," *suo-ni* "8," *ku-ni* "9."[22] Laqua and Li have a common root for "100" (*dön~dan*), which is independent of the IN root (**ratut'*).

The Kelao and Lati numerals are, in general, further removed from the IN system as reflected in Laqua and Li. Notable, however, are S. Kelao *mlĕn* (Bonifacy)~*nlĕ* (Lajonquière) < IN **lima'* "5," and N. Kelao *vleu* < IN **walu'* "8," which show retention of the liquid phoneme *l.* Kelao *pu~bu,* Lati *pu* "4" reflect an *a* > *u* vocalic shift, which is especially characteristic of Lati; in the table below, the contrast with the Laqua *e~ö* vocalism is made clear:

	IN	Li	Laqua	Lati
four	**'ə(m)pat*	—	*pe*	*pu*
five	**lima'*	*ma~pa*	*mö*	*ng(u)*
father	**bapa'*	*fa~ba*	*pe*	*pu*
eye	**mata'*	*sa*	*te*	*m-tšu*

The variability reflected in the Kadai numerals appears also in other aspects of the vocabulary. Scarcely any roots prevail everywhere, and there are a number of confusing "partial equations," yet many significant features emerge. One of the most notable of these features is the regularity shown in the roots for "dog," "pig," and "horse," the first two with Thai affinities, the last with Annamite:

		Laqua	S. Li	N. Li	S. Kelao	N. Kelao	Lati
dog	**hma* (Thai)	*mă*	*ma*	*pa*	*χmă*	*mu*	*mu*
pig	**hmu* (Thai)	*mu*	*mau*	*pau*	*χmüă*	*ma*	*me*
horse	**ngüa* (Ann.)	*rre*	*nga*	*ka*	*ngüă*	*niau*	*ngo*

Note the N. Kelao and Lati shift *u* < *a* in *mu* < **hma* "dog," and N. Li *p-* < *m-*, *k-* < *ng-* (vide supra). These loan-words, if such they be, must be of some antiquity, in view of the selective nature of the distribution (there is no trace of the prominent Thai-Chinese root **ma* "horse"), as well as the note-worthy equation S. Kelao *χm-* = Thai *hm-*, the latter a reconstructed phoneme (surd *m*) not found in any of the modern Thai languages. In the same general class belongs the correspondence between Laqua *k'ăi*, Li and N. Lelao *k'ai*, Lati *ka* "fowl," and the Thai root **kăi*; contrast the earlier stratum reflected in the series IN **manuk* "fowl, bird," Laqua *nuk*, S. Kelao *ńie* "bird," and Thai **nŏk* "bird."

The following group of comparisons, arranged roughly according to natural lexical divisions, is intended to serve as an index of the Kadai-Indonesian relationship:

1. Laqua *vuon* (Lajonquière *mo ven*) "sun," Li *ven* "day," *sa ven* "sun" ("eye of the day"), S. Kelao *du vuă* "sun," IN **wari'* "day, sun" (IN medial *-r-* < Laqua and Li *-n*). Cf. also N. Kelao *vlei* "sky," which shows a contrasting type of development (**wari'* < **wli* < *vlei*).
2. S. Li (*pa*) *pŭn* "rain," IN **'ə(m)bun* "atmospheric precipitate" (Tagalog *'ambon* "fine rain").
3. Li *nom~nam*, IN **danum* "water."
4. Laqua *păi*, Li *pei~fei*, S. Kelao *p'i*, N. Kelao *bai*, Lati *pie*, IN **'apuy* "fire." For the Li development (**'apuy* > **api* > *pei*), cf. Li *ngei*, IN **tangit'* "weep"; Li *nei*, IN **'ini'* "this."
5. Laqua *pung*, IN **bunga'* "flower."
6. Laqua *kă-dăŭ*, Li *ău*, IN **tawu'* "man (homo)."
7. Laqua *pe*, Li *fa~ba*, S. Kelao *ă-ba*, Lati *pu*, IN **bapa'* "father." For the vocalism, see the analysis above.
8. Laqua *ru* (Lajonquière), S. Li *dau*, N. Li *fo~o*, IN **'ulu'* "head." The Li development has been **'ulu'* > **wlu* > *du~o*, exactly paralleling IN **walu'* > Li *du~au* "8."
9. Laqua *δam*, S. Kelao *lă so*, N. Kelao *ma sang*, Lati *a-sa*, IN **d'a(m)but* "hair." The original palatal initial has everywhere been assibilized: **d'a(m)but* > **d'am* > *δam* and *sang~sa~so*. For the Laqua initial δ-, cf. Laqua *δău*, IN **hud'an* "rain."
10. Laqua *te*, Li *sa*, N. Kelao *dau*, Lati *m-tšu*, IN **mata'* "eye." Li appears to have developed a sibilant initial through aspiration (**mata'* > **m-t'a* > *sa*); see the

discussion below of the Thai root **ta*. S. Kelao perhaps retains the root in the compound *du vuă* "sun," paralleling Li *sa ven*, Malay *mata hari* ("eye of the day"), yet this dialect also has *du die* "moon," *du dĕ* "star." The picture is further complicated by the evidence from Lati, which has *m-tšu* "eye," *m-tšu* "month" (on different tones as recorded by Bonifacy), but *m-tšua* "moon" and simply *tšua* "star."

11. Laqua *rö*, Li *yăi~t'ăi*, N. Kelao *rau*, Latu *lu*, IN **talinga'* "ear." For Laqua rö < **talinga'*, cf. Laqua *rre*, Annamite *ngüa* "horse" (Laqua *-ö* < *-a*; see the analysis above). The Li forms point to an original **ńăi* or **niăi* with palatalized nasal initial, whence S. Li *yăi* (through further palatalization) and N. Li *t'ăi* (*n-* > *t-* is the regular N. Li shift). This reconstruction is supported by two outside comparisons, one with Thai (Dioi) and the other with IN, as shown in the table below (Central Li from Savina, Shaved Head Li and White Sand Li from Stübel):

		S. Li	C. Li	Shaved Head	White Sand
ear	**talinga'* (IN)	*yăi*	*t'ăi*	*t'ai*	*džai*
finger	**niang* (Dioi)	*yeng*	*tleăng* *t'eng*	*t'ĕng*	*džing*
snake	—	*ya*	*t'a*	*t'a*	*dža*
yellow	**kuning* (IN)	*yĕng* *hieng*	—	*t'ĕng*	*džiang*

12. S. Li (*hai*) *p'en*, S. Kelao *du pio*, N. Kelao *bang*, IN **'ipən* "tooth."
13. Li *k'ok*, N. Kelao *k'au*, IN **kaki'* "foot."
14. Lati *tšu*, IN **t'ut'u* "breast."
15. Li *dat~tlat*, IN **darah* "blood." For the final, cf. Li *p'uot~fuot*, IN **puluh* "10."
16. Laqua *nen* "fat" (n.), S. Kelao *nuă*, Lati *m-ngă* "oil," S. Kelao *nu χmüă*, Lati *m-ngă me* "fat" ("oil of pig"), IN **miñak* "oil"~**məñak* "fat."
17. Laqua *küön*, N. Li *k'an*, S. Kelao *kă mŏn~mŏn kă* (*mŏn* perhaps for *mo* "rice"), N. Kelao *ka*, Lati *k'o*, IN **ka'~*ka'ən~*ka'i* "eat." Note also Laqua *ngăm* "drink," IN **pangan* "eat" (cf. Lati *k'o* "eat," also "drink").
18. Li *sop~sap*, Lati (*ngua*) *so*, IN **rabi* "night." Cf. also Javanese *sĕrap* "twilight,"[23] and the equation Li *s-* = Thai *hr-* analyzed above.
19. Li *ngei*, IN **tangit'* "weep."
20. Laqua *tie*, IN **matay~*patay* "die." For the vocalism, cf. Laqua *te*, IN **mata'* "eye."
21. Li *diep~fiep* "raw," IN **huḍip* "live." Note "raw" = "green" = "alive" a semantic association appearing also in the Thai root **ďĭp* (vide infra). For the medial diphthong in Li, see the analysis above.
22. Laqua *dăm*, Li *dŏm* "black," IN **'i(n)təm* "black," **ḍəḍəm~tidəm* "dark."

23. Laqua *nin,* S. Li *yĕng~*hieng, N. Li *t'ĕng,* S. Kelao *t'e ni,* N. Kelao *nyi,* Lati *a-hni* (recorded as *an hi*), IN **kuning* "yellow." Li **ńeng* or **nieng* > *yĕng ~t'ĕng* (vide supra); **nieng* < **kuning,* with medial diphthong as in No. 21.
24. Li *tik~tok,* IN **ḍikih~* 'ə(n)ṭik~* 'iṭik* "small."
25. Laqua *k'ău,* S. Li *hau,* Lati *ku~kui,* IN ** 'aku'* "I." Li *hau* < **k'au* (vide supra).
26. Li *nei,* IN ** 'ini'* "this."

In addition to the above, a number of significant correspondences exist within the Kadai stock itself, thus serving to tie the group together. The more important of these lexical agreements are listed below:

1. Laqua *nen,* Li *ńan* "moon." Possibly related to IN **bulan* "moon"; thus, **bulan* > **wulan* > **dan* (paralleling **walu'* > *du* "8," ** 'ulu'* > *du* "head") > *ńan* (through assimilation to the final nasal). Cf. Li *nuk,* IN **bəḷuk* "monkey."
2. Laqua *mŏn* "sky," *mön dŏng* "thunder," S. Kelao *mön düă* "rain," Lati *m-bo* "sky" (Lajonaquière *büön*).
3. Laqua *mön dŏng,* Li *pa dang om,* S. Kelao *zü dŏng* "thunder."
4. Laqua *hŏng,* S. Kelao *ngă-ye* "water" (but *zŏng ngŭă* "drink"); N. Kelao *u,* Lati *i* "water"; cf. N. Kelao *du,* Lati *m-ti* "earth." S. Kelao retains the element *u* in the compounds *u ngĕ uă* "tears" (Lati *i m-tšu), i lă-pu* "milk" (Lati *i tšu*). A possible comparison exists with IN **wayər* "water."
5. Laqua *dăm,* Li *sam* "fruit"; S. Kelao *mă,* Lati *mi* "fruit."
6. Laqua *pĕŏ* < ** plŏ,* Li *da~tla* "fish"; S. Kelao *lü,* Lati *li* "fish." Cf. Thai **pla* "fish."
7. Laqua *kăuk,* Li *hău* < **k'ău,* S. Kelao *pă-ku,* Lati *kui* "horn." Cf. Thai **k'ău* "horn."
8. Laqua *mă~măi,* Li *mei,* S. Kelao *mu* (*vĕ*), Lati *mia* "female, mother." Cf. Thai **me* "mother."
9. Laqua *mon,* S. Li *mom,* N. Li *păm~băm* "mouth." Cf. Annamite *mŏm* "muzzle, snout."
10. Laqua *mun,* Li *müöm~püöm* "beard" (cf. the treatment of nasal finals in the foregoing example). Cf. Thai **mŭm* "beard" (only in the northern Thai speeches: Dioi *mum,* Tho *kang mum,* Nung *mom.*
11. Li *p'a möü,* N. Kelao *mau* "hand." Cf. Thai **mü* "hand."
12. Laqua *nie,* S. Li *yeng,* N. Li *t'eng~tleăng* "finger" (Li **nieng,* vide supra). Cf. Dioi (Thai) *niang* "finger."
13. S. Kelao *plă,* Lati *pio* "blood"; cf. S. Kelao *ple u,* Lati *p'i* "die."
14. Laqua *ði,* S. Kelao *ðü tu* "urine."
15. Laqua *ńung,* Li *ńau,* S. Kelao *ńu,* N. Kelao *nyö,* Lati *a-ńu* "salt." Cf. IN ** 'uyah* "salt."
16. Laqua *yeu,* S. Kelao *hă,* N. Kelao *a,* Lati *ho* "meat, flesh." Li has the puzzling forms *mam~am.*
17. Laqua *măi,* Li *mau~pau* "year."
18. N. and S. Kelao *vu,* Lati *vu* "go."

19. S. Li *müön,* N. Li *püön~pöü,* S. Kelao *χm,* N. Kelao *mu* "come." Cf. Malay *mari,* Cham *mörai~mai* "come," indicating the development IN medial *-r-* > Li *-n,* as in **wari'* > *ven* "day."
20. Laqua *neng,* Li *dĕng~t'lĕng* "red." Cf. Thai *đeng* "red."
21. Laqua *mi,* Li *mü~măü* "thou." Cf. Thai **maü* "thou," Annamite *măi* "thou" (pejorative).

We have, finally, to consider the nature of the affinity of Kadai and Indonesian with the Thai group of languages, spoken over a wide area in southern China, Siam, French Indochina, Burma, and Assam. The Thai family includes Ahom, Khamti, and Shan, in the west; Siamese and Lao, in the south; White Tai and Black Tai, in the east; Nung and Tho, in the northeast; and Dioi, in the north. Despite the geographical extent of this group, the several languages are closely interrelated, and thus rather detailed reconstructions of the parent speech can be made. The earliest systematic investigation in this field was Maspero's study of the Thai tonal system.[24] This study was supplemented by several brief articles by G. Coedès and J. Burnay,[25] but no comprehensive review of the problem appeared until almost a quarter of a century later, when K. Wulff published his monumental work on Chinese and Thai.[26] The writer has further extended the analysis undertaken by Wulff and has filled in certain lacunae in that scholar's work,[27] so that our present knowledge of Thai phonology may be regarded as reasonably complete.[28]

As regards the affiliations of Thai, the generally accepted view has been that Chinese and Thai constitute a single "Eastern" division of the Sino-Tibetan or Indo-Chinese stock, in opposition to the Tibeto-Burman or "Western" division. It is this view that has been developed by Maspero, Wulff, and, most recently, R. Shafer (largely on the basis of Wulff's work),[29] and that has given rise to attempts at direct Siamese-Tibetan comparisons, such as those of O. Schrader.[30] The writer must plead guilty on the same charge, though in modified form.[31] Almost alone among students of the Thai languages, Coedès and Burnay have evinced a healthy skepticism of the dogma of a Chinese-Thai relationship. Conrady, a pioneer in Far Eastern linguistics, sought to connect Indo-Chinese, including Thai, with the Austric stock (Mon-Khmer, Khasi, Munda, et al.) established by Schmidt, in terms of a "common substratum" (*gemeinsame Unterschicht*).[32] Wulff, apparently under the influence of Conrady, has attempted to demonstrate the existence of infixes in Siamese, which he compares with those characteristic of the Austric languages. Of Conrady's proposed Indo-Chinese-Austric grouping, Wulff makes the following assertion:

> The similarity of the formations [infixes] in both languages [Siamese, Javanese] rests not on chance, since the relationship of Malayo-Polynesian with Indo-Chinese, which Conrady sought to show with insufficient means, is certain [sicher].
>
> (*cit. supra,* p. 17, note 1)

Maspero, in a review of Wulff's work,[33] has convincingly dismantled the thesis of Thai infixation, and with it much of the Conrady-Wulff hypothesis. A similar hypothesis has been brought forward by J. Przyluski in the well-known account in *Les Langues du Monde.*[34] Przyluski suggests that Thai is transitional between Sino-Tibetan (Chinese and Tibeto-Burman) and Austric, yet offers no support for this view, other than a few comparisons of demonstrative pronouns in Siamese, Annamite, Khasi, and Palaung.

The writer's conclusions differ significantly from any of the above. The thesis presented here holds that Thai has a truly genetic linkage with Kadai and Indonesian rather than with Chinese and Tibeto-Burman (Sino-Tibetan), but has undergone extensive modification under Chinese influence. A similar view was propounded many years ago by Gustav Schlegel, in a highly unsystematic and unscientific fashion.[35] Schlegel was unaware of the existence of the Kadai group, but pointed out many analogies with Malay, and in general seems have been on the right track. The writer has developed the present hypothesis entirely independently of Schlegel, and largely as a by-product of his own synthesis of the Kadai stock.

The writer has long been aware of the fact that the lexical resemblances between Chinese and Thai are of a restricted range which fails to support the generally held view of a genetic relationship between the two languages. A careful analysis of the material assembled by Wulff, in addition to his own supplementary material, has made this fact still clearer. The primary lexical agreements lie in the numerals, especially from "3" to "10" and "100," a few words for parts of the body, certain animal names, and a number of terms for cultural objects and the like. Let us examine these loosely defined categories in some detail.

The Thai numerals from "3" to "10" and "100" are in fairly close agreement with the Chinese: Thai **sam,* Ch. **sâm* "3"; Thai and Ch. **si* "4"; Thai **ha,* Ch. **nguo* "5"; Thai **hrŏk,* Ch. **liuk* "6"; Thai **tšĕt,* Ch. **ts'iĕt* "7"; Thai **pet,* Ch. **pwat* "8"; Thai *kău,* Ch. *kiəu* "9"; Thai **sĭp,* Ch. **źiəp* "10"; Thai **pak,* Ch. **p∧k* "100." Thai ordinarily agrees with Chinese as opposed to Tibeto-Burman, yet shows interesting variations in the direction of the latter, e.g. Thai **ha,* TB **l-nga* "5," with *h- < ng-* as in Thai **han,* Ch. **ngan,* TB **ngan* "goose"; Thai **hrŏk,* TB **d-rug* "6." The Chinese forms for "7," "9," and "10" illustrate the diphthongization characteristic of that language, the *-ău = -iəu* equation being particularly well attested: Thai **k'ău,* Ch. **k'iəu* "hill"; Thai **k'ău,* Ch. **kiəu* "pigeon"; Thai **kău,* Ch. **g'iəu* "old"; Thai **gău,* Ch. **g'iəu* "owl" (Tibeto-Burman has final *-u* in this series). On the basis of the above phonetic evidence, the borrowing of this numeral system must be assigned to an early period antedating the *r- > l-, -a > -uo* shifts in Chinese. The Thai forms are still irregular, however, with **pet* "8" rather than **puăt,* and **sĭp* "10" rather than **zĭp,* and cannot be reconciled on any scheme of genetic relationship (for the latter, cf. Thai **suk,* Ch. **źiuk* "ripe."

The regular Thai numerals for "1" and "2" are **hnüng* and **song,* respectively, which appear to be remnants of the original Thai numeral system. The corresponding

Chinese terms, significantly enough, appear only in the combinations **sĭp ĕt* "11" (Ch. **'iĕt* "1"), and **ńi sĭp* "20" (Ch. **ńźi < *ńi* "2"). In addition, a basic root **džau* "20" is preserved in Lao and the western Thai languages, and **roi* "100" is found in Siamese, Lao, and some of the eastern Thai languages.[36] The evidence from the numeral system, therefore, cannot be held to speak in favor of the theory of a genetic Thai-Chinese relationship.

The common roots for parts of the body are as follows: Thai **χen* "arm," Ch. **kien* "shoulder"; Thai **veng* "leg," Ch. **γieng* "shin, shank"; Thai **fa,* Ch. **pa* "palm of the hand"; Thai **eu,* Ch. **'iäu* "waist, loins." In this group belong also Thai **hnong,* Ch. **nuong* "pus"; Thai **ńiau,* Ch. **nieu* "urine, urinate." Thai **nga* "tusk, ivory," Ch. **nga* "molar tooth," must be considered in relation to the loan-word for "elephant" (infra), the regular Thai roots for "tooth" being **k'riau* and **văn.* Similarly, Thai **p'iu* "cuticle, epidermis" is connected with Ch. **piu* "skin, hide," but the regular Thai root, **hnăng,* "skin, hide," has no Chinese correspondence. Basic roots for parts of the body such as "eye," "nose," etc. are significantly lacking in this list.

The group of common roots for animal names is equally enlightening in its exclusiveness. Here we find Thai and Ch. **ma* "horse," and the associated roots: Thai **an,* Ch. **'ân* "saddle"; Thai **k'i,* Ch. **g'yię> k'i* "ride (a horse)." These correspondences strongly indicate that the Thai peoples borrowed the horse-complex directly from the Chinese. This group also includes Thai **džang,* Ch. **ziang* "elephant,"[37] Thai **nguă,* Ch. **ngiəu* "bull, ox, cow";[38] Thai **t'ŏ* (restricted distribution), Ch. **t'uo* "hare"; Thai **kăi,* Ch. **kiei* "fowl";[39] Thai **p'rüng,* Ch. **p'iwong* (equivalent to **p'üong* "bee";[40] and, from the above discussion, Thai **han,* Ch. **ngan* "goose (wild)";[41] Thai **k'ău,* Ch. **kiəu* "pigeon"; Thai **gău,* Ch. **g'iəu* "owl"; perhaps also Thai **ngüăk* "crocodile, dragon, siren," Ch. **ngâk* "crocodile." Significantly lacking are roots for "dog," "fish," "bird," "snake," and the like.

The fourth and last of the groups mentioned above includes Thai **ngön,* Ch. **ngiĕn* "silver" (Tibetan *dngul*); Thai **gram* (restricted distribution), Ch. **lâm < *glâm* "indigo" (Tibetan *rams*);[42] Thai **tšiă,* Ch. **t'šie* "paper"; Thai **hmük,* Ch. **mək* "ink"; Thai **băi,* Ch. *b'āi* "cards (for playing)." Here also may be placed Thai *gĕm,* Ch. **iäm < *giäm* "salt," and Thai **guăn,* Ch. *χiuən* "smoke"; it should be noted that the regular Thai root for "salt" is **klüa* rather than **gĕm.* It is apparent that no great importance can be attached to this group of roots.

The above lists of the principal Thai-Chinese correspondences have been carefully drawn up, and should give an accurate picture of this aspect of the problem. There are, to be sure, additional correspondences, some of which have been cited above, but these hardly affect the picture as a whole. Below, by way of contrast, are listed the basic Thai-Indonesian correspondences on which our conclusions have been built. That these are truly basic correspondences as compared with those between Thai and Chinese is sufficiently clear even after a cursory inspection of the material.

1. Thai **wăn* "day," *ta wăn* "sun" ("eye of the day"); IN **wari'* "day, sun"; Laqua *vuon,* Li *ven* "day, sun."
2. Thai **ɓlüăn* "moon"; IN **bulan*; Laqua *nen,* Li *ńan.*
3. Thai **đau* "star"; IN ** 'a(n)daw~*ha(ŋ)g'aw* "sun"; Li *tšěm drau* "star."
4. Thai **fŏn* "rain" ("fine rain," as opposed to **hra* "heavy rain, shower"); IN ** 'ə(m)bun* "atmospheric precipitate" (Tagalog *'ambon* "fine rain"); Li (*pa*) *pŭn* "rain."
5. Thai **năm* "water"; IN **danum*: Li *nom~nam.*
6. Thai **vai* "fire"; In* *'apuy;* Laqua *păi,* Li *pei~fei.*
7. Thai **na* "rice-field"; IN **bəna'* "low-lying land, flooded land"; Laqua *ne,* Li *na~ta* "rice-field."
8. Thai **nŏk* "bird"; IN **manuk* "fowl, bird"; Laqua *nuk* "bird."
9. Thai **răng* "nest"; IN **t'aḷang* (Toba-Batak, Javanese, Malay, Dayak *sarang*).
10. Thai **rüă* "boat"; IN **paḷahu* (Malay *pərahu* "prau"); Li *da.*
11. Thai **tu* "door" (often in composition with **pak* "mouth, opening"); IN **pintu';* Laqua *tu.*
12. Thai **hruă* "head"; IN ** 'ulu'~*hulu';* Laqua *ru,* Li *du~o.*
13. Thai **ta* "eye"; IN **mata';* Laqua *te,* Li *sa.*
14. Thai **đăng* "nose"; IN ** 'ug' ung~* 'ig'ung* (Malay *hidong,* Cham *idung*); Laqua *tang.*
15. Thai **văn* "tooth"; IN ** 'ipən;* Li (*hai*) *p'en.*
16. Thai **pot* "lungs"; IN **put'uh* "heart" (Tagalog *puso',* Toba-Batak *pusu'* "heart," Javanese *pusuh* "lungs").
17. Thai **gĭng* or **grĭng* "body"; IN **daging* "body, flesh."
18. Thai **đuk* "bone"; IN **ta(n)duk* "horn"; Li *drü~füök* "bone."
19. Thai **lüăt* "blood"; IN **darah:* Li *dat~tlat.*
20. Thai **măn* "fat, oil"; IN **mińak* "oil"~**məńak* "fat"; Laqua *nen* "fat," Lati *m-ngă* "fat, oil."
21. Thai **pu,* "grandfather"; IN ** 'ə(m)pu'* "grandfather, grandchild" (reciprocal term).
22. Thai **đăm* "black"; IN ** 'i(n)təm* "black," **ḍəḍəm~*tiḍəm* "dark"; Laqua *dăm,* Li *dŏm* "black."
23. Thai **sŏm* "sour"; IN ** 'at'əm.*
24. Thai **ɓot* "blind"; IN **buta'.*
25. Thai **tai* "die"; IN **matay~*patay;* Laqua *tie.*
26. Thai **đĭp* "raw, green, alive" (Ahom has the doublet forms *dip* "living, to be alive," *lip* "unripe"); IN **huḍip* "live"; Li *diep~fiep* "raw."
27. Thai **kĭn* "eat"; IN **ka'~*ka'ən~*ka'i;* Laqua *küön,* Li *k'an.*
28. Thai **tŏt* "flatus ventris"; IN ** 'ə(n)tut~* 'u(n)tut~*kə(n)tut;* Li *t'uot.*
29. Thai **ku* "I" (pejorative); IN ** 'aku';* Laqua *k'ău,* Li *hau.*
30. Thai **ni* "this"; IN ** 'ini';* Li *nei.*

In the above set of correspondences, the most obtrusive single feature is the development of monosyllabic roots in Thai from the disyllabic roots typical of

Indonesian. It must not be assumed that all the roots involved were originally disyllabic, since in some instances monosyllabic roots can be postulated for proto-IN itself, e.g. **ṭə ṁ~*ḍəm* "black, dark" (No. 22), **tay* "die" (No. 25), **ka'* "eat" (No. 27), **tut* "flatus ventris" (No. 28), and cf. **danum* "water" with ** 'inum* "drink," apparently from a root **num.*[43] The real criterion here lies in comparison with Thai and Kadai, as well as with the more remotely related Mon-Khmer languages, e.g. IN **mata',* Thai **ta* < **m-ta* (vide infra), Mon-Khmer **mat* (Annamite *mắt*) "eye," where **mata* is the only feasible reconstruction for the parent stock. The material assembled in this paper lends itself to the view that the majority of Thai-Kadai-Indonesian roots were disyllabic rather than monosyllabic, and that Thai and Kadai have undergone extensive phonetic reduction. The writer has elsewhere called attention to a parallel reduction in the aberrant Cham dialect spoken on the island of Hainan, and to similar phenomena in the standard Cham speech of the mainland, e.g. Cham *bulan~lan* "moon," *apuĕi~puĕi* "fire" (through aphaeresis).[44] This aspect of Thai-Kadai phonology, therefore, calls for no especial demonstration.

The varying types of phonetic development shown by the Thai roots under consideration are in part explicable in terms of stress variations. Kadai offers an excellent instance of this in the bifurcate development shown by S. Li *du,* N. Li *au* "8," from IN **walu',* where we must reconstruct as follows: **walú'* > **wlu* > *du, *wálu'* > **wau* > *au.* Similarly, for Thai we must postulate shifts of the type: **danúm* > **năm* "water," **pintú'* > **tu* "door," **matá'* > **ta* "eye," but **pút'uh* **pot,* "lungs," **məńak* > **măn* "fat, oil," **búta'* > **bot* "blind," etc. The stress seems normally to have been on the last syllable, but certainly not always so.

The finals of this group of roots present fewer problems than the initials. Among the vocalic finals, we have Thai *-a* = IN *-a* (Nos. 7 and 13); Thai *-u* = IN *-u* (Nos. 11, 21, and 29); Thai *-i* = IN *-i* (No. 30); Thai *-au* = IN *-aw* (No. 3); Thai *-ai* = IN *-ay* (No. 25). Thai **văi,* IN ** 'apuy* "fire" suggest a simple *-ăi* = *-uy* equation, perhaps via an intermediate form *-uei* (cf. Cham *apuĕi),* yet Thai has both *-ui* and *-uei* series, the latter of some importance. Two independent bits of evidence indicate that Thai **văi* was developed from a root **vi,* probably via an intermediate form **viei,* thus paralleling the development shown by Thai **kăi,* Ch. *kiei* "fowl" (see note No. 39). Firstly, Dioi and a group of related dialects[45] have the form *fi* rather than the regular **făi* (contrast Dioi *kăi* "fowl"); secondly, Li has *pei~fei,* from **pi~fi* (cf. the discussion above), rather than **pai~fai* (contrast Li *k'ai,* Thai **kăi* "fowl"; Li *lai,* Thai **klăi* "far"). The lone possible analogy here is furnished by Thai **hăi,* Li *ngei,* IN **tangit'* "weep," with Thai *h-* < *ng-* as discussed above.

Nos. 10 and 12 must be considered in relation to each other. These two comparisons are, admittedly, uncertain ones, but the parallelism between them, extending even into Kadai (Li), has led to their inclusion in our list of correspondences:

	Indonesian	Thai	Kadai (Li)
boat	**paḷahu*	**rüă*	*da*
head	**hulu*	**hruă*	*du*
	** 'ulu*		

The aspiration in these roots seems to have played a role in the *l* > *r* shift. For the final of Thai **rüă* "boat," a possible parallel exists in IN **'at'u'* "dog," Thai **süă* "tiger."

The consonantal finals are regular for the most part. Final *-r,* which is lacking in Thai, is replaced by *-n* (**wari'* > **wăn* "sun, day"), as in loan-words from Khmer or Pali, e.g. Siamese *k'ănun* < Khmer *k'nur* "jack-fruit." In this connection, cf. Siamese and Lao *pun,* IN **'apur~*kapur* "lime" (probably a loanword in these southern Thai speeches). Final *-h,* also foreign to the phonemic system of Thai, is represented by *-t,* as in Kadai (**darah* > **lüăt* "blood"). Final stops and nasals are preserved in Thai, with the exception of *-t'* > *-t* (No. 16). The most likely instance of interchange of nasal finals is furnished by IN **rumah,* Li *düön,* Thai **rüăn* "house," perhaps via the forms **ruam* > **ruan.*

Short medial vowels are predominant in the Thai roots under consideration, and must be regarded as characteristic of these basic roots as a group. IN medial *-a-* is represented by *-ă-* (Nos. 1, 2, and 9); cf. also IN **balakang,* Thai **hlăng* "back" (n.). Thai ordinarily has *-ă-* for IN medial *-ə-* (Nos. 15, 20, and 22), yet has *-ŏ-* in one instance (No. 23); for the latter, cf. In **tirəm,* Siamese *hoi irŏm* "oyster," undoubtedly a loan-word in Siamese (in composition with the Thai root **hoi* "shellfish"). IN medial *-u-* is represented by *-ă-* (Nos. 5 and 14) and *-ŏ-* (Nos. 4, 8, and 28), as well as by *-u-* (No. 18).[46] After the labial stop initials *p-* and *b-,* Thai has *-o-* rather than *-ŏ-* for IN medial *-u-* (Nos. 16 and 24). An additional equation is furnished by Nos. 17 and 26, yielding IN medial *-i-* = Thai *-ĭ-.* The medial vowel of Thai **kĭn* "eat" (No. 27) cannot be satisfactorily explained on the basis of our present knowledge, though the contrast with the *-a-* vocalism of Li is matched by Thai **đĭn,* Li *dăb~fan* "earth," perhaps related to IN **tanah~*tanəh* "earth, land."

The treatment of initial consonants in Thai presents a number of interesting features. The first of these to come to the writer's attention is the peculiar aspiration of the Thai roots for "eye" (No. 13) and "die" (No. 25) in the Tho-Nung group of dialects.[47]

	IN	Laqua	Siamese	Tho	Nung
eye	**mata'*	*te*	*ta*	*t'a*	*t'a~ha*
die	**matay*	*tie*	*tai*	*t'ai*	*t'ai~hai*

With one partial exception, these are the only roots so treated in Tho and Nung,[48] hence this phenomenon cannot be explained in terms of Thai itself. On the basis of Indonesian, however, we can postulate a development of the type: **mata'* > **m-ta* > **m-t'a* > *t'a~ha; *matay* > **m-tay* > **m-t'ay* > *t'ai~hai,* with secondary aspiration after the nasal prefix.[49] The Li form *sa* "eye" can be explained along the same lines. The remarkable parallelism shown in the treatment of these two roots constitutes perhaps our most significant single piece of evidence for a Thai-Indonesian linkage.

The reconstruction of initial *ƀl-* for Thai, as in the root **ƀlüăn,* represents a new advance in Thai phonology. As ordinarily reconstructed, Thai has initial *bl-, pl-,*

p'l, and *br-, pr-, p'r,* but neither *ƀl-* nor *ƀr-.* The typical Thai initial *đ-* series shows the following equation: Siamese and Lao *đ-* = Ahom, White Tai, Tho, Nung, Dioi *d-* = Shan and Black Tai *l-* = Khamti *n-* (vide supra). Three roots, however, diverge from this equation in the direction of the initial *ƀ-* series, and in one of these roots *bl-* is actually preserved in the archaic Ahom language, hence we can safely reconstruct all three roots with initial *ƀl-*:

a. Siamese and Lao *đok* "flower," but Ahom *blok,* Tho *biok,* Nung *beok* Black and White Tai *bŏ* < *ƀok,* Shan *mok* < *ƀok.*
b. Siamese and Lao *đüăn,* Ahom *dön,* Shan *lön* "moon," but Tho and Nung *büön* "month," Black Tai *büăn,* White Tai *bön* "moon."
c. Siamese *đi,* Ahom, Tho, Nung *di,* Shan *li* "bile," but Lao and White Tai *bi,* Dioi *di* "animal bile"~*bi* "human bile."

Reconstructions: **ƀlok* "flower," **ƀlüăn* "moon," **ƀli* "bile."[50] Note that initial *ƀl-,* which is of labial type, is best preserved before the labial vowel *o,* and worst preserved before the front vowel *i;* also that Black and White Tai preserve *b-* in all three roots. Initial *ƀr-* cannot be reconstructed for Thai, and may be represented simply by *ƀ-;* cf. Thai *ƀŏm* "ripen fruits," IN **pəḷəm* "ripen fruits artificially" (Toba-Batak *porom,* Malay *pəram*), perhaps via a form **pərəm.*[51]

Some interesting equations appear among the stop consonants, especially in the labial series. Thai ordinarily has *t* < *t* (Nos. 11, 13, 25, and 28), and *đ* < *d* (Nos. 3, 18, and 26), while the correspondence shown in Thai **đăng* IN ** 'ug'ung~* 'ig'ung* "nose" (No. 14) must be considered in connection with the IN doublet forms ** 'a(n) daw~*ha(ŋ)g'aw* "sun" (No. 3). Thai **đăm* "black" must therefore be equated directly with IN **ḍəḍəm~*tiḍəm* "dark" rather than with ** 'i(n)təm* "black" (No. 22). The palatal stop (*t'*) of IN is represented in Thai by *s-* as an initial (No. 23), but by *-t* as a final (No. 16).[52] The velar correspondences are regular: Thai *k* = IN *k* (Nos. 27 and 29); Thai *g-* = IN *g* (No.17). In the labial series, however, we find two types of correspondences, viz. Thai *ƀ* = IN initial *b-* (Nos. 2 and 24), Thai *p* = IN initial *p-* (No. 16), but Thai *f* = IN medial *-b-* (No. 4), and Thai *v* = IN medial *-p-* (Nos. 6 and 15). Thai **pu,* IN ** 'ə(m)pu'* "grandfather" (No. 21) would seem to run counter to this scheme, but in this instance IN has the doublet roots **tumpu'* "forefather, sir" and **pu'* "sir," the latter evidently the basic etymon from which Thai **pu* was derived. This explanation of Thai *f* and *v* as secondary phonemes deirved from medial labial stops clears up one of the most abstruse aspects of Thai phonology. Li (southern dialect) and the Kadai languages in general have preserved the labial stop in these roots:

	Indonesian	Li	Thai
rain	** 'ə(m)bun*	*pŭn*	**fŏn*
fire	** 'apuy*	*pei*	**văi*
tooth	** 'ipən*	*p'en*	**văn*

It is a striking fact that, in the above set of comparisons, initial *ƀ* and *đ* appear to the exclusion of the sonant stops *b* and *d*. An examination of the stock of Thai roots assembled by the writer shows an overwhelming predominance of basic roots with initial *ƀ* and *đ*, some of the most important of which are listed below:

Initial *ƀ*-: **ƀon* "arum," **ƀli* "bile," **ƀau* "bridegroom," **ƀö* "butterfly," **ƀěk* "carry (on shoulders)," **ƀŏt* "cloud" (v.), **ƀiă* "cowrie shell," **ƀău* "crucible," **ƀo* "pit, well, mine," **ƀlok* "flower," **ƀĭn* "fly" (v.), **ƀět* "fish-hook," **ƀaü* "leaf," **ƀău* "light (not heavy)", **ƀuă* "lotus, water-lily," **ƀa* "mad," **ƀa* "shoulder," **ƀok* "speak," **ƀok* "tube," **ƀuei* "cocoanut spoon," **ƀan* "village," **ƀat* "wound," **ƀong* "hole," **ƀot* "blind," **ƀlüăn* "moon."

Initial *đ*-: **đüăt,* "boil" (v.), **đěk* "child," **đĭn* "earth," **đăp* "extinguish," **đüă* "fig," **đŏng* "forest," **đi* "good," **đam* "handle" (n.), **đon* "high, hill," **đong* "kind" (n.), **đăi* "ladder," **đu* 'l'ook," **đoi* "mountain," **đong* "parents of in-laws," **đeng* "red," **đăng* "shield" (n.), **đŏm* "smell" (v.), **đut* "suck," **đet* "sun, sunshine," **đap* "sword," **đăi* "thread," **đăng* "pillar," **đuăng* "fish-trap," **đŏng* "winnowing instrument," **düăn* "worms," **đăi* "obtain, be able."

In contrast to this impressive array, the sets of roots with initial *b* and *d* seem restricted indeed. With initial *b*- we find **bě* "goat" (but **ƀě* in Lao and White Tai), **be* "raft," **bu* "mountain" (but Siamese has *b'u,* as in loan-words) **băn* "seed, kind," **brŭk* "tomorrow," **bra* "large knife," **bi* "fat, big," **bi* "elder sibling" (perhaps etymologically connected with the foregoing), **bo* "father," while with initial *d*- we find **dong* "belly," **drai* "sand, gravel," **dak* "leech," **diăng* "true, correct," **dang* "road," and **do* "weave." The contrast is so marked that one is tempted to conclude that roots with *ƀ*- and *đ*- belong to the older Thai-Kadai-Indonesian stratum, and roots with *b*- and *d*- to one or more younger superimposed strata, including Chinese loan-words such as **băi* < Ch. **b'āi* "cards" (vide supra). The existence of the roots **bi* "elder sibling" and **bo* "father," with initial *b*-, does not constitute a conclusive argument against this view, inasmuch as the Thai kinship nomenclature as a whole appears to have no intimate connection with the Indonesian. The presence of initial *ƀ*- or *đ*- in a given root may even be used as supporting evidence for a proposed Indonesian comparison, e.g. Thai **ƀa,* IN **bara'* "shoulder" (**bara'* > **baa* > **đā,* contrasting with the development shown in **wari'* > **wăn* "day, sun"); Thai **đěk* "child," IN **ḍikih*~**'ə(n)ṭik*~**'iṭ ik* "small" Li *tik*~*tok* "small."

Still another problem is presented by Thai **lüăt* < IN **darah* "blood" (No. 19), apparently via a form **dlat* (*đl*- is not retained in Thai). A possible analogy here is furnished by the Thai root for "tongue," which the Li dialects treat in a parallel manner:

	Indonesian	S. Li	N. Li	Shaved Head	Thai
blood	**darah*	*dat*	*tlat*	*slat*	**lüăt*
tongue	**dilah*	*dien*	*tlien*	*slien*	**lĭn*

Yet Thai has **pla* "fish," corresponding to S. Li *da,* N. Li *tla,* Shaved Head Li *sla,* with initials as in the above series, hence no certain conclusions can be drawn.

The above discussion does not exhaust the possibilities of the complex Thai-Indonesian field, and it is possible that a more searching analysis of Indonesian material will yield further comparisons, yet it is believed that most of the important lexical correspondences have been uncovered. The writer has eliminated from the discussion certain obvious loan-words in Siamese, e.g. *muăng "mango"* < IN **maŋga.* Attention should be called, however, to the noteworthy agreement between Thai **nga* and IN **ləga'* "sesame." The Thai root **nga* is widely extended in that stock (Siamese, Lao, Shan, Ahom, White Tai, Nung), hence cannot be regarded as a recent loan from Indonesian.

It is apparent that our judgment must be based almost entirely on lexical rather than morphological analogies, inasmuch as the rather elaborate affixation system of Indonesian is not represented in Thai. We must remember, however, that the reduction of disyllabic or trisyllabic roots to monosyllabic forms, as in Thai, necessarily involves the loss or incorporation of affixed elements. Thus, if a root **tay* "die" be reconstructed for proto-IN on the basis of the doublet roots **matay*~**patay,* and the elements *ma*- and *pa*- be regarded as prefixes, the purely phonetic development **matay* > **m-tay* > **tai,* paralleling **mata'* > **m-ta* > **ta* "eye" (vide supra), necessarily entails the loss of this prefixed element. In other instances, the affix may have been incorporated in the derived form; cf. IN **miñak* "oil"~**məñak* "fat," Lati *m-ngă,* Thai **măn* "fat, oil" (No. 20), and IN **ka'*~**ka'ən*~**ka'i,* Thai **kĭn* "eat" (No. 27).

Of some interest in this connection are the traces of prefixes preserved in Siamese. Siamese prefix *kă-,* by far the most prominent of the lot, is found with a few words for parts of the body (*kă-đuk, kă-điau* "bone," *kă-đo* "male genitals," *kă-bŏ*~*bŏ* "stomach"), and with some animal names (*kă-tai* "hare," *kă-tšok* "sparrow," *kă-te* "tupaya"), but is characteristically associated with curious derived forms, e.g. *ƀong* "stick"~*kă-ƀong* "cudgel," *đong* "oscillating" ~*kă-đong* "distorted, twisted," *tšŭn* "push"~*kă-tšŭn* "touch lightly," *tŭn* "mole"~*kă-tŭn* "kind of large rat." No great importance can be attached to this prefix, yet one possible IN correspondence has been uncovered, viz. Siamese *kă-đuk* "bone," IN **ta(n)duk* "horn"; cf. the *kă-*~*tă-* interchange in *kă-bŏng*~*tă-bŏng* "mussel."

The problem of the development of tones in Thai cannot satisfactorily be handled until good material on Kadai tones is made available. As reconstructed, the Thai tonal system includes two series of tones, one connected with roots having surd initials, the other connected with roots having sonant initials (a similar system is found in Annamite and Chinese). Each of these series, furthermore, includes three tonemes, the original values of which are uncertain. It is probably significant that almost all the Thai roots having IN correspondences are associated with a single toneme, represented in Siamese by the mid-level tone (with sonant and unaspirated surd stop initials) or the high-rising tone (with other surd initials). The only exceptional roots here are **năm* "water," **ni* "this," **pu* "grandfather," and **sŏm* "sour."

Aside from the rudimentary prefixes found in Siamese, the Thai stock closely conforms to the classical type of monosyllabic, isolating languages. Maspero has successfully refuted Wulff's thesis of infixation in Siamese (see note No. 33), hence no comparison with Indonesian infixes can be made. As pointed out above, Thai agrees with Indonesian and Kadai, and sharply diverges from Chinese, in placing modifying elements after rather than before modified elements. This significant agreement in syntax contributes no little support to our Thai-Kadai-Indonesian hypothesis. Attention must also be called to the traces of a distinction between inclusive and exclusive forms for the 1st pers. pl. pronoun in Thai, as represented by the exclusive form **tu* "we" in Khamti, Lao, and archaic Siamese.[53] This distinction is paralleled in Indonesian in the forms **kita'* "we" (inclusive), **kami'* "we" (exclusive).

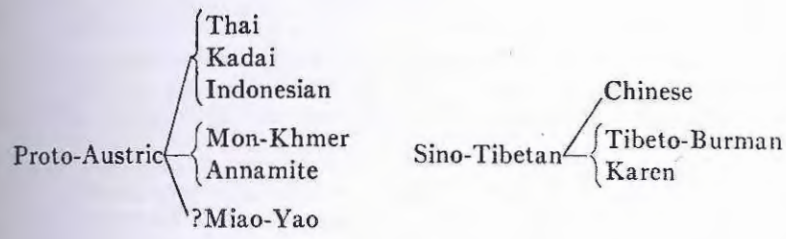
The Thai-Kadai-Indonesian hypothesis, as outlined in the present paper, bears far-reaching implications for the history of the peoples of southeastern Asia and Oceania. If we accept the view that these three linguistic stocks are genetically related, we must place the center of their dispersion somewhere in the South China area, the present home of the Kadai tribes as well as the early home of the Thai peoples.[54] On the basis of this distribution we can conclude, with a high degree of probability, that the proto-IN-speaking peoples migrated from the South China coast, perhaps via the island of Hainan, to Formosa on the north, the Philippines on the east, and Annam, Borneo, Java, Sumatra, and the Malay Peninsula on the south. The Cham and Malay linguistic areas, in southern Annam and the Malay Peninsula, respectively, surely are to be regarded as Indonesian enclaves on the Asiatic mainland, not as possible points of departure for the Indonesian migrations.

In still broader perspective, Thai-Kadai-Indonesian appears in its true light as the northern division of Schmidt's Austric superstock.[55] The archaic cleavage between Thai-Kadai-Indonesian on the one hand, and Mon-Khmer on the other, must have come about in the South China-Indochina area, with subsequent localizations of these two divisions in the north and south, respectively. The anomalous position of Malay at the present day, south of the main body of Mon-Khmer speeches, can be explained only on the basis of a seaborne migration from the islands of Indonesia. Thai and Kadai in the north, Cham in the east, and Malay in the south, show a peripheral distribution with respect to the Mon-Khmer languages. As suggested above, Cham and Malay fall into their place in this picture as intrusive Indonesian languages overlying a Mon-Khmer substratum.

Annamite, too, takes its proper place as the northeasternmost member of the Mon-Khmer stock. Annamite stands in relation to Mon-Khmer somewhat as Thai stands in relation to Indonesian. Like Thai, it has suffered much phonetic attrition, has developed a complete tonal system, and has lost its morphological apparatus of affixes. These changes must be attributed to Thai influence, in view of the not inconsiderable body of Thai roots in the language. The overwhelming majority of basic roots, however, are of Mon-Khmer rather than Thai origin. On the analogy of our analysis of Thai, there can be no question as to the genetic nature of the Mon-Khmer-Annamite relationship.[56]

With Thai, Kadai, and Annamite in their proper settings, the linguistic picture of southeastern Asia assumes definitive shape for the first time. There remains only one linguistic problem of major importance, viz. the affinities of the Miao-Yao stocks of languages, spoken throughout much of central and southern China and northern Siam and Indochina. Our material on these languages is scanty and generally poor, and almost no comparative work has been done on the group.[57] Miao and Yao are well differentiated divisions of a single stock, and each appears in a number of dialectical varieties, with Miao showing the greater variation. Pateng, spoken in the Rivière Claire section of Tonkin, is a subsidiary member of the stock.[58] Miao-Yao resembles Thai-Kadai and Annamite in its monosyllabism and tonality, and further investigation may reveal a relationship with Proto-Austric or with one of its later divisions. A final judgment here must await the reconstruction of Mon-Khmer and the assembling of more material on the Kadai languages.

The proposed classification of Southeast Asiatic languages is as follows



On the ethnological side, the Kadai group offers the most promise for future investigation. At present, our material on this group is confined to the scraps of information gathered by Bonifacy and Lunet de Lajonquière, together with Stübel's fairly extensive study of the Li tribes (see sources cited above). Stübel points out a number of Indonesian and Micronesian parallels, e.g. in weaving (cit. supra, p. 293) and basketry (id., p. 294), and expresses his astonishment at the general cultural similarity to the tribes of Formosa (id., p. 296). It may be that the Li retain certain Indonesian culture traits that have been discarded by the Sinicized Kadai tribes of the mainland. It is fairly evident, however, that the general ethnological picture of Thai-Kadai-Indonesian has been destroyed beyond repair, and that our linguistic thesis must stand or fall on its own merits.[59]

Notes

1 For some of the earlier speculation, see W. Churchill, *The Polynesian Wanderings* (Washington, 1911), largely devoted to a criticism of MacDonald's Semitic theory.

2 An outstanding example is furnished by Handy's derivation of the Polynesian Tangaloa cult from southern China (*Polynesian Religion,* Bulletin of the Bernice P. Bishop Museum, no. 34, 1927, pp. 312–330).

3 Other variants are: B'lai, K'lai, S'lai, S'ai, Hiai, Lai, Loi, Le, Dli, B'li. The Chinese character employed for this name is pronounced *li* in North China dialects, *lai* in Cantonese, and *loi* in Hoklo.

4 R. Swinhoe, *The Aborigines of Hainan* (Journal of the North China Branch of the Royal Asiatic Society, Vol. 7, 1871), pp. 25–40; J. Calder, *Notes on Hainan and its Aborigines* (China Review, Vol. 11, 1882), pp. 42–50; E. H. Parker, *The Li Aborigines of K'iung Shan* (China Review, Vol. 19, 1890), pp. 383–387; C. C. Jeremaissen, *Loi Aborigines of Hainan and their Speech* (China Review, Vol. 20, 1892), pp. 296–305; F. M. Savina, *Lexique dày-français, accompagné d'un petit lexique français-dày et d'un tableau des différences dialectales* (Bulletin de l'École Française d'Extrême-Orient, t. 31, 1931), pp. 103–199; H. Stübel, *Die Li-Stämme der Insel Hainan; Ein Beitrag zur Volkskunde Südchinas, unter Mitwirkung von P. Meriggi* (Berlin, 1937).

5 A. Bonifacy, *Étude sur les langues parlées par les populations de la haute Rivière Claire* (Bulletin de l'École Française d'Extrême-Orient, t. 5, 1905), pp. 306–323; *Étude sur les coutumes et la langue des La-ti* (*Ibid.*, t. 6, 1906), pp. 271–278; *Étude sur les coutumes et la langue des Lolo et des La-qua du Haut Tonkin* (*Ibid.*, t. 8, 1908), pp. 531–558.

6 E. Lunet de Lajonquière, *Ethnographie du Tonkin Septentrional* (Paris, 1906). Word-lists on pp. 357 (Kelao), 340 (Pen-ti-Lolo = Laqua), and 359 (Lati).

7 *Among the Tribes in South-west China* (London, 1911).

8 E. H. Parker, *Siamese Words in Hainan and China* (China Review, Vol. 18, 1889), p. 198.

9 W. Strzoda, *Die Lie auf Hainan and ihre Beziehungen zum asiatischen Kontinent* (Zeitschrift für Ethnologie, Bd. 43, 1911), pp. 193–236. Strzoda concludes, however, that "Die meisten Li-Numeralia . . . sint Rätsel und lassen sich nirgends unterbringen" (pp. 219–220).

10 *The Languages of China before the Chinese* (London, 1887). See especially his conclusions on p. 73: "In the numerals, for instance . . . similarities exist with those of some tribes of Formosa. But they are remote, and do not come from a direct relationship; they are apparently survivals of a former state of things, previous to their respective migrations, when their various ancestors had relations between themselves on the continent." An English traveller in Hainan, B. C. Henry, had somewhat earlier sought to connect the Li with the Malay on onomastic grounds (B'lai or B'lay = Malay), in his article, *The Close of a Journey through Hainan* (China Review, Vol. 12, 1883), pp. 109–124, esp. p. 115.

11 Review of Savina, *Monographie de Hainan* (1929), in Bulletin de l'École Française d'Extrême-Orient, t. 30, 1930 (pp. 436–444). Of his own Cham and Malay comparisons, however, Mus remarks: "Ces rapprochements sporadiques restent jusqu'ici de simple curiosité."

12 Review of Savina, *Lexique dày-français* (1931), in Bulletin de la Societé de Linguistique de Paris, t. 34, pt. 3, 1933, pp. 228–236.

13 Cit. supra, 1908, p. 557. Bonifacy adds the following remark: "Là paraissent s'arrêter les ressemblances entre les deux langues [Laqua and Cham], à supposer même que celles que nous signalons ne soient pas purement fortuites."

14 *Die Sprachfamilien und Sprachkreise der Erde* (Heidelberg, 1914). Schmidt places Laqua in an artificial "Eastern Thai" group.

15 *Notes ethnographiques sur les tribus du Kuey-tcheou* (Anthropos, Bd. 6, 1911), pp. 318–344; citation from p. 318.

16 None of our records of Kadai languages is satisfactory as regards transcription. Savina employs the cumbersome and inadequate *qu'ôc-ngû* system of Annamite, while Bonifacy makes use of a modified version of the same system. In the present study open vowels are written as short vowels (*ĕ, ŏ*), and the "bearded o" (*o'*) and "bearded u" (*u'*) as front rounded vowels (*ŏ, ü*).

17 Lati has retained the final stop only in *a-liep* "claw" (Thai **lĕp* "fingernail") and the Annamite loan-word *but* "pen."

18 Thai *đ-* and *ɓ-* are best reconstructed as lenis surd stop initials, since they belong in the high tonal series along with the regular surd stops (*t-* and *t'-*, *p-* and *p'-*). Li agrees with the majority of Thai languages proper in having *b-* for Thai *ɓ-*, as in Li *böü*, Thai **ɓaü* "leaf."

19 Represented in Savina's transcription by the Annamite tone *năng*.

20 For the general argument here, see the Thai discussion below.

21 *Vergleichende Lautlehre des austronesischen Wortschatzes; Bd. III; Austronesisches Wörterverzeichnis* (Beihefte zur Zeitschrift für Eingeborenen Sprachen, Bd. 19, 1930). Forms as cited by Dempwolff, with the exception of *y* for *j*, *w* for *v*, and *r* for *γ*. "Facultative" nasal infixes are enclosed in parentheses.

22 The IN system is decimal rather than quinary, yet the Formosa languages show irregular features suggestive of the latter, e.g., Sek-hwan has 5+1 = 6, 5+2 = 7, etc., and Tsuihwan and Bu-hwan have 3×2 = 6, 4×2 = 8; vide T. L. Bullock, *Formosa Dialects and their connection with the Malay* (China Review, Vol. 3, 1875, pp. 38–46).

23 Cited in R. Brandstetter, *Malaio-polynesische Forschungen; IV: Mata-Hari, oder Wanderungen eines indonesischen Sprachforschers durch die drei Reiche der Natur* (Luzern, 1908), p. 6.

24 *Contribution à l'ĕtude du systeme phonétique des langues thai* (Bull. de l'École Française d'Extrême-Orient, t. 11, 1911), pp. 153–169.

25 The most important of these are: *Note sur les tons et les initials du vieux siamois à l'époque de Sukhodaya* (Journal of the Siam Society, Vol. 21, 1927), pp. 103–117); [*v*] *et* [χ] *et leur origine* (*Ibid.*), pp. 119–126.

26 *Chinesisch und Tai: Sprachvergleichende Untersuchungen* (Det Kgl. Danske Videnskabernes Selskab., Historisk-filologiske Meddelelser, Vol. 20, pt. 3, 1934).

27 Notably in Ahom, where Wulff failed to make use of the most important lexicon on that language (Borua, *Ahom-Assamese-English Dictionary,* Calcutta, 1920); Tho, completely neglected by Wulff but for which we have two utilizable sources, viz. E. Diguet, *Étude de la langue Thô* (Paris, 1910), and Fr. Th. Gordaliza, *Estudio sobre el dialecto Thô de la región de Lang-sön* (Anthropos, Bd. 3, 1908), pp. 512–532; White Tai, for which an excellent source has recently appeared, viz. G. Minot, *Dictionnaire tăy blanc-français* (Bull. de l'École Française d'Extreme-Orient, t. 40, fasc. 1, 1940, pp. 1–237.)

28 The writer has in preparation a comparative dictionary of the Thai languages, based in large part on materials collected by the Sino-Tibetan Philology Project of the Works Progress Administration, sponsored by Prof. A. L. Kroeber of the University of California during the years 1935–40. The writer here wishes to record his indebtedness to Prof. Kroeber for having made possible this investigation of Far Eastern languages, of which the present paper may be regarded as a by-product.

29 R. Shafer, *The Vocalism of Sino-Tibetan* (Journal of the American Oriental Society, Vol. 60, 1940), pp. 302–337; (Vol. 61, 1941), pp. 18–31.

30 *Transcription and Explanation of the Siamese Alphabet* (Asia Major, Bd. 1, 1924), pp. 45–66; *Siamese Mute H* (*Ibid.*, Bd. 3, 1926), pp. 33–48.

31 Vide the occasional references to Thai problems in his articles, *Semantic Differentiation in Indo-Chinese* (Harvard Journal of Asiatic Studies, Vol. 4, 1939), pp. 213–229, and *Studies in Indo-Chinese Phonology* (*Ibid.*, Vol. 5, 1940), pp. 101–127.

32 A. Conrady, *Eine merkwürdige Beziehung zwischen den austrischen und den indochinesischen Sprachen* (Kuhn Festschrift, München, 1916), pp. 475–504); *Neue austrisch-indochinesische Parallel* (Asia Major, Introductory Volume, 1922), pp. 23–66.

33 Bulletin de la Société de Linguistique de Paris, t. 36, pt. 3, 1935, pp. 183–187.

34 Edited by A. Meillet and M. Cohen (Paris, 1924), pp. 361–384 (*Le Sino Tibétain*).

35 *Siamese Studies* (T'oung Pao, t. 2, n. s., 1902, Supplement).

36 Cf. the penetrating study by Coedès and Burnay, *Notes d'étymologie Taï, No. 1: Le nom de nombre "Cent"* (Journal of the Siam Society, Vol. 20, 1926), pp. 49–52. Coedès and Burnay identify **roi* "100" with the root **roi* "to string." They further conclude that **pak* "100" is common Thai because of the concordance of tones, yet admit the possibility of its having been borrowed from Chinese by the parent Thai speech.

37 Cf. the associated correspondence between Thai **nga* "tusk, ivory" and Ch. **nga* "molar tooth." The root for "elephant" has a restricted extension in Tibeto-Burman (Burmese *ts'ang*).

38 Ch. **ngiəu* should have produced Thai **ngău* rather than **nguă.* The latter corresponds rather to the root **ngua* "bull, ox, cow" of the Kachin-Nung-Burmese division of Tibeto-Burman.

39 For the finals, cf. Thai **găi* "who, which," Ch. **viei* "why, how, what"; Thai **k'ăi,* Ch. **k'iei* "to open." The Thai root for "egg" (**k'răi*) is independent of the Chinese roots (**luân, *d'ân*).

40 The interesting Chinese root **miĕt* "honey," an ancient loan-word from Indo-European (Sanskrit *madhu;* Old Slavic *med ŭ;* Tocharian *mit,* whence Ch. **miĕt* through diphthongization; Greek *μέθυ* "wine"; English *mead*), is not found in Thai, which makes use of the periphrasis "bee-water" (Siamese and Shan *năm p'üng*), or even equates "bee" with "honey" (Ahom, Tho) or with "wax" (Lao).

41 In view of the correspondences for "fowl" and "goose," it is somewhat surprising to find distinct roots for "duck" (Thai **pĕt,* Ch. **ap*).

42 The prototype must have been **ram-s* rather than **gram-s,* since the latter could have yielded only **grams* in Tibetan. On this line of reasoning, Chinese **lâm* < **glâm* includes a prefixed *g-* element, and the Thai borrowing can thus be dated as posterior to this prefixation, but anterior to the subsequent **grâm* > **glâm* > **lâm* development in Chinese (completed ca. 500 B. C.).

43 Cf. the remarks in R. O. Windstedt, *Malay Grammar* (Oxford, 1927), p. 20, and S. H. Ray, *A Comparative Study of the Melanesian Island Languages* (Cambridge, 1926), pp. 38 and ff.

44 See the writer's article, *A Cham Colony on the Island of Hainan* (Harvard Journal of Asiatic Studies, Vol. 6, 1941), pp. 129–134.

45 The form *fi* "fire" is cited for the Tai Yoi, Kon Yai, and To-jen dialects in W. C. Dodd, The Tai Race (Cedar Rapids, Iowa, 1923), word-lists on pp. xiv–xxi.

46 Note Thai medial *-ă-* < *-u-* only before final nasals. Thai medial *-ŏ-* < *-u-* seems to be the normal development before final stops; cf. Thai **hrŏk,* Chinese **liuk* "6"; Thai **mŏt,* Malay, Javanese, Karo *sĕmut* "ant" (cited in Brandstetter, cit. supra, p. 37). For Thai medial *-u-* < *-u-* before final velar stop, cf. Thai **luk* "anything round," IN **kəluk~*pəluk* "bend, curve."

47 Our sources for Tho and Nung are in agreement on this point, and there can be no doubt as to the reality of the phenomenon in question. An additional check is furnished by the form *t'a* "eye" cited for a dialect of Tho-Nung type by Yu Wên, *A Vocabulary of a Non-Chinese Tribe inhabiting the Tapingfu Area of Kwangsi Province, with Chinese Transliterations and Notes* (Academia Sinica, Bulletin of the Institute of History and Philology, Vol. 6, pt. 4, 1936), pp. 505–552 (in Chinese).

48 Tho and Nung *t'en* "wasp" correspond to the general Thai root **ten,* but the doublet form in initial *h-* is lacking in Nung. The regular development with unaspirated initial *t-* is observed in a long series of Thai roots, including **tăp* "liver," *tăm* "low," *tăng* "glue," *tău* "turtle," *tăt* "cut," *tĕm* "full," *tin* "foot," *tŏk* "fall," *tŏm* "mud," *tŏn* "tree trunk," *tŏt* "flatus ventris," and *tuă* "animal."

49 A good parallel here is furnished by Tibetan, which has aspirated all initial surd stop or affricate consonants after prefix *m-*, e.g., Tibetan *mtš'in* "liver", corresponding to Tibeto-Burman **m-šin.*

50 For Thai **ƀlok* "flower," cf. the subsidiary IN root represented by Bisaya *bolak,* Tagalog *bulaklak* "flower," which Brandstetter (cit. supra, p. 22) derives from a root **lak* "to unfold." No IN comparison has been uncovered for Thai **ƀli* "bile."

51 This comparison is semantically too specific to be trusted, and we should expect Thai **pŏm* rather than **ƀŏm.* IN *l* > *r* as in **pəḷahu* > **rüă* "boat," **t'alang* > **răng* "nest"; IN *ə* > *ŏ* as in ** 'at'əm* > **sŏm* "sour."

52 Dempwolff's reconstruction of *t'* rather than *s* for IN is open to criticism; cf. the review by A. Capell, in Bull. of the School of Oriental Studies, Vol. 9, 1938, pp. 459–462. Thus, IN ** 'at'əm* "sour" is represented by Tagalog *'asim,* Toba-Batak *'asom,* Javanese *'asĕm,* Malay *'asam,* Dayak *'asem,* all with initial *s.*

53 Cf. the discussion in G. Coedès, *Nouvelles notes critiques sur l'inscription de Rāma Khamheng* (Journal of the Siam Society, Vol. 17, 1923), pp. 113–120.

54 The general Thai movement southward into Indochina appears to have begun on a large scale only toward the close of the first millennium A. D. The first group of Siamese inscriptions, the Sukhodaya, are from the 13th to 16th centuries, and the famous Rāma Khamheng inscription, the earliest monument of the Siamese language, is dated no earlier than 1292; cf. G. Coedès, *Notes critiques sur l'inscription de Rāma Khamheng* (Journal of the Siam Society, Vol. 12, 1918), pp. 1–27, and *Recueil des Inscriptions du Siam; Première Partie: Inscriptions de Sukhodaya* Bangkok, 1924).

55 The writer accepts Schmidt's postulation of an Austric superstock including Mon-Khmer and Austronesian, even though this relationship has not yet been thoroughly demonstrated. In the present instance, the Austric hypothesis is useful in interpreting certain roots which Thai and Mon-Khmer have in common, notably Thai **ɣo,* Mon-Khmer **go* (Annamite *ko*) "neck." Cf. the Thai-Khmer comparisons listed in Wulff, cit. supra, pp. 68–70, and the Dioi-Khmer and Dioi-Bahnar comparisons in D. Doutreligne, *Contributions à l'étude des populations Dioy du Lang Long* (Anthropos, Bd. 26, 1931), pp. 35–53.

56 H. Maspero, *Études sur la phonétique historique de la langue annamite* (Bull. de l'École Française d'Extrême-Orient, t. 12, 1912), pp. 1–126, was so impressed by the monosyllables and tones of Annamite that he postulated a genetic kinship with Thai, even in the face of the dominant Mon-Khmer lexical element. Przyluski, in *Les Langues du Monde* (*cit. supra*), rightly breaks with Maspero on this point and classifies Annamite with Mon-Khmer.

57 Limited comparative notes on two Miao dialects are found in Yu Wên, *The Influence of Liquids upon the Dissolution of Initial Consonant Groups in the Indo-Sinic Family* (Journal of the North China Branch of the Royal Asiatic Society, Vol. 69, 1938), pp. 83–91. A brief and somewhat confused study of two Miao and two Yao dialects has been made by Fang-kuei Li, *A Yao Dialect in Ling-Yün, Kwangsi Province* (Academia Sinica, Bulletin of the Institute of History and Philology, Vol. 1, 1930), pp. 419–426 (in Chinese).

58 Vide A. Bonifacy, *Monographie des Pa-teng et des Na-ê* (Revue Indo-Chinoise, n. s., t, 10, 1908), pp. 696–706, 773–786.

59 The writer has not had access to the most recent comparative study on the Li, viz. Chungsee Liu, *Preliminary Study of the Origins of the Tribes of Hainan Island* (Meridio-Occidentale Sinense, Vol. 1, No. 1, 1940), pp. 1–23.

4

CLASSIFICATION OF THE SINO-TIBETAN LANGUAGES

Robert Shafer

Source: *Word* 11, 1, 1955, 94–111.

The present world-wide misconception that the Sino-Tibetan family of languages is divided into an eastern "Chinese-Siamese" sub-family and a western "Tibeto-Burman" sub-family[1] has resulted from two distinguished scholars working on opposite sides of the area and their almost total ignorance of the languages in their colleague's field.

Since about 300 Sino-Tibetan languages and dialects have been recorded, some division of labor has naturally occurred in their investigation. And this has followed geographical lines. In southeast Asia, great streams—the Mekong, Salween, Irrawaddy, and Chindwin—flow in a generally southerly direction, and the intermediate ridges, the monsoon forests, the Malay peninsula extending far southward, and the political division with Indo-China under France and Burma and India under Great Britain have hindered communication and have tended to compartmentize knowledge into one division facing twoard the Pacific Ocean and another facing toward the Indian Ocean.

Henri Maspero, the last great scholar to hold to the "Siamese-Chinese" division, was a product of the École Française d'Extrême-Orient at Hanoi. He published brilliant work on both the Chinese and Daic (Thai) languages. If one knows two languages one is bound to note some resemblances between them whether they are related or not.[2] Maspero noted a considerable number of parallels between Daic words and the corresponding Chinese forms—enough, he thought, to consider these languages to be closely related genetically. But Maspero knew practically nothing of the "Tibeto-Burman" languages which had been studied for some time primarily under the patronage of the British in India, and so Maspero naturally accepted "Tibeto-Burmic" as a sub-family on the authority of those who had been studying those languages.

And Sten Konow, a Norwegian scholar primarily interested in Iranic languages, was engaged by the British government in India to handle the non-Aryan

languages for the *Linguistic Survey of India*. Konow was so fully occupied in making grammatical analyses from the specimens of text of the many non-Aryan languages and dialects that he had little time for anything else.[3] Konow had very little to do with Chinese and Daic for the *Linguistic Survey*, and since the Sino-Daic numerals correspond so closely that even an amateur can see the resemblance, it was only natural for him to accept the "Chinese-Siamese" division for these languages that the knew so little about.

But Konow, the Iranist, brought to his work on the Sino-Tibetan languages the Indo-Europeanist's point of view of comparative grammar : that morphology is the primary criterion of genetic relationship. Now some Sino-Tibetan languages, like Chinese, have almost no morphology. So Konow substituted what seemed to him the closest thing to morphology—the word order—as a criterion.

Thus the "division," actually made by scholars working at opposite sides of the field and ignorant of the languages at the opposite end, was rationalized by Konow as resting upon the position of words in the sentence; he stated[4] that "the Tibeto-Burman family arranges the words of a sentence in the order of subject, object, verb, while the order in Chinese and Tai is subject, verb, object."

This statement that the fundamental division in the Sino-Tibetan family rested almost altogether upon the shift in position of one part of speech—let us say that the object precedes the verb in one division and follows it in another—is rather surprising, coming from an Indo-Europeanist. For any English-speaking beginning student of German can describe how annoying the German "displacement" of the verb is. Yet would anyone contend that German and English should be placed in different divisions of the Indo-European family because of differences in the position of the verb in the sentence?

If Sino-Tibetan languages do not have anything resembling Indo-European morphology, naturally we cannot use morphological correspondences as a primary criterion of linguistic relationship. But an alternative solution is possible. Students of Chinese will be familiar with the concept of "empty words", those which theoretically, at least, have lost their specific meaning and have come to be used as particles. Since the precise use of many of these "empty words" is still disputed even in languages studied for so long as Chinese, no special study of them in all the Sino-Tibetan languages has yet been made. But they are separate words and generally follow the same phonetic development as "full words," and since some of them are found over widely scattered parts of the Sino-Tibetan area, this writer has included a considerable number of them in his work on the phonetic development of the Sino-Tibetan languages, where they at he same time offer evidence of a common Sino-Tibetan "morphology"—as nearly as the family can be said to have one. If any one group showed no, or very few, morphological elements (pronouns, interrogatives, etc.) in common with other Sino-Tibetan languages, we should be very skeptical of genetic relationship.

But Indo-European comparative grammar is not based on morphology alone. A great part of it consists in working out the phonetic equations for the various

languages. Neither Konow nor Maspero had attempted to work out such equations and their comparisons consisted only of "look-alikes," a good many of which were erroneous. The greater part of this writer's work on Sino-Tibetan has consisted in working out such equations.[5] Only common words, numerals, parts of the body, verbs, adjectives, morphological elements, etc., have been considered and comparisons showing pronounced semantic divergence have been rejected.

A rough tally of such comparisons so far published[6] shows 216 between Bodish[7] and Chinese, 191 between Bodish and Burmese, 122 between Chinese and Daic, 101 between Burmese and Chinese, 63 between Bodish and Daic, and 38 between Burmese and Daic. Of course this will not be the final figure, but there is no reason to believe that the relative figures will vary greatly from those given above.

If we take the first two figures—216 comparisons between Bodish and Chinese but 191 between Bodish and Burmese—it would indicate that Bodish is genetically closer to Chinese than it is to Burmese. To anyone not led by the exotic appearance of Chinese characters to regard the language as a thing apart, this conclusion should not come as a surprise in view of geography and history. For while Tibet is somewhat closer to Burma than to China, to go from Lhasa to Mandalay by the shortest route one would have to cross the Himalayas and the mountain range on the Indo-Burmese frontier and cross two rather formidable rivers, the Brahmaputra and Irrawaddy. Historically we know there has been contact between China and Tibet for many centuries, but I know of no such contact between Tibet and Burma.

We may also note that in contrast to the 216 comparisons between Chinese and Bodish, we find only 122 between Chinese and Daic. That is, Chinese is considerably closer to Bodish than it is to Daic. In fact, almost since I began work on the Sino-Tibetan languages some 20 years ago, I have held that if Daic is related to the (other?) Sino-Tibetan languages, the relationship is very distant.[8] When Benedict cast doubt on the genetic relationship of Daic to Sino-Tibetan, he was merely repeating and expanding my own unpublished view and, ironically enough, in part with my own unpublished materials.

Benedict dismissed the some 200 comparisons between Daic and "other" Sino-Tibetan languages which had been adduced by Maspero, Wulff, and myself as non-basic, while he held that his own 30 Daic-Indonesian comparisons (including such cultural words as those for "rice-field" and "door") were basic.

More recently André G. Haudricourt has noted, regarding the Daic languages, that "les mots de la langue commune incontestablement proches de mots chinois sont les noms de nombres, des techniques militaires (cheval, selle, éléphant, jouget) et des techniques artisanales (métier à tisser, ouvrier, papier), bref un vocabulaire de civilisation susceptible d'emprunt. Au contraire le nom des parties du corps et le vocabulaire agricole ont peu d'affinité avec le vocabulaire chinois correspondant."[9] Haudricourt thought that the relationship of the Daic languages was to be sought in the north in the Man-Yao languages, and in the south in the Annamese-Muong. Since Haudricourt is working in these latter fields it is to be hoped that he will bring forward Daic comparisons with these groups of

languages. For it is only by searching for lexical and morphological parallels on all sides and by establishing the phonetic equations for such parallels that we can finally decide the genetic relationship of a doubtful group such as Daic.

While we cannot dismiss Haudricourt's suggestion of Chinese loan words in Daic without investigation, such precise phonetic correspondences as Burmese *k'ye`*, Siamese and Lao *k'ī`* "dung", or Lao *hnū`*, Luśei *hnu-* , breast',[10] for example, cannot by any semantic juggling be considered words of civilization likely to be borrowed. And too many of the comparisons are of that type to dismiss them off-hand. So in consideration of the very substantial number of comparisons between Daic and "other" Sino-Tibetan languages I shall continue to present Daic as Sino-Tibetan—if only as a challenge to accept such evidence as conclusive or to produce contrary evidence.

And while we have seen above that on a statistical basis Daic is closer to Chinese than to any other Sino-Tibetan group, we have also seen Chinese is not closest to Daic but to Bodish. So we should abandon the Sino-Daic division in contrast to the Tibeto-Burmic division. Rather we should set up the following mains divisions of Sino-Tibetan : Sinitic (Chinese), Daic, Bodic, Burmic, Baric, and Karenic.[11] The languages composing each division, as far as known, will be given below. Altogether we have some published material on about 300 languages and dialects of this family. Most of them are unknown even to scholars and it seems imperative to give the reader some clew to the relative position of groups in the descending scale of importance by adopting a uniform system of nomenclature; thus the descending order of the groups within a family were designated as follows:

Family (ending *-an*, as Sino-Tibetan); division (ending *-ic*, as Sinitic); section (ending *-ish*, as Bodish); branch (no specific ending); unit (no specific ending). Moreover, some of the languages are known under various names—the name they give themselves and the names other peoples give them, or under various spellings, as the Siamese spelling Dai, which is today pronounced Thai in Siam, for example, but Tai in some other languages of the division. It has been the custom for many years on the railroad maps of Europe to place the names in the language of the country, and scholars could do well to accord the same courtesy to each people, which will at the same time give the scholar a standard. This means abandoning the names given a people by its neighbors, changing Aka to Hruso, Miri to Miśing, Digaro to Taying, etc. I have also taken as standard the oldest spelling of the name of a people, usually the written rather than the spoken word, when it is known; thas the Bodish dialect Lhoskad rather than the spoken Hloke, and Dbus rather than Ü.

Sino-Tibetan family

Sinitic division [China]

MANDARIN SECTION[12] [N and W China]
WU SECTION [E coast of China] : Shanghai, Ningpo, Wenchow.
S. ANHWEI.

MIN SECTION [SE coast of China] : Foochow, Amoy, Swatow.
HUNAN.
YUE SECTION [S coastal] : Canton, Kiangsi, Hakka.

Daic division[13]

[W China, Tonkin, Kgd. of Laos, Siam, Shan St. of Burma]
SOUTH: Siamese, Lao.
NORTH: Tai noir, Tai blanc, Tho, Dioi, Ahom.
NORTHWEST: Shan, Khamti.

Bodic division

BODISH SECTION.
Bodish Branch[14].
West Bodish Unit [Kashmir] : Sbalti (Balti) ; Burig (Purik) ; Ladwags (Ladakhi) : Śam, Leh, Rong ;[15] Lahul.
Central Bodish Unit [S Tibet, N. India, Nepal, Bhutan] : Lhoskad, (Lhoke), Śarpa, Kagate, Garhwal, Spiti, Mnyamskad (Nyamkat), Dźad, Gtsang, Dbus (Ü) : Lhasa, Sikkim, Ãba (Batang), Choni, Tseku, Dartsemdo (Tatsienlu), Nganshuenkuan, Sotati-po, Paurong, Dru, Panakha, Panags, Nyarong, Ngamdo (Amdo), Khams[16].
South Bodish Unit[17] [Sikkim].
Groma (Twomowa) : Upper, Lower.
Sikkimese[18]
Dandźongka.
East Bodish Unit : Dwags (Takpa).
Tsangla Branch.
Rgyarong Branch.
Rgyarong : Pati, Wassu.
Gurung Branch [C Nepal] : Gurung, Murmi, Thaksya.
WEST HIMALAYISH SECTION [scattered, mostly W Himalaya].
NNW Branch.
Bunan.
Thebor : Sumtśu, Zangram, Sungnam, Kanam, Lippa.
NW Branch.
Kanauri : Upper, Lower, Tśitkhuli, Tukpa, Kanaśi.
Mantśati, Tśamba Lahuli, Rangloi.
Almora Branch.
Rangkas, Darmiya.
Tśaudangsi, Byangsi.
Dźanggali Branch.
Eastern Branch[19]: Thami, Bhramu.
WEST CENTRAL HIMALAYISH SECTION[20] [C Nepal]: Vayu, Tśepang, Magari.

EAST HIMALAYISH SECTION[21] [E. Nepal].
Western Branch.
Bahing, Sunwari[22], Thulung, Tśaurasya.
Dumi Unit:[23] Dumi, Khaling, Rai.
Eastern Branch[24].
Khambu Unit : Khambu,[25] Natśhereng.
Bontawa Unit: Rodong; Waling: Rungtśhenbung, Kiranti, Dungmali; Lambitśhong;[26] Lohorong,[27] Limbu,[28] Yakha.
Not definitely classified in a division; probably sections of Bodic, possibly of Burmic, certainly not of Baric:[29]
NEWARISH [C Nepal]: Newari, Pahri.
DIGARISH [NE Assam and into Tibet]: Taying (Digaro), Midu.
MIDŻUISH [in Tibet beyond NE Assam].
HRUSISH [N Assam].
Hruso (Aka) : Dialect A,[30] Dialect B[31].
DHIMALISH [Darjeeling and Jalpaiguri]: Dhimal, Ṭoṭo.
MIŚINGISH [NE Assam and into Tibet].
Miśing (Miri), Abor.
Yano.
Nyising (Dafla) : E,[32] C,[33] Tagen.
DZORGAISH (DZORGAIC?)[34] [NE Tibetan plateau and Szechuan, Kansu]: Dzorgai, Kortse, "Outer Mantse," Pingfang.

Burmic division

BURMISH SECTION.
Burma Branch [Burma].
Southern Unit: Burmese, Arakanese, Tavoy, Taungyo, Intha, Danu, Yaw.
Northern Unit: Phun (Megyaw, Samong dialects), Atśang (Ngatśang, Maingtha), Lawng (Maru), Letsi (Lashi), Tsaiwa (Atśi, Szi).
Lolo Branch [N Burma, Tonkin, SW China].
Southern Unit :[35]
Phunoi : Phunoi, Pyen, Khaskhong, Hwethom. Akha: Akha, Ako, Asong, Phana, Menghwa, Woni, Lahu, Lahuna, Lahushi, Kui.
Central Unit.
Lisu : Lisu, Lishaw, Lipha, Lipho, Kesopho, Kosopho.
Nyi, Tśökö, Weining, Ahi, Lolopho.
Northern Unit : Thongho, Pakishan, Kangsiangying, Kiaokio, Nee, Ulu, Laichau, Tudza, Nuoku.
Tonkin Unit : Mung ;[36] White, Black, Khoany.
Unclassified :[37] Manyak (Menia); Mosso, Ďion ; Duampu ; Phupha ; Nameji.
Hor Branch [E Tibet?]: Hor (Horpa)[38].
Hsihsia (Sihia) Branch[39].
MRUISH SECTION[40] [Arakan Hills].
NUNGISH SECTION[41] [N Burma]: Rawang, Metu, Melam, Tamalu, Tukiumu.

KATŚINISH SECTION [N Burma] : Katśin, Khauri ; Dźili.[42]
TŚAIRELISH SECTION[43] [SE Assam].
LUISH SECTION [SE Assam and adjoining Burma districts] : Andro, Sengmai, Sak, Kadu.
TAMAN.[44]
KUKISH SECTION[45] [Indo-Burmese frontier regions].
Southern Branch.
Śo : Sandoway, Thayetmyo, Minbu, Tśinbon, Chittagong,[46] Lemyo.[46]
Yawdwin ; Tśinbok.
Khami : S., Khimi,[47] N.
Lakher Branch.
Mara, Tlongsai,[48] Hawthai.
Sabeu.[49]
Zeuhnang.
Śandu.[50]
Old Kuki Branch.
Central Unit : Tśiru, Aimol, Purum, Langrong.[51]
Kyau.[52]
Western Unit : Hrangkhol, Biate, Hallam ; S Luhupa.[53]
Kolhreng (C. Peripheral) Unit : Kolhreng, Kom, Tarao.[54]
Lamgang (Southern) Unit : Lamgang, Anal.[55]
Langel.[56]
Central Branch.
Luśei Unit: Luśei (Dulien dialect, Ngente dialect), Zahao Hmar, Pankhu,[57] Bom.
Haka Unit: Haka (Lai), Śonśe, Taungtha,[58] Bandźogi.
Kapwi Unit.[59]
Northern Branch.
Thado, Ralte;[60] Śiyang (Siyin), Vuite (Paite).
Luhupa Branch.[61]
Maring Unit :[62] Maring, Khoibu.
Tangkhul Unit : Ukhrul, Phadang; Tśamphung.[63]
Kupome Unit : Kupome, Khunggoi, Central Luhupa, Northern Luhupa.
Western Branch.[64]
Maram, Kwoireng; Kabui, Khoirao; Empeo.[65]
Northern Naga Branch.
Hlota;[66] Ao Unit : Tśungli; Longla; Mongsen, Tśangki, Khari; Tengsa, Rong (Leptśa)[67] (Sikkim), Yatśam; Yatśumi,[68] Thukumi.
Eastern Branch.
Rengma.[69]
Simi (Sema) : Keźama, Sopvoma ; Zumomi, Dayang.
Angami : Tengima ; Tśakrima : Dzuna, Kehena, Mima.
Meithlei Branch.[70]
Mikir Branch.[71]

Baric division (Assam)

BARISH SECTION.
North Central Branch.
Garo : Atśik (Standard), Kamrup, Abeng, Dacca.
Jalpaiguri Branch.
South Central Branch: Atong, Rabha, Ruga, Kontś, Tintikeya, Cooch Behar, Kotś.
Western Branch : Bodo, Metś ; Dimasa, Hodźai ; Tipura ; Lalung ; Moran.
Eastern Branch : Tśutiya.
NAGISH SECTION.
Mośang, Śangge.
Namsangia.
Banpara, Mutonia.
Tśingmegnu (Tamlu).
Angwanku (Tableng), Mulung.
Tśang.

Karenic division[72] [C and Southerly Burma].

Yeinbaw ; Karenni : Yintale, Manö, Sinhmaw Mapauk ; Pwo : Bassein, Maulmein ; Mopwa : Dermuha, Bilitśi ; Taungthu ;[73]
Zayein : Sawntung, Padeng, Banyang ; Kawnsawng ; Gheko, Bwe, Sgaw, Wewaw, Padaung, Karenbyu.

Conclusion

Sten Konow was the first scholar to attempt to cover almost the entire Sino-Tibetan field. But his routine duties on the *Linguistic Survey of India* and his grammatical analyses made from texts permitted him to make lexical comparisons almost exclusively on modern Bodish dialects, and he did not attempt to work out phonetic equations in a serious way. Thus he did not have a broad basis of comparative work for his classification, but he appears instead to have taken into consideration geographical location, grammatical similarities, and a few random lexical comparisons made by other authors or observed by himself but not published.

Konow's classification was sometimes remarkably good when correct alignment did not involve a knowledge of phonetic shifts, a knowledge Konow did not possess. Thus his classification of the West Himalayish languages is excellent, and all his Kuki-Chin languages are indeed Kukish. It was when severe phonetic shifts occurred so that parallels were disguised, as in the Naga languages, or when there were no closely related languages apparent, that Konow was most likely to go astray.

Of the Naga languages Konow confused the West Kukish and Luhupa branches and misclassified Mikir. He included Empeo, Kabui, Khoirao, Kapwi (which he confused with Kabui),[74] and Mikir in what he called the "Nāgā-Bodo group,

bridging over the difference between the characteristic features of the two forms of speech"[75] while "Mikir clearly belongs to the same group as Kachchā Nāgā,[76] Kabui, and Khoirao." But the position of Mikir could not have been as clear as he tought, for later he transferred Mikir from his Nāgā-Bodo to his Nāgā-Kuki group,[77] though still recognizing[78] Mikir "affinities with Bodo," he considered it "much more closely connected with Kuki," yet that "it occupies a somewhat independent position."

Actually Mikir occupies the most independent position of any of the Kukish languages, both lexically and phonetically. And instead of being a connecting link with Barish, as one might expect from its geographical position and as Konow thought, it is a slight connecting link with the more distant Bodish. And whether Konow included or omitted Mikir, his Nāgā-Bodo group had no more connection with Barish (his Bodo) than have any of the other Kukish languages.

Konow's transfer of Mikir to his Nāgā-Kuki group only aggravated the latter's heterogeneity, for the group already included members of three different branches : Sopvoma (really Eastern Kukish),[79] Maram and Kwoireng (really Western Kukish), and Maring, Tangkhul (Ukhrul), Phadang, and Khunggoi (really Luhupa), while the addition of Mikir brought in a fourth branch. With such diversity of languages in one supposed group, Konow's statement[80] that the Nāgā-Kuki bridges over the gulf between Angami and the Kuki languages is meaningless, for languages from three or four groups would naturally link almost any distantly related languages.

As Rong has borrowed perhaps the greater part of its vocabulary from Bodish, with some elements from surrounding Himalayish languages, and as it is separated by some 400 miles from the Ao languages to which it belongs, its separation from the latter in the *Linguistic Survey* is understandable. Yet one may classify Rong rather precisely as belonging not only to the Ao Unit of the Northern Naga Branch, but as belonging to the Tengsa subunit, and within this subunit as being slightly closer to Tengsa than to Yatśam.[81] However, Tengsa and Yatśam are often closer in form to each other than either is to Rong—for during its period of separation Rong has developed some independent phonetic peculiarities.

The correct classification of Rong presents an interesting problem for Indian anthropologists and historians : Were the Rong left behind when the Northern Naga branch (and perhaps all the Kukish peoples) migrated from the Himalayas to their present location on the Indo-Burmese border, or are the Rong a remnant left behind from a time when the Northern Naga extended clear across the Valley of Assam ?

I shall not take up here all the other languages which Konow mis-classified as the reader may note them by comparing Konow's classification with the one presented here, and I have rectified some of his errors in papers already published.

Relationships outside Sino-Tibetan

I have presented a considerable number of comparisons, with some phonetic equations, between Sino-Tibetan and Vietnamese,[82] Athapaskan,[83] Khasic and Palaungic

(as representing Austroasian).[84] I am inclined to agree now with André G. Haudricourt, however, that Vietnamese is probably basically Austroasian, and I think that the comparisons I made between Vietnamese and Sino-Tibetan should probably be added to those I made between Sino-Tibetan and two of the Austroasian languages.[85]

Yet one might infer that I was attempting to set up a macro-Sino-Tibetan family consisting of Sino-Tibetan, Austroasian, and Athapaskan. I have not had a chance to examine the evidence critically, but I believe if it were done it would show that the Sino-Tibetan and Athapaskan roots were practically all different from those in the Sino-Tibetan and Austroasian comparisons. Does this indicate that Sino-Tibetan is really some kind of mixture from two more primitive families of languages, let us say of Pacifican and Austroasian?

Notes

1 Konow carelessly stated that Daic and Chinese "form one distinct family as compared with the Tibeto-Burman forms of speech" (*Linguistic Survey of India*, 3 [1909], p. 1). Of course he meant sub-family, but the absurdity of having families within a family of languages has been repeated in the *Encyclopedia Britanica* down to the last edition, which refers to the "Tibeto-Burman family" and the "Siamese-Chinese family" (1953, v. 22, p. 187, and v. 20, p. 596 a).

Ambrogio Ballini and Carlo Tagliavini correctly refer to sub-families but these are Konow's Sino-Siamese and Tibeto-Burman (*Enciclopedia Italiana*, v. 19, pp. 46 and 129ff. The *Diccionario Enciclopedico U. T. E. H. A.* mistakenly applies "tibeto-birmano" and "siamochino" to race. All have obviously been following Konow and the *Linguistic Survey of India*.

2 Thus after 20 years the close resemblance of Old Japanese *wata* "sea" to English *water* still sticks in my memory, although I do not not believe any genetic relationship exists between the languages.

3 Personal communication from Konow.

4 *Eneyelopedia Britanica*, 11th ed. (1911), v. 26, p. 929.

5 For bibliography, see Shafer, "East Himalayish," *Bull Sch. Or. Afr. Sl.* 15 (1953), 357 n., or "Newari and Sino-Tibetan," *Studia Linguistica* (Lund, 1952), 92, n. 1, and 93, n. 3, par. 2 ff.

6 "The Vocalism of Sino-Tibetan," *Journ. Amer. Or. Soc.* 60 (1940), 302–337; 61 (1941), 18–31; "Problems in Sino-Tibetan Phonetics," *JAOS* 64 (1944), 137–143; and "The Initials of Sino-Tibetan," *JAOS* 70 (1950), 96–103.

7 Bodish refers to Old Bodish (classical Tibetan) and languages closely related to it; for definition see below.

8 In 1938 I prepared a list of words showing the lack of precise phonetic and semantic correspondence between very common words in Daic and other (?) Sino-Tibetan languages, words such as those for parts of the body, celestial luminaries, pronouns, etc., and I tried to convince Maspero that Daic was not Sino-Tibetan. It was Maspero's insistence on Sino-Daic genetic relationship which caused me to work over Maspero's and K. Wulff's comparative data and to add my own, which nearly doubled the number of comparisons and filled in to some extent the gaps Wulff's work left in many series of finals. These Sino-Daic comparisons were presented in "The Vocalism of Sino-Tibetan" and subsequent articles with no statement regarding interrelationships.

I had discussed my skepticism regarding the relationship of Daic to Sino-Tibetan both before leaving for Europe, after talking with Maspero, and after publication of my article, so that there was no basis for Benedict's statement that I followed Maspero and Wulff in setting up an Eastern Division composed of Chinese and Daic in opposition to Tibeto-Burmic (Benedict, "Thai, Kadai, and Indonesian," *American Anthropologist* 44 (1942), 588).

9 "Les phonèmes et le vocabulaire du thai commun," *Journal Asiatique* (1948), 235 ff.

10 Cited in "Vocalism" (see n. 4 above), Table 4, no. 18, and Table 6, no. 26 respectively. Many comparisons of a similar nature will be found in the essays referred to above.

11 I use Sino-Daic, Tibeto-Baric, Tibeto-Burmic always to refer to two or more divisions.

12 For some of the recorded Mandarin dialects see Bernhard Karlgren, "Études sur la phonologie chinoise," *Archives d'Études Orientales* 15 (1915), pp. 230–1. The classification of the Chinese dialects given here is Karlgren's but with certain modifications by Yuen Renn Chao. For a dialect map of China see *Shun Pao*, 60th anniversary edition.

13 Daic languages show little divergence except in phonetic development, which Maspero used in his classification, "Contribution à l'étude du système phonétique des langues thai," *Bull. Éc. Fr. Ext.-Or.* 11 (1911), 158, n. 1. Tai Noir, according to Maspero, and Ahom, in my opinion, are intermediate. A good part of the vocabularies of many of the Hainan languages and dialects is also probably Daic.

14 Progressive phonetic degeneration of Bodish dialects from west to east through the dialect of Nganshuenkuan, after which archaic aspects increase through Khams.

15 Not to be confused with the Rong which is geographically in the Himalayas and linguistically in the Northern Naga Branch of the Kukish Section (see below).

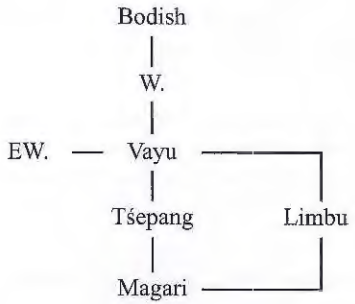
16 The Hanniu of von Rosthorn also belongs somewhere in the Central Bodish Unit; it is not a Rgyarong dialect.

17 Characterized by the shift of *-r-* to *-y-*, also a characteristic of most of Burmish.

18 Not to be confused with Sikkim in Central Bodish.

19 Thami and Bhramu, which had been separated by Konow, belong together. From the limited vocabularies of them one can only say that they are placed in West Himalayish because they appear to be closer to that group than any other.

20 A poorly defined section of which it can only be said that the languages are related to each other more than to languages of other groups; see "Classification of Some Languages of the Himalayas," *J. Bihar Res. Soc.* 36 (1950), 192ff. Interrelationships are approximately:



W. is West Himalayish, and EW. is the Western Branch of East Himalayish.

21 Konow mixed in some languages that do not belong here and omitted some that do. See my "East Himalayish," *Bull. Sch. Or. Afr. St.* 15 (1953), 356 f.

22 Phonetically degenerate compared to Bahing.

23 Diverges considerably.

24 Approaches Burmish phonetically.

25 Kulung and Sangpang are the same language recorded in different localities.

26 Tśhingtang is the same language in another locality.

27 Balali is the same language in a different locality.

28 Limbu and Yakha diverge somewhat from other East Himalayish languages.

29 None of these groups seems to possess features distinctive from Bodic or Burmic and it is a question of classifying them under one or the other. This may have to be done on a weighted statistical basis, similar in principle, perhaps, to that proposed by A. L. Kroeber and C. D. Chrétien in "Quantitative Classification of Indo-European Languages," *Language* 13 (1937), 83–103, but weighted to allow for the criticisms made by A. Meillet.

30 By Campbell. Comparatively archaic. See my "Hruso," *Bull Sch. Or. Afr. St.* 12 (1947), 184–196.

31 By the other three recorders of Hruso. Phonetically degenerate.

32 By Needham, Hamilton.

33 By Robinson.

34 This group occupies at least a portion of the Sifan plateau. It probably forms the most northeastern outpost of the Tibeto-Burmic peoples today. The vocabularies I have seen have been short and poorly recorded. Lexically these languages seem to be predominantly Tibeto-Burmic but with some peculiar features. In recent years, Chinese scholars—notably Wên Yu in *Studia Serica*—have given some attention to these languages under the name of Ch'iang, assuming that they are spoken by a people mentioned in old documents. But only one bit of the recent literature on the subject has come to my attention. Perhaps some parts of the vocabulary of Sotati-po are Dzorgaish; see Central Bodish Unit above.

35 The Northern Unit of the Burmese Branch approaches phonetically and sometimes lexically the Southern Unit of the Lolo Branch; see Shafer, "Phonétique historique des langues lolo," *T'oung Pao* 41 (1952), 191–229. In fact the transition may be said to be gradual in the loss of final consonants from the Burmese Branch to the Northern Unit of the Lolo Branch, while inversely the loss of initial sonancy is progressive from the Northern Unit of the Lolo Branch to the Burmese Branch. But the languages of the Burmish Section are remarkably uniform lexically considering the vast stretch of territory they cover.

36 Perhaps partly non-Lolo.

37 Too little data or too irregularly recorded.

38 Preserves some prefixes lost in the Burmese and Lolo Branches; somewhat degenerate regarding finals.

39 Uncertainty regarding transcription and limited vocabulary make definite classification impossible at present, but Hsihsia appears to be the most northern extension of the Burmish section.

40 See my article on "The Linguistic Position of Mru", *Journ. Burma Res. Soc.* 31 (1941), pt. 2, no. 2.

41 Closer genetically to the Burmish than to the Kukish Section.

42 More archaic in its prefixes. Only a extremely short vocabulary has been published.

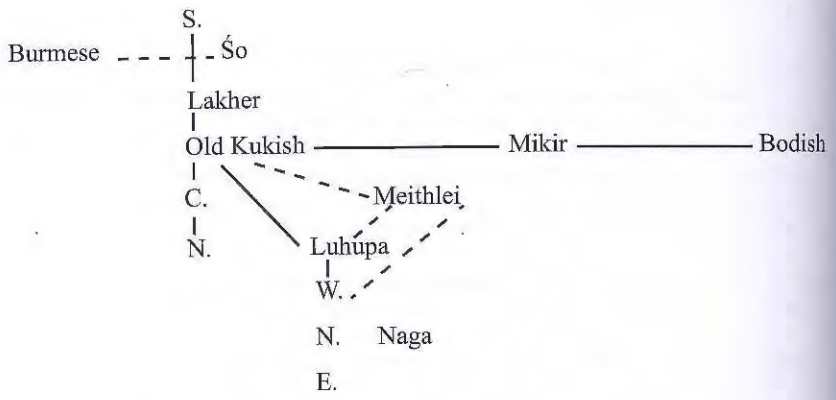
43 Records of the Luish languages vary so in extent and the ability of the recorder that it is difficult to establish interrelationships within the section. The finals of Sak are considerably altered compared to those of other Luish languages, and this is perhaps not altogether due to poor recording. It preserves medial **l* as *r*, this phoneme being

lost in the rest of the section. Certain Sak forms common to Kukish, such as those for "mother" and "bird," and not found in the other Luish languages, may be borrowings from Kukish.

44 The Taman recorded by R. Grant Brown, *JRAI* 41 (1911), 305 f., is archaic in prefixes but the phonetics of Taman is not generally clear from the brief vocabulary and probably will not be until have sufficient data to be able to climinate loan words. We may not be able to classify it until we have larger vocabularies of it and its dialects and of the surrounding languages. It appears to have been under Burmish influence but has a number of rather rare stems found in Luish, but not exclusively there, as words for "buffalo," "elephant," "horse," "salt," and "father."

45 One of the major points on which Konow's classification errs is in setting up a Naga group. The Naga languages are all Kukish except the northeastern-most, which is Baric. The proof of the first part of this statement was contained in my article on "The Naga Branches of Kukish," *Rocznik Orientalistyczny*[16] (Krakow, 1950), 467–530, and of the last part of the statement in "Classification of the Northernmost Naga Languages," *J. Bihar Res. Soc.* 39 (1953), 225–264.

The interrelationships within Kukish are approximately as follows:



Solid lines represent genetic relationship, broken lines borrowing. Burmese loan words in Śo, and Meithlei loan words in Old Kuki, Luhupa, and West Kukish, are largely due to administrative dominance of Burmese and Meithlei.

Old Kukish, taken collectively, appears to very nearly represent proto-Kukish, while phonetic degeneration, particularly in finals, has occurred to the south in Lakher and S. Kukish, and particularly in prefixes in the north in Central and Northern Kukish. But all the groups along the vertical axis are essentially the same except for varying degrees of decay and the extent of borrowing. The number of stems peculiar to individual branches along the main vertical axis are almost negligible. So I term these languages Central Core languages or Kukish proper. Kukish languages to the right of the vertical axis have decayed and diverged semantically and morphologically.

46 Slightly better preserved than most recordings of Śo.

47 A slight link with N. Khami. Data on Khimi will be found in "Khimi Grammar and Vocabulary," *Bull. Sch. Or. Afr. St.* 11 (1944), 386–434.

48 Although Parry thought he was recording the same dialect as the Mara of Savidge, there are a few phonetic differences.

49 The Lakher vocabulary of Lewin approaches the Sabeu of Parry but does not correspond exactly. It is probably somewhat intermediate between Sabeu and some of the other dialects recorded by Parry.

50 Slightly more archaic than the other Lakher recordings, as it was taken down earlier.

51 Precise classification here is doubtful.

52 Belongs here ? Eliminating borrowings from Southern Kukish, Kyau is probably close to the Western Unit of the Old Kuki Branch.

53 Provisional classification. Certainly not a Luhupa language, although it has borrowed a few words from that group, as those for "horse" and "cow," and has undergone some vocalic changes—particularly to **u* and **ui*—which are not characteristic of Old Kuki.

54 Precise classification doubtful. Probably transitional to Lamgang Unit, but in its main characteristics belongs with the Kolhreng Unit.

55 A slightly degenerate form of Lamgang; see the author's "Phonetik der Alt-Kuki-Mundarten," *ZDMG* 102 (1952), 262–279.

56 A "Central Core" Kukish language whose precise classification is doubtful. It is spoken in southern territory but probably by migrants from the Central Branch of Kukish.

57 Has some special links with Bandźogi, but probably mainly in the vocabularies of Lewin, because both languages are more archaic there than in later recordings.

58 Spoken in Southern Kukish territory but belongs essentially to the Haka Unit of Central Kukish. This is not clear from a cursory examination, as Taungtha has borrowed some words from Southern Kukish languages and has developed to some extent along individual lines.

59 Precise classification here questionable. But not to be confused with Kabui, as Konow did.

60 Has a few minor links with Vuite.

61 Probably genetically closest to Old Kuki, but not its Western Unit.

62 Diverges toward Kukish proper more than the two following Luhupa units, as already noted by Konow, p. 451.

63 Sharply divergent.

64 The subordination of Kwoireng and Khoirao noted below is not marked.

65 Butler's Empeo is a different dialect from that of Soppitt and Stewart.

66 Kyõ or Kyontsü would perhaps be a better designation.

67 Rong is slightly closer to Tengsa than Yatśam, but the latter are closer to each other.

68 Perhaps better designated Yimtśurr.

69 Perhaps better Nzong or Nzonyu. Grierson's Rengma is slightly more altered than Butler's. Ntenyi, of the northern group of Western "Rengma," actually forms a connecting link between Hlota and Simi, while the Iseni-Kotsenu of Hutton is intermediate between Ntenyi and Anyo (Eastern or Naked "Rengma"). These meagerly recorded languages cannot be more definitely placed at present.

70 In accord with the standard outlined above of using the oldest form of the native name known, I use *Meithlei* instead of the usual *Meithei*, the present pronunciation. For the Assamese call these people *Mēklē* (Konow, p. 20), which is about as close to * Meithlei as the Assamese could make (the Kukish lateral affricates are frequently written *kl, k'l*). Singh (p. 71) recorded the Thado name of the Manipuris as *Maithai*, i. e., *Mei-hlei*; and **l'l* is usually recorded as *hl* in Thado. And since **l'l* became *t'* in the language of Manipur, we must conclude that the earlier name was * *Meithlei*, which became the present *Meithei* by regular phonetic change.

71 Greatly aberrant, but definitely Kukish.

72 Lexically predominantly Burmic, but also strongly Bodic, although it is one of the most southern of the Sino-Tibetan groups. It might almost be called intermediate between Bodic and Burmic, but certain phonetic and lexical peculiarities make it preferable to regard it for the present as a separate division. Due to lack of any sharp division of the Karenic languages and the inadequate materials on some of them, the following classification is only tentative.

73 Not to be confused with Taungtha in Central Kukish.
74 P. 418.
75 P. 379.
76 I. e., Empeo.
77 Vol. 1, pt. 1, p. 66, n. 2.
78 P. 69.
79 But Konow observed (p. 451) that "Sopvomã is so closely connected with all the languages of the Western sub-group, that it might with equal propriety be classed as belonging to it as to to the Nāgā-Kuki one" and in following pages all Konow's comparisons are between Sopvoma and Angami. One may suspect that Konow discovered Sopvoma's correct classification too late to change it without inconsistency in the text, so he resorted to compromise.
80 P. 451.
81 One may get some idea of the correct classification of Rong in "Classification of Some Languages of the Himalayas," *Journ. Bihar Res. Soc.* 36 (1950), insert between pp. 173 and 174.
82 "L'annamite et le tibéto-birman," *Bull. Éc. Fr. Ext.-Or.* 40 (1940), 439–442, or "Annamese and Tibeto-Burmic," *Harv. Journ. As. St.* 6 (1942), 399–402; "Le vietnamien et le tibeto-birman," *Dân Viêt Nam*, No. 1 (Hanoi, 1948), 1–10.
83 "Athapaskan and Sino-Tibetan," *Internat. Journ. Am. Ling.* 18 (1952), 12–19.
84 "Études sur l'austroasien," *Bull. Soc. Ling.* 48 (1952), fasc. 1, pp. 133–158.
85 I make this statement on the basis of some manuscript pages sent me by M. Haudricourt. It is probable that the Manic languages are also Austroasian, judging from his "Introduction à la phonologie historique des langues miao-yao," *BEFEO* 44 (1954), 554–576.

5

NOTES ON FANG—KUEI LI'S 'LANGUAGES AND DIALECTS OF CHINA'

James A. Matisoff

Source: *Journal of Chinese Linguistics* 1, 3, 1973, 471–4.

In the first issue of *JCL* last January, there reappeared an article by Fang-Kuei Li entitled 'Languages and dialects of China' (pp. 1–13). This article was originally published back in 1937 (*Chinese Yearbook*, Shanghai), and for a long time was considered the last word on the genetic affiliations of the various languages and dialects spoken within the vast borders of China. By way of justifying the inclusion of a 36-year old article in the first issue of a new journal, the editor comments (p. 1):

> 'Although research over the past three decades has brought us a more refined understanding of the individual dialects, Li's broad outline remains essentially accurate and useful.'

This is undoubtedly true as far as the internal classification of the Chinese dialects is concerned (pp. 3–5). Furthermore, Li's discussion of the Kam-Tai languages (pp. 5–6) may be taken as perfectly authoritative—after all it was Professor Li's brilliant researches into the Kam-Sui languages which established the reality of the Kam-Tai grouping in the first place.

On the other hand, three decades of scholarship—notably by Robert Shafer and Paul K. Benedict—have made Li's internal classification of the Tibeto-Burman family (pp. 8–9) quite obsolete.[1] Even more unfortunate is the resurrection of the all-inclusive 'Indo-Chinese' grouping (p. 2), into which Li lumps Chinese and Tibeto-Burman together with Kam-Tai and Miao-Yao, largely on the basis of shared phonological and morphological characteristics (monosyllabicity and lexically contrastive tone).

1. The 'Indo–Chinese' pseudo–stock

The hypothesis of a genetic relationship between Chinese and Tai was widely accepted by scholars before 1940. Although efforts were made to document this relationship by actual lexical comparisons, the words involved were not 'core vocabulary', so that alternative explanations in terms of borrowing or diffusion were at least equally likely.[2] More impressive to students of the problem was the demonstrable similarity between the Chinese and the Tai tonal systems. As Professor Li says (p. 2), '. . . this has been considered the most powerful argument for the common origin of these languages.'

Yet recent investigations have confirmed that the development of lexically contrastive tone proceeds according to quite general phonetic principles, involving the influence of the syllable-initial and/or final consonants on the pitch of the vocalic nucleus.[3] This influence remains subphonemic until the consonantal system itself undergoes a drastic merger or loss (typically the devoicing of an initial voiced series, or the loss of a final laryngeal). At this stage the language may compensate by utilizing the formerly redundant pitch-differences (which remain after the merger or loss) for contrastive purposes. The point is, **this process of 'tonogenesis' operates independently of the particular genetic affiliations of the language in question.**[4] Conditions which favor the process (besides the consonantal perturbations we have mentioned) are that the language be basically monosyllabic in structure, and that it be in contact with other languages that are already tonal.[5]

Many examples could be adduced to demonstrate that tonal convergence and genetic relationship are totally independent things: languages may have extremely similar tonal systems and still be genetically unrelated; conversely, closely related languages may differ radically in the extent to which they exploit pitch for contrastive purposes.[6]

Reacting against the shaky 'Indo-Chinese' idea, Benedict long ago proposed a bold new genetic grouping of the languages of Southeast Asia, aligning Tai not with Chinese, but rather with Austronesian (= Malayo-Polynesian).[7] Over the last thirty years he has amassed conclusive and detailed evidence for his theory,[8] and now includes in his superstock 'Austro-Thai' not only Kam-Tai and Austronesian, but Miao-Yao as well.[9] Even though Benedict's world-view has not yet received universal acceptance—after all, very few linguists command the breadth of data that would entitle them to an opinion one way or the other—it seems safe to predict that his ideas will eventually prevail. (Nobody else is offering any serious alternatives at the moment.)

2. Subgrouping the Tibeto–Burman family

Li (pp. 8–9) divides Tibeto-Burman into 'four known divisions': (a) the 'Tibetan group'; (b) Bodo-Naga-Katchin; (c) the 'Burmese group', comprising both Burmese and Kuki-Chin; and (d) the 'Lolo group'. In a footnote the editor adds:

> 'A more general treatment of the Tibeto-Burman languages may be found in Benedict 1972.'[10]

While much remains unclear about the finer subgrouping of Tibeto-Burman, we have learned a huge amount since 1937.[11] The scheme presented by Professor Li does not reflect any of our hard-won new knowledge. The Lolo languages belong with Burmese in a tight-knit family now called Burmese-Lolo or Lolo-Burmese by Tibeto-Burmanists. Kuki-Chin belongs to a totally different branch of Tibeto-Burman from Burmese, and is now known to have very close affinities with the Naga languages, so that Kuki-Chin-Naga may be regarded as a unitary grouping. 'Bodo-Naga-Katchin' is an obsolete lumping together of languages which really belong to three separate subgroups of Tibeto-Burman: Bodo belongs to Bodo-Garo or 'Barish'; Naga, as we have said, belongs with Kuki-Chin; and 'Kátchin' or Jinghpaw is in a class by itself. In addition there are several other branches of Tibeto-Burman which Li does not mention at all.

* * *

We are now embarking on an exciting new era in Sino-Tibetan linguistics, and in Southeast Asian linguistics in general. There is an explosion of new knowledge, as modern, accurate data become available through new fieldwork and as the insights of general linguistic theory begin to penetrate this once highly esoteric field. Increasingly we will be in a position to distinguish among three types of structural similarities in the languages of the region: (a) those due to real genetic relationship: (b) those due to areal diffusion or borrowing; and (c) those due to mere chance, or parallel independent development, reflecting universal tendencies in human language.[12]

In the present context, where we are looking forward to the future, it is a disservice to a great linguist like Professor Li to republish his thoughts of thirty-five years ago with so little aggiornamento.

Notes

1 See Shafer and Benedict, 1939–41; also Benedict 1972.
2 See for example Wulff 1934.
3 See especially Haudricourt 1954, 1961.
4 See Matisoff 1970, 1973a.
5 Yet neither of these conditions is absolutely necessary. Witness the very recent development of tones in Punjabi, due to the loss of its former voiced aspirated series of initial consonants.
6 The clinching example is modern Tibetan, where some dialects have tone while others do not. See Matisoff 1973a.
7 Benedict 1942.
8 Benedict 1966, 1967a, 1967b.
9 Benedict 1968, 1973.
10 The reference is to Benedict's monumental book *Sino-Tibetan: a conspectus*, which was finally published in 1972, even though it had existed in manuscript form as early as 1942–3.

11 See Benedict 1972, pp. 4–11.
12 For a recent attempt to disentangle these various possibilities with respect to tonal developments in Tibeto-Burman, see Matisoff 1973b.

References

BENEDICT, PAUL K. 1942. Thai, Kadai and Indonesian: a new alignment in Southeastern Asia. American Anthropologist, n.s. 44.576–601.
———. 1966. Austro-Thai. Behavior Science Notes 1.227–261.
———. 1967a. Austro-Thai studies: material culture and kinship terms. Behavior Science Notes 2.203–244.
———. 1967b. Austro-Thai studies: Austro-Thai and Chinese. Behavior Science Notes 2.275–336.
———. 1968. Austro-Thai and Sino-Tibetan. Unpublished paper presented at First Sino-Tibetan Conference, Yale. October.
———. 1972. Sino-Tibetan: a conspectus. Cambridge University Press. Contributing editor, James A. Matisoff.
———. 1973. Austro-Thai. Human Relations Area Files Press, New Haven. In press.
HAUDRICOURT, ANDRE–GEORGES. 1954. De l'origine des tons en viêtnamien. Journal asiatique 242.68–82.
———. 1961. Bipartition et tripartition des systèmes de tons dans quelques langues d'Extrême-Orient. BSLP 56. 163–180.
MATISOFF, JAMES A. 1970. Glottal dissimilation and the Lahu high-rising tone: a tonogenetic case-study. JAOS 90.1.13–44.
———. 1973a. Tonogenesis in Southeast Asia. Consonant types and tone, ed. by Larry M. Hyman. Southern California Occasional Papers in Linguistics, No. 1. Los Angeles.
———. 1973b. The tones of Jinghpaw and Lolo-Burmese: common origin vs. independent development. Acta Linguistica Hafniensia, Copenhagen. To appear.
SHAFER, ROBERT and PAUL K. BENEDICT. 1939–41. Sino-Tibetan linguistics. Unpublished typescript. 13 vols. Berkeley, California.
WULFF, K. 1934. Chinesisch und Thai. Det Kunglige Danske Videnskabernes Selskab, Historisk-filologiske Meddelelser, 20.3. 260 pages.

6

SINO-TIBETAN
Another look

Paul K. Benedict

Source: *Journal of the American Oriental Society* 96, 2, 1976, 167–97.

The main findings of another look (after *Conspectus* [1972] - reviews noted) are that Sino-Tibetan is now a well-established family; Tai and Miao-Yao must still be excluded, although each has made early borrowings (especially numerals) from Chinese dialects or related languages; lexical analysis (Swadesh 100-word list) supports the taxonomic arrangement (*Conspectus*) setting Chinese apart from Tibeto-Burman, but the position of Karen remains indeterminate; the Sino-Tibetan reconstruction (*Conspectus*) remains largely unchanged despite some refinements, but recent studies have uncovered an extensive prefixation pattern (mainly *s-*, also?- and *m-*) for Archaic Chinese, radically altering the 'look' of the language in the direction of Tibetan and other Tibeto-Burman languages; finally, a review of comparative Sino-Tibetan studies reveals that data (sources) are less often at fault than scholars.*

Sino-Tibetan Studies, which had long languished in a state of academic torpor, came to life in the later 1960's, in part as a result of interest generated by the annual S[ino]*T[ibetan] conferences held each October since 1968.[1] A by-product of this *risorgimento* was the *Conspectus* (Benedict 1972; hereafter *STC*), an early (1942–43), unpublished effort brought up to date (ca. 1968–69) by the writer and by the contributing editor, J. Matisoff. As noted in the *STC* review by Egerod (1973), the annual conferences have provided "an important forum for testing the ideas put forth in the *Conspectus*," and it seems appropriate at this time to take some note of these developments, as well as of the substantive material contained in the available *STC* reviews to date.[2] There have also been new findings (below) relating to groupings within ST and to prefixation in Arch[aic] Ch[inese], both of which may require modifications in our view of the field. Additionally, it appears

that the history of ST studies, although still at a relatively early developmental stage, can shed some light on the nature of the reconstruction process itself. With these general points in mind, then, we propose to take another look at the field.

To begin with the broadest possible questions, can we now, after these many years, consider ST to be a firmly established family? If it is in fact a family, does it include Kam-Tai and or M[iao] Y[ao]? Matisoff (1973), after reviewing the first five annual ST conferences, makes the following observations:

> At the first couple of meetings some people still did not seem convinced about the genetic relationship of Chinese to Tibeto-Burman at all. After five years the voices of skepticism have been stilled, utterly.
>
> (p. 158)

> With the recent publication of Benedict's book [*STC*]. . . . the last shreds of diffidence of this matter may be cast to the winds.
>
> (fn. 14)

Well, not *utterly*—not the *very* last shreds. Miller (1974) still writes of the "*possibility* [my italics] of an original Sino-Tibetan linguistic unity"; Chang (1973) still insists, "Despite the assiduous efforts of many scholars a genetic relation (sic) for Tibeto-Burman and Chinese is still to be established"; while Lehman (1975) seems to vacillate, stating, ". . . I believe that Benedict has *all but* [his italics] proved the ST hypothesis," followed immediately by, "He has listed so many fairly obvious cognate sets . . ., that their cumulative effect alone precludes serious doubt." Despite their demurrals, the above cited writers appear in fact to regard ST as an established linguistic entity (cf. the remarks on Chang, below) and there is now a firm consensus backing ST as a fully demonstrated language family. The key point here, as widely recognized today (cf. the discussions in Downer 1971, Egerod 1973, Matisoff 1973bis), is the amount of "core" vocabularly shared by the languages under consideration. In the present case, although Lehman stresses this feature of cognate sets (see *cit. supra*) both Chang and Miller indicate some dissatisfication in their *STC* reviews. Chang finds that the semantic matchings are in part "farfetched" but his examples miss the mark: T[ibeto] B[urman] 'word, speech' and Ch. 'song', overlooking the significant Lahu gloss 'sing', (*STC*: fn. 487); TB 'hot' and Ch. 'ill', overlooking the extensive TB evidence, e.g., W[ritten] T[ibetan] *tsha* 'hot; illness' (*STC*: 27); TB 'nose' and Ch. 'self', overlooking the clear evidence for the 'nose' > 'self' semantic shift in Ch. *dz'i*^{a} (*STC*: fn. 417). Miller places heavy and repeated emphasis on the lack of any semantic association *in Chinese* of Ch. *k'o*^{b} 'bitter' with *kâ*^{c} 'liver', from P[roto] ST **ka* and *[prefix] *ka-n* (cf. Garo *k'a* 'bitter', *bi-ka* 'liver', with aspiration conditioned by initial position, as in Ch.), as if the linguist would anticipate the retention of a semantic link in entirely dissimilar, 'frozen' forms of this type (here dating from the pre-Archaic period)! The other example (TB 'tree, wood, firewood' and Ch. 'firewood') selected by this reviewer for illustration is equally disastrous for his cause (cf. Matisoff 1975).

Actually, the *STC* understates the case for basic lexical agreement between TB or T[ibeto K[aren] and Chinese. A continuing search through the comparative material, greatly aided by new findings regarding prefixed **s-* in Chinese (below), has brought to light a host of cognate sets in addition to those described in *STC*. The vast majority of the basic Ch. kinship terms and body part words can now be shown to have TB cognates, generally with little if any semantic change—so much so that it is the *lack* of a known cognate set in any given 'slot' (e.g., 'father's sister' or 'cheek') that is to be noted. The cognate sets become less numerous only as we approach the less basic areas of the lexicon, e.g., the color words are only partially represented ('white' but not 'black'; 'red' but not 'green' or 'yellow'; see *STC*: Appendices I and II). As pointed out by Miller (1974), the *STC* employs the phrase 'core vocabulary' without attempting to define it, and it is indeed difficult to place this concept within a precise framework. By way of illustration—and without any intent of raising glottochronological issues—a survey of the TB (and TK) and Chinese cognate sets (earliest reconstructed levels) from the widely utilized Swadesh 100-word list of basic (= 'core') roots[3] yields the following figures: 59 cognate pairs without significant semantic shift; 12 cognate pairs with significant semantic shifts ('nose' > 'self'; 'tree' > 'firewood'; 'burn' > 'fire' [the latter pair both paralleled in TB]); 29 pairs without apparent cognation. This tabulation, based almost entirely on material already available in *STC* and very much on the conservative side, indicates that almost three-quarters of these Swadesh roots are represented by cognate sets—a very sizeable proportion, indeed. In this connection—and again without wishing to make an issue of possible dating—attention is called to the curious coincidence presented by Burling and Bhattacharya's paper (1956) on lexicostatistic dating of the Bodo-Garo split, with the Swadesh list there showing 51 [certain] cognates, 8 doubtful cases [51 + 8 = 59] and 41 non-cognates (yielding a time span of almost 1,900 years). The Bodo-Garo figure is somewhat too low, since the authors failed to recognize certain cognate pairs (for Bodo *ga-ham*, Garo *nam-a* 'good', marked as non-cognate, see *STC*: fn. 250), but in any event the comparison is instructive, since Bodo-Garo (Barish) is one of the most closely knit TB groups.

To return to our second question above, ST now appears to have attained full status as a language family—the world's second largest in number of speakers after Indo-European—but can it be defined (as in *STC*) as including only TB, Karen and Chinese? The traditional view, embodied in a classical paper by F. K. Li (1937), grouped Tai (later Kam-Tai) and MY with ST, mainly because of their monosyllabism and their congruent tonal systems. This view has generally been favored by Chinese scholars, e.g., Li's article was reprinted in the first issue (January, 1973) of the *Journal of Chinese Linguistics* with the editorial comment, ". . . Li's broad outline remains essentially accurate and useful,"[4] while K. Chang includes both Kam-Tai and MY forms in articles on ST words for 'needle' (1969) and 'iron' (1972). A change now appears to be taking place here, however, e.g., the emminent scholar, Chou Fa-kao, in a long review article of *STC* (1972), expresses some doubt on this general point: "As to whether or not the Tai and Miao-Yao languages are related to Sino-Tibetan, we still must wait for deeper research" (transl.

from p. 165). The European scholars (notably Egerod 1973, 1974; Forrest 1973; Haudricourt 1954, 1973) as well as the American (notably Bodman 1967, 1975; Matisoff 1973bis) generally are in agreement with *STC* in keeping Kam-Tai apart from ST, without necessarily endorsing the writer's view on Austro-Thai.[5] The MY languages, which have attracted much less attention from linguists, are still included in ST in recent papers by Chang (*op.cit.*) while G. B. Downer (1971), another authority on these languages, reaches the following cautious conclusion: ". . . the relationship with Sino-Tibetan . . . may eventually be shown to be a genetic one." Like R. Shafer, who had earlier (1964) attempted to set up phonological correspondences between PTB and P[roto] MY, Downer was especially impressed by the similarities in the numerals between these two language groups. As pointed out in *STC* (fn. 14), the MY numerals for 'four' and 'six' through 'ten' closely resemble TB (rather than Chinese) numerals, but the forms for 'four' are the product of a startling convergence (below) while the remaining numerals were early loans into PMY. Continuing analysis in this area, with the aid of Purnell's (1970) reconstruction of PMY, strongly indicates that these PMY loans were from a distinct (TB or ST) language (Arch.-LMY), quite different from the early Chinese dialect (Arch.-LPT) which gave rise to the parallel P[roto] [T[ai] loans. The table below presents the pertinent reconstructed forms; PT +**et* and +**(h)ńii* used only in comp. ('11', '20', etc.), the general PT roots here being **(h)nɨŋ* 'one' and **sooŋ* 'two', the latter a back-loan from Arch.-LPT **so·ŋ* = Arch./Anc. *sŭŋ/ʂɔŋ*[d] 'pair', from PAT *[*pa*]*ts*[*a*]*ŋ*, id. (> **saaŋ* through regular shifts; Arch.-LPT regularly has medial **o·* rather *u* or *ŭ* for earlier long **a·*); the PMY 'look-alike': **pl̥ei* 'four' is from PAT **śu(m)p/l̥/at* (Atayal [Formosa] also shows this infix); the bracketed forms under PST/PTB are the conjectural prototypes.

	PT	*Arch.-LPT*	*PST/PTB*	*Arch.-LMY*	*PMY*
1	+**et*	**i̯ĕt*	**s-kat*	—	**a/i*
2	+**ńii*	**ńiəy*	**g-nəy(/s)*	—	**(ə)war*
	+*hńii*	**hńi̯əy*	*[*s-nəy*]	—	
3	**saam*	**sâm*	**g-sum*	—	**pua*
4	—	—	**b-ləy*	—	**pl̥ei*
	**sii*	**si̯əy*	*[*s-ləy*]	—	
5	—	—	**l-ŋa*	—	**pra*
	**haa*	**hŋa*	*[*s-ŋa*]	—	
6	—	—	**d-ruk*	**dluk*	**dluʔ*
	**hrok*	**hri̯ôk*	*[*s-ruk*]	—	**tluʔ*
7	**cet*	**tsi̯ĕt*	**snis*	**zńi̯ə(t)*	**zńia*
					**zńia(t)*
8	—	—	**(b-)g-ryât*	—	—
	**peet*	**pi̯at*	*[*s-b-ryât*]	—	—
		—	*[*s-ryât*]	**hi̯at*	**hyat*
9	**kaw*	**ki̯əw*	**d-kəw*	**gyuə*	**gyua*
10	—	—	**gip*	**gyi̯əp*	**gyiap*
	**sip*	**ź i̯əp*	*[*s-gip*]	—	—

Notes on table

'three': Arch.-LPT **sâm* is the likely (dialectical) source of the irregular Anc. *sâm*, for the anticipated **sə̂m*, corresponding to Arch. *səm* < PST *[*g-*]*sum*.

'four': the characteristic *s-* prefixation of Chinese is well illustrated here and in 'two', 'five', 'six', 'eight' and 'ten'; the indicated shifts are generally of 'cluster' rather than 'prefixial' type (see below).

'five': here and in 'six', Arch.-LPT simply replaced the earlier prefix with **s-*; PT **haa* < **hŋaa*, but Ong-Be (Kadai language on Hainan) has *ŋa*.

'six': Arch./Anc. [*g*]*li̯ôk/li̯uk* shows replacement of prefix **d-* by **g-*, as in WB (*khrok*) and Kachin (*kruʔ* < **kruk*), possibly in relation to **g-sum* 'three' ('six' is two 'three's'); the reconstruction is based on the identity of [*g*]*li̯ôk/li̯uk*[e] (from 'six' series: GSR-1032) and *gli̯ôk/li̯uk*[f] (from GSR-1069, which has well-attested initial *gl-*) 'grain sown late and ripening early' (contra *STC*, which cites *mli̯ôk/mi̯uk*[g] 'accord' = *m-kli̯ôk/mi̯uk*; see below).

'seven': note the contrasting developments in this numeral root, reconstructed with the rare initial **sn-* cluster; Arch.-LPT parallels the regular Ch. (*ts i̯ĕt*) development, with the anticipated *-t* for final **-s* but with unaspirated initial, as shown by the PT loan (**ćet* for **tset*, lacking in PT); Arch.-LMY, on the other hand, shows an entirely different type of handling of the initial cluster, with secondary voicing[6]; this is clearly reflected in the PMY loan, which can readily be reconstructed on the basis of the parallel PMY root for 'heavy' (Purnell reconstructs separate roots for PY [**sia*] and PM [**syoŋ*] and does not even cite the YHN form!) (low tone noted by ')[7]:

	PMY	*YCR*	*YHN*	*MCF*	*MWN*	*MPT*	*MSY*
heavy	**hnia*	*hnia*	*hni*	*hńoŋ*	*hńai*	*hńa*	*hńaŋ*
seven	**zńia*	*sìa*	*ńi*	*syòŋ*	—	—	—
	**sńia(t)*	—	—	—	*syai*	*sya*	*syaŋ*

The Miao forms show typical secondary nasalization (and some unvoicing of the initial); the MWN tone points to a PM/PMY final stop doublet such as **sńia*[*t*], suggesting a basic **zńiə* ~ **zńi̯ət* doublet for Arch.-LMY (PMY regularly drops final **-s* or replaces with *-i*).[8]

'eight': again the two 'Archaic' sources show contrasting lines of development, with Arch.-LPT nearly paralleling the regular (Ch. *pwăt*) line, both with secondary unvoicing of the 'new' initial **b-* by prefixed **s-*, but with early loss of the **-r-* (**s-byat* < **s-bryât*) rather than the **-y-* (**s-bwăt* < **s-bryât*); the same feature is found in the early loan into Min-chia (*pi̯a*) and is faithfully reflected in PT, which regularly has **ee* < **ia* (**peet* < **piat*); Arch.-LMY, however, shows simple replacement of the original **g-* prefix by **s-*, followed by **sry-* > **hi̯-* (PMY, as reconstructed by

Benedict [1975], has initial **sr-* as well as **hr-*, hence would have maintained a prototype of either kind).

'nine': again a contrast, with Arch.-LPT showing the regular development (Ch. *kjəw*) (ambiguous for *j*, not reflected in PT); Arch.-LMY shows secondary voicing after the prefix, found also in WT (*dgu*), as well as palatalization of the initial and diphthongization of the final; the PMY initial **gy-* is clearly indicated here by PY **dua* ~ PM **jua* (the MY languages tend to shift velars to dentals/palatals).

'ten': still another contrast, with Arch.-LPT probably paralleling the regular form (Ch. *d'jəp*), although PT shows typical secondary unvoicing of the initial (cf. Mak [Kam-Sui] *zip*); the Arch.-LMY form shows the anticipated palatalization and diphthongization; the PMY forms are kept apart by Purnell but can be fitted into a single scheme under the reconstruction PMY **gyiap*; cf. the following table, showing parallel vocalism in the PMY root for 'cloth' (low tone ˋ, high tone ˊ):

	PMY	*YCR*	*YHN*	*MWN*	*MCF*	*MPT*	*MSY*
ten	**gyiap*	*tsìep*	*śap*	*dyau?*	*kyŭ?*	*kàu?* *ćàu?*	*kàu?* *ćàu?*
cloth	**ntia*	*dia*	—	—	*thyu*	*ntau*	*ntau*

The protoypes labeled 'Arch.-LMY' hardly look Chinese and probably should be classified as TB, yet they do show several characteristic Ch. features, including *l* for **r* as well as extensive palatalization and diphthongization. An interesting trio of roots, all 'calendrical' in nature, would at first glance appear to be early loans from ST (in whatever form): PMY **hnɔi* 'sun/day', PST **[s-]nəy* (cf. Stau *śńi*; fn. 6); PMY **hla* 'moon/month', PTB **s-gla* [initial uncertain] but WT *hla*, P[roto] K[aren] **hla*; PMY **śńiaŋ* 'year', PST **s-ni·ŋ*.[9] The 'sun/day' correspondence can hardly be ruled out on phonological grounds (**-ɔi* < **-əy* is not unlikely for PMY); it implies an Arch.-LMY prototype **hnəy* or **hn̥əy*, with PST **s-n-* > **hn-* contrasting with PST **sn* > **zń-* (in 'seven', above) — but the 'year' correspondence, with PST **s-n-* > **sń-* (see below), *certainly* reflects an early loan, hence we must remain in doubt here. The 'moon/month' correspondence, a very obvious similarity which has attracted the attention of many scholars, including both Shafer and Downer, is almost certainly another 'look-alike', since there is an excellent PAT (Formosan) correspondence: **q[i]las* "white/moon" (a semantic association repeated in two other PAT roots), with regular sound shifts into PMY: **q-* > **h-*, **-s* > **-~ *-i*, the latter explaining the YHN doublet *la* 'moon' ~ *lai* 'month' as well as the curious added MCF gloss 'white (of egg)'! For 'year' the PMY reconstruction clearly is **śńiaŋ* (contra *STC*: fn. 251; for the initial, cf. 'seven', above), although here again Purnell failed to unite the roots; note the parallel PMY root for 'tree' (**ntiaŋ*),[10] with YNH maintaining medial **-ia-* but YCR simplifying to **-a(a)-*:

	PMY	*YCR*	*YHN*	*MCF*	*MWN*	*MPT*	*MSY*
year	**śńiaŋ*	*hńaaŋ*	*hńiaŋ*	*hnau*	*syau*	*syoŋ*	*syoŋ*
tree	**ntiaŋ*	*dyáŋ*	*dyíaŋ*	*tau*	*ntau*	*ntoŋ*	*ntoŋ*

The indicated Arch.-LMY prototype here would be the anticipated **śńiəŋ*. The marked tendency for PMY medial **-ia-* to be simplified to **-aa-* after initial **ń-* also serves to explain another early loan, viz. PMY **ńaan* 'silver', for **ńian*, pointing to an Arch.-LMY prototype **ŋiən* < PST **[d-]ŋul*, similar to the regular form: Ch. *ŋiɛn/ŋiĕn*^{k} (PMY **ń-* < **ŋ-* is regular shift); a possible doublet here, suggesting Arch.-LMY **ŋuan* (with PST **-u-* retained), is represented by the irregular L[ing] C[hun] Y[ao] *ńwan* < PMY **ń[ua]n*. The parallel loan from Arch.-LPT is PT **ŋə(ə)n* (with the rare, barely 'naturalized' vowel **ə*), suggesting that the two lending 'Archaic' sources had the same form here: **ŋiən*.

The above material is of great interest for the light that it throws upon early ST dialects/languages but it scarcely constitutes evidence for a genetic link-up of the Tai or MY languages with ST. In point of fact, the exclusiveness of the correspondences, even in the numerals, speaks very much *against* the genetic hypothesis, e.g., the PST/PMY correspondences involve the numerals *above* rather than *below* 'five'; even in Tai, where the resemblances extend below 'five', the loans for 'one' (**et*) and 'two' (**ńii* ~ **hńii*) are used only in composition, hence do not 'count' in the Swadesh 100-word list score (see below). One does not really have to make any analysis of 'core' vocabulary in connection either with Tai or with MY since there are virtually no substantial ST resemblances (other than those noted above) in this basic part of the lexicon, hence Shafer felt compelled to compare PST **s-na* 'nose' with PT **hna* 'face' (< PAT **q/ndza[q]ais* 'face/forehead'), while Downer was driven to citing PMY **mbrui* 'nose', Ch. *b'jəd/b'jĭ*^{h} (Downer 1971); Yao [YCR] *kyu* 'dog', Ch. *ku/kən*^{m} (Downer 1973) even while conceding that the PMY root has initial **kl-* [YHN *klo*]! The real problem here has always been why anyone (including the writer once upon a time) has ever seriously taken the Kam-Tai and/or MY languages to be true 'blood cousins' of ST, given the almost total lack of any basic ties in the respective lexicons.

In turning now to ST itself, still another basic question arises: is there support for the *STC* scheme of setting off Chinese from TK as a unit, and then of setting off Karen from TB? Although the latter point has always been in considerable doubt (cf. *STC*: fn. 350), the need for separating Chinese from the other ST languages has scarcely been questioned. Recently, however, N. Bodman (1973) suggested that Chinese stands in a special relationship to Tibetan: "It is also my subjective impression that the group comprising Tibetan and its near relatives is closer to Chinese than are many other groups of the Tibeto-Burman languages" (*op.cit.*: 386). Although the arguments adduced in that paper, from phonology and morphology, did not appear conclusive the suggestion did seem to have some merit inasmuch as the writer himself, like Bodman, had long been

impressed by the numerous special lexical links between Chinese and Tibetan. With this in mind—and primarily looking for support for the *STC* scheme—the writer drew up a table of cognate sets from the Swadesh 100-word list (fn. 3) for M[andarin] (Peking), P[wo Karen] (Moulmein), T[ibetan] (Lhasa = Central), B[urmese] (Rangoon), K[achin], G[aro] (Achik) and L[ushai] (Dulien), making use only of the modern spoken languages. The plan was *not* to establish glottochronology but rather to determine whether Chinese and/or Karen shared significantly more of these basic roots with any of the given TB languages. Burling (1971) pioneered in the use of this approach for four TB languages (Kachin, Maru, Lushai and Garo), dispensing with the time scale ("Let me assert with all possible vigour that I have no faith in the Swadesh list as a technique of absolute dating, or even for very accurate relative dating," *op.cit.*: 13). Although he had to rely largely on 'guess work' for the determination of cognate pairs (and many of his assignments to pairs must now be set aside on the basis of comparative phonology; cf. *STC*), it is noteworthy that he was able to find support for a special link between Kachin and Garo (credited by him to Shafer, but noted by the writer in the early 1940's: see *STC*; 7, 11).

The Swadesh 100-word list has serious disadvantages as a tool in studying ST languages, e.g., it includes such unlikely entries as 'who?', 'what?' and even 'we' while omitting many more 'basic' items, including one ('dream': PST **[r-]ma(·)ŋ*) which qualifies for membership in that select group of roots represented by a plus (= 'retained in listed meaning') in all seven languages under study (see below). It does offer the great advantage, however, of having become a standard in the field, so that one can turn to it without wondering whether or not his analysis is being distorted by an unconscious bias in his selection of 'basic' roots. An effort was made to follow the rules that customarily obtain in the field, especially as regards accepting only the 'usual' or 'most common' word under any entry, e.g., under 'all' no set has been accorded a plus, including a possible T/B set under PTB **(m-)kul* (in *STC*) and a possible T/K/L set under PTB **yoŋ* (not in *STC*): WT *yoŋs* 'all, whole'; K *yoŋ* 'all, the whole', *yoŋ-yoŋ*, id.; L *zoŋ-zoŋ* 'all' only the K word qualifies here). Entries were given a plus if the listed meaning (underlined) is retained in composition despite a semantic shift in the root; note here especially P *θe* 'skilled; (comp.) know'; L *lam* 'way, direction, place' (paralleling the basic Karen shift); (comp.) road'. The dialect specifications given above were strictly adhered to, with the result that L 'lost' a plus in the set under 'heart' (only in Ngente dialect: *niŋ*) and G a plus under 'dog' (only in 'Garo A' dialects: Koch and Ruga *kui*, Rabha *ki*). The *STC* framework was followed in determining cognation and almost all the relevant material can be found in that work.[11] As might be anticipated, however, there remain a number of problem assignments in which cognation is uncertain. It is very probable that no two scholars would find themselves in complete agreement on all these problem assignments, and in any event the number of such cases is too small to make any significant change in the findings as a whole. The main points to be noted here are outlined below for each of the seven languages under study:

BURMESE: *krai* 'star' possibly related to PTB **s-kar = *s-ka·r* (L *a·r-*) (Matisoff 1974, No. 49) but considered too vague to rate a plus; *na* 'listen' vs. *nà* 'ear' but distinct word (*krà*) for 'hear'; contrast K *nà* (low-tone) 'hear' vs. *nā* (mid-tone) 'ear', which was given a plus in relation to G *kna* 'hear' (*na-tśil* 'ear').

TIBETAN: *lo-ma* 'leaf', from **lob-ma*, as shown by West T *lob-ma*, from PTB **lâp* (K *lap*); *duma* 'to smoke', *dud-pa* 'smoke' possibly related to PTB **kəw* (Matisoff 1974: fn. 119) but considered unlikely, the T representative of this root probably being *khu-ba* 'fluid, liquid', with parallels elsewhere.

LUSHAI: *ha* 'tooth' given a plus in the general set from PTB **(s-)wa* although the initial remains unexplained (*STC*: 106).

KACHIN: *kəra* 'hair' < PTB **s-kra* given a plus in set with T *skra*, with *sam* (specialized use) an apparent early loan from B *tsham* < PTB **tsâm* (as indicated by the **ts- > s-* shift); K *ren* (< PTB **(s-)riŋ*) 'too long' (not 'long', as glossed in *STC*: 106), the regular word for 'long' (*gəlu*) showing a plus in relation to G *ro* < PTB **low*; contrast B, which has retained the first root in the general meaning 'long' (*hrań*) but has modified the latter (*lu* 'disproportionately tall'); *n-* 'not' from **m(a)-* (*STC*: 97) given a plus in this set with T and B; *ləđi* 'nose', perhaps from **[s-]na-di* as suggested by Khauri dialect *nədi*, but considered uncertain (dialect data lacking) and not given a plus in the general set from PTB **s-na*; *nra ~ nraŋ* 'bone' (fn. 11), also *nrut* < PTB **(-)rus* (*STC*: 16) and *nrut nra* (typical K couplet formation) but latter considered secondary term, the plus given for *nra ~ nraŋ* (set with G).

GARO: *baʔaŋ* 'many', from **ʔbaŋ*, directly cognate to WT *'baŋs* 'the people, the subjects' (= 'the many') < PTB **a-baŋ*, a doublet of PTB **maŋ*: WT *maŋ-ba* 'much', *maŋ-po* 'much, many', *dmaŋs* 'the common people'; L *maŋ* 'very, much' (but not 'many'); K *maŋ* 'the inhabitants of a village',[12] but no plus given any language under this entry ('many'); *mikka* 'rain' < PTB **m-ka*, cognate to WT *mkha* 'the heavens', with the basic meaning preserved at times: *mikka sim* 'to be cloudy' (sim 'black'), *mikka kim* 'to thunder' (*kim* not glossed separately); also *mikka wa* 'to rain' (*wa* < PTB **r-wa(-s)* not glossed separately); given a plus in the set with B/L/M.

PWO: *nọ* 'horn' < PK **nuŋ* has been compared (*STC*: 143) with the general TB set from **rwâ ~ *rwâk ~ *rwâŋ* (see below) but not given a plus in view of the divergent initial; *khla* 'ashes' perhaps related to the PTB root **pla* (B/G set) via a root **(-la)*, which might be analyzed as a derivative of **tap-la* (PTB **tap* 'fireplace'), especially in view of G *tapra*, Dimasa *thapla* (but Mikir *phelo* < **phla*) (*STC*: 133 and fn. 364, which cites Taungthu *pha*, possibly from **phla*), but this considered too uncertain to rate a plus; *xwe* 'full' (< PK **bray*) possibly cognate to the PTB root **bliŋ* (*STC*: fn. 377) but this also considered too uncertain to be given a plus.

MANDARIN: given a plus for 'horn': Arch./Anc. *kŭk/kɔkʷ* < **klŭk* (the cluster well attested by dialects [Yang]), with the doublet *klăk/kɔkˣ* 'deer's horn', from **kl[wâ]k* < **k-rwâk*, fitting in with the PTB **rwâ ~ *rwâk ~ *rwâŋ* series (for *STC *rwa ~ *rwaŋ*, and add Dzorgai [Thochu] *rak* 'horn' < PTB

*r[wâ]k, contrasting with *niek* 'black' < PTB **nak*); also for 'person': Arch./Anc. *ńi̯ĕn/ńźi̯ĕn*[y], a doublet of *mi̯ĕn/mi̯ĕn ~ mi̯ən/mi̯ĕn*[z] 'people' < PST **mi-n* (*STC*: fn. 428), with parallels for this shift in initial nasals; also for 'one': Arch./Anc. [*s-k*]*i̯ĕt/ʔi̯ĕt*[aa], cognate with L *khat* (contra *STC*: 94); also for 'smoke' : Arch./Anc. [*s-k*]*ien/ʔien*[ab] (*STC*: fn. 441). As regards the personal pronouns, however, present analysis indicates that the M forms for 'thou' as well as 'I' have been developed from distinct PST roots relating to 'self' (contra *STC*: 160, which handles them as doublet forms), the comparative evidence coming primarily from WT: 'I': Arch./Anc. *ŋâ/ŋâ*[ac], cognate with WT *ŋo* 'face, countenance, air, look; self, the thing itself; the self, the I' < PST **ŋâ*[13]; 'thou': Arch./Anc. *ńia/ńźie*[ae] (loan use), cognate with WT *nyid* (< **ńi-t*) 'self, same' (e.g., *ma nyid* 'the mother herself'), in more recent literature used resp. for *k'yod* 'thou, you'; WB *ńàń* (< **ńìn* < **ńi-n*) 'you' (fem.), all from PST **ńi(-t, -n)*; note the parallel development shown by PTB **ŋay* (perhaps ultimately from **ŋâ-i*): L *ŋei* 'self', WT *ŋed* (< **ŋe-t*) 'I, we' (elegant), K *ŋai* 'I' (*STC*: 65) (scored as plus for the K/M pairing).

The following tables show the roots for which cognate sets have been uncovered, arranged by number of languages represented, beginning with a select group of three roots ('die', 'name', 'fish') with plus scores in all seven languages under study (abbreviations: Ar. present in Archaic only; cp. in composition only; dl. present in another dialect only; lw. present as loanword only; rp. replaced as main word; sm. semantic difference (no plus):

	PST/PTB	*B*	*T*	*L*	*K*	*G*	*P*	*M*
(7)								
die	**səy*	+	+	+	+	+	+	+
name	**r-miŋ*	+	+	+	+	+	+	+
fish	**(s-)ŋya*	+	+	+	+	+	+	+
(6)								
eye	**(s-)myək*	+	+	+	+	+	+	sm.
moon	**s-gla*	+	+	+	+	+	+	—
kill	**(g-)sât*	+	+	+	+	+	—	+
dog	**(s-)kwəy*	+	+	+	+	dl.	+	+
two	**[g-]nəy(-s)*	+	+	+	cp.	+	+	+
ear	**[r-]na*	+	+	—	+	+	+	+
blood	**(s-)hywəy*	+	—	+	+	+	+	+
smoke	**(s-)kəw(-n)*	+	sm.	+	+	+	+	+
tail	**r-may*	+	—	+	+	+	+	+
(5)								
road	**lam*	+	+	+	+	+	sm.	—
tooth	**(s-)wa*	+	+	+	+	+	rp.	—
liver[I]	**m-sin*	+	+	+	+	—	+	—
stone	**(-)luŋ*	+	—	+	+	+	+	—
eat	**dza(-s)*	+	+	sm.	+	+	sm.	+
louse[I]	**śrik*	—	+	+	+	+	—	+

	PST/PTB	*B*	*T*	*L*	*K*	*G*	*P*	*M*
bone[I]	**(-)rus*	+	+	+	rp.	—	+	+
breast	**nəw*	+	+	+	—	—	+	+
(4)								
-nail	**m-syen*	+	+	+	+	—	—	—
meat[I]	**śa(-n)*	+	+	+	+	—	—	sm.
hand	**(g-)lək*	+	+	—	+	+	—	sm.
feather	**(-)mul*	+	sm.	+	+	+	—	sm.
fire[I]	**(s-)mey*	+	+	+	cp.	—	+	—
tree[I]	**siŋ*	+	+	+	—	—	+	sm.
know	**syey(-s)*	+	+	—	sm.	+	+	—
warm	**lum*	+	—	+	+	sm.	+	—
tongue[I]	**(-)lay*	—	+	+	rp.	+	+	—
horn	**rwâ(-k, -ŋ)*	—	+	—	+	+	—	+
rain	**(r-)wa(-s)*	+	—	+	—	+	—	+
new	**sar*	—	+	+	—	—	+	+
thou	**na(·)(ŋ)*	rp.	—	+	+	+	+	Ar.
fat, n.	**sa·w*	—	—	+	+	sm.	+	+
(3)								
neck[I]	**(-)liŋ*	+	+	+	—	—	—	sm.
sun	**nəy*	+	+	+	sm.	—	sm.	sm.
not	**m(a)-*	+	+	—	+	—	—	Ar.
round	**zlum*	+	+	sm.	+	—	—	—
sit	**( )du·ŋ*	+	+	—	+	—	—	—
star	**s-ka·r*	—	+	+	+	—	—	—
nose	**s-na*	+	+	—	—	—	+	sm.
foot	**r-kaŋ*	—	+	—	+	—	+	—
head	**(-)gaw*	—	+	—	—	+	+	—
I[I]	**ŋa*	rp.	+	—	—	+	+	Ar.
egg	**(-)təy*	—	sm.	—	+	+	+	sm.
night	**(-)ya(-n)*	+	—	+	—	—	sm.	+
full	**(s-)bliŋ*	+	—	—	+	—	—	+
stand[I]	**(g-)ryəp*	+	—	—	+	—	—	+
person	**mi(-n)*	—	+	+	—	—	—	+
leaf	**(s-)lâp*	—	+	—	+	—	—	+
fly, v.	**(s-)pur*	—	+	—	—	+	—	+
fire[II]	**(-)pwâr*	—	sm.	—	+	+	—	+
bird	**(-)tow*	—	—	sm.	—	+	+	+
(2)								
water[I]	**təy*	—	sm.	—	sm.	+	+	sm.
skin/bark	**(-)pik*	—	—	—	+	—	+	—
tongue[II]	**(-)lyak*	sm.	rp.	sm.	+	sm.	—	+
I[II]	**ŋâ(-i)*	—	sm.	sm.	+	—	—	+
liver[II]	**(-)ka(-n)*	sm.	sm.	sm.	sm.	+	sm.	+
mouth[I]	**ku*	—	—	—	—	+	—	+
white[I]	**(-)b[o]k*	—	—	—	—	+	—	+
meat[II]	**s-ńa·[k]*	—	—	—	—	—	+	+
black	**nak*	+	+	—	—	—	—	—
heart	**(s-)niŋ*	+	+	dl.	—	sm.	—	sm.

(continued)

	PST/PTB	*B*	*T*	*L*	*K*	*G*	*P*	*M*
long[I]	**(s-)riŋ*	+	+	—	sm.	—	—	sm.
one[I]	**[g-]tik*	+	+	—	—	—	—	—
give[I]	**(s-)bəy(-n)*	+	+	—	—	—	—	rp.
hair[I]	**(-)tsâm*	+	sm.	+	lw.	sm.	—	sm.
seed	**dźəy*	+	—	+	—	—	—	—
dry	**krok*	+	—	—	+	—	—	—
sleep	**yip*	+	sm.	—	+	—	—	—
ashes	**pla*	+	—	—	—	+	—	—
louse[II]	**sar*	+	—	—	—	—	+	Ar.
mouth[II]	**s-nut*	+	—	—	—	—	+	—
hair[II]	**s-kra*	—	+	—	+	—	—	sm.
that	**day*	—	+	—	+	—	—	—
knee	**(-)put(-s)*	—	+	—	+	—	—	sm.
see	**(s-)m[u, əw]*	—	—	+	+	—	—	—
stand[II]	**di·ŋ*	sm.	—	+	sm.	+	—	sm.
give[II]	**pe(k)*	—	—	+	—	—	+	—
white[II]	**(-)pwâr*	—	—	+	—	—	+	rp.
drink[I]	**(s-)əm*	—	—	+	—	—	sm.	+
one[II]	**(s-)kat*	—	—	+	—	—	—	+
water[II]	**(s-)twəy*[B]	sm.	—	+	sm.	—	—	+
bone[II]	**(-)raŋ*	—	—	—	+	+	—	—
drink[II]	*lu·(-k, -ŋ)*	—	—	—	+	+	—	—
earth	**(r-)ka*	—	—	—	+	+	—	—
hear	**(-na)*	sm.	—	—	+	+	—	—
long[II]	**(-)low*	sm.	—	—	+	+	—	—
neck[II]	**(-)twâk*	—	—	—	+	+	—	—
sun[II]	**tśar*	—	—	—	+	+	—	—
tree[II]	**(s-)bul*	—	—	sm.	+	+	—	sm.

The following table shows the number of cognate sets shared by any given pair of languages. The [Garo] bracket suggests the anomalous position of this language in the scheme of relationships. The 18/19 entry for K/P and P/K serves as a reminder that two words ('bark' and 'skin') from the Swadesh list are alternative glosses for the same root [*(-)*pik*] in these two languages.

	Burm.	*Tib.*	*Lush.*	*Kach.*	*[Garo]*	*Pwo*	*Mand.*
Burm.	—	32	29	28	20	22	17
Tib.	32	—	24	23	19	20	16
Lush.	29	24	—	23	19	23	18
Kach.	28	23	23	—	29	18/19	19
[Garo]	20	19	19	29	—	19	18
Pwo	22	20	23	18/19	19	—	15
Mand.	17	16	18	19	18	15	—

The above table lends itself to the following conclusions:

1. Chinese (Mandarin) does indeed appear to lie well outside the ordinary range of Tibeto-Burman and even of Tibeto-Karen. The 'flat' distribution (15 to 19) of the Mandarin scores also suggests that Chinese has retained a sizeable share of these basic roots but in essentially a random fashion, as consistent with the view that this language was the first to split off from the ancestral ST proto-language. The score (16) for the M/T pairing is actually on the low side and appears to preclude the idea (above) that Chinese and Tibetan have any special *genetic* relationship apart from their membership in the ST language family. Ironically, however, the very *absence* of a cognate pair, in terms of the Swadesh list scoring, can in a given case be significant, e.g. neither M nor T has the PST root *(-)*mul* 'body hair' (> 'feather' in list) in this core meaning for the very good reason that each has made the identical semantic shift to 'eyebrow' with the same medial *-*i*- doublet (*STC*: fn.'s 53, 461), hence neither gets a plus under this entry ('feather')! Again, M does not receive a plus under 'knee' since the basic root here [PTB *(-)*put(-s)*: WT *pus-mo* < **put-s*, K *ləphut*] underwent a semantic shift to 'knee-cover': Arch./Anc. *pi̯wət/pi̯uət*[ef]*~pi̯ĕt/pi̯ĕt*[eg] (cf. West T *pis-mo* for the doublet), yet the languages do share one highly specific cognate set in this area, viz. Arch./Anc. *si̯ĕt/si̯ĕt*[eh] 'knee', from **sgyit* (contra *STC*: fn. 455); WT *sgyid(-pa)* 'knee-joint, knee-hollow; calf'; cf. also Arch./Anc. *b'i̯ĕn/b'i̯ĕn*[ai] 'knee-cap' (Anal. Dict. also 'knee, leg'); WT *byin-pa* 'calf' (cf. the WT gloss for *sgyid*); also *kwɛk/kwɛk*[aj] 'hollow at back of knee', apparently from **krwak* < **r-kwak*, comparable with WB *khwak* 'concave/sunken'.
2. The above conclusion necessarily leads to a second, viz., that *after* Chinese split off from the parent ST stock it came in close contact with Tibetan for an extended period of time, of sufficient duration to permit of marked (probably mutual) linguistic influence. It would appear impossible to explain on any other basis the numerous and often detailed resemblances between the two languages, as long noted by this writer as well as by Bodman (above) and other scholars. It should be emphasized that this *antedated* the Archaic period, as can be seen from the sound shifts often involved, e.g., Tibetan shows a highly idiosyncratic pattern of kinship terms derived through *s-* prefixation and *-n*, *-d* suffixation (*STC*: fn. 284); this pattern is reflected elsewhere only in Lepcha, directly under Tibetan influence, with sporadic appearance of suffixed *-n* forms in Burmese and other languages, yet the early Chinese kinship terminology reflects this pattern in several areas, including one term tying in directly with Tibetan, viz., Arch./Anc. *sk'yi̯wət/tś'i̯uĕt*[ak] (loan use in *Erh ya*) 'sister's son (man sp.)' (contra *STC*: fn. 428; see below for reconstruction), from **s-kəw-t*, from the basic PST root **kəw* 'mother's brother' = 'the descendants of the mo's bro' (through the terminological equations ensuing from x-cousin marriage: 'si's

son' = 'mo's bro's son'), morphologically identical with WT *skudpo* 'brother-in-law' (Jäschke, but Das glosses 'wife's brother' and Desgodins has 'frater uxoris'; in x-cousin marriage = 'mo's bro's son').

3. The Karen (Pwo) scores are consistent with the hypothesis that this language split off from a parent TK stock (*STC*: 2) yet are hardly distinctive enough to rule out the view that Karen is simply another division of the TB family. The score for the P/B pairing (22) is greater than that for P/T (20) but smaller than that for P/L (23), hence this evidence tends to preclude another possible view (*STC*: fn. 350), viz., that Karen represents an early split from the Burmese-Lolo group of TB. Especially worthy of note are the low scores (18/19 and 19) for the P/K and P/G pairings, suggesting that these splits date from an early period.
4. The Garo scores are in a very low range (18 to 21), with the striking exception of the score (29) for the G/K pairing, one of the highest scores in the table. The conclusion here must be that Garo also represents an early split from the parent TB group but one that *also included Kachin*, with at least nine of the shared items being considered innovative (above table: 'fire[II]', which has M parallel, and 'bone[II], through 'tree[II]').
5. The writer has long suspected that the key 'crossroads' position of Kachin (*STC*: 5) reflects its pivotal geographical setting (N. Burma) rather than any basic linguistic makeup. Quite unlike Garo, which has only low scores apart from its high G/K pairing score, Kachin has fairly high scores for the K/T (24) and K/L (23) pairings and a definitely high score for the K/B pairing (28). These higher scores point to the existence of newer levels within Kachin, especially to a Burmese level since it is clear that the language has long been under heavy influence from the culturally dominant Burmese speech, with the result that is difficult to determine the precise status ('native' vs. loanword) of many lexical items (see the discussion of 'hair', above, under KACHIN).
6. Burmese has notably high scores for both the B/T pairing (31) and the B/L pairing (29), comparable with the K/G pairing score (29). There is a fundamental distinction here, however, in the fact that very few of these shared items can be considered innovative in any sense: perhaps 'black' and 'one[II]' for B/T (both 'heart' and 'long' involve, rather, retention of general TB roots) and 'seed' for B/L ('hair[I]' is widespread TB root). There hardly seems to be sufficient evidence here for a *genetic* grouping, say, of Burmese/Tibetan vs. Lushai, yet the L/T pairing score does seem unexpectedly low (24). The scores as a whole do strongly indicate, however, that a basic cleavage line must be recognized within TB between B/T/L on the one hand and K/G on the other, the latter ('Kachin-Garo' supergroup) probably also including the Konyak ('Naked Naga') languages as well as the obselete Chairel (*STC*: 6–7).[14]

The above analysis, in furnishing a measure of support for the general taxonomic arrangement put forward in *STC*, also shores up the 'teleo-reconstruction'

procedure adopted in reconstructing a two-tone system for PST (*STC*; fn. 494) *before* undertaking a reconstruction of the PTB system itself, since this procedure is logically dependent upon the classification. This gap has now been bridged over in part, since in addition to the earlier described correlation of the (basically) two-tone Burmese-Lolo system with Nungish tones (*STC*: fn. 494) the writer has now (1973) demonstrated an excellent correlation with the Kukish tonal system as well as a tentative rapprochement with the Tamang-Gurung-Thakali (Nepal) system, both basically of two-tone type. A massive effort (Matisoff 1974) to establish a further correlation with the three-tone Kachin system, however, has been only partially successful, but note that this language presents problems related to secondary voicing (see discussion below), which might well be implicated in those relating to tone.

The special position of Chinese in the classification is also of crucial importance in reconstructing other phonological features at the PST level, especially in relation to the vowel system.[15] The PTB/PTK system sets up in general as */aeiou/* plus vocalic length · whereas the additional vowels */əâ/* are required to handle the Chinese/TB correspondences, with further light thrown upon vocalic alternation in Tibetan (*STC*: fn. 344). Further study in this area (Benedict 1973bis and to appear) has emphasized the role of vocalic length at the PST level, with distinct Arch./Anc. reflexes not only for PST **-əy* and **-ə·y* (falling together with **-oy* and **-o·y*) and for **əw* and **-ə·w* (falling together with **-âw* and **-â·w*), as suggested in *STC* (fn. 486), but also for PST **-ây* (> *-əd/ậi*) and **-â·y* (> *-âd/âi*), with the regular Arch. final *-d* for Anc. *-i*. The important PST **â > a* (but **â· > â*) shift previously noted before dental finals (*STC*: 488) must also be recognized before labial finals, e.g., PTB **lâp* 'leaf' (see note, above, under TIBETAN) is cognate to Arch./Anc. *[l]i̯ap/i̯äp*[aq], id., from PST **lâp*. PST medial **o* fell together with PST medial **u*, yielding the same reflexes in Chinese (*STC*: fn. 479); cf. PTB **s-t[o,o·]ŋ* 'empty' (WT *stoŋ-pa*); Arch./Anc. *t'i̯ôŋ/t'i̯uŋ ~ d'i̯ôŋ/d'i̯uŋ*[am], id. < PST **(s-)toŋ*, as contrasted with PTB **s-d[o, o·]ŋ* 'unite, join' (WT *sdoŋ-ba*); Arch./Anc. *d'uŋ/d'uŋ*[an] 'together, join' < PST **(s-)do·ŋ*. Finally, as regards PST medial **e*, a length distinction can be reconstructed on the basis of distinct reflexes in Chinese, with PST **e* yielding Arch./Anc. *i̯a/i̯ä* before dentals (original or secondary) and labials but **e·* yielding *ia/ie* (contra *STC*: fn. 481, which is incomplete), e.g., PTB **s[e, e·]r* 'hail' (WT *ser-ba*); Arch./Anc. *sian/sien*[ap] 'sleet' < PST **se·r*; PTB **l[e, e·]p* 'butterfly' (WT *phye-ma-leb* [cf. *leb-mo* 'flat'], WB *lip-pra* [cf. *prà* 'flat, level'); Arch./Anc. *[l]iap/d'iep*[ap], id., from PST **le·p*. Before velar finals, on the other hand, PST **e* fell together with **i* (*STC*: fn. 481) but PST **e·* yielded Arch./Anc. *io/ie* before final *-k* (reflex before final *-ŋ* not certain), as shown by the following pair:[16]

	PST	*Arch./Anc.*	*WT*	*Lep.*	*Kach.*	*WB*	*Lush.*	*Garo*
pheasant	**(s-)re·k*	*[l]iok/d'iek*[aq]	*sreg*	*hryak*	*ri?*	*rats*	*hrit*	*rik*
sink, v.	**(s-)ne·k*	*niok/niek*[ar]	—	—	—	*nats* / *hnats*	—	—

As a final note on vocalic length in PST, it is possible that certain doublets reflect an original length distinction, e.g., *pi̯wər/pywɐi̯*[at] 'fly', *pi̯wən/pi̯uən*[au] 'fly, soar' < PST **pu(·)r*. Arch. Ch. also appears to show a trend towards loss of final velars after original (PST) long medial vowels; cf. *s-gli̯u/ṣi̯u*[av] (perhaps for **s-gli̯ug*) (C) 'number; some, several'; (B) 'to count, calculate; (tell the faults of:) reprimand', also read *s-glŭk/ṣɔk* 'a number of times', possibly cognate (P. Yang: pers. comm.) to WT *sgraŋ-ba* (also *'graŋba*) 'to enumerate, count; to upbraid, reproach' (cf. the Ch. gloss), *graŋs* 'number', from PST **s-gra·ŋ*, with shift to stop in Arch./Anc. (see discussion below). There is also evidence for loss of final **-r* after long vowel; cf. WT *mgur* 'throat, neck; voice' (resp.) (WT also has *mgul* (*-pa*) 'neck, throat'); Arch./Anc. *g'u/ɣəu*[aw] 'throat' < PST **(m-)gu·[r]*.[17]

The reconstruction of the PST consonant system, as opposed to that of the vowel system, relies primarily upon TB rather than Chinese data. The schema employed in *STC* is as follows:

PST consonants									
Labial	*p*	*b*						*w*	*m*
Dental	*t*	*d*	*ts*	*dz*	*s*	*z*	*l*		*n*
Palatal	*c*	*j*			*ś*	*ź*	*r*	*y*	*ń*
Velar	*k*	*g*							*ŋ*
Glottal					*h*				

Note: glottal stop (ʔ) is non-distinctive feature of vocalic onset.

The reviews to date have indicated relatively little dissatisfaction with this schema, apart from the setting up of a simple two-manner voiced vs. unvoiced opposition for the initial stops. As pointed out in the *STC* (pp. 20–21), within TB itself there are two 'problems' about this, one involving the presence of a small number of forms with unaspirated initial obstruents in WT (where by 'rule' they should be aspirated), the other concerned with some apparently secondary voicing, especially in Kachin and Garo. Miller (1974) expresses serious reservations about both points, while Coblin (1972–73) is primarily concerned with the first point, Bodman (1975) with the second. It should be pointed out that none of the forms (WT or other) involved fits into any other 'system' of correspondences, and this is precisely why one cannot simply set up other phonemes, as Miller apparently would like to see done. A special study of the WT forms with initial unaspirated stops, as suggested by Coblin (*op. cit.*), would probably reveal that most of them fall under category (c) of the *STC* (= prefixed forms).[18] Similarly, for the secondarily voiced stops in Kachin (see the forms tabulated by Bodman, *op. cit.*) and probably also for those in Garo, earlier prefixation appears to have played a role, as already indicated by the writer (Benedict 1973: 131); cf. Kachin *gwi* 'dog', Jili (extinct Kachin dialect) *təkwi*, Rawang (Nungish) *təgi* < PTB **kwəy*, and it seems possible that Kachin first unvoiced all or most initial

obstruents, then secondarily voiced some of them, with consequent disruption of the tonal system (cf. the remarks above). As for the extra-Tibetan evidence for a two-manner system, requested specifically by Coblin (*op.cit.*), Matisoff has reconstructed such a system for Proto-Burmese-Lolo (*STC*: fn. 76, and cf. Egerod 1973: 501), and Burling for Proto-Bodo-Garo (*STC*: fn. 21). The Chinese reflexes have always presented special problems, especially since Karlgren reconstructed both **g* and **d* (but **b* only marginal) for Arch. Ch., contrasting with *g'-* and *d'-* as initials and with *-k* and *-t* as finals. The contrast in finals has led a number of scholars, including Miller, to attempt to find 'correspondences' for both sets of stops in TB or Karen, with a special fondness for Maru, which has final *-k* for WB *-ui* < PTB **-əw*, final *-t* for WB *-e* < PTB **-əy*. These are pseudo-archaic finals, as recognized by all the linguists who have reconstructed Proto-Burmese-Lolo,[19] while the Karen (Taungthu) final *-t* cited by Miller (for 'four') is unquestionably a late numeral suffix, appearing also in the Taungthu forms for 'five' (*ŋat*) and 'nine' (*kut*) as well as in 'four' (*lit*) (*STC*: 131). The reconstruction of a separate set of voiced stops for PST, as Miller and others seem to prefer, does in fact involve an 'insuperable difficulty' (contra *STC*: 186, cited by Miller, *cit. op.*) inasmuch as we would be left with a bizarre system with the glides **w* and **y* appearing *only in medial position*, being required there to handle PST roots with medial glides, e.g., TB **gwa* 'fox'; Arch./Anc. *g'wo/ɣuo*[bd] < PST **gwa*. The *STC* solution, which recognizes **w* > *g* and **y* > *d* shifts for Arch. Ch. both in initial and final position, obviates this difficulty, and the parallelism with Maru simply furnishes further support for this line of reconstruction, accepted at least in part by Bodman (1975).

The recognition of an extensive **s* (cluster or prefix) pattern for Arch. Ch. (Li 1971; Bodman 1973, 1974; Benedict 1974, 1975bis, Yang 1975), far beyond that previously recognized (*STC*: fn.'s 419, 457 and cf. Bodman 1969), has vastly complicated the problem of establishing reflexes for stop and other consonants of the language. The writer (cit. supra) has hypothesized that in order to account for the bewildering variety of reflexes (*xie-sheng* and comparative TB evidence) it is necessary to set up a basic distinction between a cluster (non-syllabic) and a prefixial (syllabic) combination, e.g., *sn, sk* as opposed to *s-n, s-k*. As shown above (text and fn.'s 6 and 9), the cluster **sn* fell together with **st*, both yielding Arch./Anc. *ts* or *ts'*, whereas the prefixial **s-n* yielded Arch./Anc. *[s-]n/n*, as suggested by the Proto-Min high tonal series reflex (**hn*, perhaps better reconstructed as **sn*), with *ts* or *ts'* as an alternative reflex (actually the more frequent reflex in the comparative material; see the table in Benedict 1974). To further complicate matters, there is also evidence for secondary voicing of the prefixial variety (**s-n* = **sən-* > *dz'*; see fn. 8), paralleling a similar reflex shown by the dental stop combination; cf. PTB **s-tay* 'navel' (Chepang *toy* < **[s-]tay*, Garo *ste*, Kachin *dai* ~ *śədai* [second. voicing]), Arch./Anc. *dz'iər/dz'iei*[be], id. (Bodman 1973); the cluster **sd* produced unvoicing (**sd* > *ts*), as might be anticipated; cf. PTB **s-di·k* 'scorpion' (WT *sdig-pa* 'scorpion', *sdig-srin* 'crab,

crawfish'; Lushai *ti·t* < **[s-]ti·k* 'scorpion'; Garo *na-tik* 'shrimp' < 'centipede fish [*na*]'); Arch./Anc. *tsək/tsi̯ək*[bf] 'scorpion'; PTB **s-dep* (WT *sdeb-pa* 'join, unite'); Arch./Anc. *tsi̯ap*[bg] 'connect' (Bodman). Before non-dental nasals, the cluster **s* combinations yielded *x* clusters: **sm* > *x(i̯)m/x(i̯)w* (see fn. 12 for example, from PST prefixed **s-*); **sŋ* > *xŋ/x* (cf. 'goose', cited in *STC*: fn. 419); also **sń* > *xń/t'*; cf. PK (Taunghtu) **hńa* < **s-ńa* 'red'; PTB **(s-)ńa*: WB *ńa* 'to respect' (= 'to show shame'), *hńa* (< **s-ńa*) 'to be considerate'); Arch./Anc. *xn̥i̯əg/t'i*[bh] 'shame, disgrace' (= 'red [face]') (*ńi̯əg/ńźi*[bi] 'ear' is phonetic) < PST **(s-)ńa* (with 'intensive' rather than 'causative' prefixed **s-*). The 'gap' in the nasal series was filled by a newly formed *xn/t'*; cf. the minimal pair under GSR-94: *sn̥i̯o si̯wo*[bj] 'coarse raw silk', also read *xn̥i̯o t'iwo*[bj] 'to season' (with palatalized initial, probably from an original **s-* causative form such as **s-na*). The prefixial type was retained before non-dental nasals as well, again with evidence supplied by Proto-Min; cf. PTB **s-mel* 'face' (Lushai *hmel*); Arch./Anc. *[s-]mi̯an mi̯än*[bk] (Proto-Min **hm-* = **sm-*), id.; WT *so-ma* 'hemp' (Jäschke cites Skr. *soma*), Arch./Anc. *[s-]ma/ma*[bl] (Proto-Min **hm-* = **sm-*), id. (an early loan into Chinese, as shown by the vocalism, since PST **-a* yields *-o* after labials); PTB **(s-)myək* 'eye' (see fn. 16; the PBL root is **(s-)myak*), Arch./Anc. *[s-]mi̯ôk/mi̯uk*[as] (Proto-Min **hm-* = **sm-*), id.; PK **hńa* < **s-ńa·[k]* 'flesh/meat', Arch./Anc. *[s-]ńi̯ôk/ńźi̯uk*[bm] (Proto-Min **hn-*, for **hń-* = **sń-*), id.; but note Arch./Anc. *s-ŋ/s*, paralleling *s-g/s* (below), as shown by GSR-67: *s-ŋo/suo*[bn] 'gather sheaves', with *ŋi̯o/ŋi̯wo*[bp] 'fish' as phonetic (the 'gap' then filled by a newly formed *[s-]ŋ/ŋ*, with corresponding [rare] Proto-Min **hŋ-* = **sŋ-*). The complex picture presented by these nasal combinations is tabulated below:

	m	*n*	*ń*	*ŋ*
**s (cluster)*	*x(i̯)m/x(i̯)w*	*ts/ts*	*xń/t''*	*xn/x*
[pattern fill]	—	*xn/t'*	—	—
**s- (prefixial)*	*[s-]m/m*	*[s-]n/n*	*[s-]ń/ń*	*s-ŋ/s*
	smi̯w/si̯w	*sn̥i̯/si̯*	*sń̥i̯/si̯*	*sŋi̯/śi̯*
[dial. variant]	—	*dz'/dz'*	—	—
[pattern fill]	—	—	—	*[s-]ŋ/ŋ*

As indicated above, the dental stop shifts (**st*, **sd* > *ts*, **sth* > *ts'*, **s-t*, **s-d* > *dz'*) appear to have been in large part completed in the pre-Archaic period, but there is scattered *xie-sheng* evidence of 'pattern filling' through newly formed *st*, *s-t* and the like; it also seems that dental-initial forms with medial *-w-* (whether primary [< PST **-w-*] or secondary) underwent a special development in Arch./Anc., accounting for such bizarre-appearing series as GSR-11 (*d'wâ/d'uâ*[bp] phonetic),

with the triplet form *t'wâ/t'uâ* ~ *stwia/swię* ~ *st'wia/xywię*[bq] 'shred (= chopped up/minced) meat'; cf. PTB **s-tw[â]* 'chop/mince' Lepcha *tyót* < *stot* 'chop, mince with numerous pieces as meat', from **s-t[wâ]-t*; WB *t'wa* (<**[s-t] wa*) 'mince with a knife' < PST **(s-)twâ*; also the enigmatic doublet under GSR-171 (*t'wân/t'uân*[br] phonetic): *t'i̯wad/tś'i̯wäi* ~ *st'i̯wăd/xi̯wvi*[bs] 'snout'; cf. also under GSR-576 (isolated form) *s[t']i̯wər/świ*[bt] 'water'; PTB **twəy*, id. < PST **(s-)twəy*. In the parallel form with voiced initial the 'regular' shift occurred in at least one basic root; cf. PTB **s-d[w]â-t* 'sit' (WT *sdod-pa*, pf. *bsdad*), Arch./Anc. *dz'wâ/dz'uâ*[bu], id. < PST **s-dwâ*.

The comparative evidence relating to labial stops shifts after **s* (cluster or prefixial) is rather meagre, and it is difficult to determine whether the apparent voicing changes (paralleling those with dental stops) had already been completed in the pre-Archaic period, the *xie-sheng* being of little value here, e.g., PTB **(s-)bəy* 'give' (WT *sbyinpa* 'give; gift', with partially lapsed function of the nominalizing *-n* suffix), Arch./Anc. *[sb]i̯əd/pji*[bv], id. < PST **(s-)bəy*; PTB **ba·k* 'bat (animal)', Arch./Anc. *[sb]i̯uk/pi̯uk*[bw], id. < PST **(s-)ba·k* (with PST 'animal prefix' **s-*), contrasting with PTB **p(w)a* 'father', PK **ba* ~ **pha* (Taungthu), id., from **(s-)p(w)a* (regular loss of medial **-w-*), Arch./Anc. *[s-p']i̯wo/b'i̯u*[bx], id. < PST **(s-)p(w)a*; PTB **pw[â]* 'grandmother' (WB *ăphwà* ~ *ăbhwà*), Arch./Anc. *[s-p']wâ/b'uâ*[by] 'old woman' (>'grandmother') < PST **(s-)pwâ* (see above for prefixed **s-* with kinship terms); see also below for (secondary) *s-b/ṣ*, paralleling *s-bl/ṣ*. As in the dental stop series, medial *-w-* appears to have produced a special shift: **sphw* > *xw*, paralleling **sthw* > *xw* but completed in the pre-Archaic period; cf. especially PTB **pwâr* ~ **(s-)bwâr* 'burn; fire', Arch./Anc. *xwâr/xuâ*[bz] 'fire', from **sphwâr* < PST **(s-)pwâr* ~ **(s-)bwâr*.

In contrast to the situation as regards dental and probably also labial stops, the **s* element with velar stops was maintained well into the Archaic period, giving rise to two distinct series of reflexes, one for cluster **s* (shift to dental/palatal stop) and the other for prefixial **s-* (simple loss of stop or replacement by ʔ or *x*). The evidence here, which is unusually extensive in all respects, comes both from the *xie-sheng* and from comparative TB material. There is also very intriguing confirmation from Proto-Min, which (unlike Arch./Anc.) maintained a distinction between aspirated voiced stops (**bh-*, **dh-*, **gh-*) from PST initial voiced stops, and unaspirated stops (**b-*, **d-*, **g-*) from PST prefixed voiced stops, paralleling a similar distinction in the corresponding unvoiced stops. Norman (1973) has reconstructed additional 'softened' stops (written **-b*, **-p*, etc.) for Proto-Min primarily on the basis of special reflexes (continuants or zero) in the northwestern dialects of Kienyang and Kienow, with the suggestion that prefixation of some kind was the underlying factor in the development. The same dialects also exhibit two different reflexes (*k-* and *x-*) for Proto-Min **k-* as reconstructed by Norman, who states simply, 'The origin of this split is unclear'. If we postulate prefixial **s-* to account for the Proto-Min features, however, we will at the same time have an excellent reconstruction to explain the above split, with a remarkable parallelism

with the Arch./Anc. development (*x* for *x* reflex, and ø for ʔ reflex), as shown in the following table ('prefix' = 'other than **s-* or [below] **a-*):

	Arch./Anc.	*Proto-Min*	*Foochow*	*Kienyang*	*Kienow*
**k = *kh-*	*k'/k'*	**kh*	*k'*	*k'*	*k'*
*prefix + *k-*	*k/k*	**k*	*k*	*k*	*k*
**s-k*	*s-k/ʔ*	**s-k*	*k*	*k/ø*	*ø*
**s-kh*	*s-k'/x*	**s-kh*	*k*	*x*	*x*
**sk-*	*sk/t*	**t*	*t*	*t*	*t*
**skh-*	*sk'/t'*	**th*	*t'*	*h*	*t'*
**g = *gh-*	*g'/g'*	**gh*	*k'*	*k'*	*k'*
*prefix + *g-*	*g'/g'*	**g*	*k*	*k*	*k*
**s-g*	*s-g/s*	**s*	*s*	*s*	*s*
**sg = *sgh-*	*sg'/d'*	**dh*	*t'*	*h*	*t'*

The comparative material for Proto-Min prefixed **s-* before stops is fairly limited, especially for labials and dentals, but does tend to support the suggested reconstruction:

**s-b*[*oʔ*] 'thin': Arch./Anc. *b'âk/b'âk*[ca]; cf. PTB **s-b*[*ə*]*k*: WT *sbeg-pa* 'lank, thin'; PST **s-bâk* ~ **s-bək* (for the doublet, cf. 'cough' in *STC*: fn. 482).

**s-b*[*iŋ*] 'vase'; Arch./Anc. *b'ieŋ/b'ieŋ*[cb]; cf. WB *phyàṅ* (< **phyìŋ* < **[s-b]iŋ*) 'large open-mouthed pot'; PST **s-bi·ŋ*.

**s-p*[*uí*] 'to fly'; Arch./Anc. *pi̯wər/pywe̯i*[st], from PST **(s-)pur* (see above for doublet in *-n*), with additional evidence for prefixed **s-* in this word from early texts (Yang 1975) but with 'intensive' rather than the 'causative' function seen in WT (*'p'ur-ba* 'fly', *spur-ba* 'make fly').

**s-d*[*uŋ*] 'bronze'; Arch./Anc. *d'uŋ/d'uŋ*[cc]; Cf. WT *sdoṅ-ba* 'unite, join' (> 'alloy'); PST **s-do·ŋ* (see above for the vocalism), or perhaps from an earlier **(s-)luŋ* < **(s-)lo·ŋ* < **(s-)looŋ* < *(s-)lu*[*y*]*aŋ* as suggested by the AT evidence (Benedict 1975: *Introduction to Glossary*).

**s-k*[*ow*] 'dog'; Arch./Anc. *ku/ku̯*[m]; PST **(s-) kwəy* (with 'animal prefix' **s-*).

**s-k*[*oŋ*] 'jar'; Arch./Anc. **kŭŋ/kɔŋ*[cd] 'earthen jar' (not in GSR), a doublet of *s-kuŋ/ʔuŋ*[ce] 'jar', [*s-k*]*un/ʔuŋ* ~ [*s-k*]*i̯uŋ/ʔi̯woŋ*[cf] 'pitcher, jar'; cf. PK **kɔŋ* (< **-koŋ*) 'earthen jar' (possible loan from Ch.); PST **[s-]ko·ŋ* (with vocalism as in 'bronze', above).

**s-kh*[*âŋ*] (< **s-kh*[*ân*]) 'liver'; Arch./Anc. *kân/kân*[c]; cf. PTB **ka* 'bitter' (Bodo-Garo **b-ka* 'liver'); PST **(s-)ka*(*-n*) (for vocalism, see *STC*: fn. 488) with 'body part prefix' **s-*).

**s-kh*[*iu*] 'type of leek'; Arch./Anc. *ki̯ôg/ki̯u̯*[cg] 'Allium, leek, onion'; cf. PTB **s-k*[*a, a·*]*w*: WT *sko-tse* 'wild onion', Kanauri *go-ze* 'wild garlic' (app. loan from Tibetan); Kachin *săkau* (free form) ~ *gau* (comb. form) 'onion, leeks',

Maru *khau* ~ *hau* (< **s-khau*) 'onion' (app. loan from Kachin); PST **s-kâw* (see above for vocalism).[20]

The evidence for **s* (cluster or prefixial) with velar stops comes primarily from the *xie-sheng* but is supported also by TB comparisons. The prefixial type presents much less of a problem, e.g., GSR-122 has *k'i̯u/k'i̯u* ~ *s-ku/ʔu̯*[ch] 'conceal' as phonetic, with the likely cognate WB *khui* (< **[s-k]ui* < **[s-k]əw*) 'take shelter, refuge' (= 'conceal oneself') < PST **s-kə·w*; note that this series includes *k'i̯u/k'i̯u*[ci] 'body, person', *s-ki̯u/ʔi̯u*[cj] 'to warm the body (as a mother a child)', with 'causative' **s-* prefix contrasting with the 'body part prefix' **s-* found in the cognate WT *sku* 'body' (WB *kuiy* = *kui* 'body' < **[-]kəw*), from PST **(s-)kə·w*. The numerous Arch./Anc. series with initial ʔ (GSR writes ·) are for the most part to be reconstructed with initial **s-k*; cf. the following trio showing 'animal prefix' **s-* (Arch. forms bracketed when the series contains no initial velars):

[*s-k*]*i̯əŋ/ʔi̯əŋ*[ck] 'eagle, falcon'; cf. WT *skyiṅ-ser* 'eagle, vulture' (*ser* for *gser* 'gold'); PST **s-kiŋ* (contra *STC*: fn. 225; for the vocalism, see *STC*: fn. 476).

[*s-k*]*i̯wăn/ʔ-i̯wɒn*[cl] 'Mandarin duck'; cf. WT *skyarpo* 'snipe, wood-cock', *skyar-mo* 'heron', *tśhu-skyar* 'duck, bittern' (*tśhu* 'water'); PST **s-ky*[*w*] *a·r*; cf. also the apparent doublet: *kwân/kuân*[cm] 'heron' < **[-]kwa·r*.

[*s-k*]*ian/ʔien*[cn] 'lizard'; cf. WT *skyin-gor*, id. (not analyzed); PST **s-ke·n* ~ **s-ki*[·]*n*.

The **s-k* reconstructions obtained in this manner are often confirmable through doublet forms in other series (cf. 'jar', above), e.g., GSR-253 (the 'lizard' series, above), with phonetic [*s-k*]*an/ʔan*[co] 'tranquillize', includes [*s-k*]*at/ʔat*[cp] 'pull up' (< **[s-k]an* + suffix with obstruent initial), a doublet of *ki̯an/ki̯än* ~ *ki̯ăn/ki̯ɒn*[cq] (tone B) 'pull up' and *ki̯an/ki̯än* ~ *ki̯ăn/ki̯ɒn*[cr] (tone B) 'lift' (etym. same word) as well as *s-k'i̯ăn/xi̯ɒn*[cs] (tone A) 'lift' (with the anticipated tone B > A shift after *s-*); the same series (GSR-253) also includes [*s-k*]*ian/ʔien*[ct] (tone C) 'to rest, repose' (doublet of the phonetic, above), the same word as [*sk*]*ian/ʔien*[cu] (tone C) (loan) 'to rest, at ease', the phonetic of GSR-243 (see below for 'to swallow' from this series), a doublet of [*s-k*]*ân/ʔân*[cv] 'peace, peaceful, tranquil', the phonetic of GSR-146, which includes both [*s-k*]*an/ʔan* ~ [*s-k*]*ân/ʔân*[cw] (tone C) 'late', cognate to *kân/kân*[cx] (tone C) 'sunset, dark, evening', and [*s-k*]*ât/ʔât*[cy] 'root [= 'saddle'] of the nose' (< **[s-k]ân*; cf. 'pull up', above) and **[s-k]ân/ʔân*[cz] 'saddle' (not in GSR); cf. Lushai *kɔn* (< **[-]kâ·n*) 'shallow valley, saddle of a hill', *hna·r-kɔn* 'the indentation [= 'saddle'] above the nose [*hna·r*]'; PST **(s-)kâ·n*.

As illustrated in the concatenation above, the relationships often are of a complex nature and frequently include significant links between prefixial and cluster types (as in 'swallow'); cf. *kuŋ/kuŋ*[da] 'father' (> 'grandfather'), phonetic in two cognate words: **s-kuŋ/ʔuŋ*[db] 'old man, sir; father' (> 'grandfather') (not in GSR) and *skyi̯uŋ/tśi̯woŋ*[dc] 'father-in-law' (< 'grandfather' through teknonymy, a feature of Chinese kinship terminology) (listed separately as GSR-1189); cf. PTB *k*[*a,a·*]*ŋ* 'father, grandfather'; PST **(s-)ka·ŋ*. The basic phonetic series here (GSR-1173) also includes *s-kuŋ/uŋ*[ce] 'jar' (above), tying in with GSR-1184, which includes

[*s-k*]*uŋ/ʔuŋ* ~ [*s-k*]*i̯uŋ/ʔi̯-woŋ*[cf] 'pitcher, jar' (above) as well as [*s-k*]*i̯uŋ/ʔi̯-woŋ*[dd] 'to cover, conceal'; cf. PTB **s-k[u,u·]ŋ*: WT *skuŋ-ba* 'to hide in the ground; to bury, inter; PST **s-ku·ŋ*. This phonetic series (GSR-1173) should be expanded to include not only GSR-1189 ('father-in-law', a bove) but also GSR-1190 (as recognized by Li 1971), which includes *s-gi̯ôŋ/si̯uŋ*[de] (tone A) 'high', the same word as [*s-g*]*i̯ôŋ/si̯uŋ*[df] (tone A). id. (listed separately as GSR-1012) (both graphs have 'mountain' as signific); cf. PTB **s-g[a,a·]ŋ*: WT *sgaŋ* 'projecting hill or spur on side of mountain'; WB *khaŋ* (<*[*s-g*]*aŋ*) 'knoll, rising ground' (obs.), *khaŋ-ruì* 'strip of high ground, spur of mountain' (*-ruì* app. for (*ă-*)*ruì* 'bone') PST **s-ga·ŋ* (for the vocalism, see *STC*: fn. 488). The recognition of **s* combinations clarifies the phonetic role of certain elements in many characters; cf. the following:

sgaŋ/i̯aŋ[dg] 'sheep' (GSR-732), phonetic in *ki̯aŋ/ki̯aŋ*[dh] 'family name' (GSR-711) and *k'i̯aŋ/k'i̯aŋ*[di] 'Western tribe' (GSR-712), also in *kăŋ/kʌŋ*[dj] (tone A) 'soup' (GSR-747; the graph is said to include a 'lamb'); the last is cognate to *s-k'i̯aŋ/xi̯aŋ*[dk] (tone A) 'beef soup' as well as [*sk'*]*âŋ/t'âŋ*[dl] (tone A) 'hot liquid' (note tone A throughout; see below for the last two series).

[*s-g*]*i̯og/si̯ău*[dm] 'laugh' (GSR-1150), to be added to GSR-1141: [*s-k*]*i̯og/ʔi̯äu*[dn] 'bend' phonetic.

s-gɔ/ṣa[do] 'breaking voice' (GSR-1236c), to be added to GSR-36: *g'ɔ/ɣa*[dp] 'summer' phonetic.

skyi̯ər/tśi[dq] 'fine cloth' (GSR-1237k), to be added to GSR-549: *s-k'i̯ər/xi̯ei*[dr] 'thin, rare' phonetic.

sk'əm/t'i̯âm[ds] '*covet*' (GSR-645), to be added to GSR-651: *ki̯əm/ki̯əm*[dt] 'now' phonetic; cf. WT *skam-pa* ~ *rkam-pa* 'to long for; longing', app. a doublet ('to thirst for') of *skyem-pa* 'to be thirsty', *skyems* 'thirst' (resp.), from PTB **s-kəm* (as shown by the vocalic alternation; see *STC*: fn. 344); PST **s-kəm* (also cited in part in Yang 1975).

The reconstruction of the cluster type of **s* presents no difficulties when the phonetic series has only a single entry of this kind, especially if a TB comparison is available; e.g., *sk'i̯ĕŋ/t'i̯ĕŋ*[du] 'red', in an otherwise velar-initial series (GSR-831: *ki̯ĕŋ/ki̯ĕŋ*[dv] phonetic); cf. PTB **s-kyeŋ*, id. PST **s-kyeŋ*; also *sgyi̯ĕn/źi̯ĕn*[dw] 'kidney', in an otherwise verlar-initial series (GSR-368: *k'i̯ĕn/k'i̯ĕn*[dx] phonetic); cf. PTB **s-gal*: WT *sgal-pa* 'small of the back', Chepang *gal*, Lushei *kal* 'kidney'; PST **s-gal*. In some instances only a TB comparison is available; cf. *sgyi̯əp/źi̯əp*[dy] 'ten', phonetic in a series (GSR-686) with only one other palatal-initial form; cf. PTB **gip*, id.; PST **(s-)gip*. An intermediate type of case is illustrated by GSR-496, which has *sky'i̯wət/tś'i̯uĕt*[ak] 'go out, come out, bring out', also read *sky'i̯wəd/tś'wi* 'bring out, take out'; PTB **ku(-s)* 'take up or out, lift up' (*STC*: fn. 281); PST **(s-)ku(-s)*, as phonetic in a series evenly divided between dental- or palatal-initial forms and velar-initial forms; cf. also GSR-672: *g'ɛm/yăm*[dz] 'small pit' phonetic, which includes five velar-initial as opposed to nine dental- or palatal-initial entries, yet the initial velar in this series is 'proved' by the doublet *sg'əm/d'i̯âm*[ea] (tone B) 'lotus flower', *g'əm/g'i̯âm*[eb] (tone B), id. (without the anticipated tone shift); also GSR-758: *sgyi̯əm/źi̯əm*[ec] phonetic, with only three

velar-initial as opposed to ten dental- or palatal-initial entries, yet the initial velar is well supported by the comparative TB data (*STC*: fn. 464). Finally, certain series entirely lacking velar-initial entries are to be reconstructed with **s* cluster initials, e.g. GSR-686 ('ten', above) and cf. the following:

[*sk'*]*ien/t'ien*[ed] 'heaven' (tone A), a doublet of **[s-k']ien/xien*[ee] (tone A), id. (poss. phonetic; not in GSR); cf. PTB **(m-)ka(-n)*: WT *wkha* 'heaven', *nam-mkha* 'heaven, sky'; Magari *nam-khan* 'sun', Garo *mikka* = *mka* 'rain; (comp.) sky' (with 'collective' plural **-n*; see *STC*: fn. 428); PST **(s-, m-)ka(-n)*; this word ('heaven') is phonetic in a series (GSR-361) which includes only [*sk'*]*ən/t'ən* ~ [*sk'*]*ien/t'ien*[ef] (tone A) 'to swallow', a doublet of [*s-k*]*ien/ʔien*[eg] (tone A) 'gullet' (= 'the swallower'), (tone C) 'to swallow' (from earlier suffixed form) (see also 'smoke', above, from this GSR-370 series); cf. PTB **(s-)kyəw(-t)*: WT *skyugaŋ* 'a gulp, draught', *skyud-pa* 'to swallow'; WB *kyuik-khyá* (< **kyui-khyá*) 'to swallow, gulp'; PST **(s-)kyəw(-n, -t)* (see *STC*: 98–103 for these dental suffixes).

[*sg*]*aŋ/i̯aŋ*[eh] (tone A) 'South side' is phonetic in a series (GSR-720) which includes [*sg*]*aŋ/i̯aŋ*[ei] (tone A) (etym. same word) 'South side of a slope; North bank of a river; light; the light cosmogonic principle [*yang*]', cognate to *s-k'i̯aŋ/xi̯aŋ*[ej] (tone C) 'to face, turn towards; drawn towards; (face side:) window; (the facing side:) South; (the time 'turning towards' the speaker:) a little while ago' (from earlier suffixed form) and *s-k'i̯aŋ/xi̯aŋ*[ek] (tone C) (etym. same word) 'window facing North; turn towards'; cf. Garo *skaŋ* 'before' [both spatial and temporal senses] < PST **s-kaŋ*; this series (GSR-720) also includes [*sk'*]*âŋ/t'âŋ*[el] 'hot liquid' (see above under 'soup'; also [*sg*]*aŋ/i̯aŋ*[em] 'to roast, heat'; cf. PTB **(-)gaŋ* (*STC* cites **kaŋ*) 'to roast, toast, broil, fry' < PST **(s-)gaŋ*; also [*sg*]*aŋ*[en] (tone A) 'lift, raise; winnow', etc., a doublet of *s-k'i̯əŋ/xi̯əŋ*[eo] (tone A) 'lift, raise; rise', etc. and also used as loan for 'hawk', a doublet of [*s-k*]*i̯əŋ/ʔi̯əŋ*[ck] 'eagle, falcon' (above)[21], and as loan for 'metal frontlet of a horse', a doublet of [*s-k*]*i̯aŋ/ʔi̯aŋ*[ep] 'strap on breast of horse or ox' (establishing [*s-k*] for this series: GSR-718) and of [*sk*]*i̯əŋ/ʔi̯əŋ*[ev] 'breast, breast-plate', etc. (from the 'eagle' series: GSR-890, above); the GSR-718 series also includes [*s-k*]*âŋ/ʔâŋ*[ew] (loan) 'full, ample'; cf. PTB **(s-)gâŋ* ~ *(s-)kâŋ*: WT *gaŋ(-ba, -po)* 'full', *sgaŋ-ba* 'to grow or become full', *skaŋ-pa* ~ *skoŋ-pa*, pf. *bskaŋs* 'fulfill'; Kachin *koŋ* 'to be full, extended [udder with milk]; to be full, well-developed [fruit]' WB *kàŋ* 'to form [= 'become full'], as fruit in earliest stage' < PST **(s-)kâŋ* (WT also has medial *-e-* doublets, from earlier medial **-ə-* forms).

[*sg*]*u/i̯u*[ex] 'agree' is phonetic in a series (GSR-125) which includes three entries for which initial velars are to be reconstructed:

[*sk'*]*u/t'i̯u*[ey] (tone A) 'steal' (with the regular tone B•> A shift after *s-*), a doublet of *k'u/k'i̯u*[ez] (tone C) 'rob' (from earlier suffixed form) as well as *k'u/k'i̯u*[fa] (tone B) 'strike, lay hold of' and *s-ku/ʔəu*[fb] (tone B) 'beat' (from the 'body' series: GSR-122, above); cf. PTB **r-kə·w*[b] 'steal' < PST **(s-, r-)kə·w*[B].

[*sk'*]*u/t'i̯u*[fc] (tone A) 'to slight, despise' (tone B > A shift after *s-*), a doublet of *s-k'u/xəu*[fd] (tone B ~ C) 'disgrace, insult' as well as *s-k'u/xi̯u* ~ *k'u/k'əu*

(both tone C) ~ *ku/kn̥u*[fe] (tone B) 'revile, disgrace' and *s-k'u/xn̥u ~ g'u/ɣəu*[ff] (both tone C, from earlier suffixed forms, the latter with second. voicing; see fn. 21) 'revile, insult, disgrace'; cf. PTB **a-k[u, əw]*: WT *'khu-ba* 'to offend, insult' (cited also in Bodman 1974) < PST **(s-,a-)k[u,ə·w]*.[22]

[*sg*]*u/i̯u*[fi] 'sheep'; cf. PTB **s-g[u, əw]*: Kachin *săgu* 'sheep' < PST **s-g[u, ə;w]*.

PTB has prefixed **a-* = *ʔa-* (non-phonemic vocalic onset for initial wovels) as a morphological feature of considerable prominence, the unstressed form (*ʔə-*) frequently yielding *ʔ*, notably in WT (*a-chung*; see *STC*; fn. 339), WB (*auk-myit* = glottal accent; see *STC*; fn. 260), Kachin and Garo (*STC*: fn. 78).[23] Karen shows this prefix in vestigial form (*STC*: 130 and fn's 352, 354) but Chinese superficially appears to lack this element, apart from a late but possibly related stressed form found with kinship terms and kin numeratives (*STC*: 156 and fn's 422–24). Scholars have long called attention to the frequent alternation of nasals and stops in the language, both as initials and as finals (see *STC*: 156–57 for the latter). A closer study of the *xie-sheng* reveals that the overwhelming majority of instances of interchange in final position involved replacement of nasal by stop (voiced or unvoiced), apparently through assimilation to an earlier suffix, indicating that most suffixes in the language had an initial obstruent (cf. forms for 'pull up' and 'root of nose', above). In initial position, however, the common replacement was that of stop by nasal, and here it seems that we must posit a series of preglottalized stops for the Archaic stage, thus bringing it very much into alignment with TB languages, notably WT and Ka̱chin. A parallel development has taken place in Tibetan, the Khams dialect having prenasalized obstruents as reflexes for the preglottalized (*a-chung*) obstruents (both voiced and unvoiced) of WT. The further step to nasal is found in Vietnamese and Shan, which have initial *m-* and *n-* for the earlier **ʔb-* and **ʔd-* of Mon-Khmer and Tai, respectively. None of this, of course, means that anything of the kind occurred in Arch./Anch., yet a considerable body of evidence stands in favor of the hypothesis. We cannot be certain whether, as in Tibetan, both voiced and unvoiced stops were subject to preglottalization but the evidence as a whole favors this view (we write *ʔt/n* or *ʔd/n*, etc. for Arch./Anc. where cognate forms are available, otherwise [*ʔt,d*]*/n*, etc.). The *xie-sheng* themselves are highly informative, e.g., *ʔgiog/ŋieu*[fj] (tone A) 'high', a doublet of *g'i̯og/g'i̯äu*[fk] (tone A), id., is phonetic in a long series (GSR-1164) with more stop-initial than nasal-initial entries; note particularly the doublet ('triplet') *ʔdi̯an/ńi̯än*[fl] (tone B) 'trample', *ʔdi̯an/ńi̯än ~ ʔdiən/nien ~ d'iən/d'ien*[fm] (all tone B) 'tread, trample' (note the same tone throughout), written not once but *twice* with characters from series (GSR-201 and GSR-453) which otherwise have only stop-initial entries! The interpretation of certain series presents difficulties, notably GSR-152 with *t'ân/t'ân*[fn] 'sigh' as phonetic but with entries almost exclusively with nasal initials (*n-* or *ń-*); a close study of this series shows conclusively that the stop initial is basic and that the nasal initials were all derived from preglottalized stops, making this a distinctively 'preglottalized' series, just as other series were characterized by initial *s-* (cluster or prefixial; see above); cf. the following cognate groups, with ties especially to GSR-147, another series with original final *-r* (> *-n* in most entries):

(GSR-152) *t'ân/t'ân*[fn] (above) 'sigh' (phonetic in this series; the graph has 'mouth' and 'distress'); *ʔtân/nân*[fo] (tone A) 'difficult', (tone C) 'difficulty, calamity' (this character used as loan for 'ample' and 'expel' below, both with final *-r*); (GSR-147) *t'ân/t'ân*[fp] 'eshausted, fagged out'; *tân/tân*[fq] 'exhaust'; *tân/tân ~ târ/tâ*[fr] 'disease, suffering, distress'; (GSR-148) *tân/tân*[fs] 'suffering, distress'; cf. WT (medical texts) *thor-pa ~ ʔthor-pa* 'smallpox' (*dmar-thor* 'measles'), *thor-bu* 'a class of diseases, comprising dyepepsy and cutaneous disorders' < PST **(a-)târ*.

(GSR-152) *ʔtâr/nâ*[ft] 'rich, ample, fine'; GSR-147) *tân/tân*[fu] 'ample' (this character also glossed as 'single, simple', see below, and as 'exhaust', see above); (GSR-350) [*ʔt*]*âr/nâ*[fv] 'much, ample; fine, beautiful'; cf. WT *thormo* 'the growing fat [= 'ample/rich/fine'] of cows, goats, in consequence of sterility' < PST **(a-)târ*; for the final, cf. *tân/tân*[fu] (above) 'single, simple'; WT *thorbu* 'single, separate' < PST **(-)târ*.

(GSR-152) *ʔtâr/nâ*[ft] (above; add. gloss) 'expel demons and noxious influences'; cf. WT *gtor-ma* 'strewing-obliation', an offering brought to malignant demons (whence the Lepcha loan: *tor* 'an offering and ceremony of casting out evil spirits'), also the related *'t'or-ba* 'to be scattered, dispersed' *'dor-ba* 'throw or cast away; esp. to throw out, eject' < PST **(a-)târ ~ *(a-)dâr*.

(GSR-152) *ʔdiər/niei* (< **ʔdi̯âr*) ~ *ʔdâr/nâ*[fw] 'pickled meat with bones in it'; cf. WT *sdor ~ rdor* 'that which gives relish to food, seasoning, condiment, esp. *t'ug-sdor* that which gives substance to soup [*t'ug*], viz. meat; *ts'a-sdor* 'salt [*ts'a*] and meat; spice' < PST **(s-,a-)dâr*.[24]

(GSR-152) *ʔt'i̯an/ńźi̯än*[fy] (tone A) 'burn'; (GSR-217) [*ʔt'*]*i̯an/ńźi̯än*[fz] (tone A,) id. (etym. same word); (GSR-147) *t'i̯an/tśi̯än*[ga] (both tone B) 'to heat; to flame'; cf. Lepcha *tor* (< **tar*) 'to be burned, as house' < PST **(a-)tar*.

(GSR-152) *ʔt'i̯an/ńźi̯än*[fo] (above; add. reading) 'respectful' [= 'fearful']; *ʔtan/nan ~ ʔt'i̯an/ńźi̯än*[gb] 'to fear'; (GSR-147) *t'i̯an/tśi̯än*[gc] 'battle, to fight; to fear' < **(a-)tar* (no known TB cognate; contra *STC*: fn. 461).

The case for reconstructing preglottalized stops for the Archaic stage is greatly strengthened when TB comparisons are available, as in four of the above five roots; cf. also the sizeable GSR-1031 series, which contains only one nasal-initial entry, viz. *ʔtiõk/niek*[gd] 'hungry for, desirous; hungrily, dissatisfied'; cf. WT *ltogs-pa* 'hunger; hungry' < PST **(-)t[o]k*; also the very long GSR-139 series (27 entries), which contains only three nasal-initial entries, including *ʔkân/ŋân*[ge] 'wild dog'; cf. Kachin *tśyăkhyon* 'fox, wolf or wild dog', from **-khon* (second. palatalized by the prefix) < PST **(-)kâ·n*; also GSR-405, with mainly stop-initial entries but with one group of nasal-initial forms, including *ʔpi̯ĕt/mi̯ĕt*[gf] 'silent, quiet; secret; near, close', a doublet of *pi̯ĕd/pyi*[gg] (< **pi̯ĕt* + suffix) 'secret' and *pi̯ĕd/pyi*[gh] 'to shut, close' (etym. same word); cf. WB *pit* 'shut up, close, stop up' < PST **(-)pit*; including also *ʔpi̯ĕt/mi̯ĕt*[gi] 'to wipe a vessel clean'; cf. WT *'phyi-ba ~ 'phyid-pa* 'to wipe, blot out', *phyi-bdar* 'clean, cleanse; (comp.) wash [face]' (*bdar* 'rub'), *byi-dor* 'the wiping, cleaning' (an app. second. form); Lushai *phiʔ* (< **phis* < **phi-s*) 'to wash (the face)' < PST **(a-)pi(-s,-t)*.

Another possible source of confirmation is supplied by Chinese-Tai loanwords, in either direction, since PT has roots with initial *ʔ*b-* and *ʔ*d-*, including one significant early loan into Chinese ('piebald/variegated', below). PT also has nasal-initial roots on 'high' tones (< surd initials), conventionally reconstructed with initial **hm-*, **hn-*, etc., but the closely related Kam-Sui languages show two distinct reflexes here, indicating that some of these roots were preglottalized rather than preaspirated (PT or pre-PT level): PT **hma*, Sui *hma* 'dog' but PT **hm̓aay*, Sui *ʔme* 'mark'; PT **hnaaw*, Sui *hno* 'cold' but PT **hna*, Sui *ʔna* 'thick'. This is of special significance in connection with the following comparison, involving another root of 'cultural' type (see 'expel demons', above), this with Kachin as well as Tibetan cognates:

[*ʔb*]*i̯wo/mi̯u*[gj] (tone A) 'magician', from **ʔba* (regular vocalic shift), in a series (GSR-105) which contains only [*ʔb*]*i̯wo/mi̯u*[gk] (tone A) 'deceive' (etym. same word); cf. WT *ʔba-po* 'magician, sorcerer, conjurer'; Kachin *ba* ~ *dźăba* ~ *tśyăba* 'to divine', *săba* 'divination, augury' < PST **(a-)ba*.

For the above root PT has the early loan **hmo* (tone A) 'sorcery/sorcerer/magician' (note the tonal agreement), for **ʔmo* (see above; Kam-Sui cognates app. lacking), *not* from Arch.-LPT (which retains final *-a* < PST **-a*) but from the Archaic 'lineage' itself, indicating a development *ʔbi̯wo* > **ʔmi̯wo* (> PT **ʔmo*) > *mi̯u* rather than *ʔbi̯wo* > **ʔbi̯u* > *mi̯u*. Proto-Min presents a parallel situation inasmuch as the nasal-initial roots on 'high' tones also appear to have been derived in part from preglottalized forms; note especially the three velar roots (initial **hŋ- = = *sŋ*), two of which correspond to Arch./Anc. forms with initial **[k,g]/ŋ-*, viz. **[ʔk,g] ian/ŋien*[gl] 'ink-stone' (not in GSR), with phonetic *kian/kien*[gm] 'see' (phonetic in the all stop-initial GSR-241 series as well as in GSR-244: [*sk'*]*ian/t'ien*[gp] 'face'); **[ʔk,g]lăk/ŋɒk*[go] 'forehead' (not in GSR), with phonetic *klâk/kâk*[gp] 'each' (phonetic in the lengthy, almost exclusively stop-initial GSR-766 series).

The possibility of reconstructing still other (non-obstruent) preglottalized consonants for the Archaic language must be considered. PTB/PST prefixed **a-*, the ultimate source of these preglottalized forms, seems to have occurred freely before all classes of consonants, on the basis of comparative TB evidence (*STC*: 121–23), and Tibetan itself presents clear evidence for at least some of these classes, e.g., **ʔś-* < *'tś'-* (*STC*: fn. 90), **ʔr-* < *'dr-* and **ʔl-* > *ld-* (*STC*: fn. 338), also **ʔly-* > *'dź-* (*STC*: fn. 104). It is possible that the frequent alternation between initial sibilants and affricates in Chinese (cf. *STC*: fn. 455) reflects preglottalization in a manner parallel to that of Tibetan (above), e.g., *ʔsōg/tṣau*[gq] 'claw, nail', phonetic in GSR-1112, which includes the cognate *sôg/sâu*[gr] 'scratch', but a reconstruction of this kind must remain highly speculative. Similarly, there is extensive *xie-sheng* (and some comparative) evidence for affricates (*ts*, *ts'*, *dz'* and *dz* > *z*) as alternative reflexes of **s* (cluster or prefixial) combinations with velar stops, accounting for certain of the 'odd' GSR series, e.g., GSR-280, with [*s-k*]*ăt/ʔăt*[gs] 'crush under the wheels' and [*sk*]*ăt/tṣăt*[gt] 'a slip, a strip', the latter for the anticipated **[sk]ăt/tăt*; one possible reconstruction here is **[ʔsk]ât/tṣăt*, paralleling the

above example from GSR-1112, but again there appears to be no possibility of confirmation. In addition, there is a modicum of *xie-sheng* evidence suggesting that preglottalized nasals, perhaps along with *ʔl-* (see below), occurred marginally at the Archaic level, giving rise to homorganic surd stops, with velar shift to glottal stop also a possibility; cf. *ŋɔ/ŋa*[gu] 'tooth', phonetic in *ʔŋɔ/ʔa*[gv] 'raven, crow'. These shifts would necessarily have had to precede the preglottalized stop > nasal shifts, e.g., *ʔm-* > *p(')-*, followed by *ʔb-*, *ʔp-* > *ʔm-* > *m-* (for this development, see 'magician', above); cf. *mi̯wo/mi̯u*[gw] 'not, no', from **ma* (regular vowel shift); PTB **ma*, id. < PST **ma* (establishing the initial **m-* as original), phonetic in GSR-103, with exclusively initial *m-* or *xm-* entries apart from *ʔmi̯wo/p'i̯wo*[gx] 'lay the hands on', etc.; contrast the following GSR-104 series, with [*ʔp,b*]*i̯wo/mi̯u*[gy] 'martial' as phonetic, containing only *ʔbi̯wo/mi̯u*[gz] 'parrot', from **ʔbya*, app. cognate to PTB **bya* 'bird' < PST **(a-)bya*, and *pi̯wo/pi̯u*[ha] 'tax', from **pya*; cf. WT *dpya* 'tax, duty, tribute' < PST **[d-]pya*.

Both prefixial (rarely cluster) *s-* and *ʔ* are to be reconstructed before Archaic stop + *l* clusters, with the same sound shifts as noted before simple stops, with evidence of frequent doublet formation, e.g., *s-*[*gy*]*i̯ər/śi*[hb] 'excrement', also read as *s-*[*k'*]*i̯ər/xyi* 'groan', pointing to an old doublet (cf. Bodman 1973); cf. PTB **(s-)klay*, id. < PST **(s-)kləy* (with 'body part prefix' **s-*); the first form shows secondary voicing (see fn. 21), found also in *s-*[*gy*]*i̯ôg/śi̯u*[hc] 'head'; cf. PTB **s-kra* 'head hair' (also 'head' in Tamang-Gurung-Thakali [Nepal]; see *STC*: fn. 487 for vocalism) < PST **s-kra*. In both these examples, as in many others, the *xie-sheng* series involved is very brief or is otherwise ambiguous as regards reconstruction of medial *-l-* (< PST **-r-*, **-l-*), but it appears that by the Archaic period this medial element had already been replaced (**-ly-* > *-i̯-*) in many roots through palatalization, a characteristic feature of this language; contrast *s-gli̯ug/śi̯u*[av] 'number' (above) and *s-k'lŭk/xɒk* ~ *s-k'luk/xuk*[hd] 'vomit'; cf. WT *skyug-pa* (< **sklug-*), id. < PST **s-klu·k* (also cited in Bodman 1974), with *s-k'i̯əm/xi̯əm*[he] 'set forth, display'; cf. WT *'grem(s)-pa*, pf. *bkram* 'to put or lay in order; to spread out, display' < PST **(s-,a-)krəm*; also *s-*[*k'l*]*o/xuo*[hf] 'tiger' (isolated under GSR-57), an early loan from Mon-Khmer (*khla* ~ *kla* forms) via PBL **(k-)la* (*STC*: fn. 83) with 'animal prefix' **s-*, phonetic in [*sk'y*]*i̯o/tś'iwo*[hg] 'dwell, stay; keep still; to place; a place' (a doublet of *ki̯o/ki̯wo*[hh] 'sit down; dwell; settlement; repose'), with a basic doublet loan **s-klâ* (with final *-â* rather than the regular *-o* for PBL **-a*) reflected in the use of the same graph as phonetic in [*s-k'*]*ia/xji̯e*[hi] (< **[s-k']i̯ât*) 'earthenware vessel', from **s-k'yâ* < **s-k'lâ*. It appears, however, that the most common combination of this kind was Arch. *s-gl*, yielding Anc. *ṣ-* (rarely *s-*), but these forms are generally well 'disguised' in apparently unrelated series ('grove/forest', below) as well as in series with Anc. *s-* ~ *l-* alternation (GSR-498, -878, -975) or even with Anc. *ṣ-* or *s-* alone; cf. the following:

s-[*gl*]*i̯əm/ṣi̯əm*[hj] (tone A) 'dense trees, thicket, grove' (GSR-664), the prefixed doublet of *gli̯əm/li̯am*[hk] (tone A) 'forest, grove' (GSR-655); cf. Garo *-grim* 'suffixed added to a noun meaning grove, thicket', *bol-grim* 'forest' (bol 'tree') < PST **(s-)grim*.

s-[*g*]*l̥iwət/ṣiuĕt* ~ *s*-[*g*]*l̥iwəd/ṣwi*[hl] 'go along, follow', etc. (GSR-498); cf. WT *sgrod-pa* = *'grod-pa* 'to go, travel' < PST *(*s*-)*grot*; also, from the same series, [*g*]*l̥iwət/l̥iuĕt*[hm] 'fat around intestines'; cf. WT *grod-pa* 'belly, stomach' < PST **grot*.

s-[*g*]*l̥iəg/ṣi*[hn] 'recorder, scribe; record'; cf. WT *sgrig-pa* 'to lay or put in order, arrange; to put or fit together, join; to compile, write books'[25] < PST **s-grik*.

s-[*gl*]*âk/sâk*[hu] 'twist a rope' (*Anal. Dict.* also 'rope, string; bind') (isolated under GSR-770); cf. WT *sgrog-pa*, pf. *bsgrags* 'to bind', *sgrog*(*s*) 'cord, rope' < PST **s-grâk*.

The corresponding combinations of velar stop and cluster *s* appear to have been rare in Archaic, although the *xie-sheng* enable us to set up some forms along these lines; cf. **s*[*g'*]*l̥iəg/dz'i*[ht] 'spittle (of dragon)' (fn. 25), also *sk'l̥iam/ts'i̯äm*[hv] 'all', phonetic in a series (GSR-613) of velar stop + *l* type, suggesting that the affricate perhaps should be considered the 'regular' reflex here, yet another series of this type (GSR-1069) has a stop reflex: *sk'l̥iôg/t'i̯u* ~ *gliôg/lieu*[hw] 'get cured, recover', a doublet of *glɔk/lâk* ~ *gliog/liäu*[hx] 'to cure', also read [*ʔk,g*]*lŏk/ŋɔk* 'music' and [*ʔk,g*] *lŏg/ŋau* 'joy, rejoice', phonetic in GSR-1125, which includes another doublet: *sgl̥iok/i̯ak*[hy] 'medicinal plant; medicine; to cure; for the reflex in the latter, cf. *sgl̥iam/i̯äm*[hz] 'salt', from another series (GSR-609) of velar stop + *l* type, cognate to PTB **gryum*, id. (contra the analysis in *STC* fn. 472).[26] The stop reflex also appears in three words with early loan relationships with PT and PMY, with support for the velar stop in the case of both 'moss' (series [GSR-976] includes *[*s-k'*]*əg/xâi*[ic] 'laugh', listed under 1240a) and 'tripod' (series [834] includes [*sk'*]*i̯ĕŋ/t'i̯ĕŋ*[id] 'red' = *sk'i̯ĕŋ/t'i̯ĕŋ*[du] 'red' [same tone], above); note the tones (irregular in 'tripod' as result of unvoicing):

	Proto-Tai	*Proto-Kam-Sui*	*Proti-Miao-Yao*	*Arch./Anc.*
moss	**glay*^{A}	—	—	[*sg'*]*əg/d'âi*[ie] (tone A)
peach	—	—	**glaaw*^{A}	[*sg'*]*og/d'âu*[if] (tone A)
tripod	**gliaŋ*^{A}	*(*ʔ*)*gliaŋ*^{A}	—	[*sk*]*ieŋ/tieŋ*[ig] (tone B)

Palatalization played a role here comparable with that noted for prefixial *s*- (above); cf. *g'i̯ak/g'i̯ak*[ih] 'tongue', with doublets [*sg'y*]*i̯at/dź'i̯ät*[ii] (< **sg'yi̯ak* through assim.), id. (isolated character under GSR-288) and *sp'yi̯ĕg/dź'iẹ*[ij] (tone B) 'pick up with the tongue, lick' (in GSR-867, which includes initial *g'*- entries) and **sg'yi̯ĕg/dź'iẹ*[ik] (tone B) 'lick' (same word; character under GSR-1238e) < PST *(*s*-)*g-lyak* (the doublet forms are typical reflexes of PST **-yak*); cf. PTB **s-lyak* ~ **m-lyak* 'lick; tongue'; also *[*sg'y*]*i̯ăg/dź'i̯a*[il] 'musk deer' (not in GSR); cf. WT *gla-ba*, id. < PST *(*s*-)*gla* (with 'animal prefix' **s*-; for the vocalism, see *STC*: fn. 487).[27]

Preglottalization appears at times in series of stop + *l* type, as in GSR-1125 (above), but here as elsewhere the relationships are often well hidden, with two series (GSR-768 and -788) showing only Anc. initial *ŋ*-, from Arch. *ʔgl*- (or *ʔkl*-)

but with cross-ties especially to GSR-766 and -769 (this belongs in GSR-788, since the same phonetic is involved; see below); cf. the following:

[*ʔgl*]*i̯uk/ŋi̯wok*[in] 'jade', a doublet of [*g*]*l̥i̯uk/l̥i̯wok*[io] 'precious stone, precious', a derivative of [*g*]*l̥i̯uk/l̥i̯wok*[ip] 'green'; note that the 'jade' series (GSR-1216) includes [*s-k'l*]*i̯uk/x̥i̯wok*[iq] 'disconcerted', perhaps cognate (= 'to be disturbed') to WT *dkrug-pa* = West T *śrug* < **skrug* 'to stir, stir up, agitate; to trouble, disturb, confound' < PST *(*s*-)*kru·k*.

(GSR-788) [*ʔgl*]*âk/ŋâk*[ir] 'osprey' (= 'sea eagle') a doublet of (GSR-766) *glâk/lâk*[is] 'a kind of bird'; cf. WT *glag* 'eagle, vulture' < PST *(*a*-)*glâk*.

(GSR-788) [*ʔgl*]*âk/ŋâk*[it] 'scared', a doublet both of *glâk/lâk*[is] (above) 'fear' this gloss only in Anal. Dict.) and (GSR-769) **s*[*gl*]*ɛk/ṣɛk*[iu] (< **s-gl̥iâk* or **s-grâk*) 'fear' (phonetic is *s*-[*gl*]*âk/ṣɔk*[iv] 'first day of the moon', with phonetic [*ʔgl*]*i̯ăk/ŋi̯ɒk*[iw], as in GSR-788); cf. PTB *(*s*-)*grâk* 'fear' < PST *(*s*-, *a*-)*grâk*.

(GSR-788) *s*-[*gl*]*âg/suo*[iu] (tone C) (< **s-glâk* + suffix) 'inform, complain' (another reading; cf. **s*[*gl*]*ɛk/ṣɛk* 'fear', above); cf. PTB *(*s*-)*grâk*: WT *sgrog-pa*, pf. *bsgrags* 'to call, call out; publish, proclaim, promulgate; to shout, scream'; WB krak (< *[*g*]*rak*) 'honor, glory' < PST *(*s*-)*grâk*.

(GSR-788) [*ʔkl*]*âk/ŋâk*[ix] 'beat the drum and make a noise'; cf. WT *skrog-po* 'to beat (the drum)' < PST *(*s*-, *a*-)*krâk*.

In addition to the numerous *xie-sheng* series characterized by velar stop + *l* initials, Archaic also had several series with labial stop + *l*, with evidence of cluster and prefixial *s*- as well as preglottalization. The shifts closely parallel those noted for the velar stop + *l* combinations, indicating 'neutralization' of the stop between *s*- and *l*. Prefixial *s*-, represented (along with preglottalization) in only one series (GSR-178), presents no problems; cf. *blwân/luân*[iy] 'bells on horse's trappings' (Bodman [1975bis] cites Thai *bruan* neck bells for domestic animals'), phonetic in *ʔblwan/mwan*[iz] 'Southern barbarian' and *s-blwan/ṣwan* ~ *s-bl̥i̯wan/ṣi̯wän*[ja] 'twins'; cf. also *s-b̥i̯ĕt/ṣi̯ɛt*[jb] 'lute; (loan) rustling of the wind', under GSR-411 (isolated) but with *pi̯ĕt/pi̯ĕt*[ie] (GSR-405; see above) as phonetic (Bodman [1973] cites WT *sbrid-pa* 'to flutter before one's eyes'); the great scarcity of examples of this kind indicates that the **sb-* > *p*- development (see above) had already been completed by the Archaic period. As might be anticipated, it is the cluster combinations that present difficulties, with direct evidence for the *-l-* present only in two of the series ('dragon' and 'pen', below), but the comparative TB data are useful here, along with the evidence from one early loanword from Tai; cf. the following:

bl̥i̯uŋ/l̥i̯uŋ[jd] 'dragon'; cf. WT *'brug* 'thunder; dragon' (< **'bruŋ*; fn. 11) < PST *(*a*-)*bru·ŋ*, phonetic in **b'lŭŋ/b'ɔŋ*[je] 'high edifice' (not in GSR), also in *sp'l̥i̯uŋ/t'i̯wəŋ*[jf] 'favor' and *bluŋ/luŋ*[jg] 'deaf'; cf. WT *loŋ-ba* ~ *ldoŋ-ba* (< **'bloŋ*) ~ *mdoŋs-pa* (< **mbloŋ*) 'blind'[28] < PST *(-)*blo·ŋ* (for the semantics, cf. PAT 'blind' ~ 'deaf' and [diff. root] 'blind' ~ 'mute'; see Benedict 1975: *Glossary* under DEAF).

ʔblŭŋ/mɔŋ[id] (tone A) 'mixed black and white color, variegated' (another reading; cf. *blu̯ŋ/liu̯ŋ* 'dragon', above); cf. (GSR-1201) [*ʔbl*] *ŭŋ/mɔŋ*[ii] (tone A) 'shaggy dog; particolored, motley' (same word; Cant. *p'oŋ ~ moŋ*, app. reflecting the earlier stop), also [*ʔbl*]*ŭŋ/mɔŋ*[ij] (tone A) 'particolored animal, variegated' (etym. same word), from **ʔbla·ŋ* (regular vowel shift), an early loan from PT **ʔblaaŋ* 'spotted; piebald (as horse)'; cf. also from the same series (GSR-1201) [*ʔbl*]*ŭŋ/mɔŋ*[ik] (tone A) 'great; ample' (*Anal. Dict.* also 'thick; richly abundant') (Cant. *p'oŋ* [low tone] < **b'oŋ ~ moŋ*); cf. WB *pruiŋ* (< **[b] ruiŋ*) 'full, abundant' < PST **(a-)bru·ŋ*[A]

sbl̥iwət/l̥iuĕt[il] 'writing stylus or pencil', a doublet of *bl̥iwət/l̥iuĕt*[im] 'pitch-pipe', etc. (both from 'quill') as well as of *pl̥iət/pi̯ĕt*[in] 'writing stylus', from an earlier (pre-Archaic) **sbl̥i[w]ət* (see above for **sb-* > *p-*), an early loan (with **s-* prefix added) from PAT **bu̯lut* 'body hair, fur, fibre' (contra the analysis in *STC*: fn. 474, and note WT *pir* <**pit* as a relatively late loan [Bodman 1975]).

[*sbl*]*i̯ĕŋ/i̯äŋ*[io] 'full, fill', isolated under GSR-815 except for one homophone, a doublet of [*sbl*]*i̯əŋ/i̯əŋ*[ip] (loan) 'full (sc. ear of grain)'; cf. PTB **(s-)bliŋ* 'full/fill' < PST **(s-)bliŋ* (contra the analysis in *STC*: fn. 469 for this and the following comparisons).

**[ʔbl]ĕŋ/mɛŋ*[iq] 'toad' (not in GSR), phonetic in GSR-892, which includes [*sbl*]*i̯əŋ/i̯əŋ*[ir] 'a fly'; cf. PTB *s-brəŋ*, id. < PST **-brəŋ* (with 'animal prefix' **s-*); also in [*sb'ly*]*i̯əŋ/dź'i̯əŋ*[ip] (above as loan) 'string, cord'; cf. PTB (Nungish) **a-briŋ* 'cord' < PST **(s-, a-)briŋ*.

Both cluster and prefixial *s-* are to be reconstructed for Archaic before *l* (< PST **l* and **r*), perhaps along with preglottalization, with the following typical reflexes:

**l* }	> *l/d'*	**s-l* }	> **[s-]l*
**r* }		**s-r* }	

**sl* > *s̱i*
[**sr* > *x̱i*]

The shift to *d'-* perhaps occurred only before high (or palatalized) vowels (*STC*: fn. 458); cf. PTB **lyiŋ* = **liŋ* 'field' (WT *źiŋ*, Lepcha *lyăŋ*; see *STC*: fn. 246), Arch./Anc. [*l*]*ien/d'ien*[is] 'field; to hunt' < PST **li(·)ŋ*[29]; also PTB **(s-)re·k* 'pheasant', Arch./Anc. [*l*]*iok/d'iek*[aq] < PST **(s-)re·k* (see table above), but before palatalized *a* (< PST **a* or **â*) **l(y)-* merged with **y-* (> *d̨i/i̯-*) in *i̯-*; cf. PTB **lâp* 'leaf', Arch./Anc. [*l*]*i̯ap/i̯äp*[al] < PST **lâp* (see above). As in the case of nasals (above), the prefixial *s-* is reflected in Proto-Min (**hl-* = **sl-*); cf. the variant **[s-]l* form for 'six' (above) and the following pair:

PTB **s-raŋ*: WT *sraŋ* 'pair of scales; weight; ounce'; Arch./Anc. [*s-*]*l̥iaŋ/l̥iaŋ*[it] 'two; a pair' (*Anal. Dict.* also 'ounce'; graph is picture of scales); Proto-Min **hl*[*a*] *ŋ* = **sl*[*a*] *ŋ* < PST **s-raŋ*.

PTB **(s-)rwa*: WT *sro-ma* 'nit', Kachin *tsiʔ-ru* 'louse [*tsiʔ*] egg'; Arch./Anc. [*s-*]*lwân/lwân*[iu] 'egg'; Proto-Min **hl*[*ua*]*ŋ* = **sl*[*ua*]*ŋ* < **sl*[*ua*]*n*, from **s-lwa·-n* ('collective' plural **-n*, with unanticipated vowel length; see *STC*: fn. 428) < PST **(s-) rwa(-n)*.

This schema implies that Ancient forms with initial *l-* are in most cases to be reconstructed [*m-*]*l*, [*g*]*l*, [*s-*]*l* or the like; cf. PTB **liŋ* 'neck'; Arch./Anc. *m-l̥ĕŋ/l̥iäŋ*[iv] 'neck, collar' (phonetic is *m-l̥ĕŋ/l̥iäŋ*[iw] 'command'; see *STC*: fn. 419) < PST **(m-) liŋ* (cf. PTB **m-* prefix with body-part words in WT and elsewhere); also PTB **l*[*i,e*] *p* 'turtle' (WB *lip*); PK **khliʔ*, id., from **g-li*[*p*] (see *STC*: 135 for PK **kh-* < **g-*); Arch./Anc. [*g*]*l̥iap/l̥iäp*[ix] (loan) 'a kind of turtle' (graph for the phonetic of this series [GSR-637] depicted a turtle) < PST **(g-)lep*; also PTB **la* 'salt' (Miri *ələ* < **a-la*); PK **hla* (< **s-la*), id.; Arch./Anc. [*s-*]*lo/luo*[iy] 'salty' (*Anal. Dict.* also 'rock-salt', but used in graphs as general signific for 'salt') < PST **(s-)la* (Proto-Min form not available; initial **l-* possibly maintained here before the low, unpalatalized vowel).

The regular reflex of cluster *s* + *l* appears to have been *sl/s̱i-*; cf. PTB **sryam* 'sharp' (Lushai *hriam*, Garo *sram* ~ *srem-srem*); Arch./Anc. *s*[*l*]*i̯am/s̱i̯äm*[iz], id., with 'hidden' phonetic **(s-)*[*l*]*iam/d'iem*[ii] 'tongue' (usu. read [*sg'y*]*i̯at/dź'i̯ät*; see above) < PST **(s-)lyam* [PTB **(s-)lyam* 'tongue; flame'], phonetic also in [*l*] *iam/d'iem*[ka] 'sweet' < PST **le·m* [PTB (Proto-Kiranti) **l*[*e,e·*]*m*, id.], with palatalization before PST **i*; cf. PTB **(s-)riŋ* 'long/elongate' (WT *riŋ-ba* 'long', *sriŋ-ba* 'extend, stretch, postpone'); Arch./Anc. *s*[*ly*]*i̯ĕn/s̱i̯ĕn*[kb] 'stretch, extend, prolong' < PST **(s-)riŋ*. This Archaic cluster (**sl-*) represented a 'filling in' of the pattern, the earlier (pre-Archaic) cluster having been shifted to *xl/t'-*, with support for this reconstruction from the early loanword **[xl]-iet/t'iet*[kc] (Arch. form not cited in GSR) 'iron', from PAT **(m)baxliaq* (via **-xliat* through assim.); cf. the following pair of similar (perhaps cognate) roots:

PTB **sri(-t)*: WT *srid-pa* 'existence, state of being, life; things existing, the world; also a single being', WB *hrí* (< **srí*) 'to be (in some place)'; Arch./Anc. *xliər/t'iei*[kd] 'body' (phonetic is [*s-*]*liər/liei*[ke] 'ritual vase'), a doublet contra *STC*: fn. 428) of *s*[*ly*]*i̯ĕn/s̱i̯ĕn*[kf] 'body, person' < 'being'), with suffixed **-n*[30], from PST **srə·y ~ *sri(-n,-t)*.

PTB **sri(-n)*: WT *sri* 'a species of devil or demon, a vampire', *srin-po* (fem. *srin-mo*) 'demons'; Lushai *hri* (< **sri*) 'the spirit supposed to cause sickness'; Arch./Anc. *xlia/t''iɛ̨*[kg] 'a mountain demon' (phonetic in GSR-23, with otherwise initial [*s-*]*l* forms), PST **sri(-n)*, the suffixed form perhaps represented by [*ly*]*i̯ĕn/dź'ien*[kh] 'spirit; divine, supernatural' < **[s-] rin* (from GSR-385, which includes 'stretch, extend', above, as well as [*l*] *ien/d'ien*[ki] 'lightning', app. related [loan] to Proto-Yao **liŋ*, id.).

The occasional appearance of doublet forms with Anc. *x-* in series containing initial *l/d'-* entries presents a reconstruction problem, and it is suggested that Archaic still had marginal *sr-* (~ regular *sl-*) doublets, from an earlier **sr-*, yielding Anc. *x̱i-*, as indicated by the following pair:

[*l*]*i̯ap/i̯äp* ~ **[sr]i̯ap/x̱i̯äp*[kj] 'small, insignificant', with (GSR-633) [*l*]*i̯ap/i̯äp*[al] 'leaf' (above) as phonetic; cf. PTB **srap* or **s-rap*: WT *srab-pa* 'thin, tender, fine, e.g., skin, cloth, leather, paper, clouds', WB *hrap* (< **srap*) 'to graze, pass over slightly touching; cursory, slight' < PST **(s-)rap*; also the related [*l*]*iap/d'iep*[kk] (< **l̥iap*) 'unlined (= thin) garment' (cf. the WT gloss)

as well as *s[l]iap/siep*[kl] 'bottom inlay in sole, shoe', directly comparable with WT *srab-mthil* 'inner sole, welt' (= 'the thin inlay') (*mthil* 'bottom').

[l]iet/d'iet ~ **[sr]i̯ĕd/xji̯t*[km] (< **[sr]i̯ĕt* + suffix) 'laugh', from the 'special' GSR-413 series noted in *STC*: fn. 458 (containing *[l]iet/d'iet* ~ *[l]i̯ĕt/d'i̯ĕt*[kn] 'nephew/niece', cognate to PTB **(b-)ləy* [*STC* **b-ləy*, but note Garo *-ri*, Chang *li*] 'grandchild, nephew/niece'); cf. PTB **(s-)rya(-t)*, id., for *STC* **rya-t* (note Bunan *sred* < **sryat*, Nachereng *hres* < **sryat-s*) 'laugh' < PST **(s-)rya(-t)*.

The above series (GSR-413) also contains a group of entries with initial *t-* or *tś-*, three of which have apparent TB cognates with initial **l-* or **r-*, and it is suggested that these forms might be reconstructed with initial *ʔl/t-*, paralleling the indicated *ʔn/t-* (above); cf. the following:

[ʔl]i̯ĕd/t'i̯[ko] 'heavily weighted down' (Couvreur dict. also 'heavy'); PTB **(s-)ləy* 'heavy' < PST **(s-,a-)ləy*.

**[ʔly]i̯ĕt/tśi̯ĕt*[kp] 'leech' (not in GSR); PTB **(m-) li·t* 'water-leech' (*STC*) = **l[i,i·]t* (Chepang *lit*, Kachin *lip* < **lit*) ~ **m-l[i,i·]t* (Mikir *iŋlit*, Ao Naga *melet*) ~ **s-li·t* (Lepcha *hlet-bŭ*, Lushai *hli·t*) < PST **(s-,a-)li(·)t*.

[ʔl]i̯ĕt/t'i̯ĕt[kq] 'beat, a stroke' (Couvreur dict. 'couper la moisson, couper les épis, bruit de la faucille'); PTB **ri·t* 'reap, cut, scrape, shave' (WB *rit* 'to reap, mow, shave', Miri *rit* 'to cut, as small jungle') < PST **(a-)ri(·)t*.

As a result of this extensive reconstruction of Archaic we are faced with a radically new picture of the language. It now seems certain (contra *STC*: 155) that prefixed *s-* (< PST **s-*), phonologically distinct from cluster *s* + consonant, played a morphological role of some kind, perhaps also ʔ(< PST **a* = ʔ*a-*), as indicated by the occasional doublet formations ('tread/trample', above). It should be remembered that only a certain fraction of such elements have been reconstructed to date, on the basis of *xie-sheng* and/or Proto-Min data, along with the help of early loans as well as comparative TB evidence, so that the language might well have looked something like WT, with suffixes (giving rise to the sandhi tone C) and with even more prevalent *s-* prefixation.[31] There is also some evidence that prefixed **m-* and possibly **b-*, both important TB prefixes (*STC* 110–12 and 117–21), played a role in Archaic morphology; the few recovered examples have survived simply because they happened to precede consonant clusters, e.g., *m-kl̥i̯ôk/mi̯uk*[e] 'accord' (above), *m-kl̥i̯ôg/mi̯ĕu*[kr] 'bind round', also read *kl̥i̯ôg/ki̯ĕu* 'twist' (the unprefixed form) or to come before *k(i̯)w-*; cf. the following pair:

m-kwəg/muâi̯[ks] (tones A, C) 'meat on sides of spine', the same word (originally) as *mwəg/muâi̯*[kt] (tones A, C) (GSR-947), id., from an earlier shift, paralleling *mi̯u* < *m-[k]i̯u* 'despite, insult' (fn. 22; phonetic from the same series); this word is in GSR-950, which has *s-k'wəg/xuâi̯*[ku] 'ashes' as phonetic and includes *k'wəg/k'uâi̯*[kv] 'great, extend'.

kiwat/kiwet[kw] 'sleeve', also read *mi̯ad/mi̯äi*, id., for the affixed doublet: **m-ki̯wad/mi̯wäi* < **m-ki̯wat* + suffix.

In similar fashion, an original prefixed **b-* before the same labialized initial (*ki̯w-*) probably gave rise to the enigmatic *p/ki̯wăp pi̯wɒp*[kx] (< **b-ki̯wap* through assim.) 'law', with phonetic (GSR-642) *k'i̯ab/k'i̯wo*[ky] 'go away'; cf. WB *kwap* (< **[-] kwap*) 'to bind or overlay the edge or border or anything; to clamp, make fast by binding; to enforce orders; discipline' < PST **[b-]kwap*.

It is apparent that this transformation in the 'look' of Archaic, bringing it closely into line with WT and with TB languages generally, will lead to many changes, possibly some of a profound nature, in our ideas about the grammar of the language and, indeed, in the field of early Chinese studies generally. It is not clear at this time, however, whether such changes will require any basic modifications in the reconstruction of PST itself, any more than the recent advances in the reconstruction of PBL, carried out particularly by J. Matisoff and his students, have to date necessitated any modifications in the reconstruction of PTB. Chang (1973) criticizes the *STC* for recognizing prefixed **r-* as well as **s-* at the PTB level, but he appears to have overlooked the substantial evidence (with distinctive reflexes) for prefixed **r-* throughout the TB area (*STC*: 109–10), with representation also in Karen (fn. 356) and in Chinese (fn. 419). Regarding PTB prefixed **a*, the *STC* rejection of Wolfenden's distinction between 'pronominal' and 'non-pronominal' basic forms of this prefix has been supported in a recent study by Lehman (1975bis). The WT representative of this prefix (*a-chung*) continues to create problems, e.g., B. Chang (1971) follows Wolfenden in interpreting WT prefixed *b-* as **b-* + *a-*, thereby producing unnecessary complications in her analysis of the phonology of Tibetan causatives. Egerod (1973) citing the 1952 sketch by Maspero in *Les Langues du monde*, appears to reject the *STC* distinction between noun and verb roots at the PTB level. The point is not stressed in *STC* but it is apparent, Maspero notwithstanding, that at least two main classes of free forms must be recognized at that level: verbs (negatable) and nouns (non-negatable), along with a residual class of verb/nouns (e.g., PTB **r-miŋ* 'name'); the function of a given prefix such as **s-* may in fact differ so radically in verb as opposed to noun roots that distinct morphemes (**s-*[V] and **s-*[N]) should be recognized. Egerod, in the same review, also raises the question (first suggested by Henderson 1957) of whether pronominalization might not after all be of 'native' origin. This matter is the subject of a recent paper by J. Bauman (1974), who rejects the view of Maspero and Egerod that an early influence from Indo-European might have been involved, in favor of the 'native' hypothesis. The languages affected include Gyarung, which is hardly adjacent to any obvious non-ST source of influence, and Bauman points out some apparently significant similarities in the morphemes themselves, so that the possibility of reconstructing pronominalization for at least one segment of TB must be given serious consideration.

Now that comparative ST linguistics has 'come of age', as one reviewer of *STC* has put it (Lehman 1975), it behooves us to glance back at the comparative process itself. There appears to be a widespread misunderstanding of what is needed to carry out worthwhile comparative work, at least as regards the ST field. Miller (1974) spends much of a long review article on *STC* bemoaning, in effect, the

lack of suitable instruments: "If the *Conspectus* proves anything at all, it proves that it is high time to abandon these out-moded, antiquated records of languages concerning which not enough is known to enable us to subject their data to the rigorous methodology of the comparative method" (p. 209). The same reviewer furnishes a highly imaginative account (p. 198) of the Berkeley workshop of the late 30's and early 40's in which these 'antiquated' materials were compiled, making it sound rather like a language machineshop. Having directed that workshop during the last two years of its existence, the writer was struck by the incongruity of the description, since the "sifting" of materials occurred largely within the skull of the director, the staff being employed primarily in more mundane activities, although he did have the good fortune to find one educable asistant. Although Miller's comments have been selected for citation, similar pronouncements have appeared elsewhere, indicating some general failure to appreciate the basic facts of the history to date of ST comparative linguistics. What the *STC does* prove, to paraphrase Miller, is that it was possible over 30 years ago to establish the basic framework of a widespread language family and to produce a manuscript that was still serviceable as a textbook (at Columbia) some 25 years later, while the *STC* itself has continued to play that role (Lehman 1975). Actually, an excellent reconstruction of PTB could have been turned out *as early as the turn of the century*, when good sources became available for such key languages as Tibetan (Jäschke), Burmese (Judson), Kachin (Hanson) and Lushai (Lorrain and Savidge), and an equally good PST reconstruction *as early as 1923*, with the publication of Karlgren's *Analytic Dictionary* (the later reconstruction of Archaic Ch. served only to confuse scholars, as shown above!). The deficiences in Shafer's efforts extending from the 30's into the late 60's, as in the writer's effort basically of the early 40's (the MS of *STC*), are equally to be attributed primarily to *poor linguistic thinking* rather than to *poor sources*. It is perhaps comforting to think otherwise but the facts of the matter are quite plain. The *STC* itself might be regarded as a kind of "test case," since the manuscript was the product of the early 40's but was extensively annotated over 25 years later, after the appearance of a considerable amount of new source material as well as of many scholarly studies in the field of Chinese (including a new reconstruction schema by Pulleyblank), Tibetan, Burmese-Lolo and whatever, and even major reconstructions of Proto-Bodo-Garo (Burling), Proto-Burmese-Lolo (Burling, Matisoff) and Proto-Karen (Haudricourt, Jones, Burling). A review of the "new" notes (added material), however, reveals that almost without exception the more significant new findings or conclusions had little if anything to do with any of the above but were rather in the nature of "rethinking of problems," of new insights into old problems, of things that *should have been done originally*: the reconstruction of PTB labial stop + *w* clusters (fn. 78; contra Coblin [1972–73] the medial *-w-* in similar clusters *is* contrastive in Arch. [if not Anc.] Ch. and the feature must be set up for PST, where it is only marginally contrastive, as discussed in the final lines of this footnote); the recognition of a palatal series (fn. 122) and of PTB initial **ćr-* (fn. 95) and other clusters (fn. 121) as well as of a "collective" suffixed **-n* (fn. 95); the new

interpretation of WT *a-chung* (fn. 339; Denwood [1974] quite misses the point!); the reconstruction of two new vowels: PTB **ə* and **â* (fn. 344); the voluminous new material in the footnotes of the Chinese section, entirely recasting much of that part of the book, particularly in the treatment of ST vocalism—all should have been done in the early 1940's but the significant clues were missed—and a final long footnote (494) setting up a two-tone system for PST, another "discovery" that should have been made in the early 1940's at the time that the correlation of Burmese and Karen tones was first uncovered! Only the additions to the Karen section were significantly dependent upon post-1940 publications, and much to the point here is the fact (*STC*: fn. 347) that although the new *material* was supplied by the excellent studies of R. B. Jones the new *ideas* by and large (apart from relatively minor matters such as the reconstruction of PK initial **hy-*; see fn. 371) were furnished by the reconstruction of PK brilliantly carried out much earlier (1945) by A. Haudricourt, working only with the older dictionary materials, all quite available at the time to this writer, who unfortunately was able to work out only one fragment of the schema (*STC*: 151: the unvoiced nasal initials). This all probably means that important contributions were overlooked, but it also means that 25 years of brooding over what is essentially the same corpus of material can result in major changes in one's findings, hopefully for the better. It is hoped that this account will not have the effect of discouraging the collection of new material and the compiling of bigger and better lexicons, for certainly we stand in need of these things, but rather that it will stimulate some "hard thinking" on the many seemingly insoluble problems that beset us in comparative ST linguistics.

Notes

* The author is indebted to Prof. Paul Yang of Georgetown University for help with this paper, including the appended characters.

1 Matisoff (1973) has presented a valuable account of the first five years of this conference.

2 N. Bodman 1975; K. Chang 1973; W. S. Coblin 1972–73; F. K. Chou 1972; P. Denwood 1974; S. Egerod 1973; A. Haudricourt 1973; K. Lehman 1975; R. Miller 1974; K. Sedlček 1974; R. K. Sprigg 1973.

3 The following word-list has been employed (see Burling 1971): I thou we this that who? what? not all many one two big long small woman man person fish bird dog louse tree seed leaf root bark skin flesh blood bone grease [= fat/grease] egg horn tail feather hair head ear eye nose mouth tooth tongue claw [= nail] foot knee hand belly neck breasts heart liver drink eat bite see hear know sleep die kill swim fly walk come lie sit stand give say sun moon star water rain stone sand earth cloud smoke fire ash burn path mountain red green yellow white black night warm cold full new good round dry name.

4 See the note by Matisoff in *JCL* 1:3 (1973), which points out that ". . . tonal convergence and genetic relationship are totally independent things . . . ," while emphasizing that the 'four known divisions' of TB described by Li do not fit any current ideas of TB subgrouping. The most recent position on the Tai languages taken by Li (1974) is that "the relationship [to ST] has never been definitely established."

5 Miller (1974) complains, "There is nothing at all here of the Tai languages" and "The most melancholy single aspect of the decision to eliminate any consideration of Thai from these pages . . . ," followed by "This seems a great pity . . ." and "It seems almost

perverse to have turned the back on the one group of languages [Tai] about which probably the most is known . . . ," yet this reviewer also shows that he is familiar with the writer's view that the Tai languages, along with MY, belong in an entirely distinct language stock (Austro-Thai; see *STC*: fn.'s 8, 14, also Benedict 1975), hence it is not altogether clear whether Miller at present actually rejects the traditional view of a Chinese-Tai genetic relationship.

6 Cf. the parallel development in Stau (Tibet: eastern Kham), which has *śńi* 'day' < PST *[*s*-]*nəy* but *zńi* 'seven' < PST **snis*; Gyarung lacks this voicing distinction: *śńi* 'day', *keṣńit* (< **k-snis*) 'seven'.

7 Abbreviations: YCR Chiengrai ('Highland') Yao; YHN Haininh Yao; MCF Chengfeng Miao (Eastern: 'Kanao'); MWN Wei-ning Miao (Western); MPT Petchabun Miao (Western: 'White Miao'); MSY Su-yung Miao (Western: 'Magpie Miao').

8 The Arch.-LMY development is of special interest in view of the **sən-* > *dz'-* shift shown by two body part words in Chinese: *dz'i*[a] [no Arch.] 'self' < 'nose' < PST **s-na* or **s-na·r* (*STC*: fn. 471 and p. 16); *dz'i̯ĕŋ/dz'i̯äŋ*[b] 'feelings' (Couvreur: 'sentiment de l'âme' and 'passions du cœur') < PST **s-niŋ* 'heart' (WT *snyiŋ* 'heart, mind' but basically applied to 'feelings'); in this sound shift the most likely sequence is **sn-* > **zn-* > **zd'-* > **dz'-*, the Arch.-MY form fitting nicely here (see below for more on **s-* prefix).

9 Cf. Ch. *[*s*-]*nien/nien*[i] (Proto-Min has initial **hn-*); the Shuo-wen interpretation of *ts'ien* as phonetic in this word points to a variant development: **snien* > *ts'ien/ts'ien*[j] as exactly paralleling **sn̥i̯ĕt* > *ts'i̯ĕt* 'seven' (and Arch.-LPT **tsi̯ĕt*; see Table).

10 PMY maintains PAT **-ia-*; cf. PMY **ntiaŋ* 'tree' < PAT **(n)ti(y)aŋ* 'stick/handle/post/tree' (Jav. *tiyaŋ*, Malay *tiaŋ* 'post', Fiji *ndia* "stick, handle', PT **deeŋ* ~ **theeŋ* [< **(n)tiaŋ*] 'stick, bar').

11 For the roots listed in *STC*, add under **s-gla* 'moon': G *dża* (*STC*: fn. 109); add under **lum* 'warm': L *lum* 'warm, hot'; add under **(r-)kaŋ* 'foot, leg': K *lago* ~ *lagoŋ*, id. (cf. the parallel root: PTB **kraŋ* 'mosquito' > K *groŋ* [*STC*: 71]); add under **tu.ŋ* ~ **du·ŋ* 'sit': WT *'dug-pa*, id. (for the nasal > stop shift, perhaps conditioned by vowel length, cf. WT *'brug* 'thunder; dragon' < PST *(a-)brù·ŋ*; see text, below). The following roots are not listed in *STC*: PTB **(s-)m[u, əw]* 'see': K. *mu*; L *hmu*; PTB **di[·]ŋ* 'to erect, be erect, stand(ing)': Lepcha *diŋ* 'to be erect; to stand; to be at rest'; Miri *diŋ* 'to plant (anything tall); to set up or erect'; Kachin *diŋ* 'to be straight, rectilinear', *puŋdiŋ* 'zenith, top' (*STC* cites Arch./Anc. *tieŋ/tieŋ*[n] 'top of head'); WB *tań* (< *[*d*]*iŋ*) 'to place in position, build'; Lushai *diŋ* 'to stand, be upright; to stop'; Garo *tśadeŋ* 'to stand'; cf. Arch./Anc. *d'ieŋ/d'ieŋ*[o] (tone C) 'settle, establish; fix; finish, stop; finished'; *d'ieŋ/d'ieŋ*[p] (tone A) 'settle, regulate'; *d'ieŋ/d'ieŋ*[q] (tone A) 'to stop' (etym. same word; cf. Lushai semantics); PST **di·ŋ*; PTB **lu·ŋ* ~ **lu*[·]*k* 'drink': G *riŋ*, Dimasa *luŋ* ~ *liŋ*; K *luʔ* (cf. WB *loŋ* < **l*[*u*]*ŋ* 'pour into or upon'; WT *ldug(s)-pa*, pf. *blugs* (< **-lug*) 'pour, cast'; also PTB **l*[*u, əw*] 'pour'); PTB **dźəy* 'seed': B *(ă-)tsé*, Maru *ətśit* (< PBL **dźəy*); L *tśi*; PTB **(-)raŋ* 'bone': K *nra* ~ *nraŋ* (< **m-raŋ*); G *greŋ* (< **g-reŋ* < **g-raŋ*; see *STC*: 72).

12 Cf. Arch./Anc. *măŋ/mɐŋ*[r/s] (irreg. for */mɒŋ*) 'population, people' (= 'the many', as in T and K); also *xi̯*[*m*]*waŋ/xi̯waŋ*[t] 'increase', from **s-maŋ* 'make more' < PST **(s-)maŋ*; also the app. distinct root: PTB **maŋ* 'big (elder)' (Trung [Nungish] *dəmaŋ* 'big; [comp.] older [brother, uncle]); Arch./Anc. *măŋ/mɒŋ*[u] 'eldest (of brothers, etc.)', *xi̯*[*m*]*wăŋ/xi̯wɒŋ*[v] 'elder brother, senior', from **s-maŋ*; PST **(s-)maŋ*.

13 PST **ŋa* 'I' and **ŋâ* 'self' might also be regarded as doublets, of course, but there is little if any good evidence for this alternation, and in fact these two vowels are in opposition in the PST pair: **g-ya* 'right (hand)' (*STC*: fn. 487) and **(g-)yâ(-n)* 'left (hand)' (*STC*: fn. 428). As indicated in *STC* (p. 160, based on Karlgren's work), there is evidence of pronominal inflection at an early period in Chinese, with both forms derived from 'self' in object position, but this interpretation has come under repeated attack (see the discussion in Coblin 1972–73); in any event, these were the pronouns

to survive into the modern (Mandarin) Language. A form of the 'self/I' root with suffixed *-n*, of uncertain function (cf. fn. 30), is represented by Arch./Anc. *ŋan/ŋan*[ad] 'face countenance' (cf. WT), from **ŋâ-n* (regular vowel shift).

14 Preliminary analysis of several northern TB languages in terms of the Swadesh 100-word list scoring indicates that Rawang (Nungish) lies close to Burmese (pairing score is 38), pointing to a Burmese-Lolo-Nungish supergroup (cf. *STC*: 8); also that Tamang, Chepang and Miri all fall within the B/T/L supergroup as opposed to the K/G; Lepcha, however, has unusually low scores (only 12 for the Pwo and Mandarin pairings), supporting the view (*STC*: fn. 24) that this language has a non-ST substratum.

15 Bodman (1975bis) has suggested a six-vowel scheme (without **/â/*) for Archaic, while Li (1971) has set up a four-vowel system (**/aəiu/*) for this language, but neither has presented a systematic comparison with TB vocalism.

16 PST **(s-)re·k* 'pheasant' contra *STC* (84–85), which recognizes a medial **i* ~ **ya* alternation in this root (**s-rik* ~ **s-ryak*) as well as in 'eye' (**mik* ~ **myak*); the latter is cognate to Arch./Anc. *mi̯ôk/mi̯uk*[as], id., from an earlier **mi̯ək* (Arch. lacks both **mi̯ək* and **mi̯aŋ*), from PST **(s-)myək* (see text, below, for the prefix), with the same reconstruction available for PTB, dispensing with the need for recognizing the **i* ~ **ya* medial alternation as a special PTB feature, at any rate.

17 The occasional Arch. final *-n* ~ *-g* or final *-r* ~ *-g xie-sheng* contacts are perhaps to be explained along similar lines, e.g., Anc. *mi̯wĕn*[ax] 'diligent, active; hasten quickly', with Arch./Anc. *mwəg/muâi*[ay] (< **mu·g*; see *STC*: fn. 479) as phonetic (Karlgren [GSR-1251q-s] explains as a 'synonymous' word), perhaps reflecting PST **mur* ~ **mu·r* (cf. WT *myur-ba* 'quick, swift, speedy', cited in Yang 1975); also *miər/miei*[az] (tone A) 'fawn', *ŋieg/ŋiei* ~ *mi̯ĕg/myi̯ę* ~ *mieg/miei*[ba] (all tone A), id., possibly reflecting PST **mi·r*; cf. also *ʔi̯əg/ʔi̯*[bb] 'physician; potion', apparently (contra Karlgren: GSR-958) with *ʔiər/ʔiei*[bc] 'quiver' as phonetic.

18 Note that these exceptional WT forms (*STC*: 20) generally have initial *p-*, also that WT lacks the cluster **rp-* and that *-lp* occurs only in comp. (*pags-pa* and *-lpags* 'skin'), hence *pad-ma* 'leech', from PTB **r-pat*, can be considered a regular development (!), as can probably also *pus-mo* 'knee', from **l-put-s* (cf. Kachin *lăphut*, with *lă-* probably standing for PTB **lak* 'foot/leg') and perhaps *paŋ* 'bosom, lap', from **l-paŋ* (cf. Simon 1974).

19 Burling 1967; Matisoff 1972; Shafer 1966–67; Thurgood 1974.

20 Note also WT *sgog-pa* 'garlic', an apparent early (Pre-Archaic or Archaic) loan from Chinese: *sgog* < **sgi̯ôg*, from an earlier **s-ki̯ôg* (second. voicing).

21 Note the parallelism shown by both the secondary voicing and the vocalism in the two roots for 'eagle' and 'lift' (see *STC*: fn. 476 for the vocalism). It appears that the general trend toward secondary voicing of stops after prefixed **s-* > **sə-* extended not only to dentals and labials (see text, above) but also to velars in at least one of the principal dialects making up the composite known as 'Archaic Chinese'; cf. the roots for 'facing/above' and 'despise/insult' (above); also *g'i̯ôg/g'i̯ə̯u*[eo] 'maternal uncle', from *[*sə-*]*k'i̯ôg*; cf. PTB **kəw*, id. < PST **(s-)kəw*; also *g'o/ɣuo*[ep] 'door, opening', from *[*sə-*]*k'o*; cf. PTB **(m-)ka* 'open(ing), mouth, door' < PST **(s-, m-)ka*; also *sgyi̯ok/ź i̯ak*[eq/er] 'a ladle'; *sgyi̯ok/źiak* ~ *sgok/i̯ak* ~ *skyi̯ok/tśi̯ak*[es] (the third form remains unvoiced) 'ladle; to ladle out, pour out'; cf. PTB **s-kyok*: WT *skyogs-pa* 'scoop, ladle', WB *yok* 'ladle' < PST **s-kyok* (cited in part in Yang 1975). On occasion, however, Chinese retains the unvoiced velar stop which has been secondarily voiced elsewhere; cf. Arch/Anc. *[sky]i̯ôŋ/tśi̯uŋ*[et] (tone A) 'locust'; cf. WB *kyuiŋ* ~ *gyuiŋ*, id. < PST **(s-)kyu(·)ŋ*[B] (Chinese shows the anticipated tone B > A shift after *s-* < 'animal prefix' **s-*).

22 Cf. also *mi̯u/mi̯u*[fg/fh] 'despise, insult, disgrace', probably from an earlier **m-ki̯u* (isolated under GSR-138 but with apparent phonetic element *mwəg/muâi*[ay] from GSR-947, indicating a labial rather than (velar initial); see text below for evidence pointing to occasional retention of prefixed **m-* before *kl-* and *k(i̯)w-*.

23 PTB prefixed **a-* (zero consonant initial) thus contrasts with **m(a)-*, **b(a)-*, etc. The alternative view of regarding this PTB prefix as **ʔ* is ruled out simply by the fact that **/ʔ/* need not be reconstructed for the proto-language, e.g., it does not occur as a final, although this phoneme has been reconstructed for daughter-languages, notably for PBL (*STC*: fn. 76 by J. Matisoff).

24 The **s-* prefixed form (WT *sdor*) is also represented in Chinese; cf. Arch./Anc. *tsiər/tsieiᶠˣ* 'pickled food', from **sdiər* (cf. text, above, for this shift) < **sdi̯âr* (the palatalized doublet, corresponding to *ʑdiər/niei*) < PST **s-dâr*.

25 The third series noted in the text (GSR-878) has as phonetic *[g]lieg/lietʰᵒ* 'a pair', from a doublet (**gri·k*) of this root (cf. Malay *pasaŋ* 'pair,' also 'fit together') (cf. WT *ʔkrig-pa* 'twins', from the related root with **kr-*); it has a doublet *[g]li̯əg/li̯iᵉʰᵖ* 'twins' (< **grik*) under GSR-979, the phonetic of which is *[g]li̯əg/li̯iᵉʰᵠ*, also read *[s-k']* *li̯əg/xji* 'to crack, split', a doublet of *k'[l]i̯ăk/k'i̯ɒkʰʳ* 'crack, crevice' (GSR-787, containing also *s-k'[l]i̯ăk/xi̯ɒkʰˢ* 'fear' < PST **s-krâk*; see *STC*: fn. 430), the series (GSR-979) also including the puzzling **s[g']li̯əg/dʐ'i* 'spittle (of dragon)' (listed under GSR-1237q).

26 It is probable that many TB/Chinese cognate pairs remain 'disguised' as a result of the ambiguity of reconstruction (see text, above); cf. the following *possible* correspondences, both from brief *xie-sheng* series: (affricate initial) [sg]*i̯aŋ/zi̯aŋⁱᵃ* (< **dzi̯aŋ*) 'elephant'; cf. WT *glaŋ* 'ox; elephant' < PST **(s-)glaŋ* (with 'animal prefix' **s-*); (stop reflex) *[sk]i̯ôg/t'i̯ɨuⁱᵇ* 'wrist, elbow' (*Anal. Dict.* also 'forearm'); cf. PTB **ku[u, əw]*: WT *khru* 'cubit [measure from elbow to tip of middle finger]'; Digaro *lă-kɪau* < **-kru* 'elbow' (*lă-* < PTB **lak* 'arm/hand'); Garo *kru* 'a span or measure of length between the thumb and middle finger' < PST **(s-)krəw* (with 'body-part prefix' **s-*).

27 Further evidence for an earlier velar stop in this series (GSR-807) is supplied by the phonetic: *[sg'y]i̯ăg/dʐ'i̯a* 'shoot with bow; archer' ~ *[sg'y]i̯ăk/dʐ'i̯äkⁱᵐ* 'hit with bow and arrow'; cf. WT *rgyag-pa* 'to throw, cast, fling', *mda rgyag-pa* 'to shoot arrows (*mda*)' < PST **(s-, r-)gyak* (Bodman [1975bis] suggests **ryaks* for the Arch. form).

28 For WT *md-* < **mbl-*, cf. *mda* 'arrow' < PTB **mblo*, contra *STC* (111 and fn. 313), which sets up the root simply as **bla* and assigns WT *mda* and Kachin (Khauri dial.) *niŋda* (< **mda*) to a distinct root **m-da*; the Khauri form appears to be a loan from Tibetan (standard Kachin has *păla* < **bla* whereas Jili [extinct dial.] has *măla* < **mbla*); note also Magari *mya* < **mbla*; WB *hmra* < **s-mbla*, Kha Li (Southern Loloish, *ka-mla* < **-mbla*; the STC suggestion (fn. 469) that Arch./Anc. *di̯ək/i̯əkⁱʰ* 'to shoot with arrow and string attached' is cognate is hardly supported by the *xie-sheng* series involved (GSR-918), although the possibility can scarcely be excluded at this stage of study.

29 PMY has **i̯iŋ* 'field (wet, rice)', apparently an early (pre-Archaic) loan from Chinese (PMY regularly has initial **l-* for Ch. *l-*). WT also has *liŋs* 'a hunting or chase in which a number of people are engaged', tagged as an early loan from Chinese (antedating the final **-i·ŋ* > *-ien* shift) by the unpalatalized initial before **i* (contrast *źiŋ* 'field') and the suffixed *-s* (very characteristic of these early loanwords)

30 For suffixed **-n* with this body-part word, cf. *ŋan* 'face' < **ŋâ-n* (fn. 13); also Hakka *lin* 'penis', cognate to PTB **(m-)li*, id., as well as to PK **lin* 'vagina' (a semantic interchange also found in AT), offering additional evidence for the existence of this suffix at the PST level.

31 The vast efflorescence of the **s-* prefix pattern, to include numerals and most kinship terms as well as other lexical categories, certainly appears to have been a Chinese innovation. A considerable share of these roots have been well 'disguised' by phonological shifts (see text), especially where initial dentals are concerned, e.g., PTB **ta* 'father/grandfather': WT *ʔa-ta*, WB *thà-thà* 'father' but Digaro *-ta*, Chepang *to* < **ta*, Miri *ətɔ* < **a-ta* 'grandfather'; Arch./Anc. *tso/tsuoᵏᶻ* 'grandfather, ancestor', from **sto* < **s-ta*; *ts'ien/ts'ienᵗ* 'thousand' (above), from *st'ien* < **st'ieŋ*, an early loan (via **s-th[r]iaŋ*) from PAT **(k-)triaŋ*: PT **[t]hriaŋ* 'thousand (Ahom Khamti Shan), million (Dioi)'.

References

Bauman, J. 1974. "Pronominal Verb Morphology in Tibeto-Burman," *Ling. of the Tibeto-Burman Area*, Vol. 1, No. 1, 108–55.

Benedict, Paul K. 1967. "Austro-Thai Studies: 3. Austro-Thai and Chinese," *Behavior Science Notes* 2, 275–336.

———. 1972. *Sino-Tibetan: A Conspectus* (contr. ed. J. Matisoff), Cambridge: Cambridge University Press.

———. 1972bis, "The Sino-Tibetan Tonal System." In *Langues et techniques, nature et société*, Vol. 1, Paris: Editions Klincksieck.

———. 1973. "Tibeto-Burman Tones: with a Note on Teleo-reconstruction," *Acta Orientalia* 35, 127–38.

———. 1973bis, "The Proto-Sino-Tibetan (PST) Reconstruction." Paper presented at the Sixth Sino-Tibetan Conference, La Jolla, October.

———. 1974. "The Chinese *s-* Orgy." Paper presented at the Seventh Sino-Tibetan Conference, Georgia State University, October.

———. 1975. *Austro-Thai: Language and Culture*, New Haven: HRAF Press.

———. 1975bis, "The Chinese *s-* Orgy: Further Adventures and Misadventures." Paper presented at the Eigth Sino-Tibetan Conference, Berkeley, October.

Bodman, N., 1967, "China: Historical Linguistics," *Current Trends in Linguistics* 2.

———. 1969, "Tibetan *sdud* 'folds of a garment', the character [ ]ᵃ and the **st-* Hypothesis," *BIHP, Academia Sinica*, 39.2.

———. 1973. "Some Chinese Reflexes of Sino-Tibetan *s-* Clusters," *JCL* 1, 383–96.

———. 1974. "Some Random Observations on Paul Benedict's 'The Chinese *s-* Orgy." Paper presented at the Seventh Sino-Tibetan Conference, Georgia State University, October.

———. 1975. Review of *Sino-Tibetan: A Conspectus* (Benedict), *Linguistics: An International Review*, 149, 89–97.

———. 1975bis, "Tibeto-Burman Correspondences to the Chinese *teng* (Divisions 1, 2, 3, 4) and the Concept of 'Primary Yod' in Sino-Tibetan." Paper presented at the Eighth Sino-Tibetan Conference, Berkeley, October.

Burling, R. 1967. Proto-Lolo-Burmese. *IJAL* Vol. 33, No. 2, Part II, Bloomington, Indiana.

———. 1971, "The Historical Place of Jinghpaw in Tibeto-Burman," *OPWSTBL* Vol. II, Urbana: University of Illinois: Dept. of Linguistics.

Burling, R., and P. C. Bhattacharya, 1956, "Lexico-Statistic Dating of Garo-Bodo Separation," *Journal of the University of Gauhati* 7, 67–73.

Chang, B. 1971. "The Tibetan Causative: Phonology," *Academia Sinica: BIHP* 42, 623–765.

Chang, K. 1969. "Sino-Tibetan Words for *needle*." Paper presented at the Second Sino-Tibetan Conference, Columbia University, October.

———. 1972. "Sino-Tibetan 'iron': **qhleks*," *JAOS* 92, 436–46.

———. 1973. Review of *Sino-Tibetan: A Conspectus* (Benedict), *JAS* 32, 335–37.

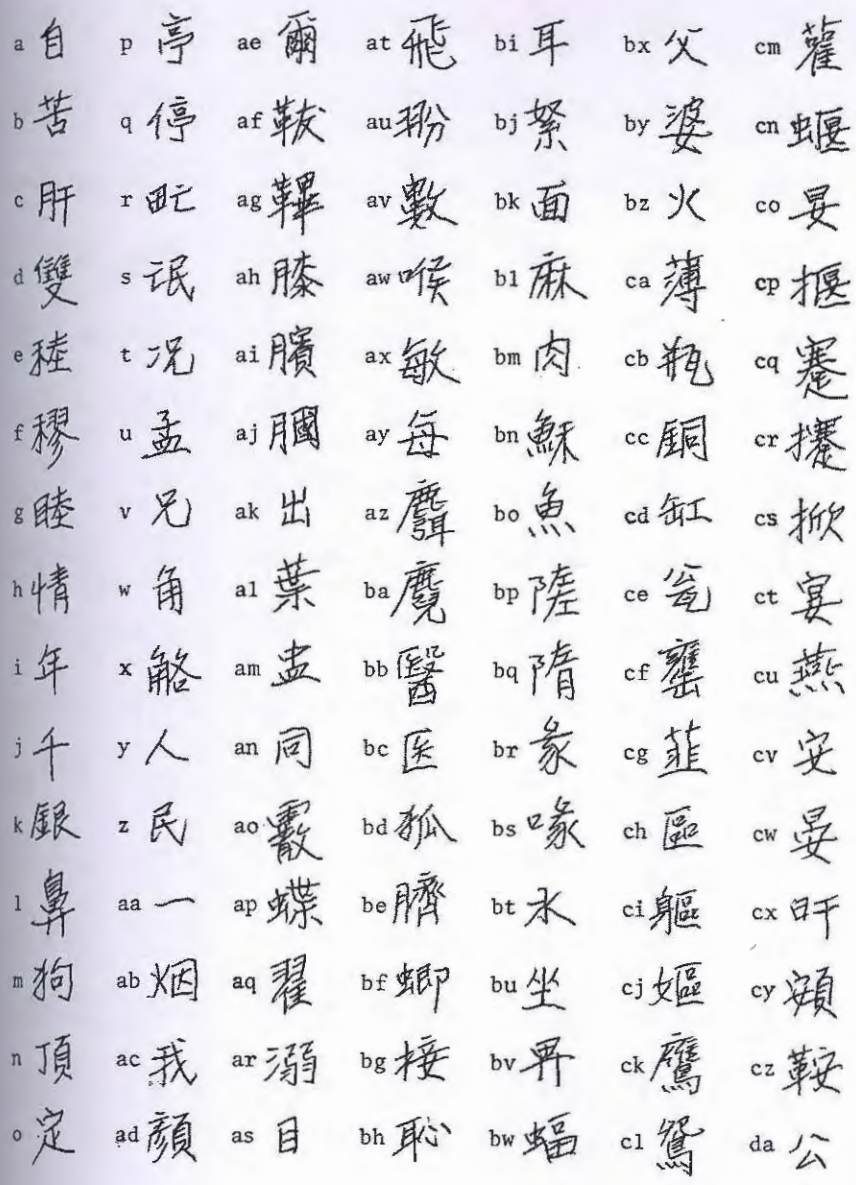
Chou, Fa-kao, 1972. "Archaic Chinese and Sino-Tibetan" (in Chinese), *Journal of the Institute of Chinese Studies of the Chinese University of Hong Kong*, Vol. V, No. 1, 159–237.

Coblin, W. S. 1972–73. Review of *Sino-Tibetan: A Conspectus* (Benedict), *Monumenta Serica* 30, 635–42.

Denwood, P. 1974. Review of *Sino-Tibetan: A Conspectus* (Benedict), *JOS*, 261–62.

Downer, G. B. 1971. "The Further Relationships of the Miao-Yao Languages." Paper presented at the Fourth Sino-Tibetan Conference, Indiana University, October.

———. 1973. "Strata of Chinese Loanwords in the Mien Dialect of Yao," *Asia Major* 18, 1–33.
Egerod, S. 1973. Review of *Sino-Tibetan: A Conspectus* (Benedict), *JCL* 1, 498–505.
———. 1974. "Sino-Tibetan Languages," *Encyclopaedia Britannica*, 15th ed.
Forrest, R. A. D. 1973. *The Chinese Language*, 3rd ed., London: Faber and Faber and Faber Ltd.
Haudricourt, A. 1954. "Les origines asiatiques des langues malayo-polynésiennes," *Journal de la Société des Oceanistes* 10, 180–83.
———. 1973. Review of *Sino-Tibetan: A Conspectus*, *BSLP* 68, fasc. 2, 494–95.
Lehman, F. K. 1975. Review of *Sino-Tibetan: A Conspectus* (Benedict), *Lang.* 51: 215–19.
———. 1975bis, "Wolfenden's Non-pronominal *a-* Prefix in Tibeto-Burman: Two Arguments from Southern Chin and Some Proposed Semantic Correlates," *Ling. of the Tibeto-Burman Area*, Vol. 2, No. 1, 19–44.
Li, Fang-kuei. 1937. "Languages and Dialects of China," *Chinese Yearbook*, Shanghai; reprinted in *JCL* 1 (1972), 1–13.
———. 1971. "Studies on Archaic Chinese Phonology" (in Chinese), *Tsing Hua Journal of Chinese Studies*, New Series 9.1–2.
———. 1974. "Tai Languages," *Encyclopaedia Britannica*, 15th ed.
Matisoff, J. 1972. *The Loloish Tonal Split Revisited*, Berkeley: University of California, Center for South and Southeast Asia Studies, Research Monograph No. 7.
———. 1973. "The Annual Sino-Tibetan Conferences: The First Five Years, 1968–1972," *JCL* 1, 152–62.
———. 1973bis, "Notes on Fang-kuei Li's 'Languages and Dialects of China'," *JCL* 1, 471–74.
———. 1974. "The Tones of Jinghpaw and Lolo-Burmese: Common Origin vs. Independent Development," *Acta Ling. Haf.* 15.2, 153–212.
———. 1975. "Benedict's Sino-Tibetan: A Rejection of Miller's Conspectus Inspection," *Ling. of the Tibeto-Burman Area*, Vol. 2, No. 1, 155–72.
Miller, R. 1974. "Sino-Tibetan: Inspection of a Conspectus," *JAOS* 94, 195–209.
Norman, J. 1973. "Tonal Development in Min, "*JCL* 1, 222–38.
Purnell, H. 1970. *Toward a Reconstruction of Proto-Miao-Yao*, Ithaca: Cornell University, unpublished Ph.D. dissertation.
Sedláček, K., 1974, Review of *Sino-Tibetan: A Conspectus* (Benedict), *ZDMG* 124, 205–6.
Shafer, R. 1964. "Miao-Yao," *Monumenta Serica* 23, 398–411.
———. 1966–67. *Introduction to Sino-Tibetan*, Wiesbaden: Otto Harrassowitz.
Simon, W. 1974. "Loss of *l* or *r* in Tibetan Initial Consonantal Clusters," *BSOAS* 37, Part 2, 442–45.
Sprigg, W. K. 1973. Review article of *Sino-Tibetan: A Conspectus* (Benedict), *Asia Major* 19, Part 1, 100–6.
Thurgood, G. 1974. "Lolo-Burmese Rhymes," *Ling. of the Tibeto-Burman Area*, Vol. 1, No. 1, 98–107.
Yang, P. 1968. *Modern Dialects and the Reconstruction of Archaic Chinese (Initial Consonant Clusters)*, Washington, D. C.: Georgetown University (mimeographed).
———. 1975. "Prefix **s-* in Proto-Chinese." Paper presented at the Eighth Sino-Tibetan Conference, Berkeley, October.



db 翁 dq 絺 ef 吞 eu 鞅 fj 堯 fy 鬱 gm 靦
dc 姒 dr 希 eg 咽 ev 膺 fk 喬 fz 燃 go 顎
dd 擁 ds 貪 eh 昜 ew 盎 fl 蹻 ga 燀 gp 各
de 崧 dt 今 ei 陽 ex 俞 fm 跲 gb 戇 gq 叉
df 嵩 du 脛 ej 鄉 ey 偷 fn 𠯿 gc 戰 gr 搔
dg 羊 dv 巠 ek 向 ez 寇 fo 難 gd 怒 gs 軋
dh 妻 dw 腎 el 煬 fa 扣 fp 嘽 ge 豻 gt 札
di 羌 dx 臤 em 揚 fb 歐 fq 殫 gf 密 gu 牙
dj 羹 dy 十 en 興 fc 媮 fr 癉 gg 祕 gv 鵝
dk 膕 dz 召 eo 昜 fd 呫 fs 癙 gh 閟 gw 無
dl 湯 ea 莒 ep 产 fe 詬 ft 儺 gi 盜 gx 撫
dm 笑 eb 菌 eq 勻 ff 詢 fu 嘽 gj 巫 gy 武
dn 夭 ec 甚 er 杓 fg 侮 fv 那 gk 誣 gz 鵡
do 嗄 ed 天 es 汐 fh 佣 fw 鬻 gl 硯 ha 賦
dp 夏 ee 祆 et 螽 fi 羭 fx 龖 gm 見 hb 屎

hc 首 hr 隙 ig 鼎 iv 朔 jk 旄 jz 銛 ko 輕
hd 殼 hs 虢 ih 膞 iw 并 jl 聿 ka 甜 kp 蛭
he 廠 ht 鼕 ii 舌 ix 弓 jm 律 kb 申 kq 挃
hf 虎 hu 凍 ij 舐 iy 繞 jn 筆 kc 鐵 kr 繆
hg 處 hv 僉 ik 咭 iz 鑾 jo 盈 kd 體 ks 腋
hh 居 hw 瘳 il 麝 ja 孿 jp 繩 ke 豐 kt 脢
hi 虛 hx 樂 im 射 jb 琴 jq 黽 kf 身 ku 灰
hj 森 hy 藥 in 玉 jc 必 jr 蠅 kg 离 kv 恢
hk 林 hz 鹽 io 琭 jd 龍 js 田 kh 神 kw 袂
hl 牽 ia 象 ip 綠 je 龐 jt 雨 ki 電 kx 法
hm 膟 ib 肘 iq 頊 jf 龐 ju 卵 kj 僕 ky 去
hn 史 ic 咍 ir 鸑 jg 聾 jv 頟 kk 襟 kz 祖
ho 麗 id 賴 is 雒 jh 弋 jw 令 kl 靨 la 辛
hp 弊 ie 苔 it 愕 ji 龙 jx 欖 km 𠯿
hq 弊 if 桃 iu 憩 jj 龍 jy 鹵 kn 姪

7

ON MEGALOCOMPARISON

James A. Matisoff

Source: *Language* 66, 1, 1990, 106–20.

Introduction: the war of metaphors

1. Joseph H. Greenberg's *Language in the Americas (LIA;* 1987) has been greeted with dismay by many specialists in Amerindian linguistics[1] (cf. Chafe 1987, Campbell 1988, Adelaar 1989), and defended by the author in a reply to Campbell in *Language* 65.1 (1989).

There is no denying that G's central thesis—that all the native language families of the Western hemisphere (except for Eskimo-Aleut and Na-Dene) are genetically related, descending from a common ancestor, 'Proto-Amerind', spoken about 11,000 years ago (335)—has a certain romantic sweep to it. Such a far-reaching claim would seem to require pretty convincing evidence to back it up. Yet G expresses his disdain for the conventional sort of historical comparison, so 'obsessionally' preoccupied with sound correspondences and asterisked reconstructions.[2] Instead he values the kind of evidence provided by the more 'powerful' method of 'mass comparison', whereby wordlists and grammatical paradigms from many languages are grokked simultaneously to see whether any root-words or affixes look alike. Languages which share a number of resemblant vocabulary items are then deemed to be genetically related, unless these forms are obvious loans (22–23).

So powerful is this method, G believes, that it yields valid results even with the worst data, and to any desired time-depth: 'The method of multilateral comparison is so powerful that it will give reliable results even with the poorest of materials. Incorrect material should have merely a randomizing effect' (29). Errors do not greatly concern G: 'Although I have exercised great care, it would be miraculous if, in handling such a vast amount of material, there were no errors of fact or interpretation' (1989:112). To this G's critic Adelaar (a Quechua specialist) rejoins that 'This is highly questionable if one looks at the quality of the data G presents where the number of erroneous forms probably exceeds that of the correct forms' (1989:253). The consensus among professional Amerindianists seems to be that G has NOT in fact 'exercised great care' in the selection or utilization of his materials, and that he has ignored the results of the best recent research on many topics.

The most damaging accusation against *LIA* is that its methodology and data are so inadequate that they are incapable of distinguishing between similarities that might be due to genetic relationship and those which are due to pure chance (see §2.4 below).

Yet in a way all this criticism of *LIA* misses the point. Some scholars are impervious to criticism because they are totally convinced that their underlying thesis is correct, regardless of any disconfirmatory factual 'details' that might be wrong.[3] In G's case this conviction derives from a larger belief system that far transcends the narrow confines of the New World, but embraces our whole planet and our whole species. Although G provisionally recognizes about 15 'major linguistic families' in the world (Table 11, 337), it is clear that in his heart he believes he can show that all human languages descend ultimately from a single common ancestor (62), since 'there is no theoretical limit to the depth at which classification can be carried out when the number of languages examined is large' (28–29). *LIA* ends on this very note: 'The ultimate goal is a comprehensive classification of what is very likely a single language family. The implications of such a classification for the origin and history of our species would, of course, be very great' (337).

Proto-Amerind is only the tip of the iceberg. We are promised another book shortly (G, forthcoming) that will prove that Indo-European itself belongs to a much vaster language family called 'Eurasiatic', which includes Japanese, Ural-Altaic, and Eskimo-Aleut (332). For many linguists, such views fall more into the category of religious beliefs than scientifically testable hypotheses, about on a par with claims that 'all languages have the same underlying deep structure' and 'the position of the stars at the moment of our birth determines our character'.

Since *de dogmatibus non disputandum est,* we must concede that *LIA* cannot really be judged by conventional criteria. Much of the controversy surrounding the book is better considered to be a dialectic between sets of opposing personality types and intellectual bents: battles of images, metaphors, temperaments.

Certain traditional dichotomies ('theory-oriented' vs. 'data-oriented' or 'lumpers' vs. 'splitters') do not quite succeed in capturing the differences between the opposing camps. G has his theories as well as his data; and many of G's critics are themselves lumpishly inclined, and would like nothing better than to succeed in demonstrating a genetic relationship between two families previously thought to be unrelated. I have made a slightly more nuanced tripartite distinction among three sorts of linguistic comparison ('micro-', 'macro', and 'megalo-'), according to the putative closeness or remoteness of the genetic relationship being investigated, and have observed that 'different species of maniacs' are attracted to each variety (Matisoff 1976:258; 1979:38).[4] G seems to scorn those who work at submegalo levels of classification as focusing 'on a limited group determined by accidents of expertise' (4).[5] He has no taste for tiny details.

Still another relevant dichotomy is between armchair- or library-scholars on the one hand (let us call them 'fauteuillistes' for short) and fieldworkers on the other. G is definitely a library-type scholar, justifiably proud of his Sitzfleisch in having gone through 'well over 2000 sources' during the preparation of *LIA.* The

fieldworker *is* only too bitterly aware of the inadequacies and messiness of his own material, *how* the accidents of elicitation cause one form rather than another to be recorded *in* one's notes during a too-brief stay in some village. The armchair linguist sees *only the* list of words on the page, authoritative and invariant, artificially neat.

All this having been said, if I were to pick a single word to describe G's apparent motivation in doing megalocomparison, it would have to be COLUMBICUBICULOMANIA—a compulsion to stick things into pigeonholes,[6] to leave nothing unclassified. G gives the impression that the highest intellectual activity is the act of classification itself, regardless of the nature of the evidence upon which the classification rests: 'Basically, the wrong question has been asked, namely, when are languages genetically related? . . . What should be asked is, how are languages to be classified genetically?' (3). He is fascinated with the astronomically high number of mathematically possible ways there are to classify a relatively small set of objects (6). Once a single overarching classification has imposed order on this chaos, the really interesting part of the work is apparently over. Let others worry about the trivial details of classification at lower taxonomic levels.

Methodological rigor in a messy world: should megalocomparison be easy or hard?

2.1. Mixed languages, diffusion, complexity, and messiness

G refers with scant respect to believers in mixed languages like C. Loukotka (38), and seems uncomfortable with notions like linguistic areas, or the diffusion of lexical and grammatical traits across genetic boundaries. For anybody who has worked in a hothouse diffusional environment like Southeast Asia, it seems self-evident that big chunks of even a language's basic vocabulary[7] and grammar can quickly and easily be remodelled under the influence of any languages with which it comes in contact, whether or not such contact languages were genetically related to it in the first place.[8] Borrowing, conflation/contamination/blending, folk etymology, semantic slippage, calquing, backloans—all kinds of messy phenomena complicate matters. These are not marginal, nor, I submit, anything to be ashamed of.[9]

It is not even necessary to travel to exotic climes to convince oneself of the rapidity with which the genetic pathways of words can become obscure. It can happen even with neologisms, before one's very eyes, in less than a generation, in one's native language. How many opinions are already current among the American Volk on the etymology of, say, *T-shirt*?[10] If we are so unsure about T-shirts, how can we be so sure about etymological connections that purportedly go back 11,000 years?

2.2. Strictness vs. laxity and the consequences of error

G likes to caricature the traditional historical linguist as a timid soul who dares not venture beyond the lowest level of comparison, considering languages 'pairwise',

only two at a time, afraid to tackle huge piles of languages at once. In fact, however, sober-minded scholars have shrunk from megalocomparisons not because they are so difficult, but because they are so easy.[11] When the number of languages being considered is large, when their relationship (if any) is remote, and the criteria for sound correspondences are lax, it is not very hard to find 'phono-semantic lookalikes'—forms which more or less resemble each other both in sound and in meaning.

G's faith that multiple errors will somehow have a randomizing effect makes it easy to ignore criticism on any particular point. Other metaphors may be invoked, however—instead of 'cancelling each other out', errors may be COMPOUNDED. The controllers of Voyager II didn't dare trust that any little errors in its trajectory would 'cancel each other out'. A tiny error on the cosmic scale grows to parsecs before you know it. Other images for the dynamics of multiple errors include the Two Drunks Supporting Each Other, the House of Cards, and the information-processing slogan 'Garbage in, garbage out'. (We shall return to this issue in connection with 'teleoreconstruction' in §4.2.)

2.3. Subgrouping without reconstructions

Many still remember the brilliant plan proposed by Sen. Aiken (R-Vt.) to bring the Vietnam War to an end. 'Let us,' said the Senator, 'simply declare that we've won, and bring the troops back home.' In a way G's thesis, if accepted, would have a similar effect, by cutting off any further discussion about WHETHER any Amerindian language was related to any other. The only interesting issues remaining would be HOW a language fit into the family as a whole. SUBgrouping then becomes the only meaningful activity for Amerindianists.

But what has been gained? G admits cheerfully enough that 'the internal subgrouping of the 11 Amerind groups remains largely unknown' (64). Isn't having no idea of HOW the language families of the Western Hemisphere relate to each other pretty similar to not knowing WHETHER they are related to each other or not? And how will Proto-Amerind ever be subgrouped without reconstructions of the individual subgroups?

It has often been pointed out[12] that genetic nonrelationship can never be proven. G calls this profound observation 'uninteresting', and counters by saying that you CAN 'disprove erroneous claims of relative degrees of relationship, e.g., 'Nahuatl is closer to Swahili than to Pima', since 'Nahuatl and Pima . . . belong to a valid genetic group (. . . Uto-Aztecan) that does not include Swahili' (5). Yet one can easily imagine a more subtle Gedankenexperiment, where the nature of the genetic grouping, if any, is not obvious in advance. Suppose we take three languages (A,B,C), all quite different from each other (though similarities exist between any dyad of them, as well as among all three), spoken in widely separated geographic areas, and none of whose further genetic affiliations are known with certainty. How can we tell whether A is more like B than A is like C or than B is like C? G seems to imply that mere eyeballing is always enough to determine the

subgrouping of entire languages, just as it suffices to establish the relative phonological resemblance among given lexical items: 'Is a form A more like B than it is like C? Given, for example, *pan / fan / ezuk,* who would hesitate?' (5). But how can this bravado be sustained in the case of even slightly more difficult examples, such as *pan / fan / vin* or *bol/biŋ/bo?* Or *èr / erku / duo?*[13] Doesn't one sometimes have to hesitate a bit?

G, perhaps making a virtue of necessity, belittles the importance of phonological reconstruction: '. . . anything approaching a complete and highly convincing reconstruction on the basis of recurrent correspondences is in general possible only with languages so closely related that it is unnecessary anyway' (33). What then is 'necessary' and what is sufficient in historical linguistics? For G, creating the overarching classification is the only intellectual act of any interest. The details can be left to the drones working amid the pigeon-droppings in the lower cubicles.

Quite a different mind-set is displayed in the *Dravidian Etymological Dictionary (DED)* of Burrow & Emeneau (1961). Although this great work contains almost 6000 cognate sets, it does not offer any reconstructions either—but this is simply because, after many decades of toil, the authors felt they had not yet resolved all the problems in the reconstruction of the vowels.[14] In every other respect the *DED* presents a maximal contrast to *LIA:* the sets of forms presented are truly cognate, and any irregularities are at least identified and possible explanations are suggested. Errors are at a minimum. Of course the time depth of Proto-Dravidian is much shallower than G's Proto-Amerind, and the validity of the Dravidian classification could just as well have been established by looking at 100 sets instead of 6000, so from G's point of view the whole book was unnecessary in the first place—though it is sure to come in handy when G tries in some future book to subsume Dravidian under a larger family!

Given G's aversion to reconstruction, it is curious that he still operates with the notion of COGNATES: 'The broad approach advocated here does not require the reckless positing of risky and uncertain etymologies. All that is needed is to show decisively more cognates than those of any rival hypothesis' (37). But how are we to decide which etymologies are totally convincing and which are 'risky and uncertain', if we arrived at them by naked eyeballing in the first place? How is a cognate different from a chance lookalike? How do we judge among competing etymologies for the same word? Do lots of erroneous cognate identifications outweigh a few solid ones? If we put in ENOUGH garbage does good stuff eventually come out?

2.4. Distinguishing chance similarities from relationship

G seems to be tilting at a straw man when he attacks '. . . the simplistic, but widely held, assumption that after a not very long period the resemblances between two related languages become indistinguishable from chance. This would be true only if there were just two languages in the world' (1989:109). We are in fact supposed

to be talking about VERY long periods indeed! Nobody is claiming that all traces of genetic relationship can be obliterated in a short time.

Everybody has a favorite list of accidental lexical resemblances (e.g. Thai *faj, taaj, rim*/Eng. *fire, die, rim*). Manipulators of these curiosities may now add G's 'final piece of preliminary evidence' (57) for the reality of the Amerind family, the 'Amerind etymology "hand, give, take"' (Table 9, 58), containing beautiful proto-Sapiens forms like Mayna (Andean) *mani* 'arm', Akwa'ala (Hokan) *man* 'hand', etc. Many scholars have attempted reductiones ad absurdum of megalocomparison by taking any two languages at random and finding large numbers of 'cognates' between them, often to very amusing effect (e.g. Callaghan & Miller 1962). Campbell (1988:602–3) tries the same gambit with Finnish and Amerind, but G is distinctly not impressed: 'I would never compare Finnish in isolation. If Finno-Ugric and the larger Uralic group to which it belongs were not already recognized, I would have discovered them' (1989:111). Can it be that G is missing the point here? THIS CHANCE COMPARABILITY OPERATES EQUALLY WELL BETWEEN ANY TWO PROTO-LANGUAGES CHOSEN AT RANDOM. One could undertake systematic comparisons between, e.g., Proto-Sino-Tibetan and 'Proto-Amerind', and find a dozen or so really snazzy lookalikes. (I will refrain from listing the ones I have found, for fear of providing grist for anybody's mill.[15])

G does seem to recognize the difference between accidental and genetic relationship: 'Of course there are occasional resemblances [between Amerind and Na-Dene and Eskimo-Aleut], attributable either to accident or to common membership in some still deeper grouping' (61). But he then goes on to give a wildly fantastical example (he calls it 'instructive') of an item which seems to him to be a good candidate for the Proto-Sapiens lexicon: an etymon meaning 'hand; finger; to point (with the finger); one', attested by Amerindian forms like Karok *ti:k* 'hand; finger', Yagua *tiki* 'one', Eskimo *tik-iq* 'index finger', as well as by Proto-Indo-European **deik-*'point', Proto-Sino-Tibetan **tik* 'one', and Nilo-Saharan forms like Maba *tek,* Fur *dik* 'one'. 'It may be more widely distributed for aught I know . . . it is likely that some of these [resemblances] reflect a common inheritance from a very extensive family—which may even be proto-Sapiens' (62).

To soften up the reader for the lexical presentations to come, G first devotes some space to morphological comparisons (44–57), discussing 'a few widespread grammatical markers . . . often involving shared irregularities' (44), e.g. the pervasive Amerindian pattern of pronominal nasal morphemes, *-n-* for 1st person and *-m-* for 2nd person. He returns to this topic in his Reply to Campbell: 'This distribution cannot be explained either by borrowing or chance. The borrowing of first- and second-person pronouns is very rare' (113). Again one wonders how G can be so sure about rejecting alternative possibilities. As far as chance goes, one could point to the quite similar and equally pervasive pattern of nasal morphemes found in those Tibeto-Burman languages with pronominal agreement systems, i.e. *-ŋ-* for 1st person, *-n-* for 2nd person—reduced from the full forms of the personal pronouns, PTB **ŋa-y* ('I; me') and **naŋ* ('thou; thee'), respectively. (The correspondences Amer. *-n-*/TB *-ŋ-* and Amer. *-m-*/TB *-n-* even show a nice

parallelism—in both cases the Amerindian nasal is 'one place further to the front' in point of articulation than the ST one!) As for the 'unborrowability' of 1st or 2nd person pronouns, one need only point to modern Thai students' slang, where the English-derived pronouns *ʔaj* 'I' and *juu* 'you' have steadily been gaining ground, partly out of exotic chic and partly as a welcome way of avoiding the elaborate distinctions built into the native pronominal system (see Cooke 1968:11, 15). At a more ancient time-depth we could mention the polite 1st-person pronouns for male speakers in Burmese (*cun-to;* 1st syll. < Old Bs. *kywan*) and Cambodian (*khñom*), which both originally meant 'slave'. Here the cultural notions of hierarchy that diffused throughout 'Hinduized' SE Asia are responsible for the parallel grammaticalizations of 'slave' to '1st person pronoun'—though the Burmese and Khmer words themselves are of course not cognate, and one is not borrowed from the other.

In fact, the more languages one looks at, the more accidental resemblances one will find in the phonological shapes of semantically similar functors and affixes—perhaps even more accidental similarities than in the case of root morphemes. It seems impressionistically that formatives with long-lived sounds like nasals and *-s-* are particularly abundant in the world's languages—perhaps because these sounds are best equipped to withstand the attrition brought about by high textual frequency.[16]

In any event, functors or grammaticalized morphemes are no more immune to accidental phonological similarities than are root morphemes.

2.5. Distinguishing areal from genetic relationship

G's methods lack the subtlety to distinguish similarities that are due to typological and areal factors from those that reflect genetic relationship. One can only imagine what he would do when confronted with a complex linguistic area like Southeast Asia, a region that is home to several quite distinct but highly ramified language families that have undergone mutual influence for millennia. One of the most striking areal features of 'Sinospheric' SE Asian languages is monosyllabicity and elaborate tone systems.[17] Tai, Hmong-Mien (Miao-Yao), and Vietnamese all have Chinese-type tone systems and thoroughgoing monosyllabicity, and all share a good-sized lexical component (including some core vocabulary) which corresponds more or less regularly in consonants, vowels, and tones. If this were brought vividly to G's attention he would probably accept it at once as conclusive evidence for the genetic relationship of all these languages (especially since he would have little to fear from contradictory morphological evidence, in view of the rudimentary morphology of Sinospheric languages). Yet there is overwhelming evidence that Vietnamese belongs to the Mon-Khmer (Austroasiatic) family (quite unrelated to Chinese), while Tai and Hmong-Mien, while perhaps ultimately related to each other, are more plausibly grouped with Austronesian than with Chinese.[18] Austronesian and Austroasiatic have a number of apparently deep grammatical features in common, including a causative morpheme in *-p-* and infixation as an ancient morphological process, yet share practically no

core vocabulary. In any case, the chief interest of SE Asian diachronic linguistics lies in attempting to unravel the threads of diffusional vs. genetic vs. accidental factors in the rich areal fabric. Were we to cut the Greenbergian knot and assume from the outset that all these language families are ultimately related, we would be led into terrible dead ends, and much of the interest would go out of this fascinating field.

Some proposals for remote linguistic relationships in Asia: the case of Japanese

3. For G, 'GENETIC RELATIONSHIP IS PLAINLY TRANSITIVE, so that if Baltic is related to Slavic, and also to Germanic . . . then Slavic must be related to Germanic . . .' (26; emphasis mine). Sometimes, however, this principle seems to amount to accepting and combining all genetic claims that have ever been made by anybody: if Linguist A suggests that language families X and Y are related, and Linguist B suggests that families Y and Z are related, that means ipso facto that X and Z are related—regardless of the independent validity of either dyadic grouping.[19] It is as if any two genetic hypotheses involving the same language automatically reinforce each other, permitting G to make ever broader groupings.

Yet as the proverb goes, a chain is no stronger than its weakest link. If either the XY or the YZ grouping is shaky to begin with, the XYZ chain can be no stronger. Consider, e.g., Japanese—arguably the world's most culturally important language whose genetic affiliations are still controversial. At least 7 theories have had their adherents: (1) HELIOPARTHENOGENESIS. The Japanese people descend from the Sun Goddess and were dropped down onto the Japanese islands along with their language at the beginning of time. (2) JAPANESE-DRAVIDIAN. The respected Japanese linguist Ohno Susumu has written a book (1980) attempting to demonstrate genetic relationship between Japanese and Tamil.[20] (3) JAPANESE-TIBETO-BURMAN. The eminent Japanese linguist Nishida Tatsuo has argued rather for a genetic relationship between Japanese and another SOV language family, Tibeto-Burman (1978), even attempting morpheme by morpheme comparisons of TB post-verbal particle strings with Japanese inflectional suffixes. (4) JAPANESE-AINU-KOREAN-CHINESE. Johannes Rahder spent years compiling pages of putative Chinese, Korean, and Ainu cognates to Japanese words (see, e.g., 1956/1959), comparing forms from modern Chinese dialects with selected syllables of semantically similar Japanese words. (5) JAPANESE-KOREAN-ALTAIC. Building on the work of S. Martin, who compared Japanese and Korean (1966), R. A. Miller develops this genetic theory in a supremely confident book with the question-begging title *Japanese and the Other Altaic Languages* (1971). (The membership of Korean itself in the Altaic family is still highly controversial.) (6) JAPANESE-AUSTRONESIAN. The simple CVCV structure of Japanese, as well as the geographic unity of the Japan-Ryukyuan-Taiwan string of volcanic islands, has long suggested a relationship between Japanese and Austronesian (Labberton 1924, Solomon 1974, Kawamoto 1977/1978), and this idea has recently been elaborated in stupefying detail by Benedict (forthcoming), who now

considers Japanese to belong to his 'Austro-Tai' family, which also includes Tai and Hmong-Mien (see below). (7) G believes that Japanese is but one of 9 subgroups of EURASIATIC, a great family which also includes Eskimo-Aleut and Indo-European.

Leaving aside (1) for the moment, we may apply G's 'principle of transitivity' to the above theories, removing any apparent contradictions among them by assuming they are ALL valid—i.e. that Dravidian, Sino-Tibetan/Tibeto-Burman, Altaic, Ainu, Austro-Tai, and Indo-European are all related to Japanese and to one another. We can be sure that it is by such leaps of transitivity that G will progress toward his ultimate goal of linking up all human languages into the Sapiens family.

Greenberg and Benedict compared and contrasted

4. It is instructive to compare G to the foremost megalocomparativist now working on East and SE Asian languages, Paul K. Benedict (B), whom G cites with approval (62, 336). B, like G, views himself as something of an outsider who uses common sense to beat the specialists at their own game. Both are armchair linguists and voracious wordlist consumers.[21] Both have anthropological backgrounds and were decisively influenced by A. L. Kroeber. And both are always absolutely convinced that they are right. Yet while G's chief intellectual pleasure seems to lie in the act of classifying or pigeonholing, B's central metaphor is rather the jigsaw puzzle. His favorite cry of etymological triumph is 'Fits perfectly!' (Never mind if a piece of the puzzle has to be rammed in by force now and then.)

The crucial difference between G and B is that B is in fact a brilliant practitioner of the traditional comparative method. When the scale of comparison is not too vast—i.e. the macrolevel as represented by Proto-TB or Proto-ST—B's instincts are sure, and he is usually right. The problems arise at the 'megalolevel', as represented by Austro-Tai-Japanese (Benedict 1975; forthcoming). Here B's desire to make all the puzzle pieces fit together neatly causes him to attempt the impossible: to give the impression of micro-rigor when dealing with (even poorly recorded) megalo-data. Even at the megalolevel B pays homage to the traditional trappings of the comparative method. He claims he is establishing 'regular' correspondences, he provides asterisked reconstructions bristling with parentheses, slashes, and brackets. Sometimes these are so complex and arbitrary that one feels like calling them 'pseudo-micritizing devices'—notational attempts to make the speculative seem rigorous. To put it another way: G sweeps difficulties under the rug, while B offers an instant 'explanation' for any difficulty. B can only cry 'Touché!' to G's remark, 'For those who see reconstruction as proof, there are so many quite legitimate ways of explaining what are apparently irregular correspondences that there is no empirical way of disproving a reconstruction' (10).

4.1. Pseudo-micritizing devices

One way of ensuring apparent 'regularity of correspondence' is to reconstruct proto-forms that are so complex canonically (e.g. containing long consonant

clusters, or even several syllables) that no given combination of proto-entities is likely to recur very often—thus obviating counterexamples. I like to call this 'proto-form stuffing'.[22]

In B's conception, the atonal polysyllables of Proto-Austro-Tai (PAT) were simplified in the Tai-Kadai and Hmong-Mien branches into tonal monosyllables, presumably under the overwhelming areal influence of Chinese. As one can imagine, with these complex polysyllables for proto-forms and these simple monosyllables for reflexes, the etymological possibilities are endless.[23]

If we add to this a rich apparatus of bracketings to optionalize virtually every portion of the proto-form, we are then home free. Thus a huge panoply of reflexes can be accommodated under B's PAT etymology RED/DARK-COLORED/REDDEN/SHAME(D), reconstructed thus (1975:361): **iɣaŋ*; **(q/)b/iɣaŋ*; **(q/)m/iɣaŋ*; **iɣa(ʔ)*; **m(a)/iɣa(ʔ)*; **[i]ɣa(ʔ)/i*; **m(a)/iɣa(ʔ)/i*. There is, for instance, no problem in identifying as cognate Proto-Mien **ʔńay* 'ashamed' and Proto-Tai **hmliaŋ* 'rust', the former via **q/mryay* and the latter via **q/mriyaŋ* < **q/m/(i)yaŋ*.[24] Japanese, with its short and phonologically simple morphemes, is also gratifyingly easy for B to fit into the Austro-Tai picture, as with HAIR/BAST/HEMP/BEARD/EYEBROW (B forthcoming: 233–5): PJse/AT **[qa-](n)tsa(m)bo[t,c]* 'bast; hair; feather > pre-Jse. **[q]a-sa[wo]* > Jse. *asa* 'hemp'; HAND/FIVE (235–7): PJse/AT **(ka-)lima* > pre-Jse. **yi[ma]* > Jse. *i* (as in *i-tu-tu,* 'with reduplicated numeral suffix').

The height of Benedictine megalocomparative ingenuity is reached in the concept of SPLIT COGNATES, i.e. cognates that have reflexes of at most one given proto-phoneme in common, since they descend from different syllables of a polysyllabic etymon. This is a powerful reconstructive tool indeed[24], used to good advantage in etymologies like PAT **[wa]kləwm[a]* 'dog' > Proto-Tai **hma,* but > Proto-Hmong-Mien *klu* (B 1975:272–3); or PJse/AT **(m)ba(ŋ)ʔiwak,* which yields the Japanese 'split doublets' *uo* 'fish' (< **uwo* < **iwo* < **ʔiwak*) and *wani* 'crocodile' (<**baŋi[wak]*) (219–20).

4.2. 'Teleoreconstruction' and conflicting megalogroupings

G claims that 'The validity of Amerind as a whole is more secure than that of any of its stocks' (59)[25], and maintains that phenomena which are inexplicable on a lower taxonomic level often become clear when the scale of analysis is enlarged: '[A stepwise comparative] procedure appears to be very virtuous, but in fact is an illusion. The reconstruction will itself be a poorer approximation to the truth if it is confined to a restricted group . . . In fact, many phenomena of narrower groups can only be understood historically by outside evidence from within the broader stock' (36). This pronouncement comes quite close to B's notion of TELEORECONSTRUCTION (B 1973)—the method of leaping back to the level of the proto-language, without being deterred by all the detailed problems which arise in the individual subgroups. B feels in fact that it is often easier to reconstruct at a higher rather than a lower taxonomic level—e.g. Proto-Tibeto-Burman (PTB) is 'easier' to reconstruct than its best-studied subgroup, Proto-Lolo-Burmese (PLB);[26] he

claims that the relationship of Japanese to the 'rest of' Austro-Tai (AT) is easier to demonstrate than are the interrelationships of its previously posited branches (Austronesian, Tai-Kadai, and Hmong-Mien).

There is certainly something to be said for a telic approach when one is operating at, say, the macrolevel, where the validity of the language family is obvious to all. But even here it is an illusion to think that the task of reconstruction gets easier when the genetic relationship becomes more remote. It is certainly harder to reconstruct PST than PTB, even though the relationship of Chinese to TB is not in doubt. It is much more difficult to go even one further step up the taxonomic tree from Proto-Tai to Proto-Tai-Kadai, let alone from there back to Austro-Tai or Austro-Tai-cum-Japanese. When a megalo-construct is already highly shaky and speculative, the very ease with which it may be further elaborated should be suspect.

For me the ultimate problem with proposals of super-remote genetic relationship is the impossibility of choosing among conflicting ones. I have mentioned the various theories about the affiliations of Japanese (§3), where G and B do not see eye to eye. The two also differ on the further genetic affiliations of Sino-Tibetan, though they both agree it is distinct from Austro-Tai. While B for once ventures no further affiliations for Sino-Tibetan, G is inclined to accept Sapir's hypothesis of a relationship between ST and Na-Dene (G 332; Campbell 1988:593). The eminent Sinologist and historian E. G. Pulleyblank claims rather that ST's true connections are with Indo-European (1978). G endorses Schmidt's old idea (1906) of a genetic connection between Austronesian and Austroasiatic (Mon-Khmer plus Munda), and favors adapting the theory to encompass all of B's Austro-Tai as well as Austroasiatic. B once believed in Schmidt's theory himself (the 'Austric' theory; see Diffloth 1985), but lately refuses to recognize anything more than an old contact relationship between Austronesian and Austroasiatic.

Conclusion

5. It is no doubt a noble idea to try to find a common genetic origin for all of the world's languages, just as a common biological origin can be demonstrated for all extant 'races' of *Homo sapiens.* G appeals to biological genetic evidence to back up his linguistic genetic claims: 'I would like to emphasize the fact that my linguistic classification shows an almost exact match with genetic classification by population biologists and with fossil teeth evidence' (1989:113). But dental consonant correspondences are much better indicators of linguistic relationship than are dental fossils. There is no linguistic correlate of DNA to furnish irrefutable proof of genetic relationship. People of any genetic makeup can speak any language. There seems little reason to believe that the New World was populated by a discrete series of exactly three waves of migration, each racially and linguistically uniform. Why not conceive of this population movement as a millennial percolation of people speaking a variety of originally unrelated languages, some

of which were subjected along the slow migration routes to prolonged mutual contact and even creolization?

Looking on the positive side, one has to admire G's intellectual vigor and daring. No one could ever accuse him of timidity or fuddyduddyism. Perhaps *LIA* will actually prove to be a boon to Amerindian linguistics because of the intensity of the counter-research it will engender. In discussing his methodology G sets up a dichotomy between the traditional comparative method of 'looking at many forms across only two or a few languages' and his own method of 'looking at few forms across many languages'. Surely what is needed is a combination of the virtues of both approaches: looking at as many well-recorded and well-analyzed forms from as many languages as possible. It is also a good idea to organize the data by semantic field as well as by phonological shape.[27]

The computer will be crucial to handle the etymological information explosion of the future, but machines will never be able to do all the work for us. There will still doubtless be room for gut feelings, intuitions, temperamental quirks, and professional rivalries, even in the Computer Age. In making etymological 'judgment calls' there will never be a substitute for hands-on human experience in a given language or language family.

Editor's note

Because of the importance to historical linguistics of the issues raised by Joseph Greenberg, his supporters, and his critics, the editor decided to solicit a discussion of these issues by a scholar whose qualifications included extensive experience in comparative linguistic research and a presumed lack of prior bias either for or against Greenberg's views. The following essay is the result.

Notes

1 My own ignorance of the Amerindian data (which I have managed to retain despite my Berkeley formation) ensures a certain objectivity, which is perhaps why I have been asked to contribute this Discussion Note. At least I cannot be accused of an axe to grind based on an 'accident of expertise' (*LIA,* p. 4)!

2 Until very recently linguists in the People's Republic of China have held a similarly unflattering view of the reconstructive enterprise, likening it to 'painting ghosts' (*huà guǐ*)—i.e. trying to lend a specious reality to something imaginary.

3 In Oriental linguistics, one thinks of such great figures as Bernhard Karlgren and Paul K. Benedict.

4 Very roughly speaking, MICROCOMPARISON can be practiced on close-knit families like Romance, Loloish, or Tai, with a time-depth of not more than about 2000 years. MACROCOMPARISON is appropriate for farflung but demonstrably valid groupings like Indo-European or Sino-Tibetan, with time-depths of up to about 6000 years. MEGALOCOMPARISON takes on any more remote relationship, where sound-correspondences are not regular and putative cognates are few, so that chance rivals genetic relationship as the explanation for perceived similarities.

5 The micromaniac might retort that real expertise in a relatively limited area is the best training for more farflung endeavors, and that G would have done well to acquire such expertise before writing a book like *LIA*.

6 I plead guilty to this awful neologism (< Lat. *columba* 'dove; pigeon' and *cubiculum* 'small compartment'), though perhaps 'pre-columbicubiculomania' is more appropriate for the prehistoric New World. Personally I have come down on the side of those who prefer the continuum to the pigeonhole: 'Pigeonholing is merely a heuristic attempt to make a continuum appear discrete' (Matisoff 1976:258).

7 An extreme case of relexification is the Bai (= Minjia) language of NW Yunnan, spoken by well over a million highly Sinicized people. While Bai definitely belongs somewhere in the Sino-Tibetan family, its precise genetic classification is difficult, since some dialects have undergone replacement of as much as 75% of their lexicon by Chinese (see Zhao, forthcoming). Sorting out the inherited vs. the borrowed components of the Bai lexicon is an intricate business. Needless to say, languages do not have to be genetically related for one to relexify the other, if the contact is intense enough.

8 It is as obvious to the areal linguist as to the creolist that no aspect of phonology, grammar, or semantics is immune to borrowing or contamination. Asian languages furnish many spectacular examples, e.g. the monosyllabization and tonalization of the Austroasiatic language Vietnamese under Chinese influence; the acquisition of SVO order by the Karenic Branch of Tibeto-Burman (under Tai and Mon influence); the loss of tone in frontier Chinese dialects of Gansu under Altaic influence; the development of elaborate verbal morphology in many minority languages of India under Indo-Aryan influence; and parallel grammaticalizations of root-verbs to aspect-markers throughout the region (Matisoff 1988).

9 In a well-written but rather sarcastic passage (11–16), G lists ten 'options' comparativists use to salvage an etymology which shows an irregular sound correspondence, of which one is invoking 'dialect mixture' (16). By rubbing the noses of comparativists in irregular cases (as if there were something shameful or unnatural about them), G is casting aspersions on the comparative method itself, thereby justifying his disinterest in regular correspondences or asterisked reconstructions. Missing from G's discussion is the obvious point that only the idealized assumption of regularity makes it possible to identify 'irregularities' in the first place.

10 The position of the sleeves with respect to the torso resembles the letter T, and that is apparently the 'true' etymology. Yet the variant spelling *tee-shirt* suggests a sporting connection with golf tees (my own previous belief). Still other members of the Volk (like my wife) believe the T stands for *tennis*.

11 Campbell uses similes like 'shooting fish in a barrel' (603) or 'playing poker with all the cards wild' (605), and would probably not object to 'taking candy from a baby'.

12 For instance by the great French comparativist of SE Asia, André-Georges Haudricourt.

13 These are the Mandarin, Armenian, and Latin forms for 'two'.

14 Personal communication, M. B. Emeneau (September 1989).

15 List furnished upon request to adults 18 years old or over. The amateur linguist and pioneer Tibeto-Burmanist Robert Shafer (1952) seriously entertained the possibility of a relationship between ST and American Indian languages, and Edward Sapir carried on a lively correspondence on this subject with Berthold Laufer in 1920.

16 I. Goddard invokes a 'universal tendency for primary grammatical morphemes to consist of a single, unmarked (phonetically commonplace) segment' to explain the widespread appearance of *-n-* as a 1st person marker (1986:202).

17 I refer to the Chinese and Indian areas of linguistic/cultural influence in Southeast Asia as the 'Sinosphere' and the 'Indosphere'.

18 See Benedict 1975, Hashimoto 1976; for opinions to the contrary see Wulff 1934, Li 1976, Yan 1983, Dong et al. 1984.

19 See, for instance, *LIA* p. 74: 'As is so often the case, they were both right . . . If we combine the two chains of suggested relationships, we have, then, Panoan, Tacanan, and Moseten on the one hand and Mataco, Guaicuru, Charruan, Lule, and Vilela on the other.'

20 So obvious is this relationship to Ohno that he does not feel the need of using Proto-Dravidian reconstructions, but rather compares Japanese to Tamil directly. It must be admitted that the typological similarities between Japanese and Dravidian are indeed striking, and there are even a fair number of phonological lookalikes in grammatical morphemes.

21 B is said to be able to memorize a wordlist after a single scanning.

22 Weidert 1987 furnishes many good examples of this strategy within the confines of the monosyllable, e.g. in his PTB reconstruction **mrgsla* for 'moon', where the monstrous cluster *mrgsl-* is set up to account for a unique correspondence of initial consonants—despite the fact that this flies in the face of the proto-syllable canon, and in general shows a lack of 'proto-Sprachgefühl' (see Matisoff 1982). (The standard reconstruction of this etymon is **s-la* or **s-gla*.)

This is not to deny that there may be cases of 'unique but regular' correspondences, as G insists (9). Thus one might well accept B's revision of the reconstruction of the initial cluster of PTB 'blood' from **s-hw-* to **s-hyw-*, since this does not strain the proto-syllable canon too much, and a parallel cluster **kyw-* is attested in the root for 'yam' (1972:51).

23 Cf. Campbell's discussion of 'short forms and unmatched segments' (1988:600).

24 To be fair, we can certainly imagine a protoform-with-respect-to-the-future like **taxi-cab* developing into *taxi* in Lg. A but into *cab* in Lg. B; or a protoname like **Elizabeth* becoming *Lisa* in Lg. A but *Betty* in Lg. B. But it will be all to the good if future linguists can uncover further cognates with shared phonemes (e.g. *tekep* or *Libby*).

25 G uses a false analogy to illustrate this point: 'For example, there is no doubt concerning just which languages belong to the vast Austronesian family, but subgrouping has proved difficult and has not led to any generally accepted result' (59). While it is true that Austronesianists are still not in total agreement on some details of Proto-AN reconstruction, the various subgroups of the family (e.g. Proto-Polynesian) are very well reconstructed indeed, and thousands of valid cognate sets have been established at various taxonomic levels, the vast majority of which illustrate recurrent, regular sound correspondences.

26 Personal confession, numerous occasions. We know enough about PLB to see how much remains to be explained.

27 This is the approach, for instance, of Kaufman's *Otomanguean Etymological Dictionary* and Diffloth's ongoing *Mon-Khmer Etymological Dictionary* project, as well as of my *Sino-Tibetan Etymological Dictionary and Thesaurus* project at Berkeley.

References

Adelaar, Willem F. H. 1989. Review of Greenberg 1987. Lingua 78.249–55.

Benedict, Paul K. 1972. Sino-Tibetan: A conspectus. Contributing editor, James A. Matisoff. Cambridge: Cambridge University Press.

———. 1973. Tibeto-Burman tones, with a note on teleoreconstruction. Acta Orientalia 35.127–38.

———. 1975. Austro-Thai language and culture, with a glossary of roots. New Haven: Human Relations Area Files Press.

———. Forthcoming. Japanese/Austro-Tai. [300-page ms version, 1985.] Ann Arbor: Karoma Press.

BURROW, THOMAS, and MURRAY B. EMENEAU. 1961. Dravidian etymological dictionary. Revised Edition. London: Oxford University Press.

CALLAGHAN, CATHERINE A., and WICK MILLER. 1962. Swadesh's Macro-Mixtecan hypothesis and English. Southwest Journal of Anthropology 18.278–85.

CAMPBELL, LYLE. 1988. Review article on Greenberg 1987. Lg. 64.591–615.

CHAFE, WALLACE. 1987. Review of Greenberg 1987. Current Anthropology 28.652–3.

COOKE, JOSEPH R. 1968. Pronominal reference in Thai, Burmese, and Vietnamese. (University of California Publications in Linguistics, 52.) Berkeley & Los Angeles: University of California Press.

DIFFLOTH, GÉRARD. 1985. What happened to Austric? University of Chicago, MS.

DONG WEIGUANG; CAO GUANGQU; and YAN XUEQUN. 1984. Hanyu he Dong-Taiyu de qinyuan guanxi [Genetic relationship between Chinese and Dong-Tai languages]. Computational Analysis of Asian and African Languages (Tokyo) 22.105–21.

GODDARD, IVES. 1986. Sapir's comparative method. New perspectives in language, culture and personality, ed. by William Cowan, Michael K. Foster, & Konrad Koerner, 191–210. Amsterdam & Philadelphia: John Benjamins.

GREENBERG, JOSEPH H. 1987. Language in the Americas. Stanford: Stanford University Press.

———. 1989. Classification of American Indian languages: A reply to Campbell. Lg. 65. 107–14.

———. Forthcoming. Indo-European and its closest relatives: the Eurasiatic language family. Stanford: Stanford University Press.

HASHIMOTO, ANNE OI-KAN. 1976. Substratum in Southern Chinese—The Tai connection. Computational Analysis of Asian and African Languages (Tokyo) 6.1–10.

KAUFMAN, TERRENCE S. 1990. Otomanguean etymological dictionary. University of Pittsburgh, MS.

KAWAMOTO, T. 1977/1978. Towards a comparative Japanese-Austronesian. Bulletin of the Nara University of Education 26.23–49 (Part I), 27.1–24 (Part II).

LABBERTON, D. VAN H. 1924. Preliminary results of researches into the original relationship between the Nipponese and the Malay-Polynesian languages. Journal of the Polynesian Society 33.244–80.

LI FANG KUEI. 1976. Sino-Tai. Computational Analysis of Asian and African Languages (Tokyo) 3.39–48.

MARTIN, SAMUEL. 1966. Lexical evidence relating Korean to Japanese. Lg. 43.185–251.

MATISOFF, JAMES A. 1976. Austro-Thai and Sino-Tibetan: An examination of body-part contact relationships. Genetic Relationship, Diffusion, and Typological Similarities of East and Southeast Asian Languages, ed. by Mantaro J. Hashimoto, 256–89. Tokyo: Japan Society for the Promotion of Science.

———. 1979. Problems and progress in Lolo-Burmese: Quo vadimus? Linguistics of the Tibeto-Burman Area 4:2:11–43.

———. 1982. Proto-languages and proto-Sprachgefühl. Linguistics of the Tibeto-Burman Area 6.2:1–64.

———. 1988. Areal and universal dimensions of grammatization in Lahu. Paper presented at the Symposium on Grammaticalization, University of Oregon. [To appear in Elizabeth Traugott & Bernd Heine, eds., Approaches to grammaticalization, Vol. II. Amsterdam & Philadelphia: John Benjamins.]

MILLER, ROY ANDREW. 1971. Japanese and the other Altaic languages. Chicago: University of Chicago Press.

NISHIDA, TATSUO. 1978. Chibetto-Biruma-go to Nihongo [Japanese and the Tibeto-Burman languages]. Nihongo no keitoo to rekishi [The lineage and history of Japanese] (Kooza Nihongo 12), ed. by Ohno Susumu & Shibata Takeishi, 229–300. Tokyo: Iwanami Shoten.

OHNO, SUSUMU. 1980. Sound correspondences between Tamil and Japanese. (Gakushuin Series of Treatises 8.) Tokyo: Gakushuin Educational Foundation.

PULLEYBLANK, EDWIN G. 1978. Chinese and Indo-European: The phonetic role of the *kanchih* signs in the Archaic Chinese script and their relation to the alphabet, together with evidence for a genetic relationship between Chinese and Indo-European. Paper presented at the Conference on the Origins of Chinese Civilization, University of California, Berkeley.

RAHDER, JOHANNES. 1956/1959. Etymological vocabulary of Chinese, Japanese, Korean and Ainu. Part I (1956): Monumenta Nipponica (Tokyo), Monograph 16. Parts II and III (1959): privately printed, New Haven.

SCHMIDT, PATER WILHELM. 1906. Die Mon-Khmer-Völker, ein Bindeglied zwischen Völkern Zentralasiens und Austronesiens. Archiv für Anthropologie 33.59–109.

SHAFER, ROBERT. 1952. Athabaskan and Sino-Tibetan. IJAL 18.12–19.

SOLOMON, JAMES R. 1974. Japanese and the other Austronesian languages: A survey of the literature. Paper presented at the First International Conference on Austronesian Linguistics, Honolulu.

WEIDERT, ALFONS. 1987. Tibeto-Burman tonology. Amsterdam & Philadelphia: John Benjamins.

WULFF, KURT. 1934. Chinesisch und Tai. Det Kgl. Danske Videnskabernes Selskab, Historisk-filologiske Meddelelser 20.3.

YAN XUEQUN. 1983. Miao-Yao Hanyu guanxici de cengci [Strata of related words in Miao, Yao, and Chinese]. Paper presented at the 16th International Conference on Sino-Tibetan Languages and Linguistics, University of Washington.

ZHAO YANSUN. Forthcoming. [Bai-Chinese Dictionary.] Edited by Grace Wiersma.

8

COMMENT ON MATISOFF'S COMPARISON BETWEEN GREENBERG AND BENEDICT

Paul K. Benedict

Source: *Linguistics of the Tibeto-Burman Area* 14, 2, 1991, 169–70.

In his recent discussion note "On megalocomparison" (*Language* 66:1.106–20, 1990), James A. Matisoff has described me as a "megalo-comparativist" and has compared (and contrasted) my methods with those pursued by Joseph H. Greenberg. He also reports that I favor a "jigsaw puzzle" approach as contrasted with Greenberg's obsession for "pigeonholing" (columbicubiculomania). I do indeed attempt to fit linguistic pieces together (see below) but I resolutely deny being a megalocomparativist. As I see it, Greenberg and I are simply playing different games and the contrast in our approaches reflects that basic difference.

I began in the A. L. Kroeber era at Berkeley as a "splitter" but before long adopted a "lumper" role. I had taken over as director of the Sino-Tibetan Philology project and had soon come face to face with an ugly fact: as Gertrude Stein might have put it, there was no there there! A veritable "black hole" existed at the very heart of the proposed language stock under investigation, with some widespread etyma for the numerals as well as for cultural items such as "charcoal", "horse", and "saddle" but nothing resembling a core vocabulary. The reigning Orientalists of that period placed almost exclusive emphasis on features such as monosyllabism, tones and isolating characteristics, with virtual neglect of lexical evidence, e.g. the great Henri Maspero classified Vietnamese as Tai-related (hence coming under Sino-Tibetan) despite the overwhelmingly Mon-Khmer nature of its lexicon. As a good anthropologist, however, I had been reared at Harvard on diffusionist doctrine and I had no trouble at all with the idea of pretty much anything in language, monosyllabism and tones included, diffusing here and there. In keeping with this line of thought. I soon lopped off from the Sino-Tibetan stock, as gently as possible, both Tai and Miao-Yao and carefully transferred Vietnamese to its ancestral Mon-Khmer. Kroeber, the administrator of the project and himself

a pretty good anthropologist, approved of the truncation but the China lobby on the Berkeley campus was mortally offended, very much as if I had removed the Great Wall from the country.

The project survived despite the outcries of betrayal and I went on to do some "lumping", first of Tai (and Kadai) with Austronesian and later of Miao-Yao and (recently) Japanese with this duo, all under "Austro-Tai". I have also remained a "splitter", however, quite unwilling to find any genetic relationship between Austronesian and Austroasiatic (Schmidt's old "Austric"). I continue to make the significant sharing of core vocabulary a litmus test for genetic relationship and I keep looking for precise fits in this area, the more complex the better (because so much the less likely to be the product of chance). Matisoff has labeled this as "proto-form stuffing" but it must be noted that the complexities often lie within one of the constituent families itself, e.g. Proto-Austronesian "I" is represented by the doublet *(ʔu-, ʔi-)aku - *(ʔi-)a(ŋ)kən, with *ʔu- and *ʔi - as pronominal markers and (ŋ) as variable nasal increment (highly characteristic of the family and also of the stock as a whole), with cognates (often "split") in the mainland families as well as in Japanese: P-Kadai *(ʔi-)aku, P-Miao-Yao *ʔyakou < *ʔi-aku ~ waŋkon < *ʔu-aŋkən (regular shifts) and Proto-Japanese-Ryukyuan *anu - wanu - waŋu < *(ʔu-)aŋku (regular shifts) (see Benedict 1990:214-16 for details). This pronominal set also serves to provide key evidence for the presence of the *ʔu- and *ʔi- markers at the earliest (Proto-Austro-Tai) level: it has its complexities, to be sure, but I can hardly see anyone attributing it to chance.

As pointed out by Matisoff (111), it is indeed curious that Greenberg abhors reconstruction yet "operates with the notion of COGNATES". I submit that Greenberg is badly misusing the word in this context and that it is pro-reconstruction macrocomparativists like me who operate with COGNATES whereas anti-reconstruction megalocomparativists like Greenberg operate rather with LOOKALIKES or COMPARABILIA (Matisoff), perhaps now in need of an abbreviation (LOOKA's? COMP's?). I further submit that if one doesn't "buy" any given product of a macrocomparativist, e.g. is not convinced that a proto-language such as Austro-Tai ever existed, he should not promote its author to the status of a megalocomparativist but simply think of him as having failed in this engagement. I play one game and am intrigued by the other but I do keep wondering whether that other game has any rules or whether a computerized robot might not be rather better at it all. But long live both games!

References

Benedict, Paul K. 1990. *Japanese/Austro-Tai*. Linguistica Extranea. Studia 20. Ann Arbor: Karoma Publishers.

Matisoff, James A. 1990. "On megalocomparison." *Language* 66:1.106–20.

9

SINO-TIBETAN LINGUISTICS

Present state and future prospects

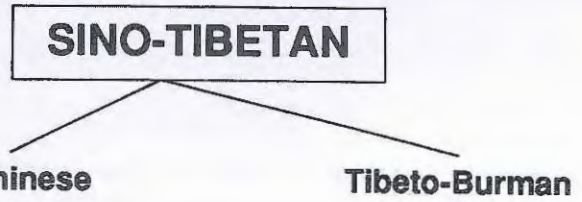
James A. Matisoff

Source: *Annual Review of Anthropology* 20, 1991, 469–504.

Introduction: what is Sino-Tibetan?

Since Sino-Tibetan (ST) is one of the greatest language families in the world—even Indo-European does not have more first-language speakers—it is sobering to realize that ST linguistics is only about 50 years old, and has been a flourishing field of inquiry for only the past 25 years. The more than 1.1 billion speakers of Sinitic (= the Chinese dialects) constitute the world's largest speech community, and scholars have been trying since the mid-19th century to situate Chinese in a wider genetic context. As the relationships between Chinese and Tibetan on the one hand, and Tibetan and Burmese on the other became obvious, vague notions of an "Indo-Chinese" family (34, 79) began to crystallize. The term Sino-Tibetan seems to have been used first by R. Shafer (177, 178), who conceived of it as a tripartite linguistic stock comprising Chinese, Tibeto-Burman (TB), and Tai (= "Daic"). Today most scholars in China take an even broader view of ST (called Hàn-Zàng in Mandarin), including not only these three branches, but Hmong-Mien (= Miao-Yao) as well. The majority view outside of China is more conservative: ST includes only Chinese (= Sinitic) on the one hand and the Tibeto-Burman languages on the other (see Figure 1). Even taking ST in its narrower sense, we are dealing with a highly differentiated language family of formidable scope, complexity, and time-depth. TB comprises hundreds of languages besides Tibetan and Burmese, spread over a vast geographical area (China, India, the Himalayan region, peninsular SE Asia), most of which is still virtually inaccessible for linguistic fieldwork, at least by foreign scholars (e.g. northeast India, Burma, Yunnan, Sichuan, Tibet, Laos, Vietnam). Only in Thailand and Nepal has vigorous international fieldwork been carried on since the 1960s.

A. Narrow View



B. Extended View

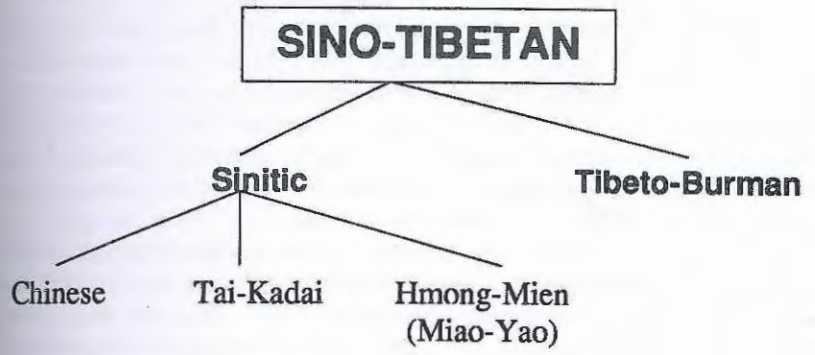


Figure 1 Two views of Sino-Tibetan

Homeland and time-depth of Sino-Tibetan

The Proto-Sino-Tibetan (PST) homeland seems to have been somewhere on the Himalayan plateau, where the great rivers of East and Southeast Asia (including the Yellow, Yangtze, Mekong, Brahmaputra, Salween, and Irrawaddy) have their source. The time of hypothetical ST unity, when the Proto-Han (= Proto-Chinese) and Proto-Tibeto-Burman (PTB) peoples formed a relatively undifferentiated linguistic community, must have been at least as remote as the Proto-Indo-European period, perhaps around 4000 BC.

The TB peoples slowly fanned outward along these river valleys, but only in the middle of the first millennium AD did they penetrate into peninsular Southeast Asia, where speakers of Austronesian (= Malayo-Polynesian) and Mon-Khmer (Austroasiatic) languages had already established themselves by prehistoric times. The Tai peoples began filtering down from the north at about the same time as the

Tibeto-Burmans (108). The most recent arrivals to the area south of China have been the Hmong-Mien (Miao-Yao), most of whom still live in China itself.

The components of Sino-Tibetan

The Chinese component

By any criterion (number of speakers, antiquity of documented written history, cultural significance, influence on other languages) Chinese ranks as one of the most important languages in the world. Yet the nonalphabetic nature of the Chinese writing system has posed unique problems for the historical linguist trying to reconstruct the phonology of earlier stages of the language, or establish a genetic connection between Chinese and other languages.

The great Swedish Sinologist Bernhard Karlgren, basing his work on the pioneering philological research of 18th and 19th century Chinese scholars, devoted some 35 years to the phonological reconstruction of the pronunciation of thousands of Chinese characters (84–87). Karlgren recognized two earlier stages of the language: 1. "Ancient Chinese" (now usually called Middle Chinese, "MC"), spoken during the second half of the 1st millennium AD, and 2. "Archaic Chinese" (now usually called Old Chinese, "OC"), spoken during the early Zhou (= Chou) dynasty at the beginning of the first millennium BC.

The reconstruction of MC is based mainly on the "rhyme-books" produced by contemporary Chinese literati, especially the *Qie Yun* (602 AD), wherein each character was given a phonetic value by glossing it with two others, the first of which had the same initial consonant as the target character, while the second had the same "rhyme" (i.e. vowel, final consonant if any, and tone) as the target character.

The tools available for the reconstruction of OC are much more indirect and tricky to use: the patterns of rhyming in the earliest Zhou texts, especially the *Book of Odes (Shi Jing)*, and the graphological structure of the characters themselves, most of which are constructed of two elements, a *radical* that gives a clue to its meaning, and a *phonetic* that gives a clue to its pronunciation. (But no more than a clue: it cannot be assumed that all the characters in a given "phonetic series" had exactly the same initial and rhyme.)

Despite the brilliant successes of Karlgren's methods, they have certain severe inherent limitations. First of all, the phonological system implied by the *Qie Yun* is forbiddingly complex, lending credence to the view that it does not represent the speech of any single dialect of the time (not even that of the Tang capital, Chang-an), but is rather *pan-dialectal,* noting distinctions made in any dialect with which the compilers (who came from various regions, as stated explicitly in the Preface) happened to be familiar. Second, there is no reason to suppose that the MC phonological system of the *Qie Yun* was the exact lineal descendant of the OC system deduced from the *Shi Jing* rhymes and the graphic structure of the characters (in the sense that, for example, Modern High German "descends from" Middle High and Old High German). Third, certain modern groups of dialects, especially the

Min dialects of Fujian (= Fukien) and adjacent regions in Southeast China, have undergone distinctive phonological developments that are impossible to trace back to the presumed MC system of the rhyme-books (see 26, 162, 163).

Despite the ingenuity of Karlgren's successors in patching up this or that aspect of his reconstructions—or perhaps because of this very ingenuity—Chinese historical phonology has until recently been in danger of degenerating into a kind of scholasticism: endless reinterpretations of the same data. For no matter how rich the material on earlier stages of a single language may be, one can only go so far by the methods of "internal reconstruction." A tripod cannot stand on a single leg. For further progress in ST comparative/historical linguistics, it is necessary to look well beyond Chinese.

The Tibeto-Burman component

The key component of ST, the branch with the most numerous and highly differentiated individual languages, is TB. The existence of the TB family was posited as early as the 1850s, when it was noticed that many words in "Written Tibetan" (WT), attested since the 7th century AD, appeared cognate to forms in "Written Burmese" (WB), attested since the early 12th century AD. British scholars and colonial administrators in India and Burma began to study some of the dozens of little-known "tribal" languages of the region that seemed to be genetically related to the two great literary languages, Tibetan and Burmese. This early work was collected in the monumental *Linguistic Survey of India* (59), three volumes of which (Vol. III, Parts 1, 2, 3) are devoted to wordlists and brief texts from TB languages.

Subsequent sporadic attempts to find cognates between Tibetan and Chinese (e.g. 181) did not get far, in the absence of any serious scheme for the reconstruction of Proto-Tibeto-Burman (PTB). It remained for the eccentric American amateur comparativist Robert Shafer to embark on a systematic project to assemble all the material then available on TB languages, and to venture a division of the family into subgroups. Much of this earlier data had been collected by officials or missionaries who had spent years living among the people whose languages they studied, and a number of the grammars and dictionaries they produced are of lasting value (e.g. 47, 64, 103, 104, 105, 113, 168). Shafer was assisted in this Depression-era WPA project by a talented junior collaborator, Paul K. Benedict, along with a motley team of half-trained indigents who spent their time combing through dictionaries and wordlists. The results were enshrined in a multi-volumed unpublished manuscript (1939–1941) called *Sino-Tibetan Linguistics* (see 6).

Shafer went on to publish his great work, *Introduction to Sino-Tibetan* (178), where he included Tai in the ST family, and offered a complex and detailed subgrouping of TB into "divisions," "sections," "branches," and "units." Despite the illusory nature of this precision, given the inadequate quality of the data then available for the various subgroups of TB, Shafer's classificatory schema has been adopted unquestioningly by many nonspecialists.

Benedict, basing his work on the same database as Shafer, arrived at different conclusions. In an unpublished manuscript entitled *Sino-Tibetan: A Conspectus*

(ca. 1941), he first of all banished the Tai languages from ST, leaving only Chinese on the one hand, and TB on the other. As for the internal subgrouping of TB, though Benedict generally followed Shafer in setting up eight major TB "nuclei," he refrained from trying to relate these by family trees (*Stammbäume*) of the traditional type, preferring to stress that many TB languages had so far resisted precise classification, and that the subgroups that could safely be established seemed to interrelate in ways too complex for a simple tree diagram.

When a revised and heavily annotated version of the *Conspectus* (henceforth *STC*) was finally published in 1972 (4), with J. Matisoff as contributing editor, this agnostic view of the internal structure of TB was retained, with most of the family pictured as "radiating out" of the geographically central Kachin (= Jinghpaw = Jingpho) language of North Burma, and the Karen languages singled out as being furthest away from this central nucleus.

STC offers close to 500 TB etymologies, as well as over 300 suggested cognates between PTB and Old Chinese. In spite of its shortcomings, its publication ushered in the modern era of ST studies, and it is now recognized as the point of departure for future work in the field.

The present state of Sino-Tibetan studies

Conferences and journals

Progress in both synchronic and diachronic ST studies since the late 1960s has been steady, both in the United States and abroad, though our knowledge of several of the subbranches of the family remains spotty. Except for the handful of "major literary" languages (Chinese, Tibetan, Burmese), and the somewhat more numerous "minor literary" ones (Xixia [= Tangut], Newari, Meithei [= Manipuri], Naxi/Moso, Yi [= Lolo], Bai [= Minjia], Pyu), no ST languages have left written texts that go back further than the early 20th century. The new era of progress has been made possible only by putting into proper relief the dozens of humble unwritten TB languages—the "poor relations" of the family, as it were. The data available for the task have been dramatically increased by an ongoing explosion of fieldwork by American, European, Indian, Japanese, Thai, and Chinese scholars. (For fairly complete bibliographies of recent research on TB through the early 1980s, see 63, 81, 195.)

Since 1968, the chief focus of scholarly activity in the field has been the annual ST Conferences, usually held in October. These have grown from small seminars (#1 had only eight participants; see 127) to full-scale conferences lasting 3–5 days and often attracting hundreds of people. At first the subject matter was limited to ST or TB phonological reconstruction; now any topic relating to the linguistics of the ST area is welcomed (including studies of Tai, Hmong-Mien, and Mon-Khmer). Now known as the *International Conferences on Sino-Tibetan Languages and Linguistics,* they have been held outside the United States at least every third year since 1976.[1] An annotated and indexed Bibliography of all 1216 papers presented at the first 21 Conferences has been prepared (92).

I Yale (1968), II Columbia (1969), III Cornell (1970), IV Indiana (1971), V Michigan (1972), VI U. C. San Diego (1973), VII Georgia State (1974), VIII U. C. Berkeley (1975), **IX Copenhagen (1976),** X Georgetown (1977), XI Arizona (1978), **XII Paris (1979),** XIII Virginia (1980), XIV Florida (1981), **XV Beijing (1982),** XVI Washington (1983), XVII Oregon (1984), **XVIII Ramkhamhaeng (Bangkok) (1985),** XIX Ohio State (1986), **XX U British Columbia (Vancouver) (1987), XXI Lund (Sweden) (1988),** XXII Hawaii (1989), XXIII U. Texas Arlington (1990), **XXIV Bangkok (1991)** [planned], XXV U. C. Berkeley (1992) [planned].

More specialized seminars and workshops (e.g. on subgroups like Loloish or Kadai) are occasionally held in conjunction with the main ST conferences, and increasingly these are being "spun off" into independent series of meetings.

More than a dozen new journals and monograph series devoted wholly or largely to ST linguistics have sprung up all over the world since the 1960s, with most of them still going strong. Some of the most significant are these: (**USA**) *Unicorn (Chi-lin),* a working-paper series produced by The Chinese Linguistics Project (Princeton; 10 issues, 1966–72); *Journal of Chinese Linguistics [JCL]* (University of California, Berkeley; from 1973); *Occasional Papers of the Wolfenden Society on Tibeto-Burman Linguistics [OPWSTBL]* (Ann Arbor/Urbana; 6 volumes, 1969–1978); *Linguistics of the Tibeto*-Burman Area [LTBA] (University of California, Berkeley; from 1974). (**France**) *Asie du Sud-est et Monde Insulindien [ASEMI]* (Paris; from 1970); *Cahiers de Linguistique Asie Orientale [CLAO]* (Paris; from 1977). (**Australia**) *Papers in South East Asian Linguistics [PSEAL]* (Canberra, Australian National University; from 1967). (**Japan**) *Tonan Azia Kenkyu (Southeast Asian Studies) [TAK]* (Kyoto; from 1963); *Computational Analyses of Asian and African Languages [CAAL]* (Tokyo Foreign Languages University; from 1975). (**China**) *Minzu Yuwen (National Minority Philology and Linguistics) [MZYW]* (Beijing; from 1979); *Zhongguo Shaoshu Minzu Yuyan Jianzhi Congshu [Minority Language Outline Grammar Series]* (Beijing: grammatical sketches of the 55 "official minority languages" of China; from about 1979). (**Thailand**) *Indigenous Languages of Thailand Research Monographs* (Bangkok: Chulalongkorn University; from 1976). (**Nepal**) *Kailash: A Journal of Himalayan Studies* (Kathmandu; from 1973). (**India**) *Phonetic Reader Series/ Grammar Series* (Mysore: Central Institute of Indian Languages; from 1969); *Compact Dictionaries of Languages of Nagaland* (Kohima: Nagaland Bhasha Parishad [Linguistic Circle of Nagaland]; from about 1972); *Resarun: Journal of the Research Department, Government of Arunachal Pradesh* (Shillong; from 1975); *Tibeto-Burman: Journal of the TB Linguistic Society* (Imphal, Manipur; from 1985).

Fieldwork and research situation, by country

The population figures in this section are mostly from *Ethnologue* (60). For more detailed discussion, including lists of minority languages, see 149.

China (including Tibet)

Total population: 1,100,000,000, of which 67 million (6.5%) are members of "minority nationalities," including Tai-Kadai (19,799,200), Tibeto-Burman (17,162,200), and Hmong-Mien (6,142,000).

Official Chinese government policy toward minority languages, paternalistic as it still is, is now quite enlightened by comparison to the contemptuous attitudes of the past. Fifty-five ethnic groups have been designated as "official nationalities" (of these 17 are TB), and several more are likely to achieve that status before long (perhaps including some of the 32 non-official TB languages so far noted by Chinese linguists). There have been lively discussions about the criteria that should be applied in order for a minority nationality and language to be granted official status (e.g. 46, 50). Mere number of speakers is not the sole criterion, and much emphasis is correctly laid on a psychological feeling of "ethnic identity." In theory, speakers of official nationality languages have the right to receive elementary education in their native tongue, and orthographies have been created for them all. On the other hand, a concerted effort is made to teach the Chinese language in every village of the country. The brightest young people of a tribal area are often sent to the nearest big city for instruction in Chinese, and are expected eventually to return to their villages to teach it.

Access to linguistic informants for foreigners remains sporadic and unpredictable, with junior scholars often faring better than more conspicuous senior professors. Many areas, especially in Yunnan and Tibet, are still totally closed to non-Chinese. Individual scholars in Beijing (especially at the Chinese Academy of Social Sciences and the Central Institute of Minority Nationalities) are usually extremely helpful, but in general government authorities are more flexible in the provinces than in Beijing.

India (including Bangladesh and Pakistan)

Total population: 750 million, of which 72% are Indo-Aryan, 25% Dravidian, and 3% (i.e. 22,500,000) other, including 5,492,858 Tibeto-Burmans.

India, by official government count, is home to 1,683 "mother tongues," of which 850 are in daily use. A much smaller though still impressive number is listed in reference 60 (pp. 460–87): 391 languages, of which 10 are extinct. Of these, no fewer than 107 belong to Tibeto-Burman, making India the undisputed heartland of the TB family. Since India's total TB population is only about 5.5 million, the average number of speakers per TB language is rather small, around 50,000.

The TB area of Northeast India has been off-limits to most foreigners for some time, largely because of unsettled political conditions. Population pressure from adjacent Indo-Aryan groups, especially the Bengalis, has intensified in recent years, leading to a series of bloody clashes as the "tribals" struggle to protect their territory against encroachment. Several TB groups of Nagaland, notably the

Bodos, are engaged in bitter battles for independence. So far, however, almost all of these ethnic groups seem to be viable, despite their minuscule populations.

Bangladesh (total population 105 million) is home to 531,000 Tibeto-Burmans, speaking 16 languages. Pakistan (total population 87.3 million) has 360,000 speakers of the Balti (= Sbalti) dialect of Tibetan.

Nepal/Bhutan/Sikkim

Total population: 16,600,000, of which 2,423,840 are TB speakers.

The dominant language family in Nepal is Indo-Aryan (IA). Besides the national language, Nepali, spoken by about 58% of the population, at least 14 other IA languages are spoken in the country. The most important non-IA family in Nepal is TB, with no fewer than 69 languages. Their average number of speakers is only about 36,000, and many have considerably fewer. Perhaps half of all TB languages in the country are in danger of extinction, due to the pressure of Nepali or other dominant languages. Fortunately Nepal remains one of the best places to do linguistic fieldwork in the Tibeto-Burman area, and there is hope of salvaging substantial samples of most of these endangered languages before they disappear. The Summer Institute of Linguistics has taken the lead in such research since the 1960s, directed first by K. Pike, then by A. Hale (see 62), but was recently expelled from the country for engaging in religious proselytization. Members of the SIL are still free to do fieldwork in Nepal as individuals, however. A German project directed by W. Winter has conducted an extensive linguistic survey of Nepal TB languages, the results of which still remain unpublished (see 209).

The small Himalayan kingdom of Bhutan (total population: 1.6 million) is peopled only by Tibeto-Burmans (1.2 million) and Indo-Aryans. Most of Bhutan's nine TB languages have respectable numbers of speakers: Dzongkha 400,000, Tsangla/Sharchopkha 300,000, Bumthapkha 250,000, Kürthöpka 100,000, Khengkha 100,000, Dzalakha 50,000, Lepcha 24,200 in Bhutan (also 23,700 in Sikkim), Chali 6,000, Tibetan 3,000 in Bhutan. (I am indebted to G. L. van Driem for these up-to-date figures, which are more accurate than those given in reference 60, pp. 434–35.)

Burma

Total population: 37,000,000, of which 28,877,000 (78%) are Tibeto-Burman and 2,776,900 (9.6%) are Tai (mostly Shans).

Burma is the world's premier TB country, the only one (besides Bhutan) where the national language belongs to the TB family. Burmese itself, with well over 20 million speakers, is far and away the leading TB language. This makes it all the more regrettable that Burma's government remains so uncompromisingly retrograde and repressive. Xenophobia has been official policy since 1964, when most foreigners were unceremoniously expelled from the country. Even today,

foreigners may normally sojourn in Burma for no more than two weeks at a time, making serious linguistic fieldwork impossible. Much of the country is no longer under the effective control of the central government, and at least six rebel movements remain powerful, four of them identified with particular minority groups: the Shans, the Karens, the Jingpho (Kachins), and the Chins. Many smaller ethnic groups are said to be suffering greatly as a result of this ongoing guerrilla warfare, especially in Shan State, where people like the Lahu are at the mercy of both warring factions.

It is thought that several dozen minority languages of Burma remain totally undocumented, most of them in the west, near the Indian and Bangladeshi borders. Even so, at least 75 TB languages of Burma are known, with an average of about 38,500 speakers. There is every reason to believe that most of them are still viable. In addition, 6 Tai-Kadai, 1 Hmong-Mien, and 14 Austroasiatic languages seem to be holding their own.

Thailand

Total population: 49,000,000, of which 45,815,000 (93.5%) are Tai speakers and 533,500 (1%) are TB groups speaking 16 languages.

The only nation in Southeast Asia never to have been colonized by a Western power, Thailand has a relatively benevolent government and a booming economy, and is perhaps the most pleasant locus for linguistic fieldwork in the whole area. Relations between the majority Tai-speaking population and the "hill-tribes" have generally been pretty good. The royal family has taken an active interest in the welfare of the nation's minority groups, whose value as tourist attractions has come to be appreciated. Yet recently resentment has been growing, partly because of the ecological damage done by the slash-and-burn agricultural techniques used in the hills. As population pressure on the land increases, Northern Tai peasants have begun to cultivate land at higher elevations; conversely, some hillfolk have been descending into the plains to experiment with wet-rice cultivation. So far these problems have not gotten out of hand, but the greatest tact and statesmanship will be required to avoid serious conflict in the future (see 155).

Laos and Vietnam

Out of Laos' small total population of 3.9 million, 2,769,000 (71%) speak Tai languages (including the national language, Lao); 1.1 million people (24%) are Mon-Khmer (MK)/Austroasiatic, speaking 57 languages, the largest number of MK languages in any country, but with a tiny average number of speakers (ca. 19,000); 175,000 (4%) are Hmong-Mien (many thousands of whom have been resettled in the United States); and only 42,500 (1%) are TB (speaking eight languages).

Laos is slowly recovering from the chaos of the Indo-China wars, and an improvement in relations with Western countries is in the offing. As in Vietnam,

there seems to be a revival of scholarly interest in minority languages on the part of Laotian scholars.

Vietnamese is the MK language with the most speakers, 55.4 million (out of a total population of 60.5 million in Vietnam). The country also includes 2.25 million Tai-Kadai, 679,000 Hmong-Mien, and 492,000 Austronesians (mostly Chamic), but only a handful of TB speakers (about 40,000, speaking 7 or 8 languages).

Tibeto-Burman languages and their subgrouping

Though the total number of TB speakers is only about 56 million, smaller than for Tai-Kadai or Mon-Khmer/Austroasiatic, the number of individual TB languages is the largest of any family in East/Southeast Asia. The relatively low overall total for TB results from the fact that its most populous language, Burmese, only has about 22 million speakers, while the number of Thai (45.5 million) and Vietnamese (55.4 million) speakers has increased rapidly in recent decades.

Language names

Of the more than 1,400 TB language names I have collected (139), many are only multiple designations for the same language or dialect. Any given language is likely to be known by several different names: its *autonym* (what the people call themselves), and perhaps several *exonyms* (what other groups call them). Some of the latter may be *loconyms* (e.g. the name of a conspicuous village where the language is spoken, or of a nearby river). Thus, the 20,000 speakers of a certain language of Nagaland call themselves and their language *Memi* (and used to call themselves *Imemai*), but they and their language are now known to outsiders either as *Mao,* or as *Sopvoma* (the name of their principal village). The 40,000 speakers of *Lotha Naga* are called *Chizima, Choimi,* and *Miklai* by the neighboring Angami, Sema, and Assamese, respectively. Conversely, quite different languages are often called by the same or very similar names: *Nung* is both a Central Tai language and a Tibeto-Burman language of the Nungish group; *Mru* is a TB language of the Kuki-Chin group, but *Maru,* also TB, belongs to the Burmish group; *Kham(s)* is both a dialect of Tibetan and a separate language of central Nepal.

Old names (*paleonyms*) now felt to be pejorative are rapidly being replaced by new ones (*neonyms*). We are now, for example, expected to say Yi, Mizo, Adi, Nishi, Karbi (instead of Lolo, Lushai, Abor, Dafla, and Mikir, respectively), even though these older terms have been enshrined in the Western literature on TB languages for a century. Nomenclatural proliferation continues apace, perhaps faster than ever before. It has recently been proposed to differentiate among approximately 25 Yi (Loloish) dialects of China by using the pronunciations of the vowels in their common autonym, e.g. Nasu, Nosu, Nusu, Neuseu, Nesu, Naso, etc (Chen Kang, personal communication, 1988).

A further complication is the fact that many language names are used in both a narrower and a broader sense, sometimes referring to one specific language

but often to a whole group of linguistically or culturally related languages. Often small or vulnerable groups will call themselves by the name of a somewhat larger or more prestigious neighbor, often hesitating to reveal their true ethnicity to outsiders. The tiny Anal people, an "Old Kukish" group of 6600 speakers in Burma and Bangladesh, "declared themselves to be Nagas in 1963" (117:379). There is even a trend in Nagaland to create larger linguistic/ethnic groups artificially by combining syllables of several individual names—e.g. *Chakhesang* (from Chokri, Khezha, and Sangtam) and *Zeliang* (from Zemi and Liangmai).

With all this in mind, my best estimate is that the TB family contains at least 250 separate languages, which may be broken down into population categories as in Table 1. For about half of the languages in the final category, accurate population figures are not available, and some of them may be in danger of extinction.

Subgrouping of Tibeto-Burman

I have elsewhere provided an account of the problems involved in attempting to subgroup TB in the light of our present knowledge (132). As a working hypothesis, I have modified the unwieldy scheme presented in *STC* (4) in several respects. For the new TB family tree that I propose as a heuristic model (and that I adopt in reference 150), see Figure 2.

Kamarupan

Benedict's Kuki-Chin-Naga, Abor-Miri-Dafla, and Bodo-Garo subgroups, spoken in Northeast India and adjacent regions of Burma, are lumped together under the purely geographical rubric of *Kamarupan* (from Kāmarūpa, the Sanskrit term for Assam). These languages constitute the center of diversification of the whole TB family. Nagaland alone, with an area of only 6350 square miles, is home to some 90 Tibeto-Burman languages and dialects. With a few exceptions, e.g. Lushai

Table 1 Speakers of Tibeto-Burman languages

Number of speakers	*Number of languages*
more than 1,000,000	9
500,000–999,000	12
250,000–499,000	11
100,000–249,000	16
50,000–99,000	16
25,000–49,000	27
10,000–24,000	44
fewer than 10,000	123

Figure 2 Subgroups of Tibeto-Burman

(104), Tangkhul Naga (12, 168), Garo (19), Tiddim Chin (77), and Bawm (176), these "Indospheric" TB languages have been poorly recorded until recently, and many are still hardly known at all.

Recent research is revealing how interesting, diversified, and important these languages are. An invaluable compendium of older data on the Naga languages is available (117), a source extensively utilized in the comparative study of the Northern Naga subgroup by W. French (48). A Weidert (205) has provided a sophisticated and data-packed study of the phonology of Kamarupan languages, marred only by its disorganized presentation and over-formalistic approach. New raw material on Kamarupan languages is becoming increasingly available in the publications of the Linguistic Circle of Nagaland (Kohima) and the Central Institute of Indian Languages (Mysore), and through the efforts of energetic scholar-administrators in Arunachal Pradesh (see 37–39, 180). Yet a great deal of work remains to be done in this area of TB, and it would be unrealistic to attempt a precise subgrouping of Kamarupan at the moment—i.e. a clarification of the higher-order relationships of the subgroups traditionally designated as Kuki-Chin-Naga, Bodo-Garo (= Barish), and Abor-Miri-Dafla (= Mirish). Several important languages seem to fall outside any of these groups—e.g. Mikir (61, 198), Meithei (193), and Mru (101). Of all these languages, the Mirish ones seem most lexically aberrant from the viewpoint of TB in general, even in its numerals (141).

Karenic

In my view the Karenic group of the Thai-Burmese borderlands should be considered just another subgroup of TB, and need not be singled out as having split off from the rest of the family at an especially early date. The argument for the special status of Karen is mostly syntactic. Alone of all TB languages (except for the heavily Sinicized Bai), Karen puts its objects after its verbs. Now that we realize that syntactic change can easily occur (either for language-internal reasons or as the result of close contact with other languages), this is a less persuasive criterion for genetic classification. Karen has been under heavy influence from Mon and Thai (both SVO languages). A tendency for the rightward "hopping over the verb" of certain nominal arguments (especially locative NPs) has also been pointed out for Northern Loloish languages under Chinese influence (207), yet there is no reason at all not to consider them to belong to "TB proper."

The Karen languages are only beginning to receive the attention they deserve. The early comparative work of R. B. Jones (83) requires fundamental revision in the light of Haudricourt's contributions (69, 74). The publication of research now in progress (e.g. 78, 185) will dramatically improve our knowledge of this key branch of TB.

Qiangic

The most exciting recent development in TB studies is the discovery of a new branch of the family, hitherto virtually unknown to Western scholars. These are the *Qiangic* languages of Sichuan. Extensive lexical and grammatical material has been collected on a dozen languages of the Qiangic group (106, 188, 190, 191). Besides Qiang, other languages in the group include Baima, Ergong, Ersu/Tosu, Gyarong (= rGyarong; see 158), Guiqiong, Muya, Namuyi, Pumi, Shixing, Zhaba. Ersu/Tosu is perhaps an indirect descendant of the extinct Xixia (= Hsi-hsia = Tangut) language, spoken in a once-powerful empire in the Tibetan-Chinese-Uighur border regions, finally destroyed by the Mongols in the 13th century. A large literature in Xixia survives, in a logographic writing system invented in the 11th century, with thousands of intricate characters inspired by, but graphically independent of Chinese, the decipherment of which is now well advanced by Japanese and Russian scholars (159, 182). It was thought at first that Xixia was a Loloish language, but it now seems more likely that it belongs to the Qiangic group.

From the limited data so far made available, the Qiangic languages seem to be of unusual interest, both synchronically and diachronically. They are characterized by initial consonant clusters comparable in complexity to those of Written Tibetan. Many of these are clearly secondary, resulting from the reduction of disyllabic compounds (see the section below on syllabic structure). Some languages of the group are tonal, while others are not, providing an ideal terrain for the investigation of the mechanisms of tonogenesis.

Kachinic

Kachinic, like Karenic, is relatively undifferentiated, consisting basically of a single language and its dialects. Kachin (= Jingpho), spoken in northernmost Burma and adjacent parts of China and India, is well known, thanks to Hanson's dictionary (64), its (unpublished) revision by Maran (116), and recent work by Chinese scholars (36). The name "Kachin" is also used loosely for various Burmish groups of Northern Burma [Atsi (214), Lashi, Maru]. Since Kachinic shows phonological and lexical similarities with several other branches of TB [Kamarupan, Himalayish, Lolo-Burmese (129)], it has been considered to be genetically central in the TB family, just as it is geographically central (4:6;22). The *Nungish* languages (90, 100, 189) seem closest to Kachinic, though it is too early to tell whether they also have a special relationship to the Qiangic group.

Himalayish

Himalayish comprises such relatively well-known languages as Tibetan, Lepcha (Sikkim; see 14, 113), and Newari (spoken in the Kathmandu valley of Nepal; see 54, 114), as well as dozens of others, some on the verge of extinction. Progress has been particularly impressive in the study of the TB languages of Nepal, especially those of the Tamang-Gurung-Thakali-Manang group (55, 152, 154); Kham-Magar (203, 204); Chepang (23); Sunwar (53); and the "Rai" or "Kiranti" languages of Eastern Nepal, which are generally characterized by complex inflectional morphology (1, 156, 196, 197, 209). The westernmost languages in the TB family, e.g. Pattani (= Manchati), belong to the Himalayish group, and are beginning to be studied by Indian scholars (179).

Himalayish languages generally preserve prefixes and initial clusters well, along with final *-s, -r,* and *-l*. Written Tibetan is consonantally the most archaic attested TB language, preserving, for example, initial clusters that had disappeared from Chinese a millennium before.

Lolo-Burmese

Burmese, attested since the 12th century AD, is one of the best-known TB languages. [Good modern grammars are available (165, 206).] The languages of the Northern Loloish subgroup (called "Yi" in China) are firmly within the "Sinosphere" (see the section below on the Indosphere and Sinosphere), and many of them have been well recorded by Chinese scholars (e.g. 49, 51, 110, 111, 216). The Central and Southern Loloish languages are spoken as far south as Thailand and Laos, where Western and Japanese scholars have had access to them since the 1960s (see 80, 94, 160, 186). More detailed comparative-historical work has been done on Loloish than on any other branch of TB (16, 21, 66, 119, 122, 123, 129, 133, 134, 148, 161, 194).

Loloish has strictly monosyllabic morphemes, few initial clusters or final consonants, often complex tone-systems, and a penchant for compounding as its chief

morphological device. The Loloish language with the most speakers and greatest dialectal differentiation is Lolo (Yi) itself, with 5 million speakers in Sichuan, Yunnan, and Guangxi, and a syllabic writing system of considerable antiquity (112). The tribal TB language that has been studied in greatest detail is Lahu (Central Loloish) (17, 120, 121, 124, 125, 130, 142, 143, 147). The Naxi/Moso language is close to the Loloish nucleus, and is of special interest because of its complex, hieroglyphic-like writing system (see 15, 123, 148, 166, 174).

Minor groups and unclassified languages

Finally, there is a tantalizing residue of TB languages that resist easy classification, seem transitional between two well-established subgroups, or have been so invaded by foreign vocabulary that their original affiliation is no longer apparent. Foremost among these is *Bai* (formerly called "Minjia"), spoken by over a million highly Sinicized people in the Dali region of Yunnan (see 43, 208, 213). The very large percentage of Chinese loanwords in Bai (for some dialects approaching 70%) has led to some rather wild speculations as to the genetic status of the language, though it is now clear that it is definitely TB. A large-scale Bai-Chinese dictionary, containing much archaic non-Sinicized vocabulary, is now in preparation by Zhao Yansun.

Genetic, areal, and typological relationships

The overwhelming cultural importance of China and India has shaped the development of the East and Southeast Asian linguistic area, but diversity is the hallmark of the region. No part of the world has a more luxuriant array of indigenous languages and cultures, and the influence of Chinese and Indian civilization has only served to enrich these, not obliterate them. Recent archaeological discoveries (see, e.g. 56–58) have demonstrated that peninsular Southeast Asia was a center of high technological advancement as early as the Pleistocene, by no means the "cultural backwater" it was once thought to be. Remarkably early dates have been determined for pottery (estimates range from 13,000 to 6800 BC), rice cultivation (3500 BC), and metallurgy (bronze by 2700 BC, iron by 1300 BC).

Sociolinguistically, the most important dichotomy in the Sino-Tibetan/Southeast Asia area is between the "great" majority languages spoken in the plains and river valleys (Chinese, Burmese, Siamese, Lao, Khmer, Vietnamese) and the hundreds of minority languages spoken in the hills and jungles (20). Members of these scattered "montagnard" groups are typically multilingual, often speaking several minority languages in addition to the coterritorial majority language. The result has been linguistic homogenization and convergence on a grand scale.

Due to long cultural contact, the Sino-Tibetan/Southeast Asian peoples have come to share a certain worldview, similar conceptual frameworks about people and nature, a sort of consensus as to what is worth talking and thinking about. Borrowing of cultural items and their associated vocabulary has proceeded in all

directions from such an early period that the original source of many words cannot now be traced (e.g. the calendrical cycle of animals used for dating events and naming people, words like *crossbow, weave, iron, needle,* etc). The areal lexicon may make semantic discriminations where Western languages do not (e.g. separate words for cooked vs uncooked rice, elder vs younger siblings, multiple verbs for drying, carrying, or cutting). As might be expected, languages spoken in the hills make careful distinctions in demonstratives and verbs according to whether the point of interest is situated above or below the speaker. This "areal semantic framework" is apparent even in the realm of conventional greetings: Languages throughout the region use formulas like *Have you eaten yet?* or *Where are you going?,* instead of 'hello' or 'good day.' This "intertranslatability" is also manifest in compound formation, where parallel lexicalizations or "areal calques" are common—e.g. *eye + foot = anklebone* (Malay **mata kaki**, Lahu **khi-mêʔ-šī**), *fly + shit = freckle/mole, pig + crazy/illness = epilepsy* (132:70); as well as in collocations of verb and object (regardless of whether the language is VO or OV)—e.g. *eat + rice* or *rice + eat, walk + road,* or *road + walk* (where English would simply have 'eat' or 'walk,' respectively).

Indosphere and Sinosphere

It is convenient to refer to the Chinese and Indian spheres of cultural influence as the "Sinosphere" and the "Indosphere" (146). Some languages and cultures are firmly in one or the other (e.g. the Munda and Khasi branches of Austroasiatic and the Kamarupan branch of TB are Indospheric; while the Hmong-Mien family, the Kam-Sui branch of Kadai, the Loloish branch of TB, and the Viet-Muong branch of Mon-Khmer are Sinospheric). Others (e.g. Thai and Tibetan) have been influenced by both Chinese and Indian culture at different historical periods. Still other linguistic communities are so remote geographically that they have escaped significant influence from either cultural tradition (e.g. the Aslian branch of Mon-Khmer in Malaya, or the Nicobarese branch in the Nicobar Islands of the Indian Ocean).

Elements of Indian culture, especially ideas of kingship, religions (Hinduism/Brahminism, Buddhism), and devanāgarī writing systems, began to penetrate both insular and peninsular Southeast Asia about 2000 years ago. Indic writing systems were adopted first by Austronesians (Javanese and Cham) and Austroasiatics (Khmer and Mon), then by Tai (Siamese and Lao) and Tibeto-Burmans (Pyu, Burmese, and Karen). The learned components of the vocabularies of Khmer, Mon, Burmese, and Thai/Lao consist of words of Pali/Sanskrit origin. Indian influence also spread north to the Himalayan region. Tibetan has used devanāgarī writing since AD 600, but has preferred to calque new religious and technical vocabulary from native morphemes rather than borrowing Indic ones.

What is now China south of the Yangtze (called "Cisyangtzeana" in 150) did not have a considerable Han Chinese population until the beginning of the current era (163, 172). In early times the scattered Chinese communities of the region

must have been on a numerical and cultural par with the coterritorial non-Chinese populations, with borrowing of material culture and vocabulary proceeding in all directions (164, 175). As late as the end of the first millennium AD, non-Chinese states flourished on the periphery of the Middle Kingdom [Nanchao and Bai in Yunnan, Xixia in the Gansu/Qinghai/Tibet border regions, Lolo (Yi) chieftaincies in Sichuan]. The Mongol Yuan dynasty finally consolidated Chinese power south of the Yangtze in the 13th century. Tibet also fell under Mongol influence then, but did not come under complete Chinese control until the 18th century.

Whatever their genetic affiliations, the languages of the ST area have undergone massive convergence in all areas of their structure—phonological, grammatical, and semantic. (An excellent general study of such phenomena is reference 192.) Hundreds of words have crossed over genetic boundaries in the course of millennia of intense language contact, so that it is often exceedingly difficult to distinguish ancient loans from genuine cognates.

The genetic position of the Tai-Kadai and Hmong-Mien families

Typologically, the Tai-Kadai languages have a strong Chinese flavor, both in their phonologies and their grammars. They are monosyllabic and tonal, and their sentence structure is similar to that of Chinese, with SVO word-order and grammaticalized verbs serving as prepositions. In this respect, Tai is in fact more similar to Chinese than is Tibeto-Burman, which overwhelmingly has SOV order. A considerable number of obviously related lexical items are common to Tai and Chinese, and these generally correspond regularly in their tonal categories. It is no wonder that Chinese linguists are universally persuaded of the genetic relationship between Chinese and Tai-Kadai, a view also espoused by a succession of Western scholars (see e.g. 118, 177, 212).

An alternative view was proposed by Benedict as early as 1942 (3), who insisted that the lexical correspondences between Chinese and Tai did not include much core vocabulary and were due to prehistoric contact. Instead he suggested that Tai-Kadai was genetically related to the polysyllabic and nontonal Austronesian (= Malayo-Polynesian) family, in a supergroup that he later dubbed "Austro-T(h)ai" (7). Just before publication of his 1975 book, he threw Hmong-Mien into "Austro-Thai" (AT) for good measure; and in the latest version of his theory he also includes Japanese (10)! Before one gets carried away by the sheer sweep of Benedict's conception, however, it must be said that his etymologies are by no means of uniform quality (see 131, 146).

The "AT hypothesis" did not at first win many adherents, and it is still highly controversial. Yet it has the great merit of downgrading tonality as a criterion for genetic relationship. As mechanisms of tonogenesis have become better understood, it has been widely accepted that tones may arise independently in genetically unrelated languages, and that tone systems are readily diffusible from one language family to another—i.e. words may be borrowed with their tones attached. (This is certainly the case with Vietnamese, which belongs genetically

to the nontonal Mon-Khmer family, but which has developed a full-blown tone system after millennia of Chinese influence (see Figure 3). The basic tonogenetic mechanisms were first explained by Haudricourt (72), but the term "tonogenesis" itself first appeared some 15 years later (122, 126; see also 153, 204).

The Tai family comprises many other languages besides Thai (= Siamese), the national language of Thailand. [It has become traditional to refer to the whole family to which Siamese belongs as "Tai" (without an *h*), reserving the spelling "Thai" for Siamese itself.] The Tai peoples (Siamese, Lao, Shan, etc) are relatively recent arrivals to peninsular Southeast Asia, having gradually percolated southward from their homeland in what is now China south of the Yangtze. The family's three subgroups, Northern, Central, and Southwest Tai, are all close to one another, implying that the period of Proto-Tai unity lay in the relatively recent past, perhaps around the first half of the first millennium AD.

F. K. Li (98) has provided the most extensive and systematic compendium of Proto-Tai (PTai) in a handbook containing reconstructions of more than 1300 Proto-Tai etyma. Because of the relatively shallow time-depth (perhaps 2000 years) that can be assumed for Proto-Tai (as opposed, for example, to PTB), its reconstruction is more solid and detailed than that of any other major language family of Southeast Asia (see 70:197). However, once one leaves the safe domain of Tai proper in search of wider affiliations, the task of the comparativist becomes difficult indeed. The closest relatives of Tai are the languages of the Kam-Sui (or "Dong-Shui") group of Southern China (see 97). Mainland Chinese linguists are now actively pursuing research into hitherto unknown languages of this group—e.g. Mulao (201) and Maonan (99). Yet the comparison of PTai and Proto-Kam-Sui presents many unsolved problems, which reach monumental proportions when we try to bring into the picture the motley languages lumped together under the rubric "Kadai." These latter languages, spoken in isolated backwoods pockets in Southeast China (*Lakkia* and *Gelao*), the island of Hainan [*Be* (= *Ong-be*) and *Li* (= *Hlai*)], and Northern Vietnam (*Lati* and *Laqua*), are obviously related somehow to Tai proper, yet the correspondences are highly irregular, bespeaking a long period of independent evolution as well as massive influence and contamination from Chinese. Until very recently the material available on these languages was pitifully inadequate for the intricate comparative work required to relate them to Tai proper. Now, however, there has been an avalanche of reliable new data recorded for the Kadai languages spoken on Chinese territory. (For Gelao, see 75; for Be, see 68; for Lakkia, see 73, 184; and for Hlai, see 144, 167.) The Kadai languages of Vietnam are still virtually a total mystery and are likely to remain so for some time. All in all one could well claim that there is no more complex and tricky problem in all of Southeast Asian historical linguistics than the reconstruction of Proto-Tai-Kadai.

The wider affiliations of Hmong-Mien (HM) are even more obscure than those of Tai-Kadai, but now that so much new material on HM dialects is coming out of China (e.g. 115, 171, 199, 200), we can be sure that our perspective on this language family will soon be radically refined. The Hmong and the Mien peoples

Figure 3 Chinese tonal influence on Tai-Kadai, Hmong-Mien, and Vietnamese

(better known by their Chinese names, Miao and Yao) have existed side by side with the other ethnic minority groups of Southern China for thousands of years, managing to preserve their cultural and linguistic identity in spite of the eventually overwhelming numerical proponderance of Han Chinese south of the Yangtze. The family consists of only two major groups, Hmongic and Mienic, though each is highly diversified into a large number of "dialects," some of which are so different from each other that they could well be considered separate languages (187).

Like Tai-Kadai, the Hmong-Mien (HM) languages are typologically very Chinese-like, monosyllabic, and highly tonal (24, 25). Languages with as many as 12 tones have been reported. The consonantal and vocalic systems of HM languages are also of uncommon richness. It is frequent to have 7 or 8 positions of articulation for obstruents. Complex or "marked" consonants (postvelars, prenasalized stops and affricates, voiceless sonorants, etc) abound, and the vowel systems of these languages are also complex. Wang Fushi (199) hypothesizes no fewer than 121 initial consonants and consonant clusters for Proto-Hmong, as well as 32 proto-finals. In general, the Hmong languages preserve initials better than Mien, while Mien preserves finals better than Hmong (44). The task of the Hmong-Mien comparativist is thus not impossible, though serious historical work on the family is still in its early stages (71, 171, 183). Good dictionaries of HM languages are available (e.g. 11, 76, 102, 109, 170, 215).

The Chinese influence on Mien is even stronger than it is on Hmong. More than 50% of the lexicon in some Mien dialects are Chinese loanwords. In other aspects of culture as well the Mien are highly Sinicized. Many Mien have been literate in Chinese, and Mien religion has been shown to derive from Chinese Taoism (93).

The Vietnam War caused a worldwide diaspora of the Hmong-Mien peoples. Many Miao and Yao cooperated with the American side during the war and were given preferential emigration opportunities from the refugee camps. Due to the tremendous influx of refugees to this country from Laos and Thailand (there are now about 80,000 Hmong and Mien in California alone), the United States is well on the way to becoming an important center of Hmong-Mien studies. Linguists resident in the United States are now doing increasingly sophisticated work on the phonology and grammar of the Hmong-Mien languages (e.g. 31, 32, 35, 82, 173), and it will not be long before American comparativists jump into the field with both feet. There has recently been founded an International Association for Yao Studies, based in Hong Kong, which has just sponsored its Third International Colloquium in Toulouse. A whole special issue of the journal LTBA, containing 14 articles, has been devoted to Hmong-Mien (Vol. 10(2), 1987).

Sino-Tibetan and areal phonology

The PTB syllable canon was quite complex, and may be reconstructed as follows:

*(P)(P) C_i (G) V (ɤ) (C_f) (s),

where P = prefix, C_i = initial consonant, G = glide (-r- -l- -w- -y-), : = vowel length, C_f = final consonant, and s = suffixal -s. Many modern TB languages (especially Sinospheric ones) have vastly simpler syllables—e.g. Lahu, with the canon

T
(C) V

where T = tone. Figuring out how these complex proto-syllables map into their simpler descendants is one of the most fascinating aspects of ST historical phonology. E.g. PTB *b-r-gyat 'eight' → WT brygad, Lahu hí; PTB *k-r-wat 'leech' → WB krwat, Magari ləwat, Lahu vèʔ).

In general we may distinguish between the polysyllabic languages of the Austronesian family, the "sesquisyllabic" languages of Mon-Khmer, and the monosyllabic ST, Tai-Kadai, and Hmong-Mien families. (The term *sesquisyllabic* introduced in reference 126, meaning "syllable-and-a-half," refers to words like Khmer /phənom/ 'mountain,' where the main part of the syllable is preceded by a "minor syllable" with shwa vocalism.) The development of tonal systems reaches its full flowering in languages of the monosyllabic type.

Historical linguistics in this area is forced to accord as much importance to tonal correspondences as to consonants and vowels. Not only do the tone systems of related languages correspond regularly, but tone systems and tonal categories may even be borrowed from language to language in a regular way, even in cases where the borrowing language was originally atonal. Tone systems are developed more elaborately in East and Southeast Asia than anywhere in the world. (Nothing in Mesoamerica or Africa approaches them in complexity.) The whole spectrum of tone-language types is exemplified in the ST area, ranging from the "omnisyllabic" systems of Chinese, Loloish, or Tai, where virtually every syllable is under a distinctive tone; to the "word-tone" systems of certain Himalayish TB languages, where tone contrasts are spread over two or more syllables (e.g. 152, 157); to the rudimentary pitch-accent systems of some Bodo-Garo or Abor-Miri-Dafla languages, where there is only a two-way contrast of low functional load between higher and lower pitch. No theory of tone and no taxonomy of tonal systems can afford to ignore the Sino-Tibetan linguistic area.

A major unresolved question is whether tonal contrasts should be imputed to the PTB/PST stage. Attempts to derive all modern TB tone-systems from a single proto-tone system have not been convincing (5), and it seems best to adopt a "polygenetic" rather than a "monogenetic" position: Tonal contrasts seem to have come and gone repeatedly and independently in the various branches of this "tone-prone" family. The most anciently attested TB language, Written Tibetan (WT), was toneless. Several modern languages have some tonal and some atonal dialects (Modern Tibetan, Qiang).

Sino-Tibetan and areal morphology

Variability and word-families: historical morphophonemics

As in any other language family, the proto-forms of ST cannot be conceived of as having been invariant in phonological shape or semantic content. Rather, ST etyma form clusters of morphophonemically related sub-roots, which have traditionally been referred to as *word families* (see, e.g., 29:173–4; 85; 202), though only recently has an attempt been made to give this concept a more precise theoretical basis (132). The "*allofams*" of a word family may differ from each other by their prefixes, by the voicing or voicelessness of their initial consonant, by their nuclear vowel, by their final consonant, and/or by their tone. These patterns of variation are not random but fall into certain well-defined classes of phenomena. Great care needs to be exercised in attributing particular forms in modern languages to the particular proto-allofam from which they descend.

Compounding and phonological bulk

Classical Chinese, with its relatively rich consonantism, was strictly monosyllabic, with the syntactic word and the phonological syllable virtually coextensive. In phonologically eroded modern dialects like Mandarin, however, most *words* are now dissyllabic, though almost all of them can still be analyzed into monosyllabic constituent morphemes. Y. R. Chao (27–29) once concocted three little stories in Classical Chinese style consisting entirely of Mandarin homophonous monosyllabic words (*shi, ji,* and *yi,* respectively, under various tones). For the Old Chinese listener, these stories would have been understandable orally, since most of the syllables were still pronounced differently. Thus the three words of the title of the story *Ten Stone Lions* (Mand. **Shí Shí Shī**) were pronounced something like ***Dyep Dyak Syər** in OC. For the stories to be understandable to a modern Mandarin listener, they would have to be recast using dissyllabic compounds or collocations to differentiate the now individually homophonous syllables (e.g. **Shí-ge** 'ten,' **Shí-tou** 'stone,' **Shī-zi** 'lion'). (An analogy is provided by those dialects of Southern American English that have merged /-in/ and /-en/ to [in], so that *pin* and *pen* must now be disambiguated by the compound forms *stick-pin* and *ink-pin,* respectively.)

In Loloish, also, the consonantal simplification of monosyllables has led to homophony on a grand scale. There are, for example, at least five Lahu morphemes pronounced **ha** that descend from once consonantally distinct PTB etyma. Besides sharing the same initial and vowel, these syllables are also *tonally* homophonous, all being under the mid-tone, unmarked in the transcription. (Many other **ha**-morphemes occur under the other 6 tones; see Table 2).

It is fascinating to observe how the various ST languages focus on different semantic features of the root noun in compound formation (what we might call the "blind-men-and-the-elephant" principle). Thus, the monosyllabic PTB root

***na** 'ear' appears compounded with morphemes meaning 'flat; leaf', or 'horn' or 'hole,' according to the individual daughter language's predilections:

***na** + ***p(r)ak** 'leaf' > Limbu (Himalayish) **nɛ-bhak** 'ear'
***na** + ***ruŋ** 'horn' > Bokar (Mirish) **nya-ruŋ** 'ear'
***na** + ***kwar** 'hole' > Rongmei (Naga) **nu-kúan** 'ear'

For the network of semantic associations involved, see Figure 4.

Syllable structure and the compounding/affixation cycle

Although the ST languages are not famous for complex morphology, there are pervasive processes of affixation evident. These affixes may be fully syllabic—e.g. ***a-** 'kinship prefix,' ***-pa/-ba** 'verb nominalizing suffix'; or they may be single consonants as far back as it is possible to reconstruct them—e.g. the dental trio of suffixes ***-t, *-n, *-s,** with a variety of derivational functions such as nominalization and collectivization (see 4:98–102; 210; 211). Sometimes they have a relatively clear meaning—e.g. the diachronically controversial pronominal agreement suffixes (41, 156, 196, 197; see below), and the paradigmatically opposable pair ***s-** 'causativizer; transitivizer; outer-directed action', and ***m-** 'stativizer; inner-directed action' (WT **mnam** 'have a smell,' **snam** 'sniff something'; see 130). Frequently, however, the meanings of affixes are elusive or obscure—e.g. ***-pa/-ma** 'bulk-providers in nouns,' or the PTB prefixes ***d- *b- *g- *r- *l-** (4:109–17).

Most interesting are the cases where modern affixes can be shown to be reduced or "cliticized" variants of once fully syllabic and meaningful root morphemes. Diphthongal vowels may represent fusions of a vowel-initial particle with the previous root-morpheme (143, 145). Final consonants are sometimes fusional

Table 2 Lahu homophonous monosyllables

	PTB	*PLB*	*Lahu monosyllables*	*Lahu disyllables*
'hundred'[a]	*b-r-qya	*ʔra[1]	ha	tê ha
'moon'	*s-gla	*s-la[3]	ha	ha-pa
'tongue'	*s-lya	*s-l(y)a[1]	ha	ha-tɛ̄
'spirit'	*s-hla	*sla[3]	ha	ɔ̀-ha
'winnow'	*g-ya(ɤp)	*ʔ-ya[1]	ha	ha ve

[a] Note that **ha** 'hundred' is not usable by itself, but must always be preceded by a numeral (e.g. **tê ha** 'one hundred'); the **-pa** in 'moon' is a meaningless suffix, ubiquitous in TB (cf WT **zla-ba** 'moon'). The **-tɛ̄** in 'tongue' looks like it once had an independent meaning, but recurs nowhere else in the language; the ɔ̀- prefix in 'spirit' (< PTB *ʔaŋ-) occurs as a bulk-provider before hundreds of Lahu roots; the particle **ve** in **ha ve** 'to winnow' is a nominalizer that occurs in the citation form of verbs (much like English *to*), serving to distinguish verbs from any homophonous nouns.

remnants of the second syllable in a compound or other collocation. Thus, Mandarin **-men** 'pluralizer' has lost its rhyme and fused its initial with the vowel of the preceding pronoun in rapid colloquial (**wǒ-men** 'we' > **wǒm, tā-men** 'they' > **tām**). The Northern dialects of Qiang (TB of Sichuan) may drop the vowel of the second elements of compounds, leading to secondary monosyllables with final consonants: 'seed' S. Qiang **zuə-pə**, N. Qiang **zəp**; 'day after tomorrow' S. Qiang **zuə-za**, N. Qiang **tśhaz** (see 9, 188). The aberrant-looking Angami Naga form **pfhə** 'bitter,' has been analyzed (136) as deriving from the fusion of an original dissyllable ***ka-ba** (where **-ba** was a nominalizing citation particle): ***ka-ba** > ***ka-wa** > ***kwa** > **pfhə**.

Conversely, it may be the first syllable of a compound or collocation that undergoes reduction, so its vowel becomes shwa and the dissyllable becomes a sesquisyllable (see above). In these cases the full meaning of the original first syllable may be lost, and "prefixization" occurs. Thus, the WB sesquisyllabic forms **pərwak** 'ant' and **səmak** 'son-in-law' derive from the Proto-Lolo-Burmese compounds ***buw-krwak** (PLB ***buw** 'bug') and ***zamak** (***za** 'son, child'), though modern Burmese speakers could not be aware of this. Similarly, the minor syllable of Siamese **sədɯ** 'navel' is derivable from **săaj** 'line; band' via the umbilical cord, while the unstressed **mə-** in many names of fruits and vegetables (e.g. **məmûaŋ** 'mango,' **məphráaw** 'coconut,' **məkhy̆a** 'eggplant') is a reduction of PTai ***hmaak** 'fruit' (98:92, 75).

It is then easy for a sesquisyllable to lose its first "minor" element, yielding secondary monosyllables that can then reacquire phonological bulk by being recompounded with fresh fully syllabic morphemes—and the cycle begins again.

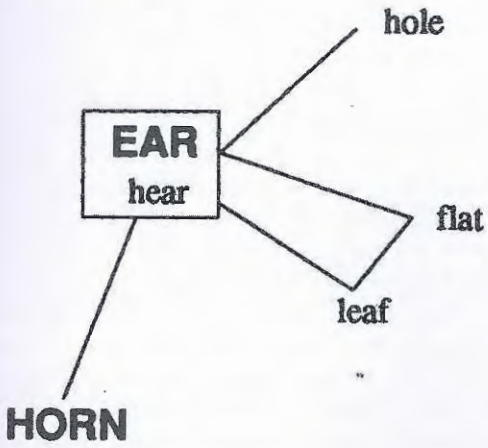


Figure 4 Semantic associations of *ear* in compound formation

ST and areal grammar

ST grammar is very different in flavor from that of Indo-European languages. On the other hand ST shares many characteristics with the grammars of other language families of East and Southeast Asia.

ST languages are "topic prominent" (96), in that NPs are freely topicalizable (movable to initial position in the clause), and (especially in the verb-final TB languages) the NPs of a clause occur in relatively free order. The VP is the dominant constituent in the clause, and sentences frequently lack "subjects." The notions of subject and object are in fact alien to ST grammar, as are such grammatical categories as active vs passive voice (91). Many TB languages (e.g. Tibetan, Newari, Akha) are ergative (often split ergative, like Limbu and Sherpa), with agents marked like instrumentals (45, 54, 196).

ST languages have a penchant for nominalizing whole sentences without embedding them into any larger unit, typically via a particle that is also used in the citation-form of verbs, and that has a relative and/or genitive function in other constructions. The connection between nominalization and relativization is often made explicit by using the same particle for both functions—e.g. Mandarin **de**, Lahu **ve** (42, 124, 125).

Aspect (not tense) is the major verbal category, so that notions like *completed action, change of state, irrealis, inchoative,* and *durative* are encoded more readily than *past, present,* and *future.* The single most satisfactory criterion for establishing that a ST word is a verb is its negatability. By this definition most words that translate as English adjectives are actually only a subclass of verbs (e.g. Lahu **qay** 'go,' **chu** 'be fat' > **mâ qay** 'not go,' **mâchu** 'not be fat').

In order to express abstract grammatical relationships, the isolating languages of the ST area have typically resorted to the specialization of full nouns and verbs. Verbs meaning GIVE, DWELL, PUT, FINISH become bleached semantically until they can serve as markers of verbal categories like CAUSATIVE, PROGRESSIVE, DURATIVE, COMPLETIVE. A noun meaning ROAD becomes a locative particle, while another meaning TOP PART turns into an accusative marker. Often a root-morpheme that has undergone grammatical specialization acquires a distinctive phonological shape, usually via destressing, sometimes assuming a special tone (e.g. Lahu **là** 'come' (verb; low-falling tone), acquires mid-tone **(la)** as it becomes a particle indicating 'motion toward', and high-falling tone **(lâ)** as a particle signalling 'non-3rd person beneficiary' (40, 147).

Verb concatenation is especially striking in the Loloish branch of TB. Lahu may juxtapose up to five verbs in a row in a single VP, one verb serving as semantic head while the others are grammaticalized or made more abstract to modify it, e.g. **ɣa qɔ̀ʔ yù tɔ̂ʔ pî** 'have to take it out for him again' ("obtain-return-TAKE-emerge-give") (52, 65, 120, 128). In Chinese and Sinospheric SVO languages like Thai, Hmong, Mien, and Vietnamese, verbs are grammaticalized into NP-markers that function like IE prepositions. Thus, "He cut the sugarcane with a knife" is literally "He use knife cut sugarcane" (31, 95, 147).

Some languages (e.g. Akha, Newari) have intricate systems of evidential particles that characterize the nature of the speaker's information—e.g. first-hand, hearsay, visual, auditory (45, 67). It is typical to have large repertoires of sentence-final particles whose function is to express emotional attitudes. These are more integrated intonationally into sentences than, for example, English interjections like *wow!,* and often occur in strings of two, three, or more. In some measure these fully syllabic morphemes serve the same function as intonation does in nontonal languages (125:380–90).

The lack of gender or number markers on ST, HM, and Tai nouns is somewhat compensated for by numeral classifiers, which serve to individuate nouns and may be said to "agree" with particular classes of nouns—e.g. flat, round, or elongated objects (67a, 169). Another characteristic way of lending phonologic/semantic substance to monosyllabic morphemes is via the morphological process of *elaboration,* which creates redundant (sometimes poetic) four-syllable expressions, of which the 1st and 3rd, or 2nd and 4th, syllables are often identical—e.g. Lahu **thī-ngɔ̂-thī-khā** 'silver and gold altars,' Burmese **cit-hrañ-lak-hrañ** 'patiently' ("mind-long-hand-long") or Tibetan **blo-gsal-lag-bde** 'intelligent; skilful' ("mind-bright-hand-apt") (125:81–88, 297–301; 140).

The historical syntax of the ST languages is coming increasingly into focus as an object of study. Because of its abundance of ancient written records, Chinese has so far received the lion's share of this attention, though the other ST languages will ultimately have just as much to teach us about the processes of grammatical change. A pioneering work in TB historical morphology is Wolfenden's early monograph (211). Good examples of more recent work in TB comparative/historical grammar are also available (e.g. 2, 40, 41, 203). Research into historical grammar depends fundamentally on progress in historical phonology and the enrichment of our conception of the protolexicon—the treasury of roots available for grammaticalization.

Several large problems remain in this area: *(a)* How did Chinese get to be SVO, when all other TB languages (except Karenic and Bai) are SOV (91, 207)? *(b)* How ancient are the pronominal agreement systems to be found in many TB languages (e.g. the Rai group of Eastern Nepal, Jingpho, Tangut), wherein the person (including 1st and 2nd person inclusive/exclusive) and number (including dual) of subject and/or object may be marked on the verb, producing agreement systems that are sometimes relatively simple (88, 89, 203) but that sometimes rival, for example, Algonkian in complexity (156, 196, 197)? Thus, Jingpho **ngai sa na ǹ-ngai** 'I will go' / **nang sa na ǹ-dai** 'you will go' / **ši sa na rà-ʔai** 'he will go' (**sa** 'go,' **na** 'future particle'). It remains to be seen whether such systems should be posited for Proto-TB, or whether they have arisen independently in the various branches of the family.

Comparative/historical phonology and semantics

The publication of *STC* (4), epoch-making as it was, marked only the beginning of the search for cognates and reconstructible etyma in the Sino-Tibetan family.

The *Conspectus,* as its name implies, was intended merely as an overview of its subject, and makes no pretensions to exhaustiveness. Even so, the number of proto-forms reconstructed is considerable. The TB part of the book contains 494 numbered cognate sets (and several dozen more unnumbered ones scattered in footnotes), as well as over 300 suggested comparisons between PTB etyma and Old Chinese forms. For an index of its TB/Old Chinese comparisons see Chou's 1972 paper (30). Other, more tentative repositories of Chinese/TB cognate identifications are also available (13, 33, 107).

Since 1972, scores of new TB etymologies have been proposed (e.g. 8, 129, 132, 133, 135, 137, 138, 145). The etymologies offered by Shafer (177) largely duplicate those of Benedict (4), but several other authors have contributed solid reconstructions at the proto-subgroup level (e.g. 48, 154). All this, however, constitutes no more than a good beginning. Whole branches and subfamilies of TB remain largely unexplored (e.g. the Abor-Miri-Dafla, Bodo-Garo, and Qiangic groups), and even the relatively well-known branches (Himalayish and Lolo-Burmese) have so far yielded up only a fraction of their etymological secrets.

It is only when one compares the stock of reliably reconstructed TB and ST roots with what has been accomplished for other language families that one realizes the full magnitude of the etymological job that remains to be done. The repertoire of reconstructed Proto-Indo-European etyma runs to several thousand items, though Indo-Europeanists have had a bit of a head start! Closer to home, it is estimated that over 4000 etyma will eventually be reconstructed for Proto-Mon-Khmer, even without counting the Munda branch of Austroasiatic (Gérard Diffloth, personal communication).

Of course progress is not merely a question of sheer weight of numbers of reconstructed forms. Even more important is the task of unraveling the myriads of interconnections—both morphophonemic and semantic—among the reconstructed items. It has already happened many times that proposed PTB or PST roots that had been considered etymologically distinct have turned out to be ultimately related. Conversely, many plausible candidates for cognacy have been shown to be unrelated. Elsewhere (132) I sketched out the theoretical framework and methodology for carrying out ST historical semantic research, and this approach has been continued for other semantic areas in subsequent articles (135, 137, 138, 140, 141). Since 1987, the *Sino-Tibetan Etymological Dictionary and Thesaurus* project (STEDT) at Berkeley, jointly funded by NEH and NSF, has been compiling a massive database of lexical material on the ST languages in an attempt to reconstruct whole semantic fields of the proto-lexicon. The goal is to produce an *Etymological Thesaurus* (151), not merely an etymological dictionary of the conventional type. The first volume, on body part terminology, is now well advanced. As our understanding of the semantic structure of the ST proto-lexicon is enriched, it should be possible to reconstruct ancient cultural patterns and thought processes (cf 18 for a similar approach to the Indo-European proto-lexicon). At the same time, the "sound-laws" for the phonological evolution of the ST languages are undergoing constant refinement.

In sum, the goal of Sino-Tibetanists should be to put ST studies on a par with Indo-European, in terms of its potential contributions to a wide range of general linguistic issues. In many ways the problems faced by diachronic and areal linguists are similar everywhere. Yet each area of the world has its own typological/cultural/historical flavor, and ST has many unique features that will be of interest to the general linguistic public. It is high time to "mainstream" ST linguistics.

Note

1 In the following list, conferences held outside the United States are in boldface.

Literature cited

1. Allen, N. J. 1975. *Sketch of Thulung Grammar.* With three texts and a glossary. Ithaca, NY: Cornell Univ. East Asia Papers, #6
2. Bauman, J. J. 1975. *Pronouns and pronominal morphology in Tibeto-Burman.* PhD thesis. Univ. Calif., Berkeley. 319 pp.
3. Benedict, P. K. 1942. Thai, Kadai, and Indonesian: a new alignment in Southeastern Asia. *Am. Anthropol.* 44:576–601
4. Benedict, P. K. 1972. *Sino-Tibetan: a Conspectus* ("STC"). Contributing ed. J. A. Matisoff. Princeton-Cambridge Ser. Chinese Linguist., #2. NY: Cambridge Univ. Press. 230 pp.
5. Benedict, P. K. 1972. The Sino-Tibetan tonal system. In *Langues et techniques, nature et société,* ed. J. Barrau, J. M. C. Thomas, L. Bernot, pp. 25–34. Paris: Klincksieck
6. Benedict, P. K. 1975. Where it all began: memories of Robert Shafer and the "Sino-Tibetan Linguistics Project." *Linguist. Tibeto-Burman Area* 2(1):81–92
7. Benedict, P. K. 1975. *Austro-Thai Language and Culture.* With a glossary of roots. New Haven: Hum. Relat. Area Files Press. 490 pp.
8. Benedict, P. K. 1976. Sino-Tibetan: another look. *J. Am. Orient. Soc.* 96(2): 167–97
9. Benedict, P. K. 1983. Qiang monosyllabization: a third phase in the cycle. *Linguist. Tibeto-Burman Area* 7(2):113–14
10. Benedict, P. K. 1990. *Austro-Tai/Japanese.* Ann Arbor: Karoma Press
11. Bertrais-Charrier, Y. 1964. *Dictionnaire Hmong (Meo Blanc)-Français.* Vientiane: Mission Catholique. Reprinted 1978, Bangkok: Assumption Press
12. Bhat, D. N. S. 1969. *Tankhur Naga Vocabulary.* Poona: Deccan Coll. Postgrad. Res. Inst.
13. Bodman, N. C. 1980. Proto-Chinese and Sino-Tibetan: data towards establishing the nature of the relationship. In *Contributions to Historical Linguistics: Issues and Materials,* ed. F. van Coetsam, L. R. Waugh, pp. 34–199. Leiden: E. J. Brill
14. Bodman, N. C. 1988. On the place of Lepcha in Sino-Tibetan: a lexical comparison. *Linguist. Tibeto-Burman Area* 11(1):1–26
15. Bradley, D. 1975. Nahsi and Proto-Burmese-Lolo. *Linguist. Tibeto-Burman Area* 2(1):93–150
16. Bradley, D. 1978. *Proto-Loloish.* Copenhagen/London: Scand. Inst. Asian Stud. Monogr. Ser. #39

17. Bradley, D. 1979. *Lahu Dialects.* Oriental Monogr. Ser. #23. Canberra: Australian Natl. Univ.
18. Buck, C. D. 1949. *A Dictionary of Selected Synonyms in the Principal Indo-European Languages: A Contribution to the History of Ideas.* Chicago: Univ. Chicago Press
19. Burling, R. 1961. *A Garo Grammar.* Monogr. Ser. #25. Poona: Deccan Coll. Postgrad Res. Inst.
20. Burling, R. 1965. *Hill Farms and Padi Fields: Life in Mainland Southeast Asia* Englewood Cliffs, NJ: Prentice-Hall
21. Burling, R. 1968. Proto-Lolo-Burmese. Spec. Publ. *Int. J. Am. Linguist.* 33(2), Part II. Issued simultaneously as *Indiana Publications in Anthropology and Linguistics #43.* The Hague: Mouton and Co.
22. Burling, R. 1971. The historical place of Jinghpaw within Tibeto-Burman. *Occas. Pap. Wolfenden Soc. Tibeto-Burman Linguist.* 2:1–54
23. Caughley, R. C. 1982. *The Syntax and Morphology of the Verb in Chepang.* Canberra: Pacific Linguist. B-84
24. Chang, K. 1966. A comparative study of the Yao tone system. *Language* 42 (2):303–10
25. Chang, K. 1972. The reconstruction of Proto-Miao-Yao tones. *Bull. Inst. Hist. Philol. (Taipei)* 44(4):541–628
26. Chang, K., Chang, B. S. 1972. *The Proto-Chinese Final System and the Ch'ieh-Yün.* Taipei: Academia Sinica Monogr., Ser. A, #26
27. Chao, Y. R. 1960/1962. *Articles on Chinese Language.* Encyclopedia Britannica
28. Chao, Y. R. 1967. *Language and Symbolic Systems.* Cambridge: Cambridge Univ. Press
29. Chao, Y. R. 1968. *A Grammar of Spoken Chinese.* Berkeley/Los Angeles: Univ. Calif. Press. 847 pp.
30. Chou, F. K. 1972. Archaic Chinese and Sino-Tibetan. *J. Inst. Chinese Stud. Chinese Univ. Hongkong* 5(1):159–237 (In Chinese)
31. Clark, M. 1979. Coverbs: evidence for the derivation of prepositions from verbs (new evidence from Hmong). *Univ. Hawaii: Work. Pap. Linguist.* 11(2):1–12
32. Clark, M. 1985. Asking questions in Hmong and other Southeast Asian languages. *Linguist. Tibeto-Burman Area* 8(2):60–67
33. Coblin, W. S. 1986. *A Sinologist's Handlist of Sino-Tibetan Lexical Comparisons.* Monumenta Serica Monogr. Ser. 18. Nettetal: Steyler Verlag
34. Conrady, A. 1896. *Eine indochinesische Causativ-Denominativ-Bildung und ihr Zusammenhang mit den Tonaccenten: ein Beitrag zur vergleichenden Grammatik der Indochinesischen Sprachen, insonderheit des Tibetischen, Barmanischen, Siamesischen und Chinesischen.* Leipzig: Otto Harrassowitz
35. Court, C. A. 1985. *Fundamentals of Iu Mien (Yao) grammar.* PhD thesis. Univ. Calif., Berkeley. 262 pp.
36. Dai Qingxia, et al. 1981/1983. *Chinese-Jingpho Dictionary (1981), Jingpho-Chinese Dictionary (1983).* Kunming: Yunnan People's Press (In Chinese)
37. Das Gupta, K. 1963. *An Introduction to the Gallong Language.* Shillong: Philol. Sect., Res. Dept., Northeast Frontier Agency
38. Das Gupta, K. 1968. *An Introduction to Central Monpa.* Shillong: Philol. Sect., Res. Dept., Northeast Frontier Agency
39. Das Gupta, K. 1971. *An Introduction to the Nocte Language.* Shillong: Philol. Sect., Res. Dept., Northeast Frontier Agency

40. DeLancey, S. 1984. Etymological notes on Tibeto-Burman case particles. *Linguist. Tibeto-Burman Area* 8(1):59–77
41. DeLancey, S. 1984. The analysis-synthesis-lexis cycle in Tibeto-Burman: a case study in motivated change. In *Iconicity in Syntax,* ed. J. Haiman, pp. 367–89. Amsterdam: Benjamins
42. DeLancey, S. 1986. Relativization as nominalization in Tibetan and Newari. Paper presented at 19th Sino-Tibetan Conf., Columbus, Ohio
43. Dell, F. 1981. *La langue Bai, Phonologie et Lexique.* Paris: Ecole des Hautes Etudes en Sciences Sociales. 174 pp.
44. Downer, G. B. 1979. Proto-West-Miao rhymes. Paper presented at 12th Sino-Tibetan Conf., Paris
45. Egerod, S. 1985. Typological features in Akha. See Ref. 195, pp. 96–104
46. Fei Hsiao-tung. 1980. Ethnic identification in China. Translated by Wang Huimin & Wu Zenfang. *Soc. Sci. China* 1:94–107
47. Fraser, J. O. 1922. *Handbook of the Lisu (Yawyin) Language.* Rangoon: Government Printing House
48. French, W. T. 1983. *Northern Naga: a Tibeto-Burman mesolanguage.* PhD thesis. City Univ. NY. 737 pp.
49. Fu Maoji. 1950. *A Descriptive grammar of Lolo.* D. Phil. dissertation, Cambridge Univ.
50. Fu Maoji. 1985. Lecture #15 on the investigation and study of national minority languages. *Minzu Yuwen* 6:52–55. (In Chinese)
51. Gao Huanian. 1958. *A Study of the Grammar of the Yi Language.* Beijing: Scientific Publishing Co. (In Chinese)
52. Genetti, C. E. 1986. The grammaticalization of the Newari verb tɔl. *Linguist. Tibeto-Burman Area* 9(2):53–70
53. Genetti, C. E. 1988. Notes on the structure of the Sunwari transitive verb. *Linguist. Tibeto-Burman Area* 11(2):62–92
54. Genetti, C. E. 1990. *A Descriptive and historical account of the Dolakha Newari dialect.* PhD thesis. Univ. Oregon. 362 pp.
55. Glover, W. W. 1974. *Sememic and Grammatical Structures in Gurung (Nepal).* Summer Inst. Linguist. Kathmandu: Tribhuvan Univ. Press
56. Gorman, C. F. 1970. Excavations at Spirit Cave, Northern Thailand: some interim interpretations. *Asian Perspect.* 13:79–107
57. Gorman, C. F. 1971. The Hoabinhian and after: subsistence patterns in Southeast Asia during the late Pleistocene and early Recent periods. *World Archaeol.* 2:300–20
58. Gorman, C. F. 1974. Modèles a priori et préhistoire de la Thailande. *Etudes Rurales* 53–56:41–71
59. Grierson, Sir G. A., Konow, S., eds. 1903–28. *Linguistic Survey of India.* Vol. III, Parts 1–3, *Tibeto-Burman Family.* Reprinted 1967 by Motilal Banarsidass (Delhi, Varanasi, Patna)
60. Grimes, B., ed. 1988. *Ethnologue: Languages of the World.* Dallas: Summer Inst. Linguist. 748 pp. 11th ed.
61. Grüssner, K.-H. 1978. *Arleng Alam, die Sprache der Mikir.* Wiesbaden: Franz Steiner
62. Hale, A., ed. 1973. *Clause, Sentence, and Discourse Patterns in Selected Languages of Nepal,* Parts I-IV. Kathmandu: Summer Inst. Linguist. and Tribhuvan Univ. Press

63. Hale, A. 1982. *Research on Tibeto-Burman Languages. Trends in Linguistics, State of the Art Report #14.* NY: Mouton and Co.
64. Hanson, O. 1906. *A Dictionary of the Kachin Language.* Reprinted 1954, Rangoon: Baptist Board of Publications
65. Hansson, I.-L. 1985. Verb concatenation in Akha. See Ref. 195, pp. 287–309
66. Hansson, I.-L. 1989. A comparison of Akha, Hani, Khatu and Pijo. *Linguist. Tibeto-Burman Area* 12(1):6–91
67. Hargreaves, D. 1983. Evidentiality in Newari. Paper presented at 16th Sino-Tibetan Conf., Seattle
67a. Hashimoto, M. J. 1977. The genealogy and role of the classifier construction in Sino-Tibetan. *Comput. Anal. Asian African Lang.* (Tokyo) 7:69–78
68. Hashimoto, M. J. 1980. *The Be Language: A Classified Lexicon of Its Limkow Dialect.* Tokyo: Inst. Stud. Lang. Cult. Asia and Africa
69. Haudricourt, A.-G. 1942–45. Restitution du karen commun. *Bull. Soc. Linguist. Paris* 42(1):103–11
70. Haudricourt, A.-G. 1948. Les phonèmes et le vocabulaire du thai commun. *J. Asiat.* 236:197–238
71. Haudricourt, A.-G. 1954. Introduction à la phonologie historique des langues miao-yao. *Bull. Ecole Française d'Extrême Orient* 44(2):555–76
72. Haudricourt, A.-G. 1954. De l'origine des tons en viêtnamien. *J. Asiat.* 242:68–82
73. Haudricourt, A.-G. 1967. La langue lakkia. *Bull. Soc. Linguist. Paris* 62(1):165–82
74. Haudricourt, A.-G. 1975. Le système de tons du karen commun. *Bull. Soc. Linguist. Paris* 70:339–43
75. He Jiashan. 1983. *Outline Grammar of the Gelao Language.* Beijing: People's Press (In Chinese)
76. Heimbach, E. E. 1969. *White Meo-English Dictionary.* Data Pap. #75. Ithaca, NY: Southeast Asia Prog., Cornell Univ.
77. Henderson, E. J. A. 1965. *Tiddim Chin: a Descriptive Analysis of Two Texts.* London: Oxford Univ. Press
78. Henderson, E. J. A. 1991. *Bwe Karen-English Dictionary.* In preparation
79. Hodgson, B. H. 1853. On the Indo-Chinese Borderers and their connexion with the Himaláyans and Tibetans. *J. Asiat. Soc. Bengal* 22:1–25
80. Hope, E. R. 1974. *The Deep Syntax of Lisu Sentences: A Transformational Case Grammar.* Canberra: Pacific Linguist. B-34.
81. Huffman, F. E. 1985. *Bibliography and Index of Mainland Southeast Asian Languages and Linguistics.* New Haven/London: Yale Univ. Press
82. Jaisser, A. 1984. Complementation in Hmong. MA thesis. San Diego State Univ.
83. Jones, R. B. 1961. *Karen Linguistic Studies: Description, Comparison, and Texts.* Univ. Calif. Publ. Linguist. #25. Berkeley/Los Angeles: Univ. Calif. Press
84. Karlgren, B. 1923. *Analytic Dictionary of Chinese and Japanese.* Paris: Librairie orientaliste Paul Geuthner. Reprinted 1974, New York: Dover
85. Karlgren, B. 1933. Word families in Chinese. *Bull. Mus. Far Eastern Antiq. (Stockholm)* 5:5–120
86. Karlgren, B. 1954. Compendium of Phonetics in Ancient and Archaic Chinese. *Bull. Mus. Far Eastern Antiq. (Stockholm)* 26:211–367
87. Karlgren, B. 1957. Grammata Serica Recensa ("GSR"). *Bull. Mus. Far Eastern Antiq. (Stockholm)* 29(1):1–332

88. Kepping, K. B. 1975. Subject and object agreement in the Tangut verb. *Linguist. Tibeto-Burman Area* 2(2):219–32
89. Kepping, K. B. 1982. Once again on the agreement of the Tangut verb. *Linguist. Tibeto-Burman Area* 7(1):39–54
90. La Polla, R. J. 1987. Dulong and Proto-Tibeto-Burman. *Linguist. Tibeto-Burman Area* 10(1):1–43
91. La Polla, R. J. 1990. *Grammatical relations in Chinese.* PhD thesis. Univ. Calif., Berkeley
92. La Polla, R. J., Lowe, J. B. 1989. *Bibliography of the International Conferences on Sino-Tibetan Languages and Linguistics,* I–XXI. STEDT Monogr. Ser., #1. Inst. Int. Stud. Berkeley: Univ. Calif. Press
93. Lemoine, J. 1983. Yao religion and society. See Ref. 155, pp. 195–211
94. Lewis, P. 1968. *Akha-English Dictionary.* Data Pap. #70, Southeast Asia Prog., Ithaca, NY: Cornell Univ.
95. Li, C., Thompson, S. A. 1974. Coverbs in Mandarin Chinese: verbs or prepositions? *J. Chinese Linguist.* 2(3): 257–78
96. Li, C., Thompson, S. A. 1974. Chinese as a topic-prominent language. Paper presented at 7th Sino-Tibetan Conf., Georgia State Univ.
97. Li, F.-K. 1965. The Tai and the Kam-Sui languages. *Lingua* 14:148–79
98. Li, F.-K. 1977. *A Handbook of Comparative Tai.* Oceanic Linguist. Spec. Publ. #15. Honolulu: Univ. Press Hawaii
99. Liang Min. 1980. *Outline Grammar of the Maonan Language.* Beijing: People's Press (In Chinese)
100. Lo Ch'iang-p'ei. 1945. A preliminary study of the Trung language of Kung Shan. *Harvard J. Asiat. Stud.* 8(3–4): 343–48
101. Löffler, L. 1966. The contribution of Mru to Sino-Tibetan linguistics. *Z. Dtsch. Morgenländisch. Ges.* 116(1): 118–59
102. Lombard, S. 1968. *Yao-English Dictionary,* ed. H. Purnell. Data Pap. #69, Southeast Asia Prog., Ithaca, NY: Cornell Univ. Press
103. Lorrain, J. H. 1907. *A Dictionary of the Abor-Miri Language.* Shillong: Eastern Bengal and Assam Secretariat Press
104. Lorrain, J. H. 1940. *Dictionary of the Lushai Language.* Bibliotheca Indica #261. Calcutta: Royal Asiat. Soc. Bengal
105. Lorrain, R. A. 1951. *Grammar and Dictionary of the Lakher or Mara Language.* Gauhati, Assam: Dept. Hist. Antiq. Stud.
106. Lu Shaozun. 1980. A brief description of the Pumi language. *Minzu Yuwen* 4:58–72 (In Chinese)
107. Luce, G. H. 1981. *A Comparative Wordlist of Old Burmese, Chinese, and Tibetan.* London: Sch. Oriental African Stud. 88 pp.
108. Luce, G. H. 1986. *Phases of Pre-Pagán Burma: Languages and History.* Oxford/NY: Oxford Univ. Press. 2 vols.
109. Lyman, T. A. 1974. *Dictionary of Mong Njua: a Miao Language of Southeast Asia.* Janua Linguarum, Ser. Practica #123. The Hague: Mouton and Co.
110. Ma Xueliang. 1949. Annotated translation of the Lolo Classic of Rites, Cures, and Sacrifices. Beijing: Academia Sinica. *Bull. Inst. Hist. Philol.* 20:577–666 (In Chinese)
111. Ma Xueliang. 1951. *A Study of Sani, a Yi Dialect.* Chinese Academy of Sciences Monogr. #2. Shanghai: Commercial Press. (In Chinese)

112. Ma Xueliang. 1986/1987. *Anthology of Old Yi Inscriptional Literature.* Chengdu: Sichuan People's Publishing Co. (In Chinese) 3 vols.
113. Mainwaring, G. B., Grünwedel, A. 1898. *Dictionary of the Lepcha Language.* Berlin: Unger Brothers
114. Malla, K. P. 1985. *The Newari Language: a Working Outline.* Monumenta Serindica #14. Tokyo: Inst. Stud. Lang. Cult. Asia & Africa. 111 pp.
115. Mao Zongwu, Meng Chaoji, Zheng Zongze. 1982. *Outline Grammar of the Yao Language.* Beijing: People's Press
116. Maran, L. *A Dictionary of Modern Jinghpaw.* Revised and enlarged version of Hanson 1906, with tones indicated. 1441 pp. (manuscript)
117. Marrison, G. E. 1967. *The classification of the Naga languages of North East India.* Vol. I, 292 pp.; Vol. II, 460 pp. D. Phil. thesis. Sch. Oriental African Stud., Univ. London
118. Maspero, H. 1952. Les langues thai. In *Les Langues du Monde,* ed. A. Meillet, G. Cohen, pp. 571–88. Paris: Cent. Natl. Rech. Sci. (nouv. ed.)
119. Matisoff, J.A.1969. Lahu and Proto-Lolo-Burmese. *Occas. Pap. Wolfenden Soc. Tibeto-Burman Linguist.* 1:117–221
120. Matisoff, J. A. 1969. Verb concatenation in Lahu: the syntax and semantics of 'simple' juxtaposition. *Acta Linguist. Hafniensia (Copenhagen)* 12(1):69–120
121. Matisoff, J. A. 1969. Lahu bilingual humor. *Acta Linguist. Hafniensia (Copenhagen)* 12(2):171–206
122. Matisoff, J. A. 1970. Glottal dissimilation and the Lahu high-rising tone: a tonogenetic case study. *J. Am. Oriental Soc.* 90(1):13–44
123. Matisoff, J. A. 1972. *The Loloish Tonal Split Revisited.* Res. Monogr. #7. Berkeley: Cent. S. & SE Asia Stud., Univ. Calif. 88 pp.
124. Matisoff, J. A. 1972. Lahu nominalization, relativization, and genitivization. In *Syntax and Semantics,* ed. J. Kimball, 1:237–57. New York: Seminar Press
125. Matisoff, J. A. 1973. *The Grammar of Lahu.* Univ. Calif. Publ. Linguist. #75. Berkeley/Los Angeles: Univ. Calif. Press. 673 pp. Reprinted 1982.
126. Matisoff, J. A. 1973. Tonogenesis in Southeast Asia. In *Consonant Types and Tone,* ed. L. M. Hyman, 71–95. Los Angeles: Univ. Calif. Los Angeles
127. Matisoff, J. A. 1973. The annual Sino-Tibetan conferences: the first five years, 1968–1972. *J. Chinese Linguist.* 1(1):152–62
128. Matisoff, J. A. 1974. Verb concatenation in Kachin. *Linguist. Tibeto-Burman Area* 1(1):186–206
129. Matisoff, J. A. 1974. The tones of Jinghpaw and Lolo-Burmese: common origin vs. independent development. *Acta Linguist. Hafniensia (Copenhagen)* 15(2):153–212
130. Matisoff, J. A. 1976. Lahu causative constructions: case hierarchies and the morphology/syntax cycle in a Tibeto-Burman perspective. In *Syntax and Semantics.* Vol. 6. *The Grammar of Causative Constructions,* ed. M. Shibatani, pp. 413–42. New York: Academic
131. Matisoff, J. A. 1976. Austro-Thai and Sino-Tibetan: an examination of body-part contact relationships. In *Genetic Relationship, Diffusion, and Typological Similarities of East and Southeast Asian Languages,* ed. M. J. Hashimoto, 265–89. Tokyo: Jpn. Soc. Promo. Sci.
132. Matisoff, J. A. 1978. *Variational Semantics in Tibeto-Burman: The 'Organic' Approach to Linguistic Comparison.* Philadelphia: Publ. Inst. Stud. Hum. Iss. 331 pp.
133. Matisoff, J. A. 1978. Mpi and Lolo-Burmese microlinguistics. *Monumenta Serindica (Tokyo)* 4:1–36

134. Matisoff, J. A. 1979. Problems and progress in Lolo-Burmese: Quo vadimus? *Linguist. Tibeto-Burman Area* 4(2):11–43
135. Matisoff, J. A. 1980. Stars, moon, and spirits: bright beings of the night in Sino-Tibetan. *Gengo Kenkyu (Tokyo)* 77:1–45
136. Matisoff, J. A. 1982. Proto-languages and proto-Sprachgefühl. *Linguist. Tibeto-Burman Area* 6(2):1–64
137. Matisoff, J. A. 1985. God and the Sino-Tibetan copula, with some good news concerning selected Tibeto-Burman rhymes. *J. Asian African Stud. (Tokyo)* 29:1–81
138. Matisoff, J. A. 1985. Out on a limb: arm, hand, and wing in Sino-Tibetan. See Ref. 195, pp. 421–50
139. Matisoff, J. A. 1986. The languages and dialects of Tibeto-Burman: an alphabetic/genetic listing, with some prefatory remarks on ethnonymic and glossonymic complications. In *Contributions to Sino-Tibetan Studies,* eds. J. McCoy, T. Light, pp. 3–75. Leiden: E. J. Brill
140. Matisoff, J. A. 1986. Hearts and minds in Southeast Asian languages and English: an essay in the comparative lexical semantics of psycho-collocations. *Cah. Linguist. Asie Orientale (Paris)* 15(1):5–57
141. Matisoff, J. A. 1986. Sino-Tibetan numerals and the play of prefixes. Paper presented at 19th Sino-Tibetan Conf., Columbus, Ohio. 104 pp.
142. Matisoff, J. A. 1988. *The Dictionary of Lahu.* Univ. Calif. Publ. Linguist., Vol. 111. Berkeley/Los Angeles/London: Univ. Calif. Press. 1436 pp., 80 plates
143. Matisoff, J. A. 1989. The bulging monosyllable, or the mora the merrier: echo-vowel adverbialization in Lahu. In *South-East Asian Linguistics: Essays in Honour of Eugénie J. A. Henderson,* ed. J. Davidson, pp. 163–97. London: Sch. Oriental Afr. Stud.
144. Matisoff, J. A. 1989. Proto-Hlai initials and tones: a first approximation. In *Comparative Kadai: Linguistic Studies beyond Tai,* ed. J. A. Edmondson, D. B. Solnit, pp. 289–321. Publ. Linguist. #86. Arlington: Summer Inst. Linguist. and Univ. Texas
145. Matisoff, J. A. 1989. Palatal suffixes in Sino-Tibetan. Paper presented at 22nd Sino-Tibetan Conf., Honolulu. 63 pp.
146. Matisoff, J. A. 1990. On megalocomparison. *Language* 66(1):106–20
147. Matisoff, J. A. 1991. Areal and universal dimensions of grammatization in Lahu. In *Approaches to Grammaticalization,* ed. E. C. Traugott, B. Heine, Vol. II, pp. 383–453. Amsterdam: Benjamins
148. Matisoff, J. A. 1991. Jiburish revisited: tonal splits and heterogenesis in Burmo-Naxi-Lolo checked syllables. *Acta Orientalia (Copenhagen)* 52. In press
149. Matisoff, J. A. 1991. Endangered languages of mainland Southeast Asia. *Diogène* (Spéc. Issue ed. E. M. Uhlenbeck). In press
150. Matisoff, J. A. *Languages of Mainland Southeast Asia.* Lang. Surv. Ser. Cambridge/NY: Cambridge Univ. Press. In preparation
151. Matisoff, J. A. *Sino-Tibetan Etymological Dictionary and Thesaurus,* Vol. I: *Body-part Nomenclature.* In preparation
152. Mazaudon, M. 1973. *Phonologie tamang.* Paris: Société d'Etud. Linguist. Anthropol. de France
153. Mazaudon, M. 1974. Tibeto-Burman tonogenetics. *Linguist. Tibeto-Burman Area* 3(2):1–123
154. Mazaudon, M. 1978. Consonantal mutation and tonal split in the Tamang subfamily of Tibeto-Burman. *Kailash (Kathmandu)* 6(3):157–79

155. McKinnon, J., Bhruksasri, W., eds. 1985. *Highlanders of Thailand.* Kuala Lumpur: Oxford Univ. Press
156. Michailovsky, B. 1988. *La langue Hayu: phonologie, morphologie, syntaxe.* Collect. Sci. Lang. Paris: Editions du Cent. Natl. Rech. Sci.
157. Michailovsky, B. 1988. Phonological typology of Nepal languages. *Linguist. Tibeto-Burman Area* 11(2):25–50
158. Nagano, Y. 1983. *A historical study of the rGyarong verb System.* PhD thesis, Univ. Calif., Berkeley
159. Nishida, T. 1964/1966. *A Study of the Hsi-hsia Language: Reconstruction of the Language and Decipherment of the Script.* Tokyo: Zauho Kankokai (In Japanese) 2 vols.
160. Nishida, T. 1966. A study of the Bisu language. *Tonan Azia Kenkyu (Kyoto)* 4(1):65–87 (In Japanese)
161. Nishida, T. 1966/1967. A comparative study of the Bisu, Akha, and Burmese languages. Part I: *Tonan Azia Kenkyu (Kyoto)* 4(3):42–68; Part II *Tonan Azia Kenkyu (Kyoto)* 4(5):52–68
162. Norman, J. 1974. The initials of Proto-Min. *J. Chinese Linguist.* 2(1):27–36
163. Norman, J. 1988. *Chinese.* Cambridge Lang. Surv. Cambridge/NY: Cambridge Univ. Press. 292 pp.
164. Norman, J., Mei, Tsu-lin. 1976. The Austroasiatics in ancient South China: some lexical evidence. *Monumenta Serica* 32:274–301
165. Okell, J. 1969. *A Reference Grammar of Colloquial Burmese.* London: Oxford Univ. Press. 2 vols.
166. Okrand, M. 1974. Na-khi and Proto-Lolo-Burmese. *Linguist. Tibeto-Burman Area* 1(1):55–97
167. Ouyang, J., Zheng, Y. 1983. *Research into the Li (Hlai) Language.* Beijing: Chinese Acad. Soc. Sci. Press (In Chinese)
168. Pettigrew, W. 1918. *Tangkhul Naga Grammar and Dictionary (Ukhrul Dialect).* Shillong: Assam Secretariat Printing Office
169. Placzek, J. 1985. The perceptual foundation of the Thai classifier system. Paper presented at 18th Sino-Tibetan Conf., Bangkok
170. Purnell, H. C. Jr. 1970. *Toward a reconstruction of Proto-Miao-Yao.* PhD thesis. Cornell Univ. Vol. I, 207 pp.; Vol. II, 237 pp.
171. Purnell, H. C. Jr., ed. 1972. *Miao and Yao Linguistic Studies.* Selected articles in Chinese transl. Chang Yü-hung, Chu Kwo-ray. Data Pap. #88, Linguist. Ser. 7. Ithaca, NY: Cornell Univ.
172. Ramsey, S. R. 1987. *The Languages of China.* Princeton, NJ: Princeton Univ. Press
173. Ratliff, M. 1986. An analysis of some tonally differentiated doublets in White Hmong. *Linguist. Tibeto-Burman Area* 9(2):1–35
174. Rock, J. F. 1963. *A Na-khi (Moso)-English Encyclopedic Dictionary,* Part I. Ser. Orientale Roma 28(1). Rome: Ist. Ital. Medio ed Estremo Oriente
175. Sagart, L. 1990. Chinese and Austronesian are genetically related! Paper presented at 23rd Sino-Tibetan Conf., Arlington: Univ. Texas
176. Schwerli, V. 1979. *A Grammar of the Bawm language.* PhD thesis. Cornell Univ.
177. Shafer, R.1966/1967. *Introduction to Sino-Tibetan.* Part I, 1966, Part II, 1967. Wiesbaden: Otto Harrassowitz
178. Shafer, R., Benedict, P. K. 1939–1941. *Sino-Tibetan Linguistics.* Bound typescript, 14 Vols. Berkeley: Univ. Calif.

179. Sharma, S. R. 1982. Loan words in Patani: problems and mysteries. *Bull. Deccan Coll. Res. Inst. (Poona)* 41:160–63
180. Simon, I. M. 1972. *An Introduction to Apatani.* Gangtok, Sikkim: Res. Dept., Arunachal Pradesh
181. Simon, W. 1929. Tibetisch-chinesische Wortgleichungen: ein Versuch. *Mitt. Sem. Orient. Sprachen Königlichen Friedrich-Wilhelms Univ. Berlin* 32(1):157–228
182. Sofronov, M. V. 1968. *Grammar of the Tangut Language.* 2 Vols. Moscow: Nauka Press (In Russian)
183. Solnit, D. B. 1985. Introduction to the Biao Min Yao language. *Cah. Linguist. Asie Orient.* 14(2):175–91
184. Solnit, D. B. 1986. The position of Lakkia within Kadai. Paper presented at the 38th Meet. Assoc. Asian Stud., Chicago. 30 pp.
185. Solnit, D. B. 1986. A grammatical sketch of Eastern Kayah (Red Karen). PhD thesis. Univ. Calif., Berkeley. 274 pp.
186. Srinuan, D. 1976. *An Mpi Dictionary,* ed. W. Pantupong. Work. Pap. Phonet. Phonol., Vol. 1. Indigenous Lang. Thailand Res. Proj. Bangkok: Chulalongkorn Univ.
187. Strecker, D. 1987. The Hmong-Mien languages. *Linguist. Tibeto-Burman Area* 10(2):1–11
188. Sun Hongkai. 1981. *A Brief Description of the Qiang Language.* Beijing: People's Publishing Co. (In Chinese)
189. Sun Hongkai. 1982. *A Brief Description of the Dulong Language.* Beijing: People's Publishing Co. (In Chinese)
190. Sun Hongkai. 1985. Peoples and languages of the Six River Valleys and their genetic classification. *Minzu Xuebao* 3:99–274 (In Chinese)
191. Sun Hongkai. 1990. Languages of the "ethnic corridor" in Western Sichuan. Transl. J. T.-S. Sun. *Linguist. Tibeto-Burman Area* 13(1):1–31
192. Thomason, S., Kaufman, T. S. 1988. *Language Contact, Creolization, and Genetic Linguistics.* Berkeley/Los Angeles/London: Univ. Calif. Press
193. Thoudam, P. C. 1980. A grammatical sketch of Meiteiron. PhD thesis. Cent. Linguist. English. New Delhi: Jawaharlal Nehru Univ. 252 pp.
194. Thurgood, G. 1981. Notes on the origins of Burmese creaky tone. Monumenta Serindica #9. Tokyo: Inst. Stud. Lang. Cult. Asia and Africa. 100 pp.
195. Thurgood, G., Matisoff, J. A., Bradley, D., eds. 1985. *Linguistics of the Sino-Tibetan Area: the State of the Art. Papers Presented to Paul K. Benedict for his 71st Birthday.* Canberra: Pacific Linguist., Ser. C, No. 87. 498 pp.
196. van Driem, G. L. 1987. *A Grammar of Limbu.* Berlin: Mouton de Gruyter
197. van Driem, G. L. 1991. *Dumi Rai Grammar.* Berlin: Mouton de Gruyter. In press
198. Walker, G. D. 1925. *A Dictionary of the Mikir Language.* Shillong: Assam Government Press
199. Wang Fushi. 1979. *Comparison of the Initials and Finals of the Miao Dialects.* Beijing: People's Press. 219 pp. (In Chinese)
200. Wang Fushi. 1983. On the dialect divisions of the Miao language. *Minzu Yuwen* 3:1–22 (In Chinese)
201. Wang Jun, Zheng Guoqiao. 1980. *Outline Grammar of the Mulao Language.* Beijing: People's Press (In Chinese)
202. Wang Li. 1982. *A Dictionary of Word Families.* Beijing: Commercial Press (In Chinese)

203. Watters, D. E. 1975. The evolution of a Tibeto-Burman pronominal verb morphology: a case-study from Kham (Nepal). *Linguist. Tibeto-Burman Area* 2(1):45–79
204. Watters, D. E. 1985. Emergent word-tone in Kham: a Tibeto-Burman halfway house. *Linguist. Tibeto-Burman Area* 8(2):36–54
205. Weidert, A. 1987. *Tibeto-Burman Tonology.* Amsterdam/Philadelphia: John Benjamins
206. Wheatley, J. K. 1982. Burmese: a grammatical sketch. PhD thesis. Univ. Calif., Berkeley. 312 pp.
207. Wheatley, J. K. 1985. The decline of verb-final syntax in the Yi (Lolo) languages of southwestern China. See Ref. 195, pp. 401–20
208. Wiersma, G. 1990. Investigation of the Bai (Minjia) language along historical lines. PhD thesis. Univ. Calif., Berkeley
209. Winter, W. 1985. *Materials Towards a Dictionary of Chamling.* Univ. Kiel. 78 pp.
210. Wolfenden, S. N. 1929. *Outlines of Tibeto-Burman Linguistic Morphology.* London: Roy. Asiatic Soc.
211. Wolfenden, S. N. 1937. Concerning the variation of final consonants in the word families of Tibetan, Kachin, and Chinese. *J. Roy. Asiatic Soc.* Part 4 (Oct.), pp. 625–55
212. Wulff, K. 1934. Chinesisch und Tai. *Danske Videnskabernes Selskab. Historisk-filologiske Meddelelser* 20(3). Copenhagen: Levin and Munksgaard
213. Xu Lin, Zhao Yansun. 1984. *Outline Grammar of the Bai Language.* Beijing: People's Press (In Chinese)
214. Yabu, S. 1982. *A Classified Dictionary of the Atsi or Zaiwa Language (Sadon dialect), with Atsi, Japanese and English Indexes.* Tokyo: Inst. Stud. Lang. Cult. Asia Africa (In Japanese)
215. Yang Dao. 1980. *Dictionnaire Français-Hmong Blanc,* ed. J. Lemoine. Paris: Comité Natl. d'Entraide
216. Yuan Jiahua. 1953. *The Folksongs and the Language of the Ahi People.* Beijing: Chinese Acad. Sci. (In Chinese)

10

THE SAL LANGUAGES[1]

Robbins Burling

Source: *Linguistics of the Tibeto-Burman Area* 7, 2, 1983, 1–32.

Background

The sub-classification of the Tibeto-Burman languages has been a subject fraught with considerable mystery. Among all the central and eastern Tibeto-Burman languages, there appear to be only four clear sub-groupings: 1. The Bodo languages of Assam, 2. Manipuri-Mizo-Kuki-Chin, 3. Lolo-Burmese, and 4. Karen.[2] Other subgroup names have been used from time to time, such as "Naga," "Naga-Bodo," "North Assam," and even "Kachin" but these seem to label little more than geographically contiguous groups for which no genuine linguistic reality has been demonstrated. Jinghpaw, Nung, Mikir, and all of the vast numbers of languages found in the arc to the north and east of the Brahmaputra river, from the north of Assam around to the Naga Hills, fall into an uncertain limbo—surely Tibeto-Burman, but not clearly or closely affiliated with any other particular Tibeto-Burman languages.

We have, of course, heard a good many suggestions, from time to time, about the linguistic sub-grouping of these languages, including an old suggestion that the Bodo languages of Assam show a special relationship both with Jinghpaw and with certain languages of the northeastern Naga Hills region. So far as I am aware, this was first suggested by Robert Shafer (1953:162) but Paul Benedict also pointed to the similarities among these languages in the *Sino-Tibetan Conspectus* (1972:6–7, hereafter referred as "STC"), and a number of others have followed suit. A number of years ago I was sufficiently impressed with the similarities between Jinghpaw and Garo (the representative of the Bodo group with which I have worked most closely), to undertake a fairly elaborate comparison between certain aspects of the two languages. I tried to show that, in spite of what appeared to me to be massive mutual borrowing between Jinghpaw and such "Kachin" languages as Maru and Atsi, evidence can still be found for an underlying special relationship between Garo and Jinghpaw (Burling, 1971). In particular, I argued that even though the phonological *system* and the kinship *system* of

Jinghpaw and Maru could be regarded as almost identical, their lexicons show striking differences. In some ways, indeed, Jinghpaw appears to be lexically more like Garo than it is like Maru.

In my earlier paper I used no data from any of the "Naga" languages, and comparisons with these languages have always been difficult because of the paucity of evidence from the Naga Hills and adjacent regions. Information has been limited to rather fragmentary vocabulary lists, such as those in Grierson (1903), and while these have offered tantalizing hints, their evidence could never be definitive. In recent years, however, a number of dictionaries of Naga languages have become available, and while these are still far from the sophisticated linguistic treatments that we would like to have, they do provide enough new data to make a reassessment of the relationship among these languages worth while. There is also one bulky attempt at a comparative treatment of the "Naga" languages, G.E. Marrison's SOAS dissertation (Marrison, 1967), which can be used to supply some otherwise missing pieces. In this paper, I survey some of the available data on the Bodo, northeastern Naga, and Jinghpaw languages, and I try to reach a judgement on the likelihood of a special relationship among them.

The final judgement about sub-grouping should rest upon a close understanding of all types of shared innovations of the sub-group and upon a detailed understanding of the phonological correspondences among the languages. In our present state of knowledge about Tibeto-Burman languages, however, we must usually be content with an examination of simpler lexical similarities. We are reduced to the following fairly obvious and simple presumptions: if a group of languages 1) share lexical items that other languages fail to share, 2) show no sign that these shared terms are due either to mutual borrowing or to the residue of a still earlier stage of the language, and 3) have similarities that go beyond those expectable by simple chance, then it is plausible to conclude that these languages shared a period of common innovation and thus form a sub-group within the larger family.

The languages that I consider in this paper, for instance, have words such as *sal, san, jan*, all meaning 'sun,' and the phonological correspondences that relate the sounds of these words are paralleled in other sets of words. Similarities of this sort can hardly be dismissed as mere chance. Whether they might be due to borrowing or to a common historical residue that happens to have been lost from other languages is more difficult to decide, but if enough lexical items pattern in the same way as the words for 'sun' it becomes increasingly difficult to attribute the similarity to anything except common innovation at an earlier common stage of the language. It is upon the basis of a few lexical similarities of this sort, similar words that are found in Bodo, eastern Naga and Jinghpaw, but not, apparently, in other Tibeto-Burman languages, that earlier suggestions about their special relationship rested. Since the distinctive word for 'sun' has been cited particularly widely as offering evidence for the special relationships among these languages, and since the word can be plausibly reconstructed as *sal*, I will refer to the group

as the *sal* languages. This paper is a survey of the lexical similarities among the languages of this group.

Sources and affiliation

There can be little doubt that the Bodo languages form a relatively unified subgroup of Tibeto-Burman, considerably more closely related to one another than to other Tibeto-Burman languages. These Bodo languages include Boro, spoken in the lower Assam valley and recently described both by Bhat (1968) and by Bhattacharya (1977). The language that, in an earlier paper (1959), I called "Kachari" is, apparently, essentially the same as Boro, and I will supplement my own data with examples drawn from Bhat (indicated with "IMSB" in the tables) and from Bhattacharya ("PCB"). I will assume that these all come from the same language, but I have not tried to reconcile the somewhat divergent transcriptions used in the various sources. I also cite a few words from the closely related Dimasa language that I have taken from Marrison (1967). Dimasa is a dialect reasonably closely related to Boro, but it is difficult to be confidant of just how similar or different they are.

Garo, spoken in the Garo Hills in the bend of the Brahmaputra river, resembles Bodo in many ways, but the two languages are by no means mutually intelligible. Boro and Garo are the best described of the Bodo languages. On Garo, I rely upon Mason, 1954, Negminza, 1972, Holbrook, n.d., as well as upon my own knowledge of the language.[3] The transcription that I use for Garo is explained in Burling (1981). It is close, but not identical, to the transcriptions used in Garo dictionaries.

I also use materials that I collected from Atong and Wanang, two languages that are fairly closely related to each other within the "Koch" group that is, in turn, coordinate with Boro-Garo. I included these data from Atong and Wanang in my comparative study of Bodo phonology (1959).

I also cite a few examples from Chutia (taken from Brown, 1895) another Bodo language spoken further to the east in the Assam valley. Examples from Chutia are listed in the tables, in the "Wanang" column, but identified as coming from Chutia. The approximate location of these languages can be seen on the map (last page). The Bodo group also includes Rabha, spoken to the north of the Garo Hills, Lalung, spoken in the middle Assam valley, and the language of Tripura which lies just north of the Chittagong Hill tracts, but I include no data from any of these languages.

The relevant eastern "Naga" languages are less well known than the Bodo languages and they have been referred to under a bewildering variety of names. Present terminology seems to have settled on six language names. Ranging from southwest to northeast these are: Chang, Phom, Konyak, Wancho, Nocte, and Tangsa (see map). Earlier names, used in Grierson or in other early sources, and copied in more recent comparative literature, include (with equivalent modern terms in parentheses): Mojung (Chang); Tamlu, Chingmengnu, and Assiringia (Pom); Tableng and Angwanku (Konyak); Banpara and Mutonia (Wancho); Namsangia and Mohongia (Nocte); Moshang and Shange (Tangsa). The best known

(or at least most often cited) of these languages has been Konyak, and the term "Konyak Languages" has sometimes been used as a general term for this group.

It seems fairly clear that these six languages show more similarities to one another than to the other "Naga" languages or to the languages of bordering regions, though even this modest claim should not be taken as fully proven. In spite of having been used as a general name for the group, for instance, Konyak appears to diverge in a number of respects from its neighbors (see conclusions, below). I know of no evidence at all, however, that would suggest that these languages are more closely related to other "Naga" languages than to Tibeto-Burman generally. The term "Naga" appears to be a purely geographical term that lacks linguistic significance. This "eastern Naga" group, in fact, straddles the modern border between Nagaland and Arunachal Pradesh (formerly North East Frontier Agency, or "NEFA"), with Chang, Phoe, and Konyak lying primarily in Nagaland, and Wancho, Nocte, and Tangsa primarily in Arunachal Pradesh. For want of a better term, however, I will continue to refer to these languages as the "eastern Naga" group.

I offer data in this paper from three of these eastern Naga languages, each described in a recent publication: Konyak, (Kumar, 1973), Nocte (Das Gupta, 1971), and Tangsa (Ngemu, 1977). Each of these small books offers a brief grammatical description and a tri-lingual word list or dictionary (with Hindi glosses, as well as English). Unfortunately, the transcriptions in these dictionaries are not everything that a linguist might ask. A number of apparent inconsistencies crop up, and it is not always clear just how letters are being used. Under the circumstances it seems best to retain the original orthographies of the sources, with all their inconsistencies, rather than make the attempt to regularize them, and this is what I have done in the tables. One must be cautious, however, about inferring too much from the precise spelling. Marrison's long tables (Marrison, 1967) include data from these eastern Naga languages along with many other "Naga" languages, and I supplement the data of the three dictionaries with items taken from Marrison's table. As in other columns, items taken from Marrison are labeled with "M". When not marked with "M", the items in these three "Naga" columns come from the dictionaries already cited.

Of all of the languages that I consider in this paper, Jinghpaw has been the most widely cited in the comparative literature. This is due, in part, to the activities of linguist-native speaker, La Raw Maran, but much of the evidence actually cited in the literature is drawn from much older sources, particularly Hanson (1917). These sources, like so much of the older Tibeto-Burman literature, suffer from a failure to indicate either glottal stops or tones, and comparative studies such as STC that rely upon these older sources cannot fill in the gaps. STC and similar works do, however, adjust Hanson's transcriptions in some other ways (*hp* becomes *ph*, *ng* becomes *ŋ*, *ă* becomes *ə*, etc.). The majority of the examples that I cite from Jinghpaw I recorded myself, and these examples are shown with glottal stops and tone marks (◜high, ¯ mid, ◝, low),[4] but in other respects I try to follow the transcription conventions used in STC, for these will probably seem most familiar to linguists. Where I have taken examples from Hanson, I show this with an "(H)".

These examples lack tone marks. The final column in the tables shows the TB reconstructions that are given in the Sino-Tibetan Conspectus (Benedict, 1972).

Phonological correspondences

It is not possible, in the present state of knowledge, to offer a detailed or definitive account of the phonological correspondences among these languages, but certain broad patterns can be sketched and these should make the significance of the lexical examples somewhat easier to judge. In the next section, as I consider particular examples, I will offer a few more detailed observations on some special problems. Here, I will point out only the broadest patterns. As in Tibeto-Burman languages generally, it is most convenient to give separate consideration to syllable initial consonants, vowels, and syllable final consonants. It would be pleasant to know more about the tones of these languages so that we could offer tonal comparisons as well, but this is not yet possible.

Initials

Initial *m-* and *n-* are well attested in all of the languages. Initial *ŋ-* exists in Jinghpaw and, presumably, in eastern Naga, but in the Bodo languages initial *ŋ* has either been replaced by *n-* (e.g. 'fish') or attached to a preceding vowel to become a syllable final instead of a syllable initial ('I', 'five').

There are several examples of voiceless aspirated stops, usually written *p-* and *k-* in the Bodo languages, but *ph-* and *kh-* in the eastern Naga languages and in Jinghpaw. These correspond to *ø-* and *h-* in Wanang. Initial *t-* (*th-*) is less well represented. There is, however, a confusing group of words in which some languages have *t-* while others have *c-*, *s-*, etc. (see Table 2b.) I will consider these later.

Correspondences for voiced and nonaspirated voiceless stops are less clear than for aspirated stops, and *g-* is particularly poorly attested. There appear to be a number of cases where *b-* and *d-* in the Bodo languages (except Wanang where these have become *p-* and *t-*) correspond to words transcribed in the Naga dictionaries with *p-* and *t-* ('today' 'snake', 'wind', 'bat' 'flower', 'fruit, 'three', 'ash', 'next', 'straight', 'live'). There are, however, a number of difficult and contradictory cases where voicing seems rather random ('Grandfather', 'tree', 'fence', 'five') and the Jinghpaw correspondences here are obscure, in spite of the fact that many apicals and bilabials do, at least, appear as apicals and bilabials in Jinghpaw as well.

Two words ('earth' and 'mother's brother') have initial *h-* throughout the eastern Naga languages and Bodo, except in Garo, where the initial is lost. This *h-* may correspond to initial *g-* in Jinghpaw. The word for 'dog' also suggests that initial *h-* in eastern Naga may correspond to an initial velar in Jinghpaw, but the evidence is by no means decisive.

Initial *c-* (a voiceless affricate) in Garo, Atong, and Wanang quite regularly appears as *z-* in Boro ('dig', 'eight', 'far', 'hundred', 'long', 'mortar', 'stand', 'thick') but the eastern Naga and Jinghpaw correspondences are unclear due to

few and contradictory examples. Both *ts-* and *th-* appear in Jinghpaw. 'Pierce', 'salt, 'sun', 'thorn', and 'urinate' have *s-* in most languages, but most often have *j-* in Jinghpaw and *h-*in Konyak.

Numerous examples are found that have initial *r-* in all languages except Konyak, where *w-* (sometimes *v-*) appears instead (e.g. 'dry', 'horn', 'sky'). Others have *l-* in all languages except Garo and Atong which have *r-*, and Bodo which sometimes has initial *r-* but where the initial is sometimes lost entirely e.g. 'drink', 'road', 'stone'). Another group ('bamboo', 'father', 'fire', 'monkey', 'pig') has initial *w-* throughout except in Boro where the initial is lost. The Konyak and Nocte dictionaries sometimes show *v-* rather than *w-* in these words but it seems doubtful to me that this represents a real difference, and I suspect that the choice between *v-* and *w-* in these dictionaries is more or less random.

Vowels.

The clearest vowel correspondences are for *i*, *a*, and *u*, and many unambiguous examples will be found in the tables where most or all languages share these vowels. All the languages appear to have vowels in the *e* and *o* positions as well, and a few correspondences relate them, but in a number of cases these mid vowels seem to be related to various sorts of diphthongs in some of the languages. The precise relationship among these will have to await better data and fuller study.

Final consonants.

These languages appear generally to have final *-p*, *-t*, *-k*, *-m*, *-n*, *-ŋ*, and *-?*, and usually *-l* and/or *-r*. For the most part the correspondences among these are straightforward. However, where other languages have *-k*, Jinghpaw regularly has a glottal stop (not indicated by Hanson or in STC), and where the Bodo languages have glottals, Jinghpaw frequently has nothing at all. These finals are regular enough so that when a putative cognate set turns up with inconsistent finals (such as an *-n* in one language that seems to correspond to an *-ŋ* in another language), one should be skeptical of the cognate relationship. Several of the Bodo languages allow final nasals and final *-l* or *-r* to be glottalized. Jinghpaw does not, and it is impossible to know whether such sounds occur in the eastern Naga languages because glottal stops are not shown in the available transcriptions.

There is one other difficult but important set of correspondences among the final consonants. In a considerable group of words, Nocte and Jinghpaw have *-n* and Tangsa has *-l*. Some Konyak examples have *-n*, but the Konyak examples are too fragmentary to give confidence. In these words, Boro, Garo, Atong, and Wanang sometimes have, respectively, *-r*, *-l*, *-r*, *-r*, and sometimes *-n*, *-l*, *-n*, *-n*. The correspondences between the Bodo languages and Naga or Jinghpaw are, unfortunately, few, but other examples of both types of correspondences occur within the Bodo languages (cited in Burling, 1959) and both correspondences seem thoroughly sound.

This account of correspondences amounts only to the most superficial sketch, but it may help the reader to judge the cognate status of the examples given in the next section. A few more detailed comments about apparent correspondences will be given as examples arise.

Lexical examples

In this section I discuss, with reference to the tables, the lexical sets that offer evidence for the relationship among these languages. The tables display examples of several sorts. Tables 1a, 1b, and 1c list examples which suggest Sal languages innovations away from general Tibeto-Burman. They constitute evidence, therefore, for a period of common innovation within a common ancestor of the modern Sal languages. Inevitably, the examples vary greatly in their reliability and this serves as the basis for assigning particular examples to particular tables. Table 1a lists those examples that seem most convincing, table 1b lists less convincing but still plausible examples, and table 1c offers a number of more problematic or questionable examples.

Tables 2a, 2b, and 2c list the representatives from the Sal languages that appear more widely in Tibeto-Burman languages. Since these words appear outside the Sal languages, as well as within it, they demonstrate little about any special relationship within this group, although in a few cases the Sal forms of do appear to be more like one another than like the forms that the words assume in other TB languages. The main purpose of listing these general TB words here is to give evidence for the sound correspondences within the Sal group, for these correspondences can help us to judge the cognate status of the words that are unique to the Sal languages. As in the case of Tables 1a, 1b, and 1c, those numbered 2a, 2b, and 2c, differ in the how convincing the examples are.

In addition to these general lists, I give two tables of terms belonging to particular semantic sets. Table 3 gives numbers, and Table 4 gives kinship terms. Discussion of a number of particular examples that occur in the tables follows.

1a. Most convincing.

Table 1a lists the cognate sets that strike me as offering the strongest evidence for common lexical innovation within this group of languages. These words appear to be excellent cognates within Sal group, but examples of plausible cognates are rare or non-existent elsewhere. Some of these sets may turn out to have cognates in other languages and such examples will then turn out *not* to provide evidence for the special relationship of the Sal languages, but collectively the evidence strikes me as impressive. Most of the examples in this set require some comment:

'Ash'. STC gives **pla* 'ashes' (137) and **tap* (18) usually meaning 'fireplace' but most, though not all, examples come from Sal languages. The

combination of the two syllables seems to be unique to the Sal group, and even examples of the individual syllables are not plentiful elsewhere.

'Burn'. STC 330 is **kaŋ* and words offered as related to it have such meanings as 'hot' or 'roast.' Words such as kam with the meaning 'burn,' appear to be unique to the Sal languages.

'Cook'. The Konyak term may be unrelated, but the remaining terms seem to be excellent candidates for cognate status.

'Cooking pot'. The final Jinghpaw glottal is expectable from final *-k* elsewhere. The initial correspondence looks convincing, and few, if any, plausible cognates of this word crop up in languages outside of the Sal group.

'Crow'. STC, pg. 99 cites a number of Sal languages in support of its reconstruction, **ka*, but outside of the Sal group offers only Tibetan *kha-tha* and Rawang *tha-kha*. Rawang *kha*, however, is elsewhere glossed as 'domestic fowl,' (Barnard, 1934, pg. 68) which makes it unlikely to be the part of *tha-kha* with the meaning 'crow.' This leaves Tibetan as the only non-Sal language example in which *kha* can be plausibly taken to mean 'crow'. In all Sal languages, moreover, bird names are commonly constructed from a syllable with the general meaning 'bird' followed by a second syllable denoting the particular species, and this pattern does not appear in the other examples offered by STC. (I am quite mystified by the fact that several bird species have more convincing cognates in TB languages than do the mammals that one would suppose must play a good deal more salient role in the lives of the people.)

'Drink'. The lack of *-ŋ* makes the Jinghpaw term a very doubtful cognate. The Konyak *y* is unexplained and may well eliminate that word too. Still, the Tangsa term seems unarguably related to the general Bodo term.

'Far'. STC offers a TB reconstruction for this term, but the only example outside of the Sal group which it offers as a putative cognate is Lushai *fa:l* which strikes me as doubtful.

'Father'. Terms for 'father' beginning with *w* are rare enough in the world's languages to invite notice. Within Tibeto-Burman, these seem to be unique. A more extensive list of kinship terms is given in Table 4.

'Fire'. Except for the word for 'sun' this has probably been the second most widely cited example of a unique Sal language innovation. Plausible cognates turn up in a few other languages with the meaning 'burn' but only in the Sal languages is it the ordinary word for 'fire.'

'Insect/worm'. This word, along with leg/foot (just below), 'hand/arm' (table 2a), and 'moon' (table 1b) form a special set, with a very peculiar pattern of initial correspondences: Garo *j-*, Atong, Wanang *c-*, Konyak *y-*, Nocte *d*, and Tangsa *j-*. A single word with this correspondence would hardly be taken seriously, but the four together cannot be easily dismissed. Of the various Sal languages, only the Jinghpaw examples for these four words seem doubtful. One's initial reaction might be that the Jinghpaw

correspondence for this initial would be *l-*, as it appears to be in some more distantly related TB languages, but it seems more likely to be *t-*. (See 'leg' below, 'hand/arm' table 2a, 'moon' table 1b.) Possible cognates of 'moon,' and likely cognates of 'hand/arm,' are found in a few TB languages outside of the Sal group, sometimes with initial *l-*, but the four words occur with particular regularity within the Sal group. The word for insect/worm appears to be distinctively Sal although the Jinghpaw word is questionable.

'Leg/Foot'. This is another member of the insect-arm-foot-moon group. Throughout the Bodo and eastern Naga languages, the word for 'foot' differs from that for 'hand' (table 2a) only in terminating with a glottal stop rather than with *-k*. An inspection of Grierson's word lists shows these two words cropping up with a bewildering range of initials, but in any single language their initials are usually alike. A few languages outside the Sal group also show possible hand/foot pairs of this sort, although the word for 'hand' seems more widespread, than the word for 'foot,' which seems more narrowly limited to the Sal languages. Outside of the Sal group the pairs are rarely as regular or as similar to each other as they are within the Sal group. If the corresponding Jinghpaw initial is *t-*, the Jinghpaw word for leg/foot cannot be counted as a cognate.

'Live/Green'. The glosses for this word are a bit variable and may merit some skepticism. 'Alive' and 'green' (in the sense of 'unripe') seem plausibly similar glosses, but while the Nocte word is glossed as 'live,' the examples suggest the meaning 'dwell' and this may eliminate the Nocte example from this set. The gloss for the Tangsa word *lungtong* is 'live' while the gloss for *tong-nga* is 'live' (v). I know of no similar words in non-Sal TB languages.

'Long'. The only non-Sal language cognate offered by STC for this word is Burmese *lu*, glossed as 'disproportionately tall.' Throughout the Sal languages it is glossed simply as 'long.'

'Mother'. Terms for 'mother' with initial *n-*, invite notice, just as do terms for father with initial *w-*. Certainly most TB languages have 'mother' terms that begin with the familiar *m-*.

'Salt.' The scattering of examples offered by STC from outside of the Sal languages seem less similar than the Sal language examples are to each other.

'Sky'. The syllable *raŋ* crops up in most of these languages as the first syllable of compounds that refer to celestial phenomena such as 'sun' and 'rain'. When *raŋ* occurs by itself, it seems always to have the meaning 'sky.'

'Sun'. As pointed out above, the words for 'sun' in these languages—*san*, *sal* or *jan*, sometimes preceded by the syllable for 'sky'—have been widely cited as offering the clearest evidence for the special status of the group. The correspondences that relate the various words for 'sun,' are reassuringly regular, except that the *hi* of the Konyak term is not an ideal cognate

for the words in the other languages and is probably unrelated. STC offers a reconstruction for this set of words but, with the exception of a doubtful term from Bahing, *tsyar*, that is glossed as 'shine', all the STC examples come from the Sal languages. Totally different words for 'sun' are found in other Tibeto-Burman languages and, with the possible exception of Konyak, this word seems to set this group of languages decisively off from others.

1b. Suggestive sets.

The sets listed in Table 1b are suggestive of relationships among the Sal languages, but they are less convincing than those listed in Table 1a, either because they are found in fewer language of the group, because they have possible cognates in other Tibeto-Burman languages, or because their sound correspondences seem less regular. Still, taken collectively, they do seem to add some support to the feeling of special similarity among the languages of this group.

'Basket'. This set appears to be sound throughout Bodo, and it is reasonable in eastern Naga. The final glottal of Jinghpaw is the regular correspondent of final *-k* elsewhere, but the more usual initial correspondence in Jinghpaw would have been *kh-*, so the *k* makes the Jinghpaw word just a bit suspect. STC offers a number of possible cognates for this term in other languages, so even though these tend, in various ways, to be a bit divergent, the word may not be unique to the Sal languages.

'Bone'. The STC reconstruction for 'bone' does not give Sal language examples, and it looks like a different word. The initial *q-* or *k-* found in the Bodo languages can plausibly be regarded as a prefix.

'Deer'. Mikir has a possible cognate, *thidzok*, and a few other languages have conceivable cognates, though they look less like the Sal language examples than the latter look like each other.

'Falcon-kite'. Burmese has words for 'vulture,' 'eagle,' etc. that look like cognates, but such words seem to be much more widely distributed in the Sal languages than elsewhere in TB.

'Moon.' This is a member of the insect-arm-leg-moon group with its odd pattern of initial correspondences. A number of TB languages have words for 'moon' with initial *l-* that are likely to be related to the Sal term, but the words seem particularly close among the Sal group.

'Navel.' The only putative cognate outside of the Sal languages offered by STC is Tibetan *lte-ba*, which seems to stretch the imagination a bit.

'Pus.' STC gives only the Jinghpaw word and Burmese *tshwe* 'decayed, crumbling, rotten' as examples. Even if the Burmese word really is a cognate, its meaning is notably different from the meaning of the Sal language examples.

'Stand.' STC gives a long list of TB languages with words meaning 'stand' that end in *-p*, but affricate initials are largely confined to the Sal languages.

The Sal languages also show more consistency in the vowel, although they are not the only languages with *a*.

'Tree.' Tables 1b and 1c each contain a word that is most often glossed as 'tree,' although plausible cognates are glossed as 'trunk' or 'firewood' in some languages. The Jinghpaw example is a conceivable, though less than ideal, cognate to either set and I list it with both.

1c. Tantalizing possibilities.

I do not consider the sets listed in Table 1c to be more than suggestively tantalizing. Some have plausible cognates in other TB languages and some have been given reconstructions in STC. Some of these sets are represented by only two or three examples, however, and in others the correspondences seem less than fully convincing. I would certainly not want to risk exaggerated claims for those listed here. Collectively they do suggest the kinds of data to which we are driven when we try to sort out the relationships among these languages. Only a few of these examples require special comment.

'Bark.' It is just barely possible that the Garo word meaning 'dog' is related to the Nocte word meaning a dog's 'bark.'

'Cut.' STC cites a few plausible cognates in other TB languages.

'Dung.' Here, again, a few possible cognates appear here and there in non-Sal languages, but they are not so frequent or regular that this word deserves to be entered into Table 2a or 2c.

'Mat.' The words for 'mat' seem more tempting as cognates than many of the sets listed in this table, but the initials are so strange as to make one cautious.

'Red.' The non-Sal examples offered by STC have meanings such as 'gold' which separate them from the Sal words which are glossed simply as 'red.' However, the correspondences among the Sal words are by no means certain.

'Sleep.' STC offers a number of words from various languages that may be related to the Sal words shown in the table, but some mean 'hide' or 'cover' instead of 'sleep,' most have no initial consonant, and most have the vowel *i*, all of which makes the Sal words seem distinctive.

2a. Most widespread cognates.

The sets listed in Tables 2a, 2b, and 2c have more reasonable looking cognates in TB languages outside of the Sal group than do those listed in Tables 1a, 1b, and 1c. This means that they do not demonstrate any special relationship among the Sal languages except in the degree to which the Sal examples may be more like each other in form or meaning than they are like the examples from other languages. The words listed in Table 2a are the Sal language representatives of some of the most widespread cognates of Tibeto-Burman. Even though these can tell us

little about any special relationship among the Sal languages, they do demonstrate the patterns of phonological correspondence that we should expect within the Sal group. Most of these sets should be self explanatory and they require little discussion. I will leave it for the reader to judge the degree to which these words seem to show specially close resemblances to one another within the Sal group.

2b. "Water" group.

The sets listed in Table 2b pose numerous problems. These words crop up variously with initial apical stops, affricates, and fricatives, and although my instincts are to want them to be cognates the patterns are so complex that we have to be skeptical about some of the sets. Because of the confusion of possible correspondences within this set of words, I have been more tolerant of diversity in this table than elsewhere, and I have included a good many words of doubtful cognates status (e.g. the words for water in the eastern Naga languages).

A few of these sets appear to be related to one another, but it is not entirely clear just what these relationships are. The words for 'egg,' in some of the Sal languages, for instance, appear to mean 'bird-water' as it does in some other TB languages. The cognate set for 'egg,' however, actually seems to be more complete than that for 'water,' since the eastern Naga words for 'water' do not appear to be related to the others. In Garo, the word for 'blood' also appears to be analyzable as having the meaning 'body-water,' though in most other languages the words for 'blood' and 'water' are different. 'Die' and 'kill' also appear to have some sort of relationship to one another.

The examples of the Bodo languages seem to exhibit an unusual number of diphthongs, and it seems probable that diphthongs had a tendency, in these languages, to turn preceding stops into affricates (Burling, 1959) but we still have a long way to go before sorting out all the relationships among these initials. All of these sets have plausible cognates in other TB languages than the Sal group alone, and STC offers reconstructions for most of them. The initials of these words demonstrate considerable confusion in languages outside the Sal group as well as within it, and about all I can do is wave readers in the direction of the table and invite them to contemplate the complexities.

2c. Less widespread cognate sets.

Table 1c lists a number of cognate sets that have representatives outside the Sal languages but that are represented in a rather scattered fashion within this group. These are considerably less widespread than the sets listed in Table 1a.

3. Numbers.

Finally, I offer two tables, each with a group of semantically related terms, one for numbers and one for kinship terms. There can be no doubt that most numbers fall into cognate sets, but numbers change in erratic ways in the TB languages and this

makes them difficult to use as a basis for judgements about sub-grouping. Adjacent numbers, for instance, seem to influence each other. 'One,' 'two,' and 'three' have picked up the prefix *-go* in Atong, but in Garo the related prefix is found only in 'two' and 'three.' Nocte has a different prefix in the same three words. It is tempting to note that the words for 'four' and 'five' in many of these languages have acquired a prefix that begins with *b*, or to note that the words for 'one' often begin with an *s* or with a plausibly related sound, and to conclude that the numbers give evidence for a special relationship among these languages, but it is very difficult to pin these similarities down. All, or virtually all, of these terms are surely cognates and most surely have cognates in other TB languages, but it is difficult to find objective criteria that can help us to decide whether these words are more similar to each other than to the words in other languages. How similar is "more similar" anyway? I will leave it to the readers to make their own judgement about the degree of similarity among the numbers.

4. Kinship terms.

Kinship terms raise other difficult problems. The kinship systems of the people who speak these languages are quite varied. The Garo and Atong are matrilineal, most of the others patrilineal. The Bodo languages other than Garo and Atong probably show considerable influence from the kinship system of the Indic speaking Assamese with whom the speakers are in close contact. The Jinghpaw have a very special kind of kinship system that, so far as we know, differentiates them sharply from any of the other groups considered here. Under such circumstances, one would hardly expect kinship terms to retain the same meaning from one language to another, even if the terms themselves, survive. Still, terms with related though not identical meanings may well be cognates. Many of these languages have cross-cousin marriage, for instance, and this makes it reasonable for the father-in-law to be called by the same term as the mother's brother. Thus it is worth looking at 'uncle' terms to see whether they might be possible cognates for 'father-in-law' terms in other languages.

When one takes an appropriately tolerant attitude toward the glosses for the kinship terms, there turn out to be a very considerable number of plausible cognates. In fact, the great majority of Jinghpaw kinship terms have plausible cognates in one or another of the Sal languages. By contrast, Jinghpaw and Maru share only one lonely kinship term, and this in spite of the fact that the Maru and Jinghpaw people have considerable intermarriage, and in spite of the fact that their kinship terminological *systems* are virtually identical (Burling 1971). Cognates to the kinship terms shown in Table 8 are, to be sure, sometimes found in languages outside of the Sal group, but the density of apparent cognates seems notably high among this group.

Conclusions

I have no doubt that a fair number of the cognates sets that I offer, even those that now seem most solid, will finally turn out to have cognates outside the Sal group,

but the collective weight of the examples I have collected seems to me to demand an explanation. I doubt if this many apparent cognates could be marshalled to demonstrate similarities for most sets of TB languages. This looks like a group of languages with some sort of historical relationship, rather than like a random collection. Chance alone does not seem to me to be a sufficient explanation for the number of similar words that unite these languages.

Nor do the similarities seem to be explainable on the basis of mutual borrowing. The languages are spread over a considerable area and, in the course of their history, they have been subjected to quite varied influences. Many Bodo languages show massive borrowing from the Indic languages, while Jinghpaw has unquestionably been strongly influenced by Burmese and by the so-called "Kachin" languages that are closely related to Burmese. The similarities among the Sal languages seem deeper, less easily attributed to borrowing, for they show up in spite of the more recent overlays of borrowing to which some languages have been subjected.

One question that inevitably arises in the course of looking at a group of languages such as these concerns their internal relationships. There can, I think, be no doubt that the Bodo languages are more closely related to each other than they are to the other languages. Nor, I think, can doubt be much greater about the special relationship among the eastern Naga languages. Numerous cognates can be found that are unique to one or the other of these groups. I have not included them in the tables, however, because they show nothing about the wider relations among the Sal languages.

Relative relationship among these two groups and Jinghpaw is more difficult to judge. As I was working on this paper, I had the feeling that Jinghpaw was often more distant from the Bodo and eastern Naga languages than the latter were from each other, but I also wondered if this might have been due to the fact that several languages are available within the other two branches, so that where one language lacks a term there is still a chance that another might supply it. If Jinghpaw lacks a term, there is no close relative to fill the gap, so it is easy to get the impression that it is harder to find cognates among Jinghpaw.

In the end it turns out that Jinghpaw offers just as many potential cognates as the other languages. For what the figures are worth, the tables contain the following number of entries for each of the eight languages (excluding the numbers, Table 3, which are virtually complete for all languages except Wanang): Bodo-113, Garo-135, Atong-66, Wanang-64, Konyak-85, Nocte-112, Tangsa-118, and Jinghpaw-138.

Of course these figures are, in large part, a function of the availability of data, but Jinghpaw appears to be at least as well represented as the other languages. I may, however, have been a bit more tolerant of questionable Jinghpaw examples, precisely because it was alone in its sub-group. The high figure for Garo is certainly an artifact my own familiarity with the language and the availability of dictionaries. Data from Atong and Wanang are limited to my very limited notes, while the Boro figure is higher because I could supplement my notes with a number of examples drawn from the Boro dictionaries.

The most curious figures are for the eastern Naga languages where Nocte and Tangsa have well over 100 entries, while Konyak has only 85. One might suppose

that these figures simply reflect the quality of the data of each of these languages, but, in fact, the Konyak dictionary is considerably fuller than the dictionaries of Nocte and Tangsa. The Konyak dictionary has between 3,000 and 4,000 entries, while the other two have only about 1,000 entries each. If the three languages are equally closely related to Bodo and Jinghpaw, one would expect the fuller Konyak data to yield more apparent cognates, but the reverse turns out to be the case.

These crude figures confirm impressions I formed while working with these eastern Naga dictionaries. I began with Konyak and I was surprised at how different it seemed from the Bodo languages with which I had worked previously. When I turned to Nocte and to Tangsa, however, I was surprised, in turn, to find so much that seemed familiar. I noted many special similarities between Konyak, on the one hand, and Nocte Tangsa on the other, so I would not want to abandon the presumption that Konyak, Nocte, and Tangsa form, along with the other eastern Naga languages, a well defined sub-group, but it may seem odd that cognates seem so much easier to find in the other languages than than in Konyak. I also find it paradoxical that Konyak is the language that has been most often cited as representative of the sub-group. The languages have even been referred to quite often as the "Konyak" group. The similarities of the eastern Naga languages with the Bodo languages might have been more obvious had comparisons been made with Nocte or Tangsa instead.

One possible explanation for the greater ease of finding cognates in Nocte and Tangsa than in Konyak is that the former two languages lie further to the northeast and, therefore closer to the Singpho country that lies just beyond—Singpho being the Assamese variant of "Jinghpaw." Here is one place where mutual borrowing among the different languages of the Sal group would have been quite possible and it could have led to an artificial boost in the number of apparent cognates in Nocte and Tangsa. An inspection of the tables does not, however, give me the impression of any particularly close or obvious layer of borrowing that gives special similarities among these languages. In the end, I have found little reason to conclude that any two of the three sub-groups of the Sal languages are more closely related to each other than to the third.

Of course we have to recognize that there are real dangers in the methods I have used in this paper. I have worked extensively with Garo and to a more limited extent with several other Bodo languages and Jinghpaw. One always finds similarities among the languages one knows best, even if they as remote as English and Chinese. Thus I have spent more time looking for similarities between Garo and Jinghpaw than between, let us say, Jinghpaw and Tibetan. Someone else with more knowledge of Tibetan and less knowledge of Garo might find similarities that point in a different direction. In the end, I can only challenge others to offer competing evidence that points to other relationships.

In the mean time, the number of similarities among these languages and, in particular, the number of common innovations that they seem to show strike me as sufficiently impressive to suggest that they have a relationship that goes comfortably beyond the minimum that we expect of all Tibeto-Burman languages. The

set of languages for which I propose the name "Sal" seems to me to be plausibly regarded as one major sub-group of the Tibeto-Burman languages.

Notes to the tables

When not otherwise noted, entries in the tables are taken from the following sources: Boro, Atong, Wanang, and Jinghpaw from my own work. Garo examples are from my knowledge of the language supplemented by Holbrook (n.d.), Mason (1954), and Negminza (1972). Konyak examples come from Kumar (1973), Nocte examples from Gupta (1973), and Tangsa examples from Ngemu (1977). I supplement these main sources from a number of others. Supplementary examples in the Boro column are taken from Bhat (1968) and labeled "DNSB", or from Bhattacharya and labeled "PCB". Examples in the Boro column from the Dimasa dialect are taken from Marrison. The examples from the Chutia language that are tucked into the Wanang column derive from Brown (1895). Items labeled "H" in the Jinghpaw column are taken from Hanson's dictionary of Jinghpaw (1917). I supplement my primary sources on Jinghpaw and the eastern Naga languages with a few items from Marrison (1967) that are labeled "M". Reconstructions in the STC column, of course, are from Benedict (1967).

Table 1a Most likely *Sal* language innovations

	BODO		*EASTERN NAGA*				*Jinghpaw*	** Sino-Tibetan Conspectus*
	Boro	*Garo*	*Atong*	*Wanang*	*Konyak*	*Nocte*	*Tangsa*		
ash	ha-to-pla	tap-pra	tap-pa-ra		tūpla (M)	tap-la	taptha (M)	dàp	* tap (18)
burn	kam	kam-	kam-	ham-		akham	kham 'hot'		* pla (137) * ka• ŋ (330)
cook	soŋ 'boil' (DNSB)	soŋ-			yang	song-dang hum 'kitchen' (hum = 'place')	song		
cooking pot	dəʔ	dik	dək		tūk (M)	tik	koti (cik)	diʔ	
crow	ʔdaw-kha (PCB)	doʔ-ka		duka (Chutia)		wa-kha	wokha	ùkhá	* ka (pg. 99)
drink	iaŋ	riŋ-	raŋ-	ləŋ-	ying		ling	lùʔ	
dry	gə-razn	razn-	razn-	ran-	wan	ran			
face/ forehead	mukaŋ 'face' (DNSB)	mik - kaŋ 'face' (mik = 'eye')			shakeng 'fore-head' (M)	khang 'forehead'	khangkang 'forehead'		
far	gə-zarn	ceʔl-	carn-	pi-cən-	jay		ajal (M)	tsān	* dzya• l (229)
father			a-wa	a-wa-par		wa	wa	wà	
finger	yaosi (Dimasa)	jak-si			yashao 'toe' (M)	dak-su	jaksi		
fire	oɪr	waʔl-	warr	war	vun	van	wal	wràn	* bar, par (220)

(continued)

Table 1a (continued)

	BODO				*EASTERN NAGA*			*Jinghpaw*	** Sino-Tibetan Conspectus*
	Boro	*Garo*	*Atong*	*Wanang*	*Konyak*	*Nocte*	*Tangsa*		
insect/ worm	yung (Dimasa)	joʔ-oŋ	coʔŋ	coŋ	yenna		jong	·jīnjāy	
leg/foot	aʔ-tiʔŋ	jaʔ-	caʔ	ca-tim	ya	da	ja	làgó	
live/green	taŋ-nəʔ	taŋ-				atong	luntong tong-nga		
long	gə-lauʔ	roʔ-	rauʔ-	pi-ləu-	low	alo	alo	gàlù	* low (279)
mother					nyu	nylong	nu	nù	
pestle		ri-mol				man	mol	thummun (H)	
rice, uncooked	maiʔ-roŋ	me-roŋ	mai-roŋ		wong	wong, vong	wong		
salt	sem (Dimasa)			shing (Chutia)	hüm	sum	sim	jùm	* g-ryum (245)
shoulder	pakná koŋ 'armpit' (DNSB)	pakre, pak = 'armpit'			phakdeang			gəpnàʔ	
sky	saraŋ 'be clear' okráŋ 'sky' (DNSB)		raŋ-wa 'rain' raŋ-san 'sun'	raŋ-øəi 'rain'	wang	rang 'sky' rang-pat 'rain'	rang 'sky' rangche 'rain'	màrāŋ, 'rain'	
sun	san	sal	raŋ-san	san	wanghi	sa:n	rangsal	jān	* tsyar (187)
wing	gaʔŋ	graŋ	ga-raŋ	ka-raŋ	yang	arang (M)	worang		

Table 1b Suggestive *Sal* language cognate sets

	BODO				*EASTERN NAGA*			*Jinghpaw*	** Sino-Tibetan Conspectus*
	Boro	*Garo*	*Atong*	*Wanang*	*Konyak*	*Nocte*	*Tangsa*		
basket	kaʔ-da	kok	kok	hok	khog	chi-khok 'small basket'	khak	káʔ	* kuk (393)
bone	be-gɛʔŋ	greŋ	ge-reŋ	ke-reŋ	wan (M)	ra:	rang	nàrā	* rus (6)
cold	'go' zaŋ (PCB)	kaʔ-sin-					rangsong	kàshūŋ	
cover (vb.)	pin (DNSB)	pin-dap-						phun (H)	
deer, sambhur		mat-cok			tok (M)	cok	cɔk (khohi)	cákyī	* d-yuk (386)
dive/sink	trəb (DNSB)	rip-						phuŋ-lip (STC)	* lip (375)
falcon/kite		doʔ-reŋ		durong (Chutia)	lungl ieng			làŋjì	* laŋ (333)
house	noʔ	nok	nok	nok	nok				
moon		ja-joŋ			linyu	da	japi	shátā	* s-la. g-la (144)
navel		ok-ste				po-te		shà-dāy	
pus		min-su					tacu	mátsūy	* tswly (183)
seed		bit-cri	caʔ-ri	ca-li	üri	khet-ali	uli	nli(H)	
stab/pierce	suʔ	suʔ-	suʔ-	su-tik-		su (M)	tasun	jú	
stand			cap-	cap-		acap	chap (M)	tsáp	* g-ryap (246)

(continued)

Table 1b (continued)

	BODO				*EASTERN NAGA*			*Jinghpaw*	** Sino-Tibetan Conspectus*
	Boro	*Garo*	*Atong*	*Wanang*	*Konyak*	*Nocte*	*Tangsa*		
tiger	mosa (DNSB)	mat-ca			shahnyak 'leopard'	sao	shah	shàrō	
today	dinəy/dini (DNSB)				teny in/ deny in		thaneh	dai ni (H)	
tree	piʔ-paʔŋ	bi-paŋ 'trunk'			pangpem 'beam'	bang	bang	phún	
wife/ woman	hin-zau 'woman'	jik 'wife' meʔ-cik 'woman'		mə-cəu 'woman'	sheko	dehiak	eshik 'wife' (M)		
yesterday	miya (DNSB)	me-ja sal (sal = 'day')				meja, maja	baja		

Table 1c Tantalizing possibilities

	BODO				*EASTERN NAGA*			*Jinghpaw*	** Sino-Tibetan Conspectus*
	Boro	*Garo*	*Atong*	*Wanang*	*Konyak*	*Nocte*	*Tangsa*		
animal	məy (DNSB)				mei (M)				
bark, of animal		a-cak 'dog'				achok			
bear					shapnyu	sap-ba	shap	tsáp	
big	dar (DNSB)	daʔi-				adong	adil (M)		
bite			gak-	kak-		kak (M)	kak		
bring	lab (DNSB)	raʔ-ba-					lamkai (M)	la (H)	
come	pay (DNSB)				pei (M)				
cut	daʔn	deʔn-	daʔn-					dàn	* dan (22)
dung	'khi (PCB)	ki				hi		khyí	* kliy (125)
garden/ fence	bari (DNSB)	ba-ri				pan	pal	màphān	
hold			rauʔ-	leu-		aloat	lo-a		
imitate		sniŋ-						sam naŋ (H)	* naŋ (334)
light wt.	'ge'seŋ (PCB)	rit-ceŋ-					achan (M)	tsāŋ	* r-ya• ŋ (328)
mat	ém (DNSB)	am		yamshu (Chutia)		ham	xam		
new						anyian	anai	nīŋ nān	
nose	goʔn-toʔŋ	giŋ-tiŋ	na-kuŋ	ne-kuŋ		kho	khung		
red	gə-zaʔ	git-cak	bi-sak	pi-sək	ütak	acak	ashang	khyēŋ	* tsyak (184)
right (hand)		jak-ra			yakja (M)	dak-cha	jak xah	khrá	* g-ya. g-ra (98)
run		kat-						gàt	

(*continued*)

Table 1c (continued)

	BODO				*EASTERN NAGA*			*Jinghpaw*	** Sino-Tibetan Conspectus*
	Boro	*Garo*	*Atong*	*Wanang*	*Konyak*	*Nocte*	*Tangsa*		
shake	samaw (DNSB)							shəmu (H)	
sleep			cəu-	cu-cəu-	shipu	ajup	jip	yɔ́úp	* ip (114)
stomach	uɔ-dəl	ok		ok			wok		* pu•k. buk (358)
suddenly		raɔ ŋ-san						làŋiətá?	
swim			huŋ-nu-	huŋ-nə-	yiang-yat	ajuk	jung-jap	phùŋ-yòt	* pyaw (176)
thin	bá (DNSB)	baɔ-	baɔ-	pa-ta-la-	pee			phà	* ba (25)
tickle		juk-juk-						kəjuk (H)	
urine	sidi (Dimasa)	su-bu				sa	sit	jit (H)	* ts(y)i(77)
vulture		so-gin					akun		
waist	jengkhong (Dimasa)	caŋ-ki-con					khing	nshaŋ (M)	
wind (air)						pong	rangpong	màpuŋ	
wolf		seɔ-el			shwo		shol		
wood/tree	bon 'firewood' (DNSB)	bol	ban	pan				phúŋ	

Table 2a Most widespread Tibeto-Burman cognates

	BODO				*EASTERN NAGA*			*Jinghpaw*	** Sino- Tibetan Conspectus*
	Boro	*Garo*	*Atong*	*Wanang*	*Konyak*	*Nocte*	*Tangsa*		
bamboo	oɔ-a	waɔ-	waɔ			wa(h)	wah	káwá	* r-wa (44)
bitter	ká (DNSB)	kaɔ-			khah	akha (M)	akhah	khá	* ka (8)
black			bi-nak	pe-nek	ünyak	anyia(k)	anyak (M)		* nak (pg. 88)
cane	raydəŋ (DNSB)	re			wei	ri	ri-sa, ri-xat	ri (H)	* rey (478)
dog			gəiɔ	kui	kui	hu	hi	gùi	* kwiy (159)
dove		doɔ-kru					wotokhu	u khrudu (H)	* kruw (118)
dream		ju-maŋ			mang	mang (M)		yɔ̀ùpmāŋ	* maŋ (82)
ear		na-cil	na-kar	na-kor		na	nah	nā	* g-na, r-na (453)
earth/ground	naɔ	aɔ-	haɔ	ha	ka (M)	ha	hah	gá	* r-ka (97)
eye	moɔ-gon	mik-	mək-	mək-	mük	mit(mik)	mik	myìɔ	* mik, myak (402)
fish	naɔ	naɔ-tok	naɔ	na	nyah	nga	ngah	ŋá	* ŋya (189)
flower/blossom	biɔ-bar	bi-bal	bar	par	jupiang	apo	pilpung	nampan (H)	* ba•r (1)
ginger	naiɔ-zeŋ	eɔ-ciŋ	ceŋ		tieng	cing	cing		
hair/feather		kim-il			-men			mūn	* (s-)mul, (r-)mul (2)
hand/arm	aɔ-kaiɔ	jak	cak	cak-toŋ	yak	dak	jak	lətáɔ	* lak (86)
heavy	risiba (Dimasa M)				yih	ali:	alih	lī	* s-liy (95)
hole	'knor (PCB)	kol				kan		nkhun (H)	kwar (350)
horn	goŋ	groŋ	go-roŋ	ko-roŋ	üwong (M)	arong	rong	nàrūŋ	* ruŋ (85)
hot/warm	alu (DNSB)						alom	lum (H)	* lum (381)
I	aŋ	aŋ-	aŋ	aŋ		nga	nga	ŋāy	* ŋa (406)

(continued)

Table 2a (continued)

	BODO				*EASTERN NAGA*			*Jinghpaw*	** Sino- Tibetan Conspectus*
	Boro	*Garo*	*Atong*	*Wanang*	*Konyak*	*Nocte*	*Tangsa*		
knee		(jaʔ-)sku				da-ku	langku		* m-ku• k (pg. 120)
laugh	mi-ni		mi-mi-	mi-ni-				mànī	* m-nwi(y) (191)
name	muŋ	bi-muŋ	bi-muŋ	muŋ	min (M)	ming (M)	ming	myīŋ	* r-miŋ (83)
nest	buthup (Dimasa M)	bi-tip					wotip	tsíp	
palm/sole	apá (DNSB)	-pa			-pha	-pha	-pha	làphàn	* pa (418)
pig	oʔ-ma	wak	wak	wak		wak	wak	wàʔ	* pak•pwak (43)
ripe	gumun (Dimasa)	min-					min-na	myīn	* s-min (432)
road	la-ma	ra-ma	ram	lam	lam	lam	lam	lām	* lam (87)
smoke		-ku	-ku-si	-cu		(van-)khu	khuh	khùt	* kuw (256)
snake	zi-bəu	cip-bu	də-bəu	tu-pu	pu	pu:	puh	pu (H)	* buw (27)
stone	oʔn-tai	roʔŋ-te	roʔŋ-tai	loŋ-tai	long	long. lo:ng	long	nàiùŋ	* r-luŋ (88)
straight	gə-təŋ	toŋ-toŋ			ting			dīŋ	
suck	səp			bu-sup-	jep, hüp	acup	sip	cùp	* dzo•p (69)
tail		kiʔ-me	di-mai	ci-mai		me(h)	me	nìŋ-mày	* r-may (282)
thorn	su, busu (Dimasa)	buʔ-su			huh	shu	suh	jú	* tsow (276)
thou	nəŋ	naʔŋ	naʔŋ	naŋ	nang	nang	nang	nāŋ	* naŋ (407)
tongue	sa-lai	sre	te-ra-pak	te-lai	yee	theli	phali	shìŋlet	* m-lay. s-lay (281)
tooth	ha-tai	wa-gam	wa	øa	pha	pa	pa	wā	* s-wa (437)

Table 2b "Water" group

	BODO				*EASTERN NAGA*			*Jinghpaw*	** Sino- Tibetan Conspectus*
	Boro	*Garo*	*Atong*	*Wanang*	*Konyak*	*Nocte*	*Tangsa*		
blood	təiʔ	aʔn-ci	təiʔ	ci		he	tahi	sày	* s-hwiy (222)
die	təi	si-	təi-	ci-			tia	sī	* siy (232)
dig	zauʔ	coʔ-	cauʔ-		sho (M)	thu (M)	thut	céʔ, thù	* tu (256)
eat	zaʔ	caʔ-	saʔ-	sa-	ha (M)	cha	sat	shá	* dza (66)
egg	dauʔ-dəi	doʔ-ci	dauʔ-dəi	tau-ci		vu-ti	woci	ùdì	
fruit	pi-taiʔ	bi-te	taiʔ	t'əi			bangti (cf. 'tree')	sì	* sey (57)
kill		soʔ-ot-		ga-tat-			tikduk	sàt	* g-sat (58)
mortar		caʔ-am				tham		thùm	* tsum (75)
oil	taw (DNSB)	to		thu (Chutie)		tian	ten	sáw	* sa•w (272)
sit		n-aoŋ-		mo-soŋ-		tong	tong	dūŋ	* tu•ŋ, duŋ (361)
sweet	dáy (DNSB)	ciʔ-				atu		dùy	* twi(y) (166)
water	dəi	ci	dəi	ci-ka	yiang	jo	jung	khàʔ	* ti(y) (55)

Table 2c Less widespread but possible cognate sets

	BODO				*EASTERN NAGA*			*Jinghpaw*	** Sino-Tibetan Conspectus*
	Boro	*Garo*	*Atong*	*Wanang*	*Konyak*	*Nocte*	*Tangsa*		
axe	ruwa(DNSB)	ru-a		loa (Chutia)	wa	wa:-ka:		ŋàwā, nìŋwā	* r-wa (441)
bat		doʔ-bak			oupak(M)	phakarang (M)			* ba•k (325)
bird						vu	wo	ù	* wa (99)
break	pay:(DNSB)	peʔ-			pai (M)				* pe (254)
buffalo						lei	li	wálõy	* lwa•y(208)
cattle/cow	məkəw(DNSB)	mat-cu			mahu	man	man	dumsu (H)	
chin						ka, ka-ra	kah	ŋìŋ-khá	* (m-)ka, (s-)ka (470)
cry	gáb(DNSB)	grap-						khràp	* krap (116)
descend					yu (M)		ju	yu (H)	* yu(w) (p. 101)
good					mai			māy	* may (300)
head	koro(DNSB)	sko				kho	kho		
leech	luwád (DNSB)	ru-at			at	wavot, sawot	sawat	wòt	* r-pat (45)
louse	teʔ-ma	tik	kə-rət	hə-rək		rit	tharik	tsīʔ	* s-rik(439)
monkey						ve	wi	wōy	* woy (314)
neck	gə-dəʔ-na	git-dok	dok-e-reŋ	tuk-ur				dùʔ	* tuk (392)
owl	daokhu (Dimasa)				nyakao(M)	vakhu (M)		ùkhú	* gu/ku (p. 164)
rat					yuh	ju-pu		yú	* b-yuw (93)
vomit	gaba (Dimasa)				phai	phat-phe	phat	nphat (H)	
weave		dok				kat: tak	khat-wat	dàʔ	* tak (17)

Table 3 Numbers

	BODO				*EASTERN NAGA*			*Jinghpaw*	** Sino-Tibetan Conspectus*
	Boro[1]	*Garo*	*Atong*[2]	*Wanang*	*Konyak*	*Nocte*	*Tangsa*		
one	[2]se	sa	goesa		ja	wan-the	ashe	láŋay	
two	[1]nōy	gin-i	goeni		ni	wan-nyi	ani	lákhôŋ	* g-nis (4)
three	[2]tham	git-tam	goetam		lum	wan-ram	atom	màsūm	* g-sum(409)
four	[1]brōy	bri	biri		peli	beli, bali	bali	màlī	* b-liy(410)
five	[1]ba	boŋ-a	banga		nga	banga	banga	màŋā	* l-ŋa, b-ŋa (78)
six	[2]do, [2]ro	dok	korok		wok	irok	tharok	krúʔ	* d-ruk(411)
seven	[1]sni, [2]si [1]ni	sin-i	sene		nyet	ingit	sanat	sànìt	* s-nis (5)
eight	[2]zad, [2]daŋ	cet	chatgik		tet	set, sat	ashat	mətshát	* b-r-gyat (163)
nine	[2]si [1]kho	sku	chiku		du	khu	akhu	càkhù	* d-kuw (13)
ten	[1]zi, [1]zu	ci-kuŋ	chaigik		pen	chi	asi	shī	* ts(y)i(y) (408)
twenty	khon (Dimasa)	kol	kol				tokom	khūn	* (m-)kul (397)
hundred	[2]zow [1]se	rit-ca-	ratja			ca-the (the='one')	shasha (she= 'one')	làtsā	* r-gya(164)

[1]Boro from PCB [2]Atong from Grierson (1903)

Table 4 Kinship terms

	BODO				*EASTERN NAGA*			*Jinghpaw*	** Sino-Tibetan Conspectus*
	Boro	*Garo*	*Atong*	*Wanang*	*Konyak*	*Nocte*	*Tangsa*		
grandmother	bə́y(DNSB)	am-bi			opi, be	vi	wi	wōy	
grandfather		a-cu				te(i)	te	jī	
mother					nyu	nyiong	nu	nù	
father	piɔ-pa	pa-gip-a waŋ 'FaYBr'	a-wa	a-wa-par	pa	wa	wa	wà	* pa (24)
MoYoSi	adə́y(DNSB)	maɔ-de						nu doi (H)	
MoBrWi/ Mo-in-law		mani 'MoBrWi ni-o-taŋ 'Mo-in-law	naiɔ	nəi	nyei	nipa		nī	* khu (225)
MoBr/ Fa-in-law	haw: (DNSB)	oɔ-bit-e 'Fa-in-law' mama-o-acu 'MoMoBr'	hauɔ	hau	kūo	hopa, hapa 'Fa-in-law'	ho, 'MoBr' hoca 'Fa-in-law'	gù 'FaSiHu'	* khu (225)
ElSi					yoinya nyayung		nakho	nā	
ElBr					pheipu	pho-pho	phokho	phù	
YoSib		no	nau	a-nau	nao		no	nāw	* na•w (271)
Cross-cousin	bənaŋ 'sib-in-law' (DNSB)	bo-niŋ (male)						nīŋ (female)	
child		bi-sa	sa-	sa-sa		cha	sa (M)	shà	* t,sa. za (59)
niece/nephew		nam-cik 'cross-niece'						nām 'cross niece, nephew	
nephew		gri						khrī	
grandchild		su			ho, hu	chu	ashuo (M)	shù	* sū(w)

Notes

1 This is a revised version of a paper that I first prepared for the XVth International Conference on Sino-Tibetan Languages and Linguistics held in Beijing, People's Republic of China, August, 1982. I am specifically indebted to Paul K. Benedict and to David B. Solnit, both of whom offered helpful suggestions about a number of my entries, and I am more generally indebted to the hosts of the Beijing conference for their warm hospitality.

2 Karen has always looked manifestly Tibeto-Burman to me, and I see no reason to omit it from the list of sub-groups of the larger family.

3 I did ethnographic and linguistic field work in the Garo Hills from 1954 to 1956 with the help of a generous fellowship from the Ford Foundation.

4 I worked with La Raw Maran in the summers of 1968 and 1969. I extend my personal thanks to him for his high-spirited help, and my institutional thanks to the Center for South and Southeast Asian studies at the University of Michigan which helped to support our work together.

Bibliography

Benedict, Paul K. *Sino-Tibetan: A Conspectus*. Cambridge: University Press. 1967.

Bhat, D. N. Shankara, *Boro Vocabulary* (With a Grammatical Sketch). Poona: Deccan College Postgraduate and Research Institute, 1968.

Bhattacharya, Pramod Chandra, *A Descriptive Analysis of the Boro Language*. Gauhati: Gauhati University, Department of Publication, 1977.

Brown, W[illiam] B[arclay], *An Outline Grammar of the Deori Chutiya Language*. Shillong: Printed at the Assam Secretariat Printing Office. 1895.

Burling, Robbins. "Proto-Bodo." *Language*. Vol. 35, No. 3, July-September, 1959, pp. 433–453.

Burling, Robbins. "The Historical Place of Jinghpaw in Tibeto-Burman." *Occasional Papers of the Wolfenden Society on Tibeto-Burman Linguistics*, Volume II. F.K. Lehman, ed. Urbana: Department of Linguistics, University of Illinois, 1971.

Burling, Robbins. "Garo Spelling and Garo Phonology." *Linguistics of the Tibeto-Burman Area*. Vol. 6, No. 1, pp. 61–81, 1981.

Grierson, G. A. *Linguistic Survey of India*. Vol. III, Part II, Tibeto-Burman Family, Specimens of the Bodo, Naga, and Kachin Groups. First edition, 1903. Reprint, Delhi: Motilal Banarsidass, 1967.

Gupta, K. Das. *An Introduction to the Nocte Language*. Shillong: North-East Frontier Agency, 1971.

Hanson, O. *A Hand-Book of the Kachin or Jinghpaw Language*. Rangoon: American Baptist Mission Press, 1917.

Holbrook, Lucy M. *A Garo Word Book* n.d. (Typescript).

Kumar, Brajbihari *Hindi Konvak English Dictionary*. Kohima: Nagaland Bhasha Parishad, 1973.

Marrison, Geoffrey E. *The Classification of the Nagā Languages of North-east India*. Two volumes. Doctoral dissertation, University of London (School of Oriental and African Studies). 1967.

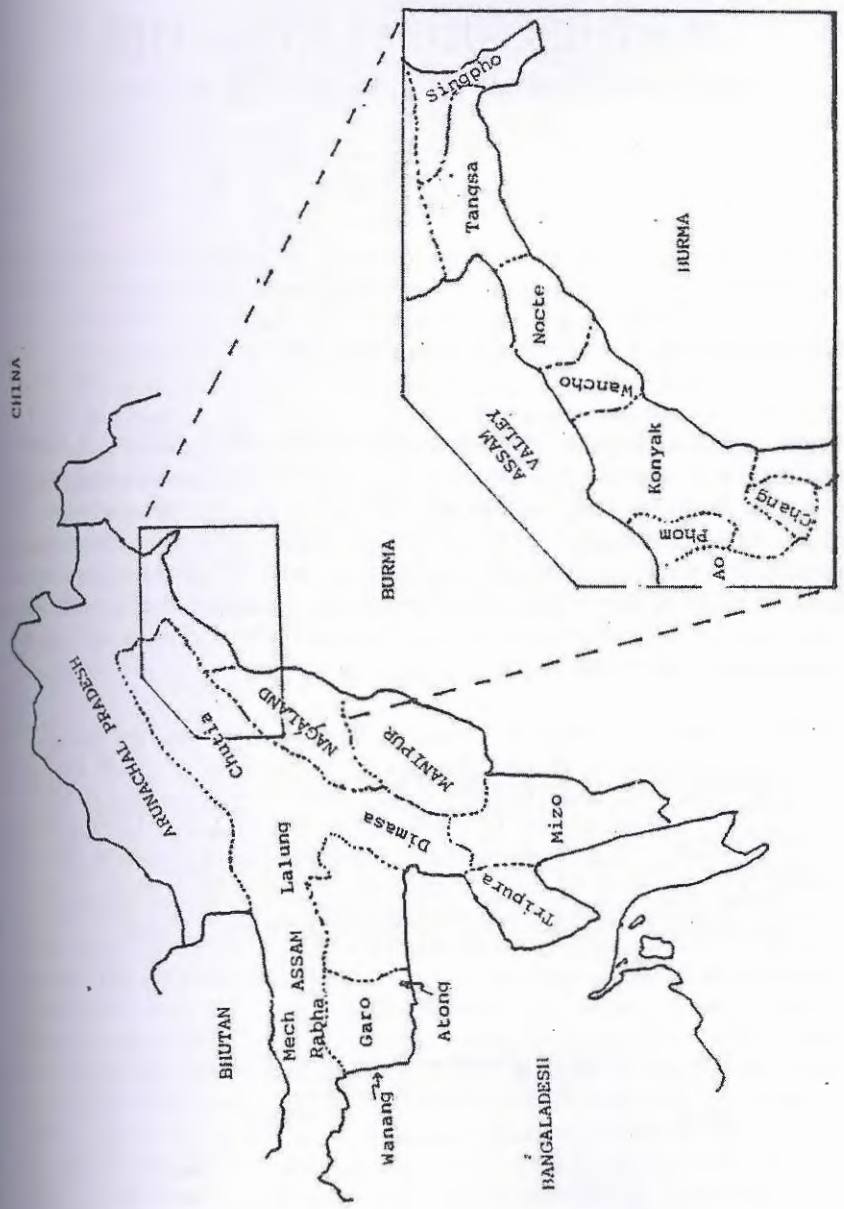
Mason, M.C. *English-Garo Dictionary*. Tura, Garo Hills: Miranda Library, 1954.

Nengminza, D. S. *The School Dictionary, Garo to English.* Tura, Garo Hills: Miranda Library, 1972.

Ngemu, T. *Moklum Language Guide*. Shillong, 1977.

Shafer, Robert, "The vocalism of Sino-Tibetan," *JAOS* 60:302–37 (1940); 61:18–31 (1941).

Solnit, David B. "Aspects of the Diachronic Phonology of the Konyak Languages". Paper presented at the 12th International Conference on Sino-Tibetan Languages and Linguistics, Paris France, October, 1979.



11

ON THE EVIDENCE FOR THE RELATIONSHIP KIRANTI-RUNG[1]

Karen H. Ebert

Source: *Linguistics of the Tibeto-Burman Area* 13, 1, 1990, 57–78.

In an article published in 1984 Thurgood presented some arguments in favor of a 'Rung' subgroup of Tibeto-Burman. In which he includes Gyarong, Kham, Chepang, Jinghpaw, Tangut, Qiang and Nungish. Thurgood also seems to suggest a closer relationship between Kiranti and Rung on the basis of agreement patterns. The paradigms compared are from Gyarong and Thulung, a Western Rai language. I shall here present material from the verbal paradigms of Southern Rai, which turn out to be much closer to Gyarong than Thulung or, as far as I can see, any language of the Rung group.

The Southern Rai languages Chamling, Bantawa, and Puma are spoken in Eastern Nepal in the area between the rivers Sun Kosi, Dudh Kosi, Rawa Khola, and Arun. Bantawa is also used as a lingua franca by Rai people of different descent in Ilam province. The languages are closely related, Puma sharing 80% cognates with both Chamling and Bantawa. The Rai data are, if not indicated otherwise, from my own fieldwork on Chamling and from materials of the Linguistic Survey of Nepal.

1. Gyarong

Table 1 Gyarong verbal paradigm (Jin et al. 1958, Nagano 1984)

Direct		*Inverse*	
1>2s	ta- -n	2>1s	kə-u- -ng
1>2d	ta- -Nch	2>1d	kə-u- -Nch
1>2p	ta- -ny	2>1p	kəu- -y
2s>3	tə- -u/-n	3>2s	tə-u- -n
2d>3	tə- -Nch	3>2d	tə-u- -Nch

Table 1 (continued)

Direct		*Inverse*	
2p>3	tə- -ny	3>2p	tə-u- -ny
1s>3	–ng	3>1s	wu- -ng
1d>3	–ch	3>1d	wu- -ch
1p>3	-y	3>1p	wu- -y
3s>3	–u	3dp>3	wu-

1.1. Agreement

The Gyarong verbal system is characterized by agreement with speech act participants (SAP) independent of semantic role. If both actants are SAP the verb agrees with patient. The pronominal suffixes reflect the common Tibeto-Burman (TB) pronouns **nga* "I". **na(ŋ)* "thou", as well as plural **i*. The other actant is marked neither for person nor for number; 3rd person is never marked for number.

There is a prefix (tə-艾 ta-艾 kə-) in all 2nd person configurations. DeLancey (1980:53) splits ta- into tə-+ a and attributes to a the function of uniquely specifying the 1>2 configuration. Although this configuration is often set apart from the other forms in TB, no such prefix is known from other languages. Therefore it seems more likely that *ta-*, realized as *tək-* before velar stops, derives from *tə-kə-* (Nagano 1984:72), *kə-*being a 1st person prefix (cf. 5.2.). *kə- -ng* in 2>1 stands for 1st person, just as *ta/tə- -N*in 1>2, 2>3, and 3>2 stands for 2nd.

1.2. Direction marking

Inverse

All inverse configurations, with an agent lower on the scale of natural viewpoint or empathy than patient[2], have the prefix *(w) u-*. Although DeLancey (1980:53) convincingly assigned the status of inverse marker to Gyarong *wu-*, Nagano (1984:71) claims that it stands for 3rd person. An inverse marker can of course easily be taken for a 3rd agent marker, and a direct marker for 3rd patient. Reinterpretations along this line have taken place in several TB languages (e.g. Limbu) - but not, I think, in Gyarong. The crucial forms are those that mark 1st/2nd and 3rd/3rd configurations. Gyarong 2>1 with the prefixes *kə-u-* obviously contradicts Nagano's analysis. Elsewhere Nagano (1984:75, 181) seeks a way out in making *wu-* a marker for non-first. This can hardly be regarded as a satisfactory solution; we would then, according to Nagano's agreement scheme PAT-AG-(verb)-PAT, expect e.g. **wu-tə-*in 2>3.

Also, it is hard to see how 3s>3 *-u* and 3ns>3 *wu-* would fit into Nagano's scheme. The distribution of *-u* and *wu-* clearly indicates that 3ns>3 counts as inverse; i.e. 3rd non-singular participants rank lower on the scale of natural viewpoint than 3rd singular. The actant hierarchy is for Gyarond and, as we shall see, at least for some Kiranti languages:

1 > 2 >> 3s > 3ns

The main break is between speech act participants and non-participants. Speaker and hearer have nearly equal status. 2>1 counts as inverse in Gyarong and Kiranti, but 1>2 forms a separate category, which may be called 'local' in analogy to the Algonquian convention.

Direct

There is synchronically no clear indication of a direct marker in Gyarong, but on the basis of comparative evidence -u in 2s>3 and 3s>3 can be identified as a trace of it. We can postulate the following pattern underlying direction marking in Gyarong:

-u	direct
(w)u-	inverse

2. Chamling

Compared to Gyarong the Chamling paradigm looks more complex, but the underlying principle is the same.

Table 2 Chamling (West) verbal paradigm[3] (past)

Direct		*Inverse*			*Intransitive*		
1>2s	-na	2>1s	ta-	-unga			
1s>2d	-na-ci	2>1de	ta-	-ac-ka			
1s>2p	-na-i	2>1pe	ta-	-i(m)-ka			
1ns>2ns	-na-ni						
or:		or:					
1de>2	-(n)-ac-ka	2d>1	ta-	-aci			
1pe>2	-(n)-um-ka						
1s>3s	-unga	3>1s	pa-	-unga	1s		-unga
1di>3s	-aci	3>1di	pa-	-aci	1di		-aci
1de>3s	-ac-ka	3>1de	pa-	-ac-ka	1de		-ac-ka
1pi>3s	-um	3>1pi	pa-	-i	1pi		-i
1pe>3s	-um-ka	3>1pe	pa-	-i(m)-ka	1pe		-i(m)-ka
1s>3ns	-ung-c-unga						
1nsi>3ns	-um-c-um						
1nse>3ns	-um-c-um-ka						
2s>3s	ta- -u	3>2s	ta-	-a	2s	ta-	-a
2d>3s	ta- -aci	3>2d	ta-	-aci	2d	ta-	-aci
2p>3s	ta- -um	3>2p	ta-	-i	2p	ta-	-i
2s>3ns	ta- -u-c-yu						
2ns>3ns	ta- -um-c-um						
					3s		-a
3s>3s	-u	3d>3	pa-	-aci	3d		-aci
3s>3ns	-u-c-yu	3p>3	pa-	-a	3p	mi-	-a

Before looking at agreement a note is in order regarding tense marking. The Chamling past marker -a appears only if there is no other suffix, e.g.

ta-id-a	he gave you
pa-id-a	they gave him

The person suffixes *-unga*, *-na* and *-ka* are used only in the past: in the non-past *-a* is replaced by the non-past marker *-e*. These suffixes (also present in the independent pronouns *ka-nga* "I", *kha-na* "you", *kai-ka* "we (pe)") should not be analysed as *(u)ng* + *a* etc.; i.e. past tense is unmarked in those forms. The vowel a before the dual marker *-ci* was originally a past marker (as it still is in Bantawa), but it has become part of the suffix.

	CHAMLING	cf. BANTAWA	
PAST	*ta-ims-aci*	*tü-ims-a-ci*	you (d) fell asleep
	2-fall asleep-d	2-fall asleep-PAST-d	
NONPAST	*ta-ims-ac-e*	*tü-im-ci*	you (d) will fall asleep
	2-fall asleep-d-NONP	2-fall asleep-d	

2.1. Agreement

We find the same pronominal suffixes in Gyarong and Chamling (for -um see below):

	GYARONG	CHAMLING
1s	*-ng*	*-nga*
1p	*-y*	*-i(m)*
2s	*-n*	*-na*
2p	*-ny*	*-ni/-i*

In both languages there is agreement with SAP in configurations involving 3rd person. Moreover both languages use a *t*-prefix for 2nd person. There is, however, a marked difference in the distribution of the 2nd person affixes. In Chamling (and in most other Kiranti languages) *-na* is restricted to 1>2. In all other configurations *ta-* functions as 2nd person marker, whereas Gyarong has the double marking *tə- -n*. Both Gyarong and Chamling uniquely specify the configuration 1>2, Chamling by *-na*, Gyarong by *ta-(<tə-kə-)*. In Gyarong 2>1 agrees with 1st person *kə- -ng)*. while the corresponding Chamling form agrees with both participants (*ta- -unga*, but cf. below under b).

Chamling has the following additional markers in comparison to Gyarong:

a) The exclusive marker *-ka*; Gyarong has an exclusive pronoun, but the feature is not marked on the verb.

b) Gyarong marks only the number of the SAP the verb agrees with: Chamling tends to mark non-singular for both participants. Competing constructions exist in 1st/2nd configurations, since only one dual or plural marker is possible in a form. From Table 2 we can derive

2d>1pe	*ta-id-i-ka* 2-give-p-e	or	*ta-id-aci* 2-give-d	you gave us
2d>1de	*ta-id-ȁc-ka* 2-give-d-e	or	*ta-id-aci*	you gave us

where the first forms agree with patient in number, the second with agent. Alternative forms also exist for the 1ns>2ns configurations, which we would expect to agree with 2nd person, as in the first of the following forms. The last form shows pure subject agreement.

1 pe>2p	*i-na-ni* give-1>2-2p	or	*i-n-um-ka* give-1>2-1pAG-e	we gave you
		or	*id-um-ka* give-1pAG-e	

c) The dual marker *-c i* is not restricted to 1st and 2nd person, but also indicates dual and plural of 3rd person patients. 1st and 2nd person agreement suffixes are repeated after the 3rd non-singular patient marker:

1s>3ns	*tyok-ung-c-unga* see-1s-3nsPAT-1s	I saw them

There is thus a tendency in Chamling to mark number of both participants, and towards agent agreement in competing configurations.

2.2. Direction marking

Direct

It is problematic to assign a synchronic function to Chamling *-u*. On the basis of comparative evidence, especially from the neighboring Bantawa, it is clear that -u has developed from a *direct marker*. This function is still recognizable in some configurations, e.g.

3s>3s	*rhaik-u*	he scolded him	cf. GYARONG	*nasngo-u*
vs.				
3p>3s	*pa-rhaik-a*	they scolded him		*vu-nasngo*
2s>3s	*ta-rhaik-u*	you scolded him		*tə-nasngo-u/n*
vs.				
3s>2s	*ta-rhaik-a*	he scolded you		*tə-u-nasngo-n*

The generalization of the suffixes *-unga* for 1st person singular and *-aci* for dual agent irrespective of semantic role shows that the direction marking is not carried through systematically. Only one informant of the Eastern dialect insisted on *-u-ci* in 2d>3 (but not in 1d>3). The vowels *u* and a have become fixed parts of the suffixes, and can even replace stem final vowels, as in

ta-ma	come	*t-unga*	I came
mu-ma	make	*m-aci*	they (d) made it

The generalization of these suffixes seems to be the result of fairly recent developments, as the comparison with Bantawa shows (cf. 3.2.). Note that the interpretation of *-u* as a 3rd patient marker would leave us with the problem of *-unga* and *-aci*; moreover it would be incompatible with the 3p>3 form without *-u*.

1st and 2nd plural have *-um* in direct, *-i* or *-im* in inverse. It is evident that *-um* combines a direct and a plural marker: *-um* < *-u* + **-iN*. [4]

Inverse

The prefix *pa-* occurs in 3>1 and 3ns>3. i.e. In the inverse configurations without t-prefix. The crucial forms for the analysis as inverse marker are

3p>3	*pa-*
3s>3s	*-u*

They demonstrate that *pa-* cannot be a 3rd agent marker, just as *-u* cannot be a 3rd patient marker. Two informants occasionally used a direct form for 3p>3 with the 3p agent prefix *mi-* and the suffix *-u*.

3p>3p	*mi-tyok-u-c-yu* 3pAG-see-DIR-3nsPAT-DIR	they saw them

The fact that the corresponding inverse form *pa-tyok-a* (in which the plurality of agent and patient cannot be marked)[5] does not have *-u* confirms our analysis of *pa-* as inverse marker. The two forms, coexisting in at least some Chamling dialects, but also in Puma and Bantawa, indicate that the switch between direct and inverse in the 3>3 configurations is not stable.

The absence of *pa-* in forms with the *t*-prefix is explained by a prefix restriction in Chamling (cf. fn.5). Like inverse *pa-*, the negative prefix *pa-* is not realized in forms with the *t*-prefix, e.g.:

3s>3s	*tyok-u*	he saw him	NEG:	***pa-tyok-aina***
2s>3s	*ta-tyok-u*	you saw him		***ta-tyok-aina***

2.3. The kha- paradigm

Chamling has an alternative paradigm for actions directed towards 1st person. In the Eastern dialect the following forms have replaced the corresponding forms of the West Chamling paradigm presented in Table 2.

3s>1	*kha-*	*-a*
3d	*kha-*	*-aci*
3p	*kha-mi-*	*-a*
2s>1	*kha-ta-*	*-a*
2d	*kha-ta-*	*-aci*
2p	*kha-ta-*	-i

The prefix *kha-* is used with a 1st person patient irrespective of number. A form like *kha-ta-tyok-a* "you saw me/us" replaces five different forms of the Western dialect.

What function are we to assign to this prefix? The *kha*-paradigm differs markedly from the rest of the inverse forms. Whereas inverse markers do not disturb the split agreement pattern, the *kha*-forms have subject agreement; the forms following *kha-* are identical with intransitive verb forms. *kha-* can thus not be an inverse marker. It functions as an indicator of 1st person patient. This implies two remarkable deviations from the canonical TB agreement pattern: (1) the introduction of a role specific marker; (2) the neglect of number of a speech act participant. A parallel development occurred in Limbu (cf. 4.1.).

3. Puma and Bantawa

In Table 3, I present the Puma and Bantawa paradigms together with Chamling. Some dual and exclusive forms which are not needed for the comparison have been left out.

The Puma data are from two verb morphology questionnaires of the Linguistic Survey of Nepal. The data were not recorded by linguists and are not always reliable; dubious and obviously wrong forms have therefore not been included in the table. The Bantawa verbal paradigm was presented by Novel Kishore Rai, a linguist and native speaker of Bantawa from Ilam. A slightly different Bantawa paradigm was recorded by Gvozdanović (1985:121) in Ilam with an informant who migrated fifty years earlier from Bhojpur. Instead of N. K. Rai's *üm-* this informant uses *mü-*. Further divergences are indicated in square brackets in Table 3.

3.1. Agreement

The *t*-prefix occurs in Puma and Bantawa in the same configurations (henceforth 'T-configurations') as in Chamling. 2nd person *-na* is restricted to 1>2. The dual marker *-ci* is again also used as a 3rd non-singular patient suffix. Like Gyarong, but unlike Chamling, Bantawa distinguishes 1p *(-in, -um)* and 2p *(-nin, -num)*.

Table 3 Southern Rai verbal paradigms (past)

	Puma	*Bantawa*	*Chamling*
DIRECT			
1s>2s	-na	-na	-na
>2p	-na-ning	-na-nin	-na-i
1pe>2s	?	üm- -a [ø- -ni-cia]	-na/-n-um-ka
>2p		üm- -a-nin ["]	-na -ni/ "
1s>3s	-ung	-ung	-unga
>3ns	-ung-c-ong	-ung-c-üng	-ung-c-unga
1pl>3s	?	-um	-um
>3ns	-um-c-om	-um-c-üm	-um-c-um
2s>3s	to- -i	tü- -u	ta- -u
>3ns	to- -i-ci	tü- -u-ci	ta- -u-c-yu
2p>3s	to- -um	tü- -a-num	ta- -um
>3ns	to- -um-c-um	tü- -a-num-c-üm	ta- -um-c-um
3s>3s	-i	-u	-u
>3ns	-i-ci	üm- -u-ci	-u-c-yu
3p>3p	(mo- -i-ci)	üm- -u-ci	(mi- -u-c-yu)
INVERSE			
3p>3s	po- -a	ü- -a	pa- -a
3d>3s	po- -aci	(ü-) -a-cu [ü- -a-cu]	pa- -aci
3s>2s	to- -a	tü- -a	ta- -a
3p>2s	ni-to- -a	üm- -a	"
3s>2p	"	tü- -a-nin	ta- -i
3p>2p	ni-to- -ning	üm- -a-nin	"

			WEST-CH.	*EAST-CH.*
3s>1s	po- -ung	(ü-) -a-ng	pa- -unga	kha- -a
3p>1s	ni-po- -ung	üm- -a-ng	"	
3s>1pi	kha- -a	(ü-) -in	pa- -i	kha- -a
3p>1pi	kha-ma- -a	üm- -in	"	kha- mi-
2s>1s	to- -nga	tü- -a-ng	ta- -unga	kha- ta-
2p>1s	kha-to- -in	tü- -a-ng-nüng	ta- -unga	kha- ta-
2s>1pe	kha-to- -a	tü- -in-ka [tü- -ni-cia]	ta- -i (m)-ka	kha- ta-
2p>1pe	kha-to- -in	" ["]	"	kha- ta-

Bantawa *üm-/mü-* is a 3p agent marker (cf. Chamling *mi -*). This marker takes different shapes in Puma:

mo-	initial
ma-	after *kha-*
ni-	preceding *po-*, *to-*

N. K. Rai uses *üm-* also in 1p>2 configurations, resulting in identity of forms with 3p>2. e.g.,

1pe/3p>2s	*üm-Dhatt-a* -hit-PAST	we/they hit you
1pe/3p>2p	*üm-Dhatt-a-nin* -hit-PAST-2p	we/they hit you (p)

Here *üm-* seems to be a generalized plural agent marker, although the possibility cannot be excluded that 1pe>2 counts as inverse in Bantawa, with a rather natural hierarchy 1(+2) > 2 > 1+3 > 3. A further peculiarity of those forms is that *üm-* has led to the suppression of the 2nd person *t*-prefix in 3>2 and of the n-suffix that marks 1>2 in Kiranti.

Like Chamling, Bantawa has competing forms for non-singular 2>1 configurations. In the form N. K. Rai gives in his paradigm the suffix agrees with the patient:

2>1pe	*tü-Dhatt-in-ka* 2-hit-1p-e	you (s/p) hit us

But in an example sentence N. K. Rai uses a different form:

khana-nin-ʔa	*üngka-n-ka*	*tü-Dhat-ni-ci*
you-p-ERG	1-p-ex	2-hit-2-ns

The interpretation of the suffixes is Rai's; the Bhojpur data suggest that a different analysis is possible:

1pe>2s	*Dhat-ni-cia*
2s>1pe	*tu-Dhat-ni-cia*

As *-ni* indicates 1p, *-cia* must stand for exclusive: *-in-ka*, *-ni-cia* and *-ni-ci* would then be variants coding the same features. Those Bantawa examples illustrate the general instability of the Kiranti verbal paradigms, especially in the 1st/2nd non-singular configurations, and the reuse of affixes for different purposes.

3.2. Direction marking

Direct

In all direct configurations Puma has *-u* or *-i* in the paradigm for *dhed-* "beat", *-i* being a variant of *-u* after alveolar consonants. With a stem final velar or bilabial we would, as far as I can gather from some sample sentences of the Survey questionnaires, have *-u* throughout.[6]

Puma shows a tendency towards generalization of *-ung* for 1st (cf. 3>1s forms). Bantawa distinguishes direct and inverse more systematically, using *-u* in all direct and in no inverse configuration except 3d>3.

	BANTAWA	cf. CHAMLING	
1s>3s	*Dhatt-u-ng* hit-DIR-1s	*caidh-unga*	I hit him
3s>1s	*ü-Dhatt-a-ng* INV-hit-PAST-1s	*pa-caidh-unga*	he hit me
2d>3s	*tü-Dhatt-a-c-u* 2-hit-PAST-d-DIR	*ta-caidh-aci*	you (d) hit him
3s>2d	*tü-Dhatt-a-ci* 2-hit-PAST-d	*ta-caidh-aci*	he hit you (d)

Inverse

The cases where Gyarong has simple *Wu-* and Chamling *pa-* are split up in Bantawa into forms with *ü-* and forms with an additional plural marker **mi* (cf. fn.4):

in the Ilam dialect:	*üm- (< ü + *mi)*
in the Bhojpur dialect:	*mü- (< *mi + ü)*

The following distribution suggests an interpretation of *ü-* as inverse marker:

3s>1s	*ü-Dhatt-a-ng*	you hit me
3ns>3s	*ü-Dhatt-a*	they hit him
3s>3s	*Dhatt-u*	he hit him

ü is a high central unrounded vowel, which reflects common Kiranti *u* or *i*. Thus 2nd *tü-* cannot have developed from *ta-*, but is most likely a combination of **ta + u*. I suppose that *tü-* was originally used in inverse configurations and then generalized to all T-configurations, just as *mü-/am-* was generalized to all 3p agent forms.

The inverse marking function of the prefix *a-*, for N. K. Rai optional in most configurations, is blurred further by the affixation *a- -u* in 3d>3 in both dialects. In a comparative context, however, the origin of *a-* in a direction system seems beyond doubt.

Puma has the *pa-* and *kha-* prefixes combined into one single paradigm. *kha-* is regularly used in configurations with a 1st non-singular patient, but only once with 1st singular. Like Chamling *pa-*, the prefix *po-* is incompatible with the 2nd person *t*-prefix, but it does combine with the 3p agent marker.

3s>2s	*to-dhed-a* 2-hit-PAST	you hit me	cf. CHAMLING	*ta-caidh-a*
3ns>1s	*ni-po-dhed-ung* 3nsAG-INV-hit-1s	they hit me		*pa-caidh-unga*

4. Other Kiranti languages

4.1. Limbu

If we compare Table 4 with Table 3 it is obvious that the Limbu verbal paradigm has much more in common with Southern Rai than Khaling Rai has.

Limbu *ke-* corresponds in distribution to the t-prefix in Southern Rai, the *n-* suffix again being restricted to the 1>2 configurations. Limbu has a further prefix a-, which stands for 1p inclusive. Neither prefix is sensitive to direction.

Table 4 Further verbal paradigms

	LIMBU/Panthare (past)	*KHALING (Rai) (nonpast)*	*RAWANG (Nungish)*
1>2			
1s>2s	-ne	-nä	-ng
>2p	-ne-ning	-nu	-ning
1pe>2s	-ne-gya	i-	-i
>2p	-ne-ci-gya	i- -ni	-i
DIRECT			
1s>3s	-ung	-u	-ngu
>3ns	-ung-si-ng	-nu	
1pi>3s	a- -um	-ki	-i
>3ns	a- -um-si-m	”	
2s>3s	kɛ- -u	i- -ü	e- -u
>3ns	kɛ- -u-si	i- -nu	
2p>3s	kɛ- -um	i- -ni	e- -ning
>3ns	kɛ- -um-si-m	”	
3s>3s	-u	-ü-	-u
>3ns	-u-si	-nu	
3ns>3s	mɛ- -u	”	
>3ns	mɛ- -u-si	”	
INVERSE			
3s>2s	kɛ- -a	i-	e-
3ns>2s	kɛ-mi- -a	”	
3s >2p	kɛ- -i	i- -ni	e- -ning
3ns>2p	kɛ-mi- -i	”	
3s>1s	-ang	i- -ngaa	e- -ng
3p>1s	me- -ang	i- -ngaa-nu	
3s>1pi	a- -a	i- -ki	e- -i
3p>1pi	a-mi- -a	”	

Table 4 (continued)

	LIMBU/Panthare (past)	*KHALING (Rai) (nonpast)*	*RAWANG (Nungish)*
INVERSE			
3s>1pe	yapmi- -a	i- -kaa	
3p>1pe	yapmi-me- -a	”	
2s>1s	ke- -ang	i- -ngaa	e- -nga
2p>1s	yapmi-ke- -a	i- -ngaa-nu	e- -sha
2s>1pe	”	i- -kaa	”
2p>1pe	”	”	”

The non-singular markers are familiar from Rai: *-si* (= Southern Rai *-ci*) indicates any dual or a 3ns patient, *me-/mi-* a 3ns agent. The verb can have more than one prefix, e.g.

3p>2s	*ke-m-su-ʔ* 2-3pAG-touch-NONP	they touch you
2p>1s/nse	*yapmi-ke-ssu-ssi -ʔ* 1(e)PAT-2-touch-2p(?)-NONP	you (p) touch me/us (Weidert & Subba: 218: glosses mine).

There is no direction marking in Limbu. All 2>3 and 3>3 (including 3ns>3) configurations have *-u*; i.e. *-u* functions as a 3rd patient marker. There is no trace of an inverse marker.

The Panchthare Limbu prefix *yapmi* - is restricted to inverse configurations with 1st person exclusive or singular patients.[7] It can substitute for the old personal affixes, but the replacement is not regularly carried through. In the paradigms, Weidert & Subba irregularly list alternative forms, e.g.

2d>1d/pe	*yapmi-ke-hip/yapmi-ke-hips-ya-ʔ*	you (d) beat us (1985:181)

In Phedappe Limbu, according to van Driem, the prefix *a-* is optionally replaced by *napmi-* in 2>1 configurations (v.Driem: 78), but in the following text example it is used instead of the 1s suffix:

aŋga	yang-in	na·pmi	pi·r-ɛʔ
me	money-ABS	1	give-IMP
give me my money!			(v.Driem: 302)

In Phedappe, the origin of the marker is transparent: *napmi* is an impersonal pronoun analogous to French *on*.

The replacement of 1st person markers in inverse configurations constitutes an interesting parallel to the Chambling and Puma *kha-* paradigms. The process seems to start with plural patients: in Limbu an impersonal form replaces 1st person exclusive affixes in inverse configurations: Puma *kha-* replaces both inclusive and exclusive affixes. While in Limbu and Puma a 1st singular suffix is affected only occasionally, the substitution has been carried through in East Chamling to all 1st patient configurations.

4.2. Khaling-Dumi

There are two prefixing languages in the Northwest of the Kiranti area: Khaling and Dumi, which have a vocalic prefix (Khaling *i-*, Dumi *a-*)[8] in all inverse forms and in 2>3 (with a further generalization to 1p>2 in Khaling: cf. Bantawa *üm-!*). The Khaling forms (given by Toba. pers. communication) are listed in Table 4.

Exactly identical patterns are found in the Nungish languages Rawang (Table 4, data from DeLancey 1980) and Trung. Rawang has *e-*, Trung *nə-* in all T-configurations and in inverse. Synchronically this strange distribution of the prefix does not make sense; it can only be understood as a collapse of a 2nd prefix with the inverse marker.

There are also some traces of a direct marker *-u* in both Khaling-Dumi and Rawang-Trung. The overall pattern of affixes is more similar between those two groups, supposed to belong to different subgroups of Tibeto-Burman, than between Khaling-Dumi and Southern Rai.

5. Eastern Kiranti in comparison

The similarities in the Gyarong and Eastern Kiranti (Southern Rai and Limbu) verbal systems is too obvious to be overlooked or to be ascribed to coincidence. In both groups we have:

a) suffixed pronominal agreement markers *-nga* and *-na*,
b) agreement with SAP (with a tendency towards subject agreement in Kiranti),
c) a dual marker *-ch/-ci*,
d) a 1p/2p marker *-i* (+nasal),
e) a *t-* or *k-*prefix for 2nd,
f) traces of a suffixed direct marker *-u*,
g) traces of a prefixed inverse marker *u-* or *pa-* (not in Limbu),
h) inverse marking of 3ns>3 (not in Limbu).

Features a)-d) are well known from other Tibeto-Burman languages; features e)-h) are not.

5.1. Direction markers

Relics of direction marking can be traced in all Eastern Kiranti languages as well as in Gyarong. The distribution, irrespective of the synchronic function, comes out as in Table 5.

In Southern Rai there is some instability in the marking of 3ns>3, as also shown in Table 1. Some Chamling speakers use a direct form in *-u* besides the common inverse *pa-*form in 3p>3. Bantawa has both ü- and -u in 3d>3. Despite those minor variations, the hierarchical ordering of 3s over 3ns is clearly recognizable in Southern Rai. An inverse interpretation of 3ns>3 configurations has not been attested for any other Tibeto-Burman language so far[9]. Both the form of the inverse marker and its distribution are, to our present knowledge, unique to the Kiranti-Rung complex and constitute the strongest argument for a close relationship.

The only Kiranti language for which a direction system has been claimed is Hayu. DeLancey ascribes a directive function to Hayu *-ko* and *-su* and relates *ə* to Nocte inverse *h*. I find the latter case rather unconvincing, especially as *-su* is restricted to configurations with a 1st singular patient. Hayu *-ko* is synchronically used as a 3rd patient marker (Michailovsky 1988:113), but like Kiranti *-u* it seems to be an old direction marker.

Direction systems that bear similarity to those of Kiranti and Gyarong are found in Kham and Chepang, both supposed to be included in Rung: e.g.

MHAI KHAM	DIRECT		INVERSE	
	1s>3s	*-ng*	3s>1s	*o- -ngu /-u-ngu*
	2s>3s	*nə- -n*	3s>2s	*o- -nu /-u-nu*

(cf. DeLancey 1980:100). The inverse configurations have either a prefix *o-* and a suffix *-u*, or *-u*-PERS-*u*, exemplifying the phenomenon of morphemes shifting position found over and over again in TB. Chepang uses *-u* and *-taa* in a way reminiscent of direct and inverse (DeLancey 1980:57f).

Table 5 Traces of direction markers

	Gyarong	*Chamling-Puma*	*Bantawa*	*Limbu*
1>2	-	-	-	-
1>3	-	-u	-u	-u
2>3	(-u)	-u	-u	-u
3s>3	-u	-u	-u	-u
3ns>3		(-u)	-u	-u
3ns>3	wu-	pa-	ü-	
3>1	wu-	pa-	ü-	-
2>1	u-	-	(*ü-)	-
3>2	u-	-	(*ü-)	-

A direction system has also been claimed for Nocte (DeLancey 1980:82 1981:641f). Nocte *h* is part of the system of local deixis. cf.:

NOCTE	*Wankhu ka-ta*	W. went
	Wankhu ka-tha	W. has come
	nga-ma ate-nang chien-ta-k	I asked him
	ate-ma nga-nang chien-tha-ng	he asked me
		(Das Gupta 1971:73,20)

Both its form and its integration into the local deictic system places Nocte *h* far away from the Kiranti-Gyarong direction system.

5.2. The t- and k-prefixes, Chamling kha-

The following morphemes are widespread in Tibeto-Burman both in pronouns and in agreement:

velars for 1st:	*ka, nga*
alveolars for 2nd:	*ta, na*

In Kiranti-Rung the stops primarily occur in prefixed position, the nasals suffixed, but the positions may be switched around.[10] The identity of Gyarong and East Kiranti prefixes could therefore be due to accident. Together with the direction marking system, however, the prefixes gain in significance for the establishment of a close relationship between Kiranti and Rung.

If we take Gyarong *kə-* to be a 1st person marker, for which there is some internal evidence (1s>1p *ka-* <**kə-kə-*, *1>2 ta-* <**tə-kə-)*[11], how are we to explain that the prefix stands for 2nd in Limbu? Limbu must have reinterpreted **k-* as a 2nd marker in a 1st/2nd configuration and generalized it to the other T-configurations. This seems plausible if **k-* originally occurred in 2>1 (Gyarong *kə -ng* 1- -1, Limbu *kə- -ang* 2- -1).

The prefix *kha-* occurs in Chamling and Puma in configurations with 1st person patients. A relation with **ka-* (present in Chamling *ka-nga* "I", *ka-i* "we (pl)") is doubtful. Apart from the aspirated initial, 1st person singular and plural are otherwise never collapsed in Kiranti, whereas *kha-* is used to indicate a 1st person irrespective of number.

Table 6 Distribution of the t- and k-prefixes

	Gyarong	*S.Rai*	*Limbu*
1>2	ta-(tə-k-)		
2>3	tə-	tə-	kɛ-
3>2	tə-	tə-	kɛ-
2>1	tə-	tə-	kɛ-

Another possibility that should be taken into consideration is that *kha-* developed parallel to the directive systems in some Kuki-Chin languages. The directive prefix *hong-/on-* in the Northern Kuki languages Sizang, Tiddim, and Paite reflects an earlier verb "come" (DeLancey 1980:170ff). The prefix indicates movement towards speaker or hearer, as in

SIZANG	*hong-pe-tu*	hi	he will give it to me
	hither-give-FUT	IND	
cf. E-CHAMLING	*kha-id-e*		"
PAITE	*ka chanai* *on-pia-in*		give me my share!
	my share	hither-give-Imper	
cf. E-CHAMLING	a-ro° *kha-id-anna*		give me my food!

Although **kha* is a widespread TB motion verb root, it is not traceable in Chamling-Puma or any of the neighboring languages. The origin of *kha-* in the Chamling-Puma 1st patient configurations therefore remains obscure.

6. Some historical speculations

The shared features between Southern Rai and Limbu make it most likely that they had some period of common history after splitting away from the other Rai peoples. This hypothesis will of course have to be confirmed with material from other domains. However, as verbal paradigms tend to be conservative, and as it is hard to imagine that people would borrow a complex paradigm of the Kiranti type, I think the postulation of an Eastern Kiranti subgroup, comprising Limbu and Southern Rai, is quite sound.

The grouping together of 'Rai' languages seems to have had a geographical or political rather than a linguistic basis. Khaling and Dumi share more than 80% cognates, but less than 35% with any other Rai language (Hansson, ms.). It could be that their history is different from that of the other Rai groups and that some of the features they share with Rai are the result of areal diffusion.

This leads us to the question of higher groupings. We have found striking similarities in the verbal paradigms of Gyarong and Eastern Kiranti on the one hand and Nungish and Khaling-Dumi on the other hand. This grouping cuts across the classification of Gyarong and Nungish into Rung and of Khaling-Dumi with Kiranti.

An independent invention of the complex verbal paradigms of Gyarong and Eastern Kiranti is most unlikely. The question is whether we are dealing with retention or innovation. As Gyarong and Eastern Kiranti have not been shown to be closely related in other respects one would probably opt for retention. But there is no evidence for direction marking of the Kiranti-Rung type anywhere outside those groups, the Northern Kuki directive system being of a different order. The direction system, together with the distribution of the t-/ k-prefixes makes it seem

likely that the ancestors of the Kiranti and the Gyarong once were at least neighbors participating in the u-/-u direction marking and the prefixing wave.

The case of Khaling-Dumi and Nungish is weaker and needs further investigation.

I will not carry my speculations further and propose a new genealogical subtree.[12] It is surely premature to claim a new subclassification only on the basis of verbal paradigms, although they must, due to their conservatism, play a crucial role in any classification. More detailed descriptions of more languages and more detailed comparisons are necessary. Discovery of further prefixing Kuki-Chin languages of the Lakher type (cf. DeLancey 1989) may change the picture completely.

Notes

1 This article is a revised version of a paper read at the 21st ICSTLL in Lund, Sweden, in 1988. For comments on an earlier version I would like to thank Scott DeLancey. Thanks are also due to the Deutsche Forschungsgemeinschaft for financial aid that made my fieldwork in Nepal possible and to Werner Winter for giving me access to the data of the Linguistic Survey of Nepal.

2 For discussion of the category 'inverse' see e.g. Comrie (1980), DeLancey (1981:641ff). The relevance of 'viewpoint' for the TB verb is demonstrated in DeLancey (1980, 1981).

3 The Chamling paradigm has been constructed from 12 elicited paradigms (no 2 of them alike) on the basis of consistency and occurrence in 100 pages of transcribed texts. Informants' errors were frequent in elicitation, but not in natural speech. The regular variations are included in the table.

4 The plural marker appears in different shapes in Kiranti, e.g.

-m	1p and 2p in Southern Rai and Limbu
-im	1p in Chamling
mi-	3p in Chamling and Limbu
üm-/mü-	3p in Bantawa
-mi	3p and 1p in Thulung, 3p in Nachering, Chourase, Mewahang
-in	1p in Bantawa
-ni	2p in Chamling, Kulung, Thulung
-nin	2p in Bantawa
-niN	2p in Puma
-ni	3p in Dumi, Koi
ni-	3p in Puma
-i	2p in Chamling and Limbu; 1p in Chamling, Dumi, Thulung

mi seems to prevail for 3p agent, *im/in* for 1p/2p.

5 (West) Chamling can have only one prefix: *pa-* outranks *mi-*, *ta-* outranks *pa-*, 3nsPAT is marked only in direct. In inverse configurations the suffix *-aci* always indicates dual agent: e.g.

pa-tyok-aci they (d) saw him/them (but not: they (p) saw them)

6 The realization of the suffix vowel in direct configurations in Puma and Chamling is:

	WEST CHAMLING	EAST CHAMLING	PUMA
after/before velar/bilabial	*-u*	*-u*	*-u*
otherwise	*-yu*	*-yi*	*-i*

e.g. WEST CHAMLING	EAST CHAMLING	
lhap-u	same	he caught him
lhap-u-c-yu	*lhap-u-c-yi*	he caught them
caidh-u-m	same	he beat us
caidh-yu	*caidh-yi*	he beat him

7 The presentation in Weidert & Subba is rather chaotic. The forms given in Table 4 are from their table on pp. 60–61. In the actual paradigms *ya·pmi-* is used irregularly for 1d/pe patient, seldom for 1s patient in combination with a non-singular agent, never in 2/3s> 1s. Sometimes variants are given, e.g. *lɛ·sya?*, *yapmilɛ·?*(197) "it burns us (de)", but only *yapmi·* in the same paradigm for the corresponding past form and for 3s> 1pe.

8 Some small intermediate groups between the Limbu and Bantawa area (e.g. Dungmall, Athpariya, Chulung, Chintang) have 2nd *a* -.

9 A 3s > 3ns hierarchy exists in Chukchi and Koryak; cf. Comrie 1980.

10 E.g. Kham: Is *nga-*, 2s *ne-*; Chepang: 2nd *-te?*, 2nd pronoun *naang-te?*; Bantawa: 1st pronoun *üng-ka*; Chamling: *ka-nga*; Kuki-Chin: 1st *ka-*, 2nd *na-*; some conservative Kuki-Chin languages moreover have *-ng* and *-te*.

11 Gyarong has *kə-* also with 3rd person in the intransitive paradigm in some totally unpredictable configurations (Nagano 1984:64f). Nagano does not comment on his irregular intransitive paradigms.

12 Also I do not believe that migrating people always split up neatly in a way that would allow linguists to draw genealogical trees. The comparative work of the Linguistic Survey of Nepal has shown that it is not possible to draw a Stammbaum for the Rai languages on the basis of lexical innovation and sound shifts. Waves of innovation have spread from different centers of diffusion. There is no reason to believe that TB peoples should have split up in a neater way in past times.

References

Comrie, Bernard. 1980. "Inverse verb forms in Siberia: Evidence from Chukchee, Koryak, and Kamchadal." *Folia Linguistica Historica* 1.1, 61–74.

Das Gupta, K. 1971. *An Introduction to the Nocte Language*. Shillong: North-East Frontier Agency.

DeLancey, Scott. 1980. *Deictic categories in the Tibeto-Burman verb*. Ph.D. dissertation, Indiana University.

———. 1981. "An interpretation of split ergativity and related patterns." *Language* 57, 626–57.

———. 1989. "Verb agreement in Proto-Tibeto-Burman." *BSOAS*, 52.2, 315–33.

Driem, George, van. 1987. *A Grammar of Limbu*. Berlin, New York, Amsterdam: Mouton de Gruyter.

Ebert, Karen H. 1987. "Grammatical marking of speech act participants in Tibeto-Burman." *Journal of Pragmatics* 11, 473–82.

Gvozdanović, Jadranka. 1985. *Language System and its Change*. Berlin, New York, Amsterdam: Mouton de Gruyter.

Hansson, Gert. ms. "Recent main Kiranti areas." *Linguistic Survey of Nepal*. University of Kiel.

Jin Peng, Tan Kerang, Qu Altang, & Lin Xiangrong. 1958. "Jiarong-yu Suomo-huade yuyin he xingtai." (Phonology and morphology of the Suomo dialect of Jyarong). *Yuyan Yanjiu* 3, 71–108.

Michailovsky, Boyd. 1988. *La Langue Hayu*. Paris: Éditions du Centre National de la Recherche Scientifique.

Nagano Yasuhiko. 1984. *A Historical Study of the rGyarong Verb System*. Tokyo.

Thurgood, Graham. 1984. "The 'Rung' languages: a major new Tibeto-Burman subgroup." In: Brugmann, C. & Macaulay, M. (eds). *Proceedings of the 10th Annual Meeting of the Berkeley Linguistics Society*. Berkeley.

Weidert, Alfons & Subba, Bikram. 1985. *Concise Limbu Grammar and Dictionary*. Amsterdam: Lobster Publications.

12

THE LINGUISTIC POSITION OF TANI (MIRISH) IN TIBETO-BURMAN

A lexical assessment*

Jackson T.-S. Sun

Source: *Linguistics of the Tibeto-Burman Area* 16, 2, 1993, 143–88.

Introduction

The obscure Tani (Mirish, Mishingish) languages of southern Tibet and Arunachal Pradesh have only recently begun to receive the attention they deserve (Chhangte 1990, 1992: Sun 1993, 1994). The aim of this paper, which is part of an ongoing project to study the phonological and lexical diachrony of these languages, is to contribute toward clarifying the linguistic position of Tani languages in the Tibeto-Burman family from the vantage-point of reconstructed Proto-Tani (hereafter PT).[1]

Section 1 surveys and contrasts exisiting views on the affiliations of Tani in Tibeto-Burman. Section 2 inspects in detail a number of Tibeto-Burman languages which have been nominated in the literature as possible close relatives of Tani. After screening out a few unlikely contestants, a pilot lexical study is conducted in section 3 to weigh the degrees of lexical affinity between Tani and the remaining candidates as compared with three control languages, Written Tibetan, Written Burmese, and Garo. The implications of the output of this study on the phylogenetic position of Tani are then discussed. In the concluding section, we consider the nature of the relationship between Tani and Digarish (consisting of two known languages: Taraon and Idu), the language group which turns out to be most akin to Tani in basic vocabulary.

1. Existing views on the place of Tani in Tibeto-Burman

The genetic affiliations of Tani with Tibeto-Burman have seldom been called into question,[2] and should now be considered *proven beyond reasonable doubt* in view of the accountability of much of the PT phonological developments in terms of PTB (Sun 1993, chapter IV).[3] However, there is no consensus yet as to how Tani interrelates with other Tibeto-Burman languages. In fact, as shown in

the following survey of the subgrouping literature, opinions diverge sharply from each other with regard to both lower-level and higher-level affiliations of Tani in Tibeto-Burman.

1.1. Konow: 'North Assam'

In the colossal Linguistic Survey of India, Tani languages, along with other little-known Tibeto-Burman languages of Arunachal Pradesh, were brought together in the so-called 'North Assam' group. This was clearly meant to be an expedient, geographical grouping, as shown in the following quote from Sten Konow, the linguist originating this term (Konow 1909:568, 569, emphasis ours):

> The North Assam group is *not a well-defined philological group* with salient grammatical features distinguishing it from other Tibeto-Burman forms of speech . . . In many important points, however, Mishmi[4] differs from Abor-Miri, and *the points of correspondence just referred to are not of an importance sufficient to prove a close connexion between the two forms of speech.*

As for higher-level connections, Konow made only a vague suggestion (Konow op. cit.:572):

> The North Assam forms of speech can be described as links which connect the Tibetan and Himalaya dialects with the languages of the Bodo, Naga, Kuki-Chin and Kachin groups.

1.2. Shafer: Mishingish (Bodic/Burmic)

The distinctness of the 'North Assam' languages is further underscored in Shafer 1955:102, where no less than four separate groups are recognized: Mishingish (= Tani), Digarish (= Taraon-Idu), Midźuish (= Kaman-Meyöl), and Hrusish (= Hruso = Aka). Shafer did not attempt a further classification but suggested that all of them are 'possibly sections of Bodic, possibly of Burmic, *certainly not of Baric'* (Shafer op. cit.:102).

1.3. Benedict: Mirish (major Tibeto-Burman nucleus)

While positing Abor-Miri-Dafla (i.e. Mirish in the narrow sense = Tani) as one of the major nuclei of the Tibeto-Burman family, Benedict (1972:5) suggests that to this division perhaps also belong not only the three Arunachal neighbors of Tani: Taraon, Kaman, and Hruso, but also the geographically more distant Dhimal group of Sikkim and Nepal. This claim, in effect, upgrades for the first time Konow's 'North Assam' from an *areal* to a *genetic* grouping. He further speculates that this group (Mirish in the extended sense) could ultimately be

linked with Kachin (Jingpo), Baric (Bodo-Garo and Konyak), Nungish, and Lolo-Burmese under the supergroup 'Burmic' (op. cit.:11). This view was soon given up. In Benedict 1976:178; fn. 14, he proposes instead that, as far as core vocabulary is concerned, Tibetan, Chepang, Tamang (i.e. Bodic), Burmese-Lolo-Nungish, Lushai (Kuki-Chin-Naga), and Miri (Tani) form one supergroup as against Kachin, Garo, Konyak languages, and Chairel (or Chakpa, a Luish language according to Bradley 1993:7).[5]

Benedict's revised view on the linguistic position of Abor-Miri-Dafla (AMD = Tani) can thus be interpreted as follows: At a lower-level, AMD is most closely related to Hruso, Taraon, Kaman, and Dhimal: these languages are allied further with Lolo-Burmese, Bodic, and Kuki-Chin-Naga, as against Kachin and Baric. It is important to note that while Benedict ventures explicit claims about possible lower-level close relatives of Tani, he agrees with Shafer that *Tani is not akin to Baric.*

1.4. Other ideas

Egerod 1974 also contains a classification of Tibeto-Burman, founded largely on Shafer and Benedict's frameworks. According to Egerod, Mirish (= Tani) is one of the major branches of Tibetic (= Shafer's Bodic); further, all of the other sections (Dhimalish, Digarish, Midźuish, Hrusish, Newarish, and Dzorgaish) left unclassified between Bodic and Burmic by Shafer are directly assigned to 'Other Tibetic'. Although further genetic subrelations among these Tibeto-Burman groups are not explored by Egerod, it is clear that, like Shafer and Benedict, he does not consider Mirish to be closely affiliated with Baric.

Matisoff 1991, DeLancey 1991 and Bradley 1993 are among the most recent statements on the genetic relationships among the Tibeto-Burman subgroups.[6] Incorporating information on the newly described Tibeto-Burman languages as well as some recent low-level subgrouping proposals, they all depart in significant ways from their predecessors. In DeLancey 1991, an expanded notion of Baric is suggested, subsuming not only Bodo-Garo and Konyak-Naga (= French's Northern Naga), but also Kuki-Naga, Kachinic (Jingpo), and Mirish. What is more, in this classification 'Mirish' includes the three Mishmi languages in addition to Tani proper, but not Dhimal (assigned to Bodic) or Hruso (not mentioned in his framework). This extended conception of Baric may be inspired by the geographically-based *Kamarupan* (i.e. Assamese Tibeto-Burman) group first proposed in Matisoff 1985b: fn. 8, where, however, the term is explicitly stated to be 'a neutral overall designation for the TB languages of NE India and adjacent areas'. Matisoff 1991:480–1 proposes a simplified heuristic subclassification model of Tibeto-Burman with seven major Tibeto-Burman subgroups including Kamarupan (again with the disclaimer that this is a 'purely geographic rubric'), under which we find Kuki-Chin-Naga, Mikir, Meithei, Mru, Bodo-Garo, as well as Abor-Miri-Dafla. Unlike DeLancey's Baric, however, Kamarupan does not include Jingpo, which is assigned to form a subgroup (Kachinic) by itself. Bradley 1993 contains a wealth of valuable new demographic and

sociolinguistic information, especially concerning the Tibeto-Burman languages of India and Burma, but is unconventional in many ways. Adopting purely geographic labels, Bradley classifies Tibeto-Burman into four major groups: Western (Bodic), Northeastern India (= Burling's Sal group plus Kuki-Chin-Naga and Luish), Southeastern (Burmese-Lolo and Karenic), and Northeastern, a tentative medley group containing not only languages which Benedict 1972 puts under Mirish (i.e. Tani, Hrusish, Dhimalish, and the Mishmi languages), but also Nungish, Qiangic,[7] and some widely divergent minor languages such as Sulung and Bugun.

It is evident that there is hardly any agreement among Tibeto-Burmanists today concerning the precise linguistic affiliations of Tani in Tibeto-Burman. While this indeterminacy reflects the immature state of higher-level Tibeto-Burman subclassification in general (Thurgood 1985, Sun 1988, Dai 1989, DeLancey 1991, Matisoff 1991),[8] the uncertainty surrounding the linguistic position of Tani and related languages in particular can be directly attributed to the shortage of comparative data essential for recovering the linguistic histories of these languages, which in turn makes definitive subclassification well-nigh impossible.

Yet, what is relatively uncontroversial is that languages of the Tani group (i.e. Shafer's Mishingish, Benedict's Mirish in the narrower sense) do form a compact unit, more closely related to each other than to any other Tibeto-Burman language. We think it is important for the clarification of this issue to assert with certainty that *no other Tibeto-Burman language known to us deserves a place on the same taxonomic level as the two major Tani subgroups (Eastern and Western Tani).* Hence, earlier proposals which subsume languages like Midu (Thurgood 1986:93),[9] Aka (Nishida 1979:77), or Sulung and Bangru (Sun 1983:267)[10] directly under Tani proper are untenable. This is not to deny, of course, that Tani may not be grouped further with other Tibeto-Burman languages in a coordinate relation under some higher Tibeto-Burman division, the topic of the next section.

2. Possible close relatives of Tani

What, then, are the *collateral relatives* of Tani proper in the Tibeto-Burman family? A number of languages have been mentioned in the literature as showing particular affinity with Tani, including Lepcha (Bodman 1988); rGyarong (Nagano 1984); Dhimal (Benedict 1972, Bradley 1993); Hruso (Benedict 1972; DeLancey 1991:431; Bradley 1993). These proposals will be considered below in light of our improved understanding of the Tani evidence.

2.1. Lepcha

The phylogenetic position of Lepcha, a Tibeto-Burman language of Sikkim, has also been highly controversial. Earlier analyses have aligned Lepcha with Naga (specifically, the 'Northern Naga' branch of Shafer 1955:106),[11] Tibetan-Kanauri

and Kiranti (Benedict 1972:7–8), and Mikir (Bauman 1976). In a valuable recent revisit to the issue, Bodman (1988) compares Lepcha with a number of Tibeto-Burman languages which are lexically most similar to Lepcha, including an unidentified variety of Adi extremely similar (if not identical) to Padam. The substantial evidence of the lexical affinity between Lepcha and Adi comprises a list of 130 cognate pairs, based on which some important Lepcha-Adi consonantal correspondences are worked out.

On close inspection, however, many such sets appear to be *common retentions* from the original Tibeto-Burman lexical stock, and do not demonstrate by themselves any special lexical relationship between Lepcha and Adi. They include the following: *blood, blow, bow (weapon), carry on back, crab, cry (weep), dig, dream, drink, eat, eye, fire, fish, five, flat, four, give, horn, male of animals, leech, otter, ripe, road, seed, shade, smell v., snake, son-in-law, star, stone, three, tongue, two, and wood.*

Furthermore, the cognacy of the following items seems doubtful:

- 'sew' Lepcha *hrap,* Adi om-*kap*: The true Tani root for 'sew' is the first element om- (< PT *fiom); the second element -kap,[12] on which the comparison is based, is a verbal particle signifying 'closure'. Thus, the precise meaning of Adi om-kap is 'sew up'. This makes Adi -kap semantically less compatible with the Lepcha form.
- 'spirit' Lepcha a-*pil,* Adi a-*bur* a-jo (listed as a-bun a-jo in Lorrain 1907:361; a typo?). The Adi form a-bur a-jo can indeed mean 'spirit', but the phonology does not match (Like Lepcha, Padam preserves -l, but the form in question ends in -r).
- 'crumb' Lepcha *p'yol,* Adi pim-*pil*: The Lepcha form, which does not mean 'crumb' at all, is an adverbial which occurs in reduplicated form p'yol p'yol (e.g. *p'yol p'yol* glo nóŋ 'to fall into pieces'). The Adi word is a compound composed of the 'grain' root PT * pim plus an element pil (< PT *p jul) which refers to small rounded objects in general and appears also in such compound words as 'grain', 'coin', 'uvula', 'clitoris', and 'kidney'.
- 'dig up' Lepcha *bɔl; byol, byul*; Adi du-*bur.* The Adi compound, which has a more specific meaning of 'dig up (earth) and make it powdery', contains the morpheme du-, the real root for 'dig' (< PT *du); the -bur element, semantically incompatible with the Lepcha forms, is a (resultative) verbal particle meaning 'so as to be powdery'.
- 'beetle' Lepcha *bŭt.* Adi je-*put.* The Lepcha word is glossed 'insect that eats and causes destruction' in Mainwaring-Grünwedel 1979:258, and seems to be derived from the verb bŭt meaning 'pulverize, decay (of tooth)'. The Adi form, on the other hand, refers to 'scarab, dung beetle' and is transparently composed of je 'excrement/ dung' plus put 'burrow/bore v.'.
- 'steep' Lepcha dóp, Adi tap-. The Adi form seems to be a resultative verbal particle which means rather 'down, become horizontal (of something upright, e.g. a tree)'.

- 'stick, adhere' Lepcha *krap,* Adi *gap.* The Adi morpheme, which appears in the compound geŋ-*gap* 'adhere/stick to', actually means 'grasp/hold' and is here used as a resultative verbal particle after geŋ-, the true root for 'adhere, stick, heal'.
- 'close (v.i.)' Lepcha *zap*; Adi a-*dap.* The central meaning of the Lepcha root zap is 'place compactly'; zap seems to take on the meaning 'close together' only in an adverbial phrase să-zŭ-să-*zap.*

The following pairs seem to involve convincing cognates; however, further comments can be added to them:

- 'divide, distribute' Lepcha ór, Adi or. The two words involve different (nevertheless related) meanings in the respective languages. The Lepcha form means 'separate (people or things) that which are close together', whereas the Adi form (< PT *hor) means rather 'distribute'.
- Lepcha *rŭm* 'god', Adi u-*rom* 'ghost': Lepcha rŭm seems to refer more generally to 'benevolent spirits' and is thus semantically closer to the Adi word, which is from PT *rom 'ghost (ancestral)' (contrast PT *ju 'evil spirits').
- 'pubic hair' Lepcha *măt,* Adi a-*mut.* Actually, the semantics of the given roots in both languages goes beyond 'pubic hair'. The Adi form goes back to PT *mut, a general 'hair' root (for both body hair and hair of head). The Lepcha root măt also appears in the compound bon-*măt* 'beard (mouth-hair)'. Also to be noted is the shared -t final, rarely found in Tibeto-Burman words for 'hair'. The cognacy of these forms to PTB *mul is dubious, as there is otherwise little evidence for the *-l > -t shift in either language. In fact, PTB *mul is directly attested in the Lepcha doublet a-*myal* 'body hair, feathers, armor', as well as in the Mising L forms nam-*mur*; soŋ-*mur* <*nap-mul; *čok-mul 'beard' (PT *čok 'chin/jaw').
- 'taboo, omen' Lepcha nyo, Adi ño. The Padam Adi form is a verb which means 'be tabooed or quarantined for religious reasons'; the Lepcha form is glossed as 'be ominous, have a bad effect'. The really remarkable fact, not mentioned by Bodman, is that both of these forms show the same variant form with -t (suffix?): Lepcha nyot; Padam-Mising L ñot!

What is surprising about Bodman's comparative list is that many cases of plausible lexical comparability between Lepcha and Padam Adi (Eastern Tani) coincide with the east-west lexical split among Tani languages, and the forms more common in Western Tani do not resemble the Lepcha forms at all. Consider the following examples:

- 'breeze' Lepcha fár, Adi a-sar. This is an Eastern Tani word; cf. Western Tani: *rji (< PTB *g-ləy).
- 'swell' Lepcha bróm; Adi pom (< PTB *(s-)bwam). This form appears to occur in Padam only; other Tani < PT *bruŋ (< PTB *bliŋ-pliŋ 'full').

- 'fear' Lepcha ro(-m), Adi le-*ro.* Milang Ta-*re*-ma; Padam-Mising L le-*ro;* aŋ le-*lo* (aŋ ='heart'); other Tani < PT *pV-so - bV-so.
- 'sky, heaven' Lepcha tă-lyaṅ, Adi ta-leŋ ~ ta-jeŋ. This is mainly an Eastern Tani form (see section 3.2.2).
- 'return, (give) back' Lepcha lót, Adi - lat. This form, another verbal particle, is used only in Eastern Tani; contrast Western Tani -kur.
- 'girdle' Lepcha a-rek, Adi mag-rek. This form is found in Padam only.

We can also contribute a few more items to the list of Lepcha-Tani comparabilia:

- Lepcha pán 'be forgetful, absent-minded', PT *mit-*pan* 'forget' (PT *mit- = 'extinguished').[13]
- Lepcha pán 'break off v.i.' vs. Lepcha fán (< *ph-?) 'break off v.t.'; Padam-Mising L ben~bet 'break off v.i.'; Padam-Mising L pen-pet 'break off v.t.'. This is one of the rare instances where Tani preserves the familiar Tibeto-Burman transitivity-based voicing alternation (cf. Xiandao Achang bio '(of thread) be broken v.i.' vs. phio 'break (thread), v.t.'; Taraon bɹɯn[53] '(of ropes) be broken' vs. phɹɯn[53] 'break (ropes)' (Sun et al. 1980:205).[14]
- 'nest' Lepcha a-*šap*; PT *sup.
- 'revolve in mind; reason' Lepcha myóṅ; PT *mɯŋ 'think'.
- 'take' Lepcha lóŋ; PT *laŋ.
- 'bowels' Lepcha tă-*klí*; PT *kri. Matisoff 1978a:214–5 suggests that these forms may originate from PTB *kləy 'excrement'. The root also occurs in compound words for 'belly' and 'navel' in Tani, but not in Lepcha.
- Lepcha mlo 'world, universe'; PT *mroŋ 'world/land/earth'.

We have shown that although Bodman's original list of Lepcha-Adi comparisons needs revision, the rather remarkable lexical tie between these languages cannot be overlooked. In addition to a few new items added to the list (further search will doubtlessly uncover more), we have also made the discovery that despite the geographical location of the present Lepcha-speakers to the west of the Tani language area, it is in Eastern Tani (particularly Padam Adi), that the more striking similarities are found. Does this mean that Lepcha and Tani are close kin on the Tibeto-Burman genealogical tree? We will defer judgement until this issue is further explored below.

2.2. rGyarong

We now turn to rGyarong, another language supposedly showing special affinity to Tani according to Nagano 1984. One of the most noteworthy claims in this work is that rGyarong in its deepest lexical stratum is more intimately related to AMD (i.e. Abor-Miri-Dafla) than to either Tibetaŋ (the traditional view) or Qiangic (a view espoused by leading Qiangic specialists of China; see for instance Sun 1982 and Huang 1991).[15] In order to demonstrate this new linguistic alignment, Nagano presents a comparative list of about a hundred core vocabulary items (mostly verbs) with which to establish sound correspondences between the GC

(i.e. lCog-rtse) dialect of rGyarong and AMD. The AMD data is taken from Yano B unless otherwise stated (actually, forms are often cited from the distinct Tagen B variety instead), interspersed with Abor-Miri forms (hereafter AM) taken from Lorrain 1907. To one's puzzlement, Ao Naga and Mikir forms are included under the AMD heading, though these languages had never been considered to belong to the AMD group. What is also peculiar is Nagano's decision to use modern lCog-rtse forms instead of reconstructed Proto-rGyarong roots in his rGyarong-AMD comparison.[16] Rather than presenting a thorough review of the rGyarong-Tani lexical connections suggested by Nagano, the following sample set of comparisons supposedly representing rGyarong-Tani *dental-stop* correspondences (Nagano 1984:142), will be examined; the highlighted segments in the GC and AMD forms therein being the proposed equations:

- 'dig': GC tuw, Yano B du-to. The Yano B form goes back to PT *du which, like the rGyarong form, is a reflex of the prevalent PTB etymon *du-tu (STC #258). This is a common TB root attested in various TB branches and cannot be regarded as evidence of a special lexical link between rGyarong and Tani.
- 'hit': GC tom, AM dem. This rGyarong form is derived from PTB *dup-dip; *tup-tip 'beat' (STC #399). The nasal-final form tom 'I shall hit' is clearly secondary (< top + ŋ). cf. the infinitive form ka-*top* from the same lCog-rtse dialect cited in Anonymous 1991 (hereafter ZMYYC): 1081 and Qu 1984: 79. Padam-Mising L dem has a more specific meaning 'beat (with a stick, etc.)' and is clearly a separate root. The true cognate with rGyarong -top 'hit' is rather PT *tup 'strike', both being reflexes of PTB *tup.
- 'big': GC kte; Yano kte. No such Yano B form exists. The real Yano B root should be just -tè, a bound morpheme occurring with classifiers. Again both forms may reflect a common PTB root *tay (STC #298).
- 'see': GC mto; Yano kâ-to. This is a misinterpretation. Instead of the real root kâ (< PT *kaŋ) 'look/see' which is mistaken for a 'prefix' (op. cit.:90), the Yano morpheme selected for comparison, -to, is an imperative marker which appears on all citation-form verbs in Bor's Yano-Tagen wordlist.
- 'straight': GC sto; AM adong. This Padam L form actually means 'long' (cf. PTB *duŋ, STC p.19) rather than 'line', contra op. cit.:143.
- 'cold': GC sytak (i.e. [ʃtak]); Yano po-teng-pa. This Yano B form actually means 'dry (of clothes)' (cf. Bengni S pɯ-tɯŋ). We fail to see any possible connection, formal or semantic, between these GC and Yano words.
- 'go': GC thal; AM to. The AM form is unknown. As far as we know, no Tani language has this form with the given meaning.
- 'put': GC tha; AM tâk. The rGyarong form exemplifies a well-attested Tibeto-Burman root PTB *ta (STC #19), with an open rhyme. The AM form, occurring in a compound *tak*-po 'put (cover) on', is semantically compatible but the fact that tak- is a checked syllable makes the connection dubious.
- 'ask (enquire)': GC tho; Yano B tao-to. Tani languages, like some other Sino-Tibetan languages, use the same verb root for both 'listen/hear' and 'ask (i.e. cause to listen)'.[17] We believe that the variant forms Padam-Mising L tau, Yano B and Tagen B tao for the meaning 'ask' may reflect the same PT root *tas. The association of the Tani and rGyarong forms, though superficially plausible, is weakened by the fact that rGyarong (lCog-rtse dialect) uses a completely separate root for the meaning 'hear/listen' ka-rəŋ-na (ZMYYC).
- 'give': GC dit; Yano ǰi. The palatal initial in the Yano B form is secondary. The real PT root should be *bi (< PTB *bəy, STC #427), which is cognate rather with the regular GC word for 'to give' wə (< Proto-rGyarong *biʔ, cf. Dashuigou rGyarong bi-).[18]
- 'arrive': GC Ndu; AM tok. The AM form tok actually means 'descend'. The real Padam-Mising word for 'arrive' should be pɯŋ (< PT *pɯŋ, attested mainly in Eastern Tani languages, cf. also Bokar OY pɯŋ).
- 'meet': GC rdo; Yano che-tok. The 'Yano' form is actually a word from Mikir, which is not even a Tani language. The real Yano B word for 'meet' is gue-ter-ra (i.e. go + ? + verbal particle of reciprocality, cf. Bokar gɯ-tum-ra:).

In short, eight ('hit', 'see', 'straight', 'cold', 'go', 'give', 'arrive', 'meet'). or two thirds, of the twelve proposed cognate sets above are probably misidentified, while the sets for 'dig', 'give', and 'big', although legitimate for setting up rGyarong-Tani consonantal correspondences, are of limited value for proving the proposed lexical affiliation since common TB roots are involved. Therefore, although Nagano starts with the sensible idea of probing deep lexical relations by focusing on a selected area of core vocabulary, namely basic verbs,[19] the forms randomly picked from modern Tani languages, unfortunately, failed to provide him with a reliable basis for comparison.

Nagano's alignment of rGyarong with Tani may strike those who have examined the structures of both language groups as quite surprising, for the two groups diverge from each other in almost every linguistic subcomponent. Phonologically, rGyarong has a much richer system of segmental contrasts. In contradistinction to the situation in Tani, aspiration is phonemic in rGyarong stops/affricates. Moreover, while Tani has only one (palatal) series of affricates, rGyarong distinguishes as many as four (dental, retroflexed, alveopalatal, and palatal). Although consonant clusters are not unknown in Tani (especially Western Tani), they cannot begin to compare in number and variety with the impressive array of consonant clusters found in rGyarong. The differences in morphosyntax are even more fundamental. Although both languages utilize considerable affixation, rGyarong is predominantly *prefixing* while the Tani languages are mainly *suffixing*. In terms of function, rGyarong boasts highly complex derivational as well as inflectional morphology, in contrast to Tani where morphological processes are much less abundant. Furthermore. rGyarong is an ergative language[20] with many head-marking features (Nichols 1986), including a system of verb agreement which indexes not only person and number, but also direction (or person hierarchy, i.e. direct vs. inverse) of the discourse participants. All Tani languages, on the other hand, display the so-called 'anti-ergative' pattern (LaPolla 1992), where agents

are generally not case-marked while a single 'object' case marks a number of semantic roles, including patients, recipients, beneficiaries, and even temporals.[21] The two languages also employ distinct verb-phrase structures. In Tani, various complements and modifiers of the verb, along with such other categories as tense, aspect, polarity, and modality, are generally expressed by a large set of postposed 'verbal particles'. This characteristic is so important in Tani that it may not be too wide of the mark to say that the study of the Tani verb phrase is largely the analysis of such verb particles. No comparable phenomenon obtains in rGyarong, where many of these categories are conveyed by verbal prefixes instead. This, in short, leaves the lexicon as the only likely linguistic sub-system in which possible close genetic ties between rGyarong and Tani can be sought.

In order to assess the assertion that rGyarong is closely affiliated with Tani in its deepest lexical core, I have examined a total of 383 basic adjectives (stative verbs) and verbs listed in ZMYYC, yielding the following comparable pairs between rGyarong (i.e. Proto-rGyarong as proposed in Nagano 1984)[22] and Proto-Tani in these two basic semantic areas (states and actions):

Table 1 Comparison of selected basic verbs in Tani and rGyarong

Gloss	*Proto-Tani*	*rGyarong*
'big'	*tə-*ta	*k-Te
'come'	*puŋ ('arrive')	*bo
'cover'	*kap	*p-Kap
'die'	*si	*syi
'dig'	*du	*duw
'dream'	*maŋ	*r-mo
'eat'	*do	*za
'exist'[23]	*duŋ	ndo
'heavy'	*ji	*li
'itch'[24]	*fak	*ya
'lean (against)'	*grəŋ	kə-nə-ŋ*grə*
'melt, thaw'	*jit	kə-*ndzi*
'ripe, cooked'[25]	*min	*ə-min
'run'	*duk-ɟuk	kə-rjjək[26]
'smell'	*nam	*nam[27]
'stand'	*rop[28]	*ro[29]
'sweet'	*ti:	*ci
'thin (of people)'	(Bokar OY gi)	kə-nə-*khi*
'vomit'	*b(r)as	kə-mə-*mphət*
'wait'	*jaŋ	ka-na-*jo*
'weep'	*krap	ka-ma-*kru*

That is, out of the 383 sets compared, only twenty-one pairs (or about 5%) show enough resemblance to be considered *probable* cognates. Furthermore, rather than revealing uniquely shared rGyarong-Tani lexical relations, the majority of such pairs (e.g. 'die', 'dig', 'eat', 'heavy', 'smell', 'ripe', 'stand', 'vomit', 'weep') involve roots widely attested in the Tibeto-Burman family.

To assess further the lexical relations between rGyarong and Tani *vis-à-vis* other Tibeto-Burman members, I conducted another sample comparison including Tibetan and Burmese, two other languages showing considerable affinity to rGyarong. The items utilized for ths pilot study are narrowed down to the seventeen verbs from the Swadesh 100 core vocabulary list:[30]

Table 2 Comparison of selected Tani verb roots with rGyarong, Tibetan, and Burmese

GLOSS	*Proto-Tani*	*rGyarong*	*Written Tibetan*	*Written Burmese*
'drink'	***tɯŋ**	*mot	'**thung**	sok
'eat'	*do	***za**	**za**	**sâ**
'bite'	*g(j)am	kha mtʃĭk khə-lət	so brgyab	kuik
'see'	*kaŋ-paŋ	mto	mthong	mrang
'hear'[31]	*tas-paŋ	*r~na	thos; rna-ba 'ear'	krâ; na
'know'[32]	***ken**	*sye	shes; **mkhyen** [hon.]	si'
'sleep'[33]	***jup**	*r-ma	nyal; gnyid	**ip**
'die'	***si**	***syi**	**si;** 'chi	**se**
'kill'	*man	*sat	gsod	phyak; sat
'swim'	*bjaŋ	*pjaw	rkyal; 'phyo	po
'fly v.'	*bjar	*N-pjam	'phur	pyaṁ
'walk'	*in	ptşe	'gro	hlyok; hrok
'come'[34]	***(ɦ)aŋ**	*bo	**yong~'ong;** 'byon	la; waŋ
'sit'	*duŋ	ni[35]	'dug; snye(s) 'recline, lean against' (?)	thuiŋ
'stand'	*dak; ***rop**	***ro**	lang; 'greng	**rap**
'give'	*bi	dit; **wə**	sprad; sbyin	**pê**
'say'	*lu; *ban	ka-rjo	bshad; smra	prô

Table 2 yields the following pairwise cognate numbers: Tani-rGyarong 4/17, Tani-Tibetan 8/17, Tani-Burmese 7/17; rGyarong-Tibetan 8/17–10/17;[36] and rGyarong-Burmese 8/17.[37] It is important to note that rGyarong has twice as many cognates with Tibetan and Burmese than with Tani, and that the rGyarong-Tani pair shows the *lowest* cognate count among all five pairs. To the extent that cognate counts derived from such a limited sample can be suggestive of the *relative* strength of lexical ties among the languages compared, rGyarong appears to be much more closely related in basic vocabulary to Tibetan and Burmese[38] than to Tani. This fact, coupled with the striking structural differences between the two Tibeto-Burman groups, makes their intimate genetic connection highly improbable.

2.3. Dhimalish

Dhimal (in Darjeeling and the Jalpaiguri area of Sikkim and eastern Terai, Nepal), and the closely related Ṭoṭo (south of the borderline between Bhutan and West Bengal) are two small languages comprising the obscure Dhimalish section of

Shafer 1955:102. The only documentation on these languages available to us are Hodgson 1847 for Dhimal and Sanyal 1955 for Ṭoṭo. The association of this group to Tani is vaguely suggested by Benedict in STC, and we quote: "Abor-Miri and Dafla make up the nucleus of the 'North-Assam' group of Konow and the Linguistic Survey of India. Aka (or Hruso) has the most points of contact with this nucleus, and *Dhimal (in Sikkim) the fewest*" (p. 6). From this statement alone it is not certain whether Benedict refers to a contact or genetic relationship. However, on the previous page (p. 7), he does consider Dhimal to be a likely addition to the Abor-Miri-Dafla (Mirish) nucleus.

A revisit to the Dhimalish sources, however, has failed to reveal too many significant points of agreement between Tani and Dhimalish. The following test comparisons, utilizing again the seventeen basic verbs from the Swadesh 100-word list, should be suggestive of the genetic distance between the two groups:[39]

Table 3 Comparison of selected basic verbs in Tani and Dhimalish

GLOSS	*Proto-Tani*	*Dhimal*	*Ṭoṭo*
'drink'	*tɯŋ	**ám**	āng
'eat'	***do**	**chá**	**cā**
'bite'	*g(j)**am**	---	**cā**-pir
'see'	****kaŋ***-paŋ	dó; **khang**	**kāng;** ting
'hear'	**tas*-paŋ	hén	hing
'know'	*ken	gé	gē
'sleep'	*jup	jim	jing-ju; jin
'die'	***si**	**sí**	**shi-pu**
'kill'	*man	shé	pāi
'swim'	*bjaŋ	nó-i	---
'fly v.'[40]	*bjar	*bhír*	*bi* -u
'walk'	*in	hi-gil	tē
'come'	*(ɦi)aŋ	lé	lē
'sit'	*duŋ	*yong*	*i-ung*
'stand'	*dak; *rop	jáp	lŏ-o; *lo* -
'give'	*bi	pí	pi
'say'	*lu; *ban	dóp	jāng

The Dhimal and Ṭoṭo words for 'eat', 'die', 'give' and 'look' are undoubtedly cognate with the PT roots. The cognacy of the Ṭoṭo form for 'stand', and the Dhimalish words for 'fly v.' and 'sit' (italicized in the table) to the corresponding PT roots are uncertain. Everything considered, we get at most 7 cognates out of 16 pairs compared, which is equivalent to the cognate figure between Tani and Burmese obtained by using the same test sample. The set for 'look/see' (PT *kaŋ. Dhimal khang. Ṭoṭo kāng) may appear to be a striking parallel between the two groups; yet, this root occurs also in many *Kiranti* languages, e.g. Bahing koŋ 'look, watch': Chamling. Bantawa khaŋ 'look, see'. Newari khan- 'see'. On the other hand. Dhimalish seems to exhibit many more lexical links with Kuki-Chin, and especially with Tibetan, as pointed out in Shafer 1950:207.

At any rate, the similarities between Tani and Dhimalish are far from numerous,[41] otherwise they would not have escaped the attention of both Konow and Shafer. It seems, therefore, futile to search for deep connections between Tani and Dhimalish, although more extensive inquiry (and with much better Dhimalish data) needs to be done to properly assess the 'points of contact' between the two groups which prompted Benedict to place them in the same subgroup.

2.4. Hrusish

The obscure Hrusish branch is named after its best-known representative, the Hruso (paleo-exonym Aka) language of West Kameng, Arunachal Pradesh. The remarkable linguistic divergence of Hruso from neighboring Tibeto-Burman languages was already noted by Konow (1909b). Shafer 1947 compares various early wordlists of 'Aka' and concludes that actually two very distinct 'dialects' of Hruso can be established: Dialect A and Dialect B. To Dialect B, or *Hruso proper,* belong most early records of 'Aka'. Shafer's Dialect A of Aka is actually a distinct language, represented only by Campbell's (1874) variety of 'Aka'. We have recently made the discovery that Shafer's 'Dialect A of Hruso' is none other than the language of the *Dhammai* (exonym: *Miji*) tribe distributed to the north of the Hruso country. For this important language, which is more conservative than Hruso proper, we are now able to consult Simon 1979, a far more ample source than any available to Shafer. There is at least one more Hrusish language in Arunachal Pradesh, namely the language of the *Bangru* tribe of North-western Upper Subansiri district.[42] Publications on the Bangru language are completely non-existent. Our limited fieldwork data on Bangru[43] reveals such striking resemblances between Bangru and Dhammai that they may even turn out to be dialects of the same language.

The lexical similarities between the Hrusish languages and Tani (especially Western Tani) are indeed notable and deserve to be carefully investigated.

2.5. Languages of the 'Mishmi' tribes

Comparable to Hrusish languages of the west, the languages spoken by the Mishmi tribes are the most important linguistic neighbors of Tani in the east. Unlike Tani or Hrusish, however, these languages by no means form a coherent unit. Instead, they fall into two distinct groups, Taraon-Idu (Shafer's *Digarish*) and Kaman (Shafer's *Midźuish*). Sun et al. 1980: 299–315, to date the only comparative study of the Mishmi languages based on accurate first-hand data, turns up remarkable differences. Of the 2477 native lexical items compared, 2089 or 84.4% are non-cognate, including quite a few core Tibeto-Burman items such as 'man (homo)', 'snake', 'sit', 'hand', 'hair', 'weep', 'know', 'buy', 'tooth', 'hear', 'rain', and 'house'. The morpho-syntactic disparity between the two groups is also considerable. For example, Kaman has pronominal verb agreement while Taraon and Idu do not: moreover, Kaman sometimes uses prefixes (e.g. taŋ$^{55-}$ 'nominalizer',

mai[55]-/mɯ[31]- 'negator', ai[53]- 'prohibitive marker') while Taraon and Idu, like Tani, always use suffixes (e.g. Taraon - ja[31] 'nominalizer', - jim[55] 'negator', - ja[53] 'prohibitive marker'). These languages, therefore, do not appear to be as intimately related to each other, contrary to what Thurgood 1985 suggests. Thus, before we even begin to compare them further with Tani (or with any other language), we must bear in mind that the alleged unity of the Mishmi languages is still an unproven hypothesis.

As stated above, most Tibeto-Burman classifications place the Mishmi languages close to the Tani nucleus. Indeed, even a cursory glance at the data shows considerable parallels between Tani and these languages (in particular Taraon and Idu), calling for more detailed exploration.

In summary, after inspecting a few alleged close relatives of Tani, we have decided to screen out rGyarong and Dhimal as improbable candidates. In the following section, the remaining languages will be further assessed by means of a more detailed lexical test.

3. Tani's next of kin: a further search

3.1. Methodological perspectives

Much doubt has been cast on the validity of lexicostatistics in historical linguistic research; Matisoff 1978a:1.14 outlines the hazards of a particular application of this method, namely the use of cognate counts in setting up subgroups among related languages.[44] However, the following statement seems quite reasonable (Thomas and Headley 1970:411, emphasis ours):

> Lexicostatistics is not a precision tool. Careful phonological reconstruction is necessary if one desires detailed information about language relationships. *Lexicostatistics is useful, however, for giving a quick general picture of language groupings.*

In fact, the authors of the preceding quote claimed that the results of their lexicostatistic analysis of Mon-Khmer internal relations can be 'presented with the confidence that the general outlines will still be standing after detailed phonological reconstruction has been done' (Thomas and Headley op. cit.). The ensuing two decades have seen considerable advances in comparative Mon-Khmer and phonological reconstruction of many Mon-Khmer subgroups (Monic, Waic, Aslian, etc.); indeed, the Thomas-Headley subgrouping framework turns out to have stood the test of time, judging by a recent authoritative statement on Mon-Khmer subclassification (Diffloth and Zide 1991).[45] Consider also the small-scale lexicostatistic study presented in Benedict 1976, where Tibetan, Burmese, Kachin, Garo, Lushai, and Pwo Karen were compared with Mandarin Chinese in terms of the Swadesh 100-word list, with the primary purpose of testing the solidarity

of the Tibeto-Burman grouping vis-à-vis Chinese and Karen. It is on the basis of this analysis that Benedict proposes the '*basic cleavage line*' in Tibeto-Burman between the Baric-Jingpo supergroup and practically all other TB groups. This hypothesis has been corroborated by a follow-up comparative study of Northern Naga (i.e. Benedict's Konyak group), leading the author to conclude with confidence that the validity of the Bodo-Garo-Northern Naga-Jingpo supergroup 'should no longer be in doubt' (French 1983:727). A key factor behind these two useful (in the sense of producing new and viable ideas, inspiring further research, and contributing eventually to growing consensus) applications of lexicostatistics is that the investigators are all specialists in the respective language families, which means that the risk of cognate misidentification was minimized, and sensible adjustments in the Swadesh wordlist could be made to fit the particular target language families. Therefore, lexicostatistical methods, if applied with due caution and without extravagant claims,[46] may still serve as *subsidiary* tools for detecting probable subgrouping patterns.

Although the non-existence of genetic relations between languages is unverifiable in principle, it is possible to ascertain whether any two given members in a group of related languages share a *particularly close* relationship. However, this cannot be done simply by listing random similarities, because alternative explanations (borrowing, areal features, shared substratum, common retention, etc.) are not ruled out. Even if regular sound correspondences in the basic vocabulary are demonstrated, the special relation between the two languages remains unproven, for such equations can, by definition, be established between any two genetically related languages anyway.[47] What we need to do, obviously, is to single out *uniquely shared linguistic features* which set these languages apart from all others, enough to 'tip the scale against any contrary hypothesis which sets the relationship merely at the level of the underlying proto-language' (Bauman 1976:26). However, sorting out the linguistic relations between Tani and its possible next of kin in Tibeto-Burman poses a currently insurmountable problem; the study of the Tibeto-Burman languages of Arunachal Pradesh and the immediate environs, among which the close relatives of Tani are most likely to be found, is still in its infancy, and we simply do not have the amount of linguistic information required for such detailed comparative analysis. What we can do at the present stage is no more than offer a *process of elimination,* which narrows down potentially promising avenues for further research.

3.2. A lexicostatistic test

A lexicostatistic study has been conducted (see the comparative table in the Appendix below) with the aim of assessing degrees of lexical affinities between Tani and four possible close relatives surviving the preliminary screening of the previous section: Taraon, Kaman,[48] Lepcha,[49] and Dhammai.[50] Written Tibetan, Written Burmese, and Garo, which have never been suspected of being

intimately related to Tani, are added as control languages. The modest objective of this pilot study is to eliminate dubious candidates according to a simple and, we trust, reasonable principle: if a language is a true next of kin of Tani, then there should at the very least be a *significantly higher* percentage of shared core vocabulary between this language and Tani than that between Tani and languages from separate major divisions of Tibeto-Burman, in this case Written Tibetan (Bodish), Written Burmese (Lolo-Burmese), and Garo (Bodo-Garo). The test wordlist used in this study is based on the CALMSEA 200-word list[5] proposed in Matisoff 1978a: 284–96. For some CALMSEA glosses, however, no PT reconstructions are presently obtainable; either because extreme internal variation precludes positing uniform PT roots (e.g. 'descend', 'bamboo', 'sweat'), or Indic loanwords are suspected (e.g. 'needle', 'silver'), or simply because the gloss is not realized by distinct roots in most Tani languages (e.g. 'twenty'). In such cases (thirty-seven in total), CALMSEA glosses are replaced with the following items, mostly body part terms and common verbs: 'angry', 'borrow', 'call/cry', 'come', 'dead body', 'count', 'do', 'door', 'dry/wither', 'duck', 'exit', 'face', 'fireplace', 'float', 'flow', 'fly (insect)', 'gall', 'grandfather', 'grandmother', 'hungry', 'kidney', 'knee', 'language', 'melt', 'nest', 'placenta', 'rot', 'seed', 'shoulder', 'soul', 'suck', 'swallow (v.)', 'take', 'think', 'tired', 'tiger', and 'wet'. The resultant compromise list, we hope, contains few glosses that are arguably not part of the lexical core of the target languages. Our cognacy judgement[52] with respect to WT, WB, and Lepcha should be relatively uncontroversial, for much is known about the historical phonology of these languages, and expert guidance is readily available from STC and various other works on Sino-Tibetan reconstruction. The same can be said of Garo, the best known of all Baric languages, not only because it was one of the principal languages on which the PTB reconstructions in STC were based, but also thanks to a series of treatises on Baric contributed by Robbins Burling (especially 1959, 1983, and 1992).[53] Cognate detection involving the other target languages is much more difficult. In the case of Taraon and Kaman, although we are lucky to have access to mutually complementary Indian and Chinese sources (the accuracy of the latter is quite impeccable), the phonological developments of these languages, especially the less conservative Taraon language, are not yet well-known.[54] Dhammai is even more troublesome in terms of data reliability and cognate identification. Furthermore, thirty-three test items are missing from the word list in Simon 1979 (the only available substantial source on this important language), although it is not clear to what extent the incomplete data may cause the *averaged* cognate percentage to be skewed.

3.3. Results and discussion

Each of the languages compared contains a number of forms of indeterminate cognacy with the corresponding PT roots. Such is the case, for instance, between

PT *kɯ 'dove/pigeon' and WT 'ang-*gu* 'pigeon'.[55] A more conservative estimate may discount these doubtful cases, a bolder count would include them all, while the cognate figure closest to reality may lie somewhere in between. These two different figures, then, represent the *range* of possible cognation between the given language and PT. Since, for example, WT shows two doubtful cognates (the other being PT *be, WT spre 'monkey') and fifty-six sound ones, the cognate ratio between PT and WT ranges from 56/200 (or 28%, conservative estimate) to 58/200 (or 29%, less conservative estimate). The much larger percentage of such uncertainty for Taraon is a function of the phonological deviancy of the language. The output of this study can be summarized in the following table:

Table 4 Cognate figures between Tani and seven Tibeto-Burman languages

	WT	*Garo*	*WB*	*Taraon*	*Kaman*	*Dhammai*	*Lepcha*
available forms	200	194	200	200	200	167	200
cognate count	56–58	46–50	54–57	59–**76**	**43**–50	43–49	47–49
percentages	28–29	24–26	27–28.5	29.5–38	**21.5**–25	26–29	23.5–24.5
average percentage	28.5	25	28	**33.75**	**23.3**	27.5	24

The output obtained from this pilot study has a number of noteworthy implications for the phylogenetic position of Tani.

First, this lexicostatistic test has indeed accomplished its unpretentious mission of *separating off problematic candidates* from among the possible close relatives of Tani. The cognate figures of PT with both Lepcha and Kaman are *lower* than those between PT and the three control languages. In particular, the PT-Kaman cognate percentage is the lowest of all figures obtained. If core vocabulary is reliable at all as an index of relative genetic distance, then these facts should constitute strong disproof of any intimate relation between either of these languages and Tani. As for the lexical similarities between Lepcha and Tani observed by Bodman 1988, alternative explanations must be sought, such as shared substratum,[56] or early contact (in southern Tibet?) of the two language groups before their migration to the present locations. In short, our findings support Bodman's conclusion that although Adi may be among the TB languages which are more similar in lexicon to Lepcha,[57] the relationship between them is not very close (Bodman op. cit.:4).

Compared with Lepcha and Kaman, Dhammai shares a higher cognate percentage with PT, yet, this figure is still lower than that between PT and WT. Although we are not well-informed enough about the linguistic structures of the Hrusish languages to say anything definite about the relation between

Hrusish and Tani, we do suspect that the similarities between them[58] may be the consequence of prolonged contact rather than exclusively shared linguistic history, and that the true roots of Hrusish may lie somewhere else in Tibeto-Burman.

Cognate percentages between PT and the three control languages run between 24 and 29. The close clustering of these figures indicates that Tani indeed forms a distinct division in Tibeto-Burman, coordinate with other major nuclei in the family. The lower Tani-Garo figure suggests that Tani is more akin to WB (Lolo-Burmese) and WT (Bodic) than to Garo (Baric), corroborating Benedict's inclusion of Miri on the non-Baric side of the 'basic cleavage line' in Tibeto-Burman. This also shows that subgrouping Tani under Baric (e.g. DeLancey 1991a) may not be advisable. Furthermore, Tani shares almost as many cognates with WB as with WT, a finding which is all the more remarkable since Lolo-Burmese and Tani (or for that matter any Arunachal Tibeto-Burman groups except perhaps Singpo) have never been known to be in close areal contact. This calls into question Egerod's decision to classify Tani directly under Tibetic (Egerod 1974).

The language that stands out with the highest cognate figure with Tani is Taraon (29.5%–37.5%, average **33.75%**). This figure, interestingly, is higher even than that of the Taraon-Kaman pair (30%–33%, average 31.5%).[59] The large gap between the more conservative (29.5%) vs. the bolder cognate estimate (37.5%), nevertheless, reflects our current inability to distinguish between true cognates, allofams, and chance look-alikes. However, as stated, we have made an attempt to uncover the elusive sound laws of this language, and our cognacy judgments, we contend, are at worst educated guesses rather than wild speculations.

4. More thoughts on the Tani-Digarish relationship

A major outcome of this study is that Digarish (Taraon and Idu) may be the Tibeto-Burman group most similar in lexicon to Tani. However, before jumping to the conclusion that Digarish and Tani are collateral relatives in Tibeto-Burman, we should be reminded that the fundamental research necessary to prove such an intimate connection has not been done, and alternative accounts of such lexical parallels cannot yet be ruled out. Since to adequately pursue this line of research would involve at least another dissertation-length study, we will have to content ourselves with suggesting a few interesting Tani-Taraon parallels in other linguistic subcomponents.

With regard to shared *peculiar* phonological innovations, the development of PTB *dz- to PT *d- is paralleled by Taraon th-; e.g. PTB *dza, PT *do, Taraon tha[53] 'eat'. Elsewhere in Tibeto-Burman, PTB *dz- usually either survives as

affricates (e.g. Mawo Qiang dzə; WB câ 'eat') or spirantized (e.g. WT za; Jingpo ja[55] 'eat').[60] Another possible example of common phonological aberrancy is the irregular *palatalized* initial in the following roots: PT *rjam, Taraon liɯŋ[53]-gie[31] < PTB *la(:)m 'fathom'; PT *rjum 'dusk/evening', Taraon liɯŋ[53] 'night', < PTB *rum ~ *rim 'dusk' (STC #401); PT *fia-; Taraon xa[31]nia[53]pɯm[55] < PTB *s-na 'nose' (STC #101).

The remarkable lexical affinities between Taraon and Tani are not restricted to content words. Some *grammatical* morphemes may also be cognate:

- 'comparative auxiliary' PT *jaŋ; Taraon joŋ[53] [61]
- 'imperative suffix' PT *to; Taraon tio[53]
- 'prohibitive suffix' PT *jo; Taraon ja[53] [62]
- 'experiential aspect marker' PT *kɯ; Taraon koŋ[35]

The morphosyntactic structures of the two groups have not been carefully explored, but some prima facie resemblances exist here as well. In both groups, pronominal verb agreement is lacking. Further, the predominant verbal morphology in both cases is suffixal. Digarish languages, like languages of the Tani group, also seem to exemplify the 'anti-ergative' case-marking type, where patient and recipient nominals receive *identical* marking while agents are seldom case-marked.

On the other hand, the differences between the two groups seem to overshadow their similarities. Apart from their overall lexical differences, many of the characteristic Tani lexical items and phonological developments (such as PTB *-a > PT *-o, and the shift of all PTB diphthongs into PT monophthongs) find no counterparts in Digarish. The overwhelming majority of grammatical morphemes in Tani and Digarish are also unrelated. From the few available syntactic descriptions, the two groups also show important disparities in morphosyntax. For instance, Digarish languages use separate existential verbs depending on the *animacy* of the subject, a distinction unattested in any known Tani languages. As stated, although some Tani languages do contrast different existential verbs, the relevant distinctions are usually *polarity* (e.g. Bengni S do: 'exist/have'; ka:-ma: 'not exist/have') or even *posture* (Apatani A da 'exist (referent standing)'; du 'exist (referent sitting)'; do 'exist (referent lying)') of the predicated nominal (Abraham 1985:70–3). Moreover, relative clauses in Taraon are formed simply by gapping, without first nominalizing the embedded clause as is usually the case in the Tani languages.[63]

In summary, even though Digarish and Tani bear some striking resemblances, their equally impressive differences make it doubtful that this relationship could be an intimate one, even if future studies could establish an exclusively shared genetic relationship between them.

Appendix[64] Comparison of 200-word core-vocabulary in eight Tibeto-Burman languages

Gloss	*PT*	*WT*	*Garo*	*WB*	*Taraon*	*Kaman*	*Dhammai*	*Lepcha*
alive	*tur	gson-pa	taŋ-	hraŋ	a^{31} suŋ^{55}	kɯ^{31} ɹǎŋ^{35}	śɯn	zu
angry[65]	*haŋ-fak	ʾkhro; ʾtshig; sdang	ka-o-naŋ	cit-chûi; mjak	khɯm^{55} mioŋ^{55}	sun^{55} dɯt^{55} lat^{55}	nen	a-mlem nók non; li; sak lyak
ant	***ruk** ~ *rup	**grog**-ma	---	pu-**rwak**	kɯ^{31} juɯ^{53}	tɛu^{31} kɹik^{53}	---	tŭk-fyil
arrow	***puk**	mda'	**#bra**	hmrâ	***pu**^{55}	a^{31} wat^{55}	nu	tsóń
ascend	*čaŋ	ʾdzeg	ga-kat	tak	tu^{31} dzi^{35} noŋ^{55}	lɯn^{55} xai^{55}	khɯn?	***hrón***
awake (v.i.)	*fiut[2]	gnyid sad	#mik-rak	nûi	dzɯ^{55} a^{55}	kɹǎŋ^{55}	phrɯ-u	ší
banana	*ko-pek	skyes-sdong	te-rik	hŋak	pha^{31} dzi^{55}	xa^{31} biuŋ^{55}	ru-ðaŋ; ru-łaŋ	-blo
bear (n.)[66]	***tum**	**dom**	map-il	wak **waṁ**	ta^{31} **ṃm^{55}**	**kum^{55}**	šu-***tsaŋ***	să-na
belly	*kri	grod khog	ok	puik	kɯ^{31} jɯŋ^{55}	dăk^{53}	rug	(tă-) băk
bird	*taŋ	bya	do?	hŋak	pia^{55}	wa^{35}	bu-zu(?)	to
bite	*gam~ *gjam	so brgyab	cik	kuik	tie^{55}	săk^{55}	tha?; šu- wrai?	tsuk; ran
bitter[67]	***ka**:-	**kha**	**ka**	**kha**'	khlai^{55}	**kha^{53}**	mu-khu?	kri
blood	***vi**:	khrag	a?n-ci	**swô**	xa^{31} ɹuai^{53}	a^{31} ɹui^{35}	žai	vi (nyo)
blow	***nut**	ʾbud	spo-	**hmut**	mɯŋ^{53}	thɯt^{55}	---	**măt ~ mŭt**
bone	*loŋ	rus-pa; gdung	greŋ	a' rûi	ɹɯ^{31} boŋ^{35}	ɕiŋ^{55} ɹak^{55}	(mɯ-) Ijaŋ	a-hrăt
borrow[68]	*nar	g yar; skyi; brnyan	ra?-cak	hŋâ; khyê	xa^{31} ŋa^{55}	a^{31} ŋat^{55}; łu^{53}	---	*nyo-lyă
bow (n.)	*rji	gzhu	*cri	lê	a^{31} lai^{53}	gaŋ^{35}	gɯ-ri?	să-li
brain	*pYk-ni	klad-pa	ta-niŋ	û-hnok	pɯ^{31} ŋɯm^{55}	nɯn^{53}	---	a-t'yak *yón; a-yăṅ*
branch	*ñak	yal-ga; gel-pa	*cek-si	a'-khak	xa^{31} ɹa^{53}	ŋkhăi^{35}	ou-dɯ-tsaŋ	a-kóŋ; a-nŭṅ
breath	*sak, ŋa	rngam	raŋ-sit	(ə-) sak	ɕɯ^{55}	ntshon^{35}	dɯ-thu	sóm
burn (v.i.)	*gu	ʾbar	kam	tok	xɹaɯ^{53}	gɹi^{35}; xɯ^{31} nai^{35}	phrjaŋ; rau?	mi dyak
buy	*rə	nyo	**bre**	way	**bɹai^{35}**	ɕip^{55}	phɯnə	par
call/cry	*grok	'grags-'grogs	o-kam; crik	hac; khaw	xa^{31} tiɯŋ^{55} a^{31}; #grã:	bɯu^{53}	then	lik
child/ son	*fio	bu	bi?-sa	sâ	a^{55} (ju^{55} a^{55})	sa^{55} wai^{53}	zu	a-kŭp
cloud	*muk ~ *mək	sprin-pa	a-ram	tim	a^{31}m^{55}	ka^{55} măi^{35}	mei-miw	-byoṅ
come	*vaŋ	ʾong	re?-ba?	la	χa^{31} naŋ^{55}	xu^{53}	dai	di; lat; t'i
count	*krɯ	brong	#chan	rañ'	ta^{31} tsai^{55}	xa^{31} tsɯt^{55}	---	frón
day[69]	*lo	nyi-ma	sal	rak	kɯ^{31}n^{53}	ŋin^{53}	wu	nyí
dead body	*si-***maŋ**	ro	**mang** gi-si	ə-lôŋ	thɯŋ^{55}	dzal^{35}	---	(a-) fŭṅ
die	***si**	ʾchi	**si**	**se**	**ɕi^{55}**	**si^{53}**	či	mak
dig	***du**[1]; *ko*[2]	**rko**[2]; ʾbru	co?[1]	tû[1]	ua^{55}	gua^{35}; ɔon^{35}	thau?[1]	du[1]; byol
do	*rjw[1]; mo[2]	byed; spyod	dak	lup; mu[2]	ba^{53}	pam^{35}	ru[1]	mat; zuk; fat
dog[70]	***ki**:	**khyi**	a-chak	**khwê**	**kuaɯ^{53}**	**kui^{55}**	ša-ži?	kă-ju (pă-li)
door	*rjap	sgo	do-ga	Taṁ-khă	ka^{31} lɯŋ^{35}	mphun^{53}	ban-phi?	(tŭṅ-) vyeṅ
dove	***kɯ**	ʾang-**gu**	do-kru	khui	pia^{53} kɹau^{35}	tɕi^{55} khiuŋ^{55}	bjun-lo	fă-wu-fo
dream	*jup-**maŋ**	rmi-lam; rmang	ju-**maŋ**	ip-mak	ja^{55} **mo^{53}**	ka^{31} muŋ^{35}	---	**móṅ**
drink	***tɯŋ**	ʾhung	riŋ	thok	tim^{35}	**tauŋ^{55}**	thuŋ	t'ăṅ~ t'óṅ, báp
dry/wither	***san**	skam-po	ra?n	khrok	***coŋ**^{35}*	sal^{53}	mɯ-khjaŋ	a-**šin**; a-**són**; a-jep
duck	***jap**	ngur-ba	do-***gep***	bhai	ma^{31} tɕi^{53} pia^{53}	kɹai^{35} pit^{55}	ŋu-so	*dam-byó
ear[71]	***ña**-ruŋ	**rna**	**na**-cir	**nâ**-rwak	kɯ^{53} naŋ^{35}	iŋ^{55}	žo?	a-nyor
eat	***do**	**za**	**ca?**	**câ**	**tha^{53}**	**tɕa^{53}; ɕa^{53}**	**tsu?**	zo; wam-mat; t'a
egg	*pɯ	sgo-nga	bit-ci; do?-ci	u'	ma^{31} na^{53}	kɹai^{55} sit^{55}	du-rin?	a-ti
eight	*pri-ñi	brgyad	cet	hrac	liɯm^{35}	i^{55} lioŋ^{35}	sɯ-gi?	kă-kŭ
excrement	*e:	skyag-pa	ki	khyê	klai^{53}	tɯ^{31} khɯi^{53}	---	ʾayít; it; e

(continued)

Appendix (continued)

Gloss	*PT*	*WT*	*Garo*	*WB*	*Taraon*	*Kaman*	*Dhammai*	*Lepcha*
exit	***len***	thon; ʾbyung	---	thwak	**leŋ**^{35} bi^{31}	xa^{55} tha^{35}	---	záń
extinguished[72]	***mit***	shi	#ki-**mit**	se	xa^{31} **mɯn**^{53}	**măn**^{53}; **#mɯt**	---	mí mak
eye	***mik***	**mig**	**mik**-ron	**myak**-seʾ	bɯ^{31} lɯm^{55}	min^{55}; **#mik**	miʔ	a-mik
face	*mik-mo:	gdong; ngo; bzhin	mik-kaŋ	myak-hna	ŋaŋ^{55}	a^{31} gul^{35}	gɯ-mjaʔ	a-mlem
fall (from a height)	*ho	lthung	gak-on	kyaʾ	blai^{55} dau^{55}; #ga-lja:	mit^{55} ti^{35} sau^{55}; #bral	dɯ-jɯn	hlat; glo; klo
far	***do***	rgyang-ring-po	ceʔl	wê	**dia**^{55}	klam^{55}	mɯ-rɯnʔ	[rŭ]
fat/stout	***ǰɯŋ***	rgyags- pa; tsho- ba	mil	waʾ	**diɯŋ**^{53}	kɯ^{31} diaŋ^{55}	zaʔ-mɯ-do	[šu]; a- tʿor ~ a- tʿyor; [ńur]
fat (n.)	***fu***	snum-pa	mit-dim	chi	ta^{31} **so**^{53}	ta^{31} **si**^{55}	thai-bau	a-šut < [**šu**]
fear	*bY-so: -*pY-30:	ʾjigs; zhed; dngang	ken-	krok; khrok	ɹai^{55}	ta^{31} si^{55} tɯp^{53}	(mɯ-) rin	[ro]
finger[73]	*ke(ŋ)	mdzub-mo	jak-si	lak-hñûi	a^{31} bɹɯŋ^{55}	duŋ^{35}	gi-tsoʔ	ká-jóm
fire	***mə***	**me**	waʔl	**mi**	na^{31}**mɯn**^{55}	**măi**^{53}	**maiʔ**	**mí**
fireplace	*ram ~*rom	(me-) thab	cu-dap	mi-lâŋ-phui	sai^{53} gɹoŋ^{53} gɹoŋ^{55}	sai^{55} gɹoŋ^{55}	loʔ	[kom]; [dap; dop]
fish	***ŋo***	**nya**	**na**ʔ-tok	**ŋâ**	ta^{31} ŋaŋ^{53}	a^{31} **ŋa**^{55}	thui; tčui	ŋo
five	***ŋo***	lnga	boŋ-a	**ŋâ**	ma^{31} **ŋa**^{35}	kɯ^{31} len^{55}	bu-ŋu	tă-ŋo
flee	***kat***^{1}	ʾbros	#ke-ne **kat**	thwak- prê; hroŋ	lɯi^{53}	lun^{55}	---	tor; tet
float	*bjaŋ	lding	#git-cho; bal-bo	po	ɹau^{55} a^{31}	jau^{53}	---	plyuń
flow	*bɯt	ʾbab; rgyug	#jo-kang; so-ol-ang	ci	#blum	#tai	---	dáń; nóń; yŭ

flower	***puŋ ~ pun***	me-tog	bi-bal	ə-**pwaŋ**ʾ	ta^{31} pɯ^{55}	**phan**^{53}	ou-boʔ	rip; [bor]
fly (n.)	*jiŋ	sbrang-bu	tam-pi	yaŋ; phrut	ta^{31} laiu^{35}	giul^{35}	bɯ-luŋʔ	sŭm-bryoń
fly (v.)	***bjar***	ʾphur	**bil**	Pyaṁ	jim^{35}	phiuŋ^{55}	gɯ-nui	lám
foot	***lə***	rkang-pa	jaʔ	khre	gɹoŋ^{53}	***pla***^{55}	**lai**	(a-) tʿoń; (a-) dyaṅ
forget	*mit-**pan**	rjed	gu-al	meʾ	we^{55} ma^{31} sa^{53}	a^{31} mlaŋ^{55}	thlaŋ	hryu; plón; myón; **pán**
four	***pri***	**bzhi**	**bri**	**lê**	ka^{31} **pɹai**^{55}	kɯ^{31} bɹɯn^{53}	b(ɯ)-li	**fă-li**
fowl	*rok	bya-de	#do-o-rang	krak	tiu^{53}	kɹai^{55}	du-zu	hik (-kŭp)
frog	*tɯk	sbal-ba	#beng-bek	phâ	pa^{31} ɹai^{53}	kaŋ^{55} khɹik^{55}	---	tă-lŭk
fruit	***ze;** *pɯ	shing- tog; ʾbras-bu	bi-**te**	ə-**si**	ta^{31} **ɕi**^{53}	sit^{53}	ou-then	[pót]
full	***brwŋ***	gang	gap	**prañ**ʾ	**blɯŋ**^{55}	**phlăŋ**^{55}	---	**a-blyăń**
gall	*pw	mkhris-pa	kaʔ-kit	sâñ-khre	thw^{31}- mwn^{55}	măn^{53}	---	*kʿi-bo
give	***bi**	ster; skur; sbyin	oʔn	**pê**	ŋŋ^{35}	**pi**^{55}	**bi(ʔ)**	**byi** (n); **bi**; bo
grandfather	***to**	mes-po	a-cu	ə-phûi; ə-bhûi	a^{31} **tia**^{55}	kuŋ^{35}	a-luw	tʿi-kuń
grandmother	***jo**	phyi-mo; ma-mo	am-bi	phwâ; ə-bhwâ	a^{31} **ja**^{55}	măi^{31} ŋu^{35}	a-žui	nyi-kuń; nyo-kuń
guts[74]	***kri**	rgyu-ma	bi-bik	u	**kɯ**^{31} **łai**^{55}	**xa**^{31} **lăi**^{35}	luŋ	tă-kli
hair (on body)[75]	*mɯt	spu	kin-i; kim-ir	ə-mwê	ɱ^{55}	bɯl^{35}	phiw	myal
hand/arm[76]	***lak**	**lag-pa**	***jak***	**lak**	a^{31} tio^{55}	***ɹau***^{53}	gi	kă; ká
have/exist[77]	***duŋ**	yod; ʾdug	**doŋ**	hrì	i^{55}; aŋ^{55}	tɕau^{53}; kam^{35}	***du***	nyi
head	***dum; *tuk**	mgo	**sko**	khôŋ	kɹu^{53}	kɯu^{53}	u	a-tʿyak; **tok**
heart (organ)	*puk	snying	Kaʔ-toŋ	hnaʾ-lûṁ	xa^{31} po^{55} tiai^{53}	lɯm^{35}	luŋ	a-lŭt
heavy	*fiit	ljid-po	#jrim	lê	waɯ^{55} a^{55}	ka^{31} lăŋ^{35}	mɯ-ljiʔ	lí; bryón-nă; glám-lă

(*continued*)

Appendix (continued)

Gloss	*PT*	*WT*	*Garo*	*WB*	*Taraon*	*Kaman*	*Dhammai*	*Lepcha*
horn	***rəŋ**	rwa	**groŋ**	khyui	ɹau^{55}	kɹăŋ^{35}	šu-žuŋ	(a-) **róṅ**
horse	*kɯ	rta	#gu-re	mrâŋ	ma^{31} ɹoŋ^{55}	pa^{31} xoŋ^{35}	šu-gro	on; *ta
hundred	***lɯŋ**	brgya	rit-ca	ra	ma^{31} ***lɯŋ**^{55}*	wa^{31} je^{53} mu^{53}	bɯ-loŋ	kʿa fă-ṅo
hungry	*kY-noŋ	ltogs; bkren	#ok-kri a-ni	cha; mwat; ŋat	na^{31} tioŋ^{53}	di^{31} iŋ^{53}	ten-či	krit
I	***ŋo**	**nga**	**aŋ**	ŋa	***xaŋ**^{35}*	ki^{53}	***ñaŋ***	kă-do; go
ill	*ki	na(d)	sa; jom	na	ŋaŋ^{35}	nat^{55}	no	dăk
insect	***pum**	ʾbu	joʔŋ	pô	ta^{31} **pɯm**^{55}	klauŋ^{55}	bi-luŋʔ	[bik]
iron	*rjok	lcags	sil	saṁ	sai^{53}	tɯ^{31} gli^{53}	sen	pŭn-jen; lăṅ-să a-lŭt
itch[78]	*fak	ʾbun; za	#mi-to; ka-kit	yâ	ma^{31} əo^{53}	phuŋ^{53}	gu-dzu	jak
kidney	*krat-pjɯl	mkhal-ma	#gi-la; ko-rong-te	kjok-kap	ɕi^{55}	ṅtɕhi^{53}	mɯ-gɯ-bauʔ	*kʿa-dok
kill	*man	gsod	soʔt	phyak; sat	se^{55}	sat^{55}	wai	sót
knee	*lə-bɯŋ	pus-mo	jaʔ-sku	dû	pha^{31} bɯm^{55}	pa^{31} pau^{35}	lai gɯ-phiw	tŭk-păt
knife	*rjok	gri	a-te	thâ	ta^{31} ɹa^{55}	sot^{35}; kɹa^{35}	vai-	ban
know[79]	***ken**	shes; **mkhyen** [hon.]	u-i	si'	ka^{31} sa^{53}	ŋit^{35}	ñi; zɯ-u	tʿyak; yă
language	*gom	skad	ku-sik	bha-sa	khi^{55} tɯ^{31} kɯ^{55}	khi^{55} lai^{55}	lau	a-riṅ
laugh[80]	*ŋil	dgod	ka-diŋ	râi	ma^{31} ɹa^{55}	kɹit^{55}	tho	tʿyăn; sak prok; zól
leaf	*nə	lo-ma	bi-jak	a'-rwak	naŋ^{35}	lap^{53}	ou-leʔ	lóp; a-nyóm
leech (land)	*pat^{1}	pad-pa	ru-at	ɦmyo'	ka^{31} pe^{53}	tɯ^{31} wat^{53}	dɯ-veʔ	-fót; šŭm-pat
left-side	*lak-ke	g.yon	jak-a-si	lak-wâi	tɯ^{31} ***kim**^{55}*	kɯ^{31} wai^{53}	su-vjoʔ	vim
lick	***rjak**	**ldag**	#cha-**srak**	**yak**	lio^{53}	lo^{53}	---	*lók
liquor	*poŋ	chang	cu	se	ju^{53}	si^{53}	čaŋ	či
listen/hear[81]	***tas**; *tas-paŋ*	nyan; thos	kin-a-	nâ-thoŋ; krâ	tha^{31} ɹɯŋ^{55}; tha^{31} tiɯŋ^{53}	ta^{55} giat^{55}; **tat**^{55}	rui	tʿyo
liver	***zin**	mchin-pa	bi-ka	a'-**sâñ**	ɹu^{55} xa^{31} tiai^{53}	blai^{31} blai^{33}	mɯ-thɯn	a-byet
look/see[82]	*kaŋ; *kaŋ-paŋ	lta; mthong; rig	ni-; nik-	krañ'; mraŋ	***ruey**^{53}*; ka^{31} tiɯŋ^{53}	thoŋ^{55}; ŋuŋ^{55}	waŋ	ṅak, ší; hyón
louse (head)	***fuk**	shig	tik	sân	***tshau**^{53}*	săl^{53}	fiʔ	*šák
man (homo)	***mi**	**mi**	man-de	lu	**me**^{35}	tsoŋ^{35}	ñiʔ	mă-ró
marrow[83]	*loŋ-kin	rkang; ngo-bo-nyid	#gheu	khraŋ-chi	ɹu^{53} su^{53}	xiŋ^{53}	---	yăṅ; sŭṅ-dăk
meat	*dɯn	sha	beʔn	(ə)-sâ	ta^{31} bɹeŋ^{53}	ɕin^{53}	šu-čuŋ	a-mán
melt	*jit ~ *jet	bzhu	#jron-gat	pyo	ji^{53}	jau^{55}; kɹăl^{55}	---	*jŭ; *šŭ
monkey[84]	***be**:	spra; ***spre*** (ʾu)	---	myok	ta^{31} min^{53}	a^{31}muŋ^{35}	šu-***bo***	să-hŭ
moon	*po-**lo**	**zla**-ba	***ja***-joŋ	**la**'	xa^{55} lo^{55}	**lai**^{53}	lu	**lă**-vo
mortar	*par	sgog-ting	caʔ-am	Chuṁ	loŋ^{35}	gloŋ^{35}	dɯ-lo	[tsam]
mountain	*di	ri	aʔ-bri	toŋ	thɯi^{55} ja^{55}	a^{31} dzau^{35}	phuŋ-	hlo; rók
mouth[85]	*nap-paŋ; gam	kha	ku-sik	pâ-cap; mê-ce'	thɯ^{31} ɹwm^{53} bum^{35}	ŋtɕhɯu^{53}	go	a-boṅ
nail	***zin**	**sen-mo**	#jak-skil	lak-**sâñ**	**a**^{31} łun^{55}	ɹău^{53} dzit^{55}	gi-thɯn	pŭn-ći
name	***mun**	**ming**	bi-muŋ	na-mañ	**a**^{31} **muŋ**^{55}	a^{31} măŋ^{55}	**min**ʔ	a-bryaṅ
neck[86]	*lɯŋ	ske; mgul; ʾjing-pa	git-dok	lañ-pâŋ	pa^{31} ṅŋ^{55}	xuŋ^{55}	---	[tok]; [liṅ]
nest	*sup	tshang	bi-tip	suik	a^{31} jɯ^{55} #pja:-sag	mphău^{53}; #ŏ-wa sap	---	-šap

(continued)

Appendix (continued)

Gloss	*PT*	*WT*	*Garo*	*WB*	*Taraon*	*Kaman*	*Dhammai*	*Lepcha*
night	***jo**	nam; mtshan-mo	wal	na'; nañ'	kɯ^{31} **ja^{55}**	ŋal^{53}	ǰaŋ-gǫu	[nap]
nine	*kjo-naŋ	dgu	sik-u	kûi	ka^{31} nɯŋ^{55}	nan^{55} mu^{53}	sɯ-thɯn	kă-kyót
nose	*ña-pum; ña-buŋ	sna	giŋ-tiŋ	hna-khôŋ	xa^{31} nia^{53} pɯm^{55}	min^{55} nioŋ^{35}	ñi	[nóm]
old (of things)	*ku~ *kju	rnying-pa	git-cam	hôŋ	me^{53}	tauŋ^{35}	mɯ-šwo	[ṅo]; sŭk-kyor
one	*kon	gcig	sa	tac	khɯn^{55}	kɯ^{31} mu^{53}	uŋ	kat
otter	***ram**	**sram**	mat-**tram**	phyaṁ	xa^{31} ɹɯŋ^{35}	**ɹam^{35}**	---	să-**ryóm**
palm	*lak-pro	lag-mthil; thal-mo	jak-pa	wâ	#a:-tjo-ka:	#rok ta-pa	gi dɯ-luŋ	[lyók]
penis	*mrak	mje	---	li	#mlõ	#jaŋ	---	t'ik
pig	*rjek	phag-pa	wak	wak	bɯ^{31} **liai^{55}**	**li^{55}**	žo	món
placenta[87]	*mam	sha-ma	---	ə-khyâŋ	a^{55} po^{55}	sa^{55} sap^{55}	---	kap-p-ŭṅ; 'ayeṅ-čót (-tyól)
poison[88]	*dɯk; *mro	dug	*bi-si	ə-chip	**thai^{53}**	**tau^{53}**	nɯ-phaŋ	[bo]; a-nyiṅ
put	*pa	'jog	don-	thâ	xa^{31} go^{55}	kɹal^{55}	rou	dya; t'o
rain (n.)	*pV-doŋ; *mV-doŋ	char	mik-ka	mûi rwa	ka^{31} ɹa^{55}	a^{31} waŋ^{55}	phrjo	so
rat	*ko-buŋ	byi-ba; tsi-tsi	*me-se	krwak	ka^{31} tɐi^{55}	si^{55} nu^{53}	---	kă-lók
red	*lɯŋ	dmar-po	git-cak	ni	ɐi^{53}	kap^{31} sal^{35}	mu-tsu	a-hyir
rice[89]	***pim**	'bras-chan	mi	tha'-mâŋ	ta^{31} **peŋ^{35}**	ɐat^{53}	an tsa-vo	nŭm-or-mo
right-side	*lak-bruk	g.yas	jak-ra	ya	tɯ^{31} tɐa^{55}	kɯ^{31} jau^{53}	ši-dzin	gyóm
ripe	*min	smin-pa	min-	chim';**hmañ'**	#ha:-mɯŋ	#shu-mm	**min**	[kru]; a-**măn**
river	*si; *buŋ	chu	ci-bi-ma	mrac	tɯ^{31} lɯu^{35}	tɯ^{31} lo^{35}	vu-do	uṅ kyoṅ
road	***lam**	**lam**	**ra-ma**	**lám**	a^{31} lim^{55}	bloŋ^{35}; **lam^{55}**	**lem**-baŋ; hlen	**lóm**

root	*pɯr; *m(j)a	rtsa-ba; rtsad	ja?-dir	ə-mrac	xa^{31} ɹai^{55}	kɹa^{53}	-khrin	a-fja; a-băn; [sáṅ]
rot	*jaŋ	ruł	so-	pup	tshɯŋ^{55} xo^{31}	ɹam^{53}	---	byót
round (globular)	***lum**	ril-ba; zlum-po	ta?m-bi?	wûiŋ; **lûṁ**	geŋ^{55} weŋ^{55} da^{55}	ga^{55} waŋ^{55} na^{55}	mɯ-dɯ-riu	a-blam; a-pům
salt[90]	*lo	tshwa	ka-ri	châ	pla^{35}	tɯ^{31} min^{55}	lu	vóm
scratch	***fiok**	'phrug; phur	ku-**ak**	**yak;** phrok~phyok	***wa^{55}***	glua^{35}	gɯ-fjʔ	hut
seed	***li**	sa-bon; son	bit-cri	myûi-ce'	ta^{31} ***plai^{55}***	xa^{31} lɯi^{35}	thei-žo	li
sell	*pruk	'-tshong	pal	rôŋ	kha^{31} ji^{55}	xa^{35}	tsuŋ-ru	ŭl
seven	*kY-**nɯt**	bdun	sin-i	khu-nac	weŋ^{53}	***nɯn^{53}***	mja?	kă-kyăk
sew	*fiom	'tshem	#sik; ko	khyup	#ru	taŋ^{55} kɹap^{55}	bɯ-ča	hrap
sharp-edged	***rat**	rno	mat	thak	***ɹa^{55}***	***kɹat^{55}***	---	lăt-let
shoot[91]	*ap	'phen	go	pac	o^{53} ja^{31}	top^{55} kap^{35}	buw	óp
shoulder	*gor-	dpung-pa; phrag-pa	pak-re	pu'-khûṁ	khɯ^{55} liɯŋ^{53} pa^{35}	a^{31} pho^{55}	pa-stuŋ	tŭk-puń
shy	*han-ñiŋ	skyeng; khrel; 'dzem	#kat-ca	hrak	#ha:-la:g-a:	#i-juk-rai	dai	uk; a-mlem glo
sit	*duŋ	sdod; 'dug	a-soŋ	thuiŋ	di^{55}	lăp^{55}	ǰuŋ?	ṅan
six	*krə (ŋ)	drug	dok	khrok	ta^{31} xɹo^{53}	kɯ^{31} tam^{53}	re?	tă-răk
skin	*pin	(1)pags-pa; ko-ba	bi-gir	ə-re	ko^{55}	uŋ^{35}	phri?	a-kap; a-t'un; a-pi
sleep[91]	*jup	nyal; gnyid-log	tu-si	ip	n^{53}	ŋui^{55}	ǰi	mik krap
smell (v.)	***nam**	**snom**	---	**nâm;** hru	**nɯŋ^{35}**	ntshiŋ^{55}	**ñen**	**n(y)óm**
smoke (n.)	*mɯ-kɯ	du(d)-ba	wa?l-ku	mi-khûi	ma^{31} khɯu^{53}	ta^{31} khɯi^{53}	thuŋ	mí-kan
snake	***bɯ**	sbrul	cip-bu	mrwe	ta^{31} **bu^{55}**	ɹuɯl^{35}	nɯ-**buw**	**bŭ**

(*continued*)

Appendix (continued)

Gloss	*PT*	*WT*	*Garo*	*WB*	*Taraon*	*Kaman*	*Dhammai*	*Lepcha*
soft	*mjak	mnyen; 'jam; snyi	#nom; ri-nok	pyo'	ñim^{55} m̥^{55}	ka^{55} miŋ^{35}	mɯ-bɯ-lja?	nŭm
son-in-law	***mak-bo**	mag-pa	#ca-wa-ri	sâ-**mak**	kɯ^{31} mu^{53}	tsa^{53}	---	**myók**
soul/spirit	*ja-lo	nyam(s)	#jaŋ-gi sil-ci; gi-sik	lip-pra	ta^{31} gɹa^{35}	ka^{31} mau^{35}	---	a-pil; [jŭm]; hyit
sour	***kruŋ**	skyur	me-seŋ	khyañ	***xɹɯ**^{55}	săl^{55}	mɯ-čuŋ	a-čór; rók-nón
spittle[93]	*kjul	mchil-ma	ku-ci	tam-twê	khɯ^{31} lai^{35}	dʑăl^{35}	že?	dyuk
stand	*dak; ***rop**	'greng	ca-deŋ	**rap**	deŋ^{35}	loŋ^{53}	gjuŋ	diṅ
star	***kar**	**skar**-ma	a-ski	kray	kha^{31} dɯn^{55}	kɯ^{31} grun^{35}	do-tsuŋ	să-hór
steal	*pjoŋ	rku	ca-u	khui	a^{31} kau^{53}	kal^{55} xɯu^{53}	tsɯ-khɯ?	tŭk-mo mat
stone	***lɯŋ**	rdo	ro?ŋ-te	kyok < klok	***phlaŋ**^{35}	**lăuŋ**^{35}	gɯ-luŋ	lăṅ
suck	*bruŋ	'jibs	op	cut; cui'	du^{55}	jip^{55}; #thət	bɯ-nu	yup; háp
sun	***ñi**	**nyi**-ma	sal	**ne**	ɹɯn^{53}	min^{35}	jo; zu?	să-tsŭk
swallow (v.)	***met**	(khyur) **mid**	**#mi**-nok	myui	blai^{53}	biap^{53}	bu-ʃui	yop; hyuł; am- mat
sweet[94]	***ti**:	mngar~dngar	ci-	khyui	ɕau^{55}	*tim*^{35}	mɯ-jaŋ	a-klyam
swidden	*rɯk	zhing-ka	a-ba	lay	kha^{31} liau^{55}	a^{31} **kuŋ**^{55}	vaw	nyót
tall	***mjo-*me**	rnga-ma	ki?-**me**	a'-**mri**	lɯ^{31} **mɯn**^{55}	a^{31} măi^{55}	---	[ší]
take	***laŋ**	'khyer; len~ long	ra?-; rim	yu	ɕi^{35}	ta^{31} lat^{55}	lu?	**lyă**; le; **lyo**
ten	***rjɯŋ; *čam**	bcu	ci-kiŋ	ɔ-chai	xa^{31} **lɯŋ**^{55}	kiap^{55} mu^{53}	**lin**	kă-tí
thick (book)	*brɯŋ	mthug	rit-ca?-	thu	bi^{31} tɕoŋ^{55}	bi^{31} tɕoŋ^{55}	---	tăṅ
thin (book)	*bY-čor	srab	ba?-	***pâ***	***ba**^{55} a^{55}	kɯ^{31} ***pa**^{35}	mɯ-dɯ-thaŋ	sap
think	*mɯŋ	sem(s); bsam	can-ci	thaŋ; câñ	ta^{31} we^{55}	ntshɯm^{55}	mjen; šu	(sak) ċíṅ
thou	***no**	khyod; khyed; **nyid** [hon]	**na?ŋ**	**naŋ**	**noŋ**^{35}	**no**^{53}	ñi	hó; a-do
three	*ñum	gsum	gil-**tam**	**sûm**	ka^{31} **sɯŋ**^{55}	kɯ^{31} **sûm**^{55}	gu-thun	**sam**
tiger	*mro (*mjo?); *paŋ-tə	stag	mat-ca	kyâ	bo^{55} da^{55}; #ta:-mja	bo^{55} da^{55}	tiŋ-graŋ	să-t'ăṅ
tired	*pe	dub; thang chad	ne?ŋ-	mô	giai^{53}; #he-ra:	ɕai^{55}; *min-jin	khaŋ-ru	pyăl
tongue	*rjo	lce	sre	hlya	thɯ^{31} liɯ^{53} na^{35}	blai^{53}	že?-ɣi	a-li
tooth	*fi:	so	wa-gam	swâ	laŋ^{35}	*si*^{55}	thu	a-fo; fo-ki
two	*ñi	gnyis	gin-i	hnac	ka^{31} n^{55}	kɯ^{31} jin^{53}	gni	nyăt; nyi
urine	***sum; *si**	gcin; (dri-) chu	su-bu	chi	kɯ^{31} ***tɕɯŋ**^{55}	tɯ^{31} ɕit^{55}	brui?	jit
village	*nam-pom; duŋ-luŋ	yul-gru; grong	soŋ	rwa	ma^{31} tiɯŋ^{55}	mɯ^{31} tăŋ^{53}	gɯ-bjaŋ	lí brom; lí broṅ; lí kyoṅ
vomit	***bat**^{2}~ ***brat**^{2}	skyug	#ci-sat; wa-kal	aṁ	***ma**^{53}	**phat**^{55}	***mu***	**mót**; hluṅ
wash body; bathe	*ɦɯr	'khru ~ 'khrud; chu rgal	a-u	khyui'	ma^{31} nɯm^{55} tsai^{53}	ta^{31} ɹɯɯl^{35} lai^{53}	---	mŭ-tŭt; mŭ-čóṅ
water	***si**	chu	**ci**	re	ma^{31} **tɕi**^{53}	a^{31}ti^{35}	vu	uṅ
weave	***čum**	'thag	dok	rak	ta^{31} tiɯ^{55} tio^{53}	tho^{55} tan^{55} tho^{55}	**čun**	t'ok
weep[95]	***krap**	ngu; shum; **khrap**	grap	ŋui	**khɹo**^{53}	ŋai^{55}	---	hryóp; prám mat
wet	*ju-jaŋ	rlon-pa	so-si	cui; cwat	puɯm^{55}	phom^{55}; #kan-sak	mu-gro?	šăl
white	*pun~ *puŋ	dkar-po	gip-bok	phru	lio^{53}	kɯ^{31} mphlaŋ^{55}	mɯ-grjaŋ	[du]
wind	***rji**	rdzi; rlung; lhag-pa	bal-wa	le	xa^{31} ***ɹɯŋ**^{55}	bauŋ^{35}	jo	sŭṅ-mŭt; so-mŭt
wing[96]	***lap**	gshog-pa; 'dab-ma	graŋ	a'-toŋ	ta^{31} loŋ^{55}	ŋkhloŋ^{35}	gɯ-či	pă-ku; pŭṅ-ku
wood	***sɯŋ**	**shing**	bol	sac	ma^{31} **sɯŋ**^{53}	**săŋ**^{35} khliŋ^{55}	u	**šaṅ**; kuṅ
year[97]	***ñiŋ**	lo; **-ning**	#bil-si	hnac	kɯ^{31} **nɯŋ**^{55}	lau^{53}	du-ren	nam (tum)

Notes

* Earlier versions of this paper, condensed and revised from Chapter V of my UC Berkeley dissertation (Sun 1993), were presented at the 26th International Conference on Sino-Tibetan Languages and Linguistics, Osaka, September 13–17, 1993, and at an institutional colloquium of the Institute of History and Philology, Academia Sinica on November 22 of the same year. Thanks are due to many scholars for providing helpful discussions and constructive comments, especially Randy J. LaPolla, Yoshio Nishi, James A. Matisoff, Paul K. Benedict, Sun Hongkai, Søren Egerod, Paul Renkul Li, Boyd Michailovsky, and Dibbon Wu.

1 A preliminary phonological reconstruction of Proto-Tani is proposed in Chapter II of Sun 1993, from which the reconstructed PT roots cited herein are taken. The Proto-Tibeto-Burman (hereafter *PTB*) reconstructions are based mainly on Benedict 1972 (hereafter *STC*).

2 The great lexical differences between Tani and other Tibeto-Burman languages (only 12.5% agreement of basic vocabulary with Tibetan and Burmese according to his calculation) has led Marrison to doubt not only the genetic affiliations of Tani with Tibeto-Burman, but also "the reality of the Tibeto-Burman language family as generally accepted . . . The Tibeto-Burman family is an unsatisfactory construct, and this whole field of investigation should be reopened" (Marrison 1988:216). My own lexical study, however, has turned up much higher cognate figures between Tani and both Tibetan and Burmese (see 5.3. below). Even if Marrison was right about the cognacy rates, his radical view on the status of Tibeto-Burman, we believe, would be hard to accept for most Sino-Tibetanists.

3 For instance, the regular sound correspondence between PTB *-əy and PT *-i is backed up by as many as eleven cognate sets, all belonging to basic vocabulary (Sun 1993: 4.3.1.2.).

4 As shown by ensuing research, the Mishmi languages do not form a coherent linguistic unit either. Rather, there is a fundamental cleavage between Digaro-Chulikata-Midu (Taraon-Idu) and Miju (Kaman). Thurgood 1985:81 claims that the Mishmi languages belong with Nungish under a supergroup 'Kaman-Nung' with 'fully substantiated lower-level genetic relationships'. We believe that this claim, which remains totally unproven, underestimates the great differences between the two Mishmi groups (for a more conservative view, cf. Sun et al. 1980:299–315).

5 Incidentally, Benedict's revised view on the special relationship between Jingpo, Bodo-Garo, and Northern Naga seems to be receiving growing endorsement (Burling 1971, 1983: French 1983). The most drastic move in this direction is taken by Weidert 1987: fn.22. where Jingpo is put directly under one of the three branches of Barish: Western Barish (= Bodo-Garo, or Burling's Garo branch); Eastern Barish-I or Arunachal Barish (= Tangsa, Nocte, Wancho); and East Barish-II (= Konyak, Phom, Chang, Khiamngan, and *Jingpo*). An alternative view groups Jingpo rather with Lolo-Burmese, forming a *'Jiburish'* subgroup on the strength of hundreds of cognates between Jingpo and Lolo-Burmese and some parallel phonological developments (Matisoff 1974). In Matisoff 1991:481, however, Jingpo (Kachinic) and Lolo-Burmese are treated as separate major Tibeto-Burman subgroups.

6 Dal 1989 and Sun 1988 are not considered here because they deal only with the Tibeto-Burman languages of China.

7 Surprisingly, Bradley's Qiangic also includes such languages as Tujia, Baima, and even Bai. Chinese scholars now tend to regard Baima as a divergent dialect of Tibetan (Zhang Jichuan and Huang Bufan, p.c.)

8 One important reason for the lack of agreement in Tibeto-Burman subgrouping may be the different criteria (often implicit) used in the various subgrouping proposals. Thus,

Thurgood puts Nung in his 'Rung' supergroup apparently on morpho-syntactic evidence only (Thurgood 1985). What is not explained is the considerable amount of shared basic vocabulary between Nung and Lolo-Burmese (STC:8: Benedict 1976: fn 14).

9 Thurgood claims that 'Even from the limited LS] sample of data, it is clear that the Chulikata Mishmi [= Midu] . . . must be subgrouped with these Adi languages rather than with the Miju language' (Thurgood 1986:93). Actually, Midu should be equated with Idu (autonyms: Idu, Midu, Dudu), which Thurgood in the same paper correctly assigns to the Taraon group.

10 Sun Hongkai's tentative inclusion of Sulung and Bangru under the Nishi-Bangni subgroup of Tani (Sun 1983:267) was done apparently at a time when linguistic data on these languages was not yet available to him. His more recent view is that Sulung and Bokar (other Tant languages are not mentioned) are distinct languages belonging to the 'Jingpo' supergroup, which also contains Jingpo, Nungish, and the Mishmi languages (Sun 1988:69).

11 This is not same as the 'Northern Naga' (or Konyak) languages of French 1983. Rather, it refers to the group of Naga proper which Weidert terms Naga-II, comprising Ao, Lhota, Sangtam, Yimchunger, and Northern Rengma (Weidert 1981: fn. 3).

12 From PT *kap 'cover'. Cf. rGyarong pkap: Jingpo mă^{31}*kap*^{31}: Dulong ta^{55}*kŏp*^{55}: *Kaman ŋkhap*^{53}: *WT 'kheb~'gebs* 'cover'.

13 Cf. Damu OY mit-pan to-mit 'forget'. Prof. Matisoff suggests that the *mit- element may reflect PTB *m-yit 'mind'. This is possible, but the normal PT 'mind/think' root is *muŋ.

14 Causativity in modern Tani is normally expressed by means of affixation (usually involving the morpheme 'do/make' mo:) rather than by stem-modification.

15 Wolfenden 1936:168 also suggested that rGyarong may be a moderately close surviving relative of Xixia (Tangut), which is now generally considered to be another Qiangic language (Sun 1988:67, Matisoff 1991: 482).

16 This is perplexing given the general principle that if two languages bear a true genetic relation, then the further back one traces their histories, the more similar they should be.

17 Cf. the Chinese parallelism: wèn 聞 'hear' vs. wèn 問 ask'.

18 Medial -w- in lCog-rtse rGyarong often comes from earlier (phonetically prenasalized) voiced stops (cf. lCog-rtse ta-wo: Dashuigou rGyarong ta-mbo 'deaf n.'; cf. Jingpo na^{31}phaŋ^{55}; Mawo Qiang bu; Queyu rni^{55}pa^{13}; Muya na^{33}mba^{24}; Nusu boŋ^{55}; WB nâ-pâng; Garo beŋ-a; Tangsa ^{3}baŋ; < PTB *baŋ; lCog-rtse tə-wro; Dashuigou rGyarong tə-ŋgro 'sinew; tendon'; cf. Mawo Qiang Qə^{4}; Xide Nosu gu^{55}tsɛ^{33}; WB a'-krô; Nusu gɯ^{55}: Dulong dɯ^{31}gɯ^{53}: WT rgyus. Dashuigou 大水溝 (previously known as Benzhen 本貞). like the better-known lCog-rtse and Suomo varieties, belongs to the Eastern dialect of rGyarong. The Dashuigou data cited herein were collected by the author in two recent field trips to western Sichuan.

19 Cf. Matisoff 1976 in which body-part terminology is chosen as the target semantic area in an exploration of shared contact vocabulary between Sino-Tibetan and Austro-Tai.

20 Patients carry no case-marking in rGyarong. In this regard rGyarong differs from languages of the 'Qiangic' group (to which rGyarong has been assigned by some Chinese scholars).

21 For more discussion, see J. T.-S. Sun 1994:4.2.

22 Unfortunately, only a limited number of Proto-rGyarong roots are proposed in Nagano 1984:133–9. Where Proto-rGyarong reconstructions are unavailable, modern (lCog-rtse) forms (unasterisked), are cited from ZMYYC.

23 The PT root also means 'sit/stay/dwell'. rGyarong uses a completely different form ka-ñi for sit/dwell'.

24 Nagano posits an open-syllable proto-form *ya; however, a lCog-rtse form -jak, with a *checked* syllable, appears in ZMYYC.
25 This PT root means only 'ripe'.
26 Cf. WT rgyug.
27 Nagano 1984 provides the lCog-rtse form nam-nam. Compare the different form kə-nə nsə nsət in ZMYYC.
28 The 'stand' meaning of PT *rop is preserved in Bokar OY. Reflexes of this root occur elsewhere mainly as an adverbial verbal particle meaning 'upwards'.
29 It is not clear why Nagano chose to reconstruct this root as an open syllable despite the lCog-rtse form ka-*rjap* (ZMYYC).
30 The main roots are italicized: cognates with PT roots are boldfaced.
31 The rGyarong root is cognate with WT *rna*-ba 'ear' and WB nâ 'ear', na 'listen.'
32 The predominant rGyarong words for this gloss are cognate with WT shes and WB si' < PTB śey (STC #182); cf. lCog-rste ka - *ʃə* (ZMYYC). Tsanla ka-nga-*syis*. Khamto ka-*syr*. Suomo ka-nə-*msyi*. Chos-kia ko-*syu* (Nagano op. cit.:109). Nagano also gives the alternative PG root *gye-s which he links with PTB *m-kyen (and which is thus supposedly cognate with PT *ken), but it is not clear what data support this reconstruction.
33 Nagano associates this rGyarong root with WT rmi < PTB *r-mwəy 'sleep'. The equation rGyarong -a <-> PTB *-əy, however, seems restricted to this single example.
34 WB waŋ means 'enter'. WT 'byung 'emerge, come, go' is listed in the cognate set for PG *bo in Nagano op. cit.: 84; however, if this rGyarong root came from PTB *byon (STC #179) as Nagano suggests, then the true WT cognate should rather be 'byon 'go, arrive, appear'.
35 This rGyarong root is linked with WT snye(s) 'lean against, lie down'; again, the equation between rGyarong -i and WT -e(s) is limited to this pair.
36 The following glosses are considered to involve rGyarong-WT cognates: 'eat', 'see', 'hear/ear', 'know', 'dle', 'kill', 'swim', 'come'. The cognacy of the pairs PG *r-ma, WT r-mi 'sleep', and PG *nyi 'sit', WT snye(s) 'lie down' is possible but uncertain. Thus, the number of rGyarong-WT cognates in this sample ranges from eight to ten.
37 The following items are judged to involve rGyarong-WB cognates: 'eat', 'fly v.', 'hear', 'know', 'die', 'kill', 'stand', and 'swim'.
38 The strong rGyarong-Lolo-Burmese lexical ties, suspected by Benedict (p.c.), is an area awaiting further investigation.
39 Data transcription follows the original sources. Probable cognates with the PT roots are boldfaced; suspicious look-alikes are boldfaced and italicized.
40 PT *bjar reflects PTB *byer. The Dhimalish forms may come rather from PTB *pur~pir, now considered a separate root (STC fn. 249).
41 This is also the impression of Dr. Sueyoshi Toba (p.c. 1993), who has been working on this Tibeto-Burman group in Nepal.
42 The Bangru (autonym Levai/lə^{31}væ^{55}/) tribe consists of about a thousand souls whose villages are distributed in the Lagong area along the Tibetan-Indian border (Anonymous 1989:248). Note the similarity between the name *Levai* and the Miji autonym *Dhammai* (/ðɯm-mai/). It is possible that the Levai represents a northeastern subbranch of the Mijis of Eastern Kameng. The name Bangru (/bɯŋ-ru/) is a Bengni exonym: cf. also the Sulung exonym of Levai: Buzwa (/bu^{33}zwa^{53}/).
43 I recorded about a thousand Bangru words from my Sulung consultant, who has a speaking knowledge of this language, during field work in Tibet in the summer of 1992.
44 The two most serious problems pointed out by Professor Matisoff being (a) How can one ensure that one's cognate identification is reliable, when detailed knowledge about the sound laws in the languages compared may be lacking? (b) How can an all-or-none

(i.e. cognate vs. non-cognate) scoring method reflect the gradient nature of phonological-semantic relationships in the lexical data?
45 Both scholars are among the world's leading Austro-Asiaticists. They have demoted Thomas and Headley's 'Malacca' (i.e. Aslian) and Nicobarese from coordinate families of Mon-Khmer to branches within Mon-Khmer, added a few minor new discoveries like Mang and Palyu (Laf), and proposed some possible higher-level divisions (Northern, Eastern, Southern, Vietic), but the basic Mon-Khmer branches remain identical to Thomas and Headley's original proposal: Viet-Muong, Khasi, Palaungic, Monic, Khmuic, Katuic, Bahnaric, Khmer, and Pearic.
46 Such as the controversial application of lexicostatistics to dating proto-languages (*glottochronology*).
47 Thus, the sound correspondences between such language pairs as rGyarong-AMD (Nagano 1984), Lepcha-Adi, and Lepcha-Nung (Bodman 1988) alone do not constitute sufficient proof that these languages are more closely related.
48 The Taraon and Kaman data are cited mostly from Sun et al. 1980 and from ZMYYC. Forms missing from these sources are supplemented from Chakravarty et al. 1963 for Taraon and Boro 1979 for Kaman.
49 Lepcha forms are taken from Mainwaring-Grünwedel 1979. Root forms (enclosed in square brackets as in the original source) are cited where available; e.g. the root (kri), rather than the suffixed adjectival form a-krím, is given for the gloss 'bitter'. Loanwords (chiefly from Tibetan) are marked with an asterisk in the dictionary: such forms are avoided herein except in the rare cases where the asterisked forms turn out to be the only ones listed for the given meaning.
50 Dhammai forms are based on Simon 1979. The sound system of Dhammai is retranscribed as follows (phonetic symbols used in the original are enclosed within parentheses):

1. Vowels: a, e, w, (i), i, o, u
2. Consonants:

p	t	ts	č (c)		k	ʔ
ph	th	tsh	čh (ch)	kh		
b	d	dz	ǰ (j)		g	
f	θ	s	š (sh)	h		
v	ð	z	ž (zh)			
m	n		ñ		ŋ (ng)	
	l					
	ɬ					
	r					
w		j (y)				

Remarks: (1) Dhammai may have contrastive vowel length and phonemic tone; neither gets marked in the main body of this source. (2) The glottal stop is a phonemic syllable coda, represented in the source by -h. (3) Dhammai has a peculiar lateral consonant symbolized by Simon as ll, which he describes as being 'articulated with the tongue rolled'. This is probably the retroflexed lateral ɭ.
51 Abbreviated from Culturally Appropriate Lexicostatistical Model for South-East Asia, this list represents Prof. Matisoff's revision of the Swadesh basic vocabulary list to make it culturally more appropriate for Southeast Asian languages.
52 Cognate identification in Tibeto-Burman is an extremely risky undertaking. Our general attitude is to be more willing to *err on the conservative side*, for our knowledge of the various languages involved (except perhaps Tibetan) is not sufficient to allow

bold speculation. In this study, forms are treated as cognate only if they are considered to descend from one and the same *proto-allofam* (i.e. variants of the same proto-word-family, Matisoff 1978a:17). Thus, WB klok~kyok and PT *lɯŋ 'stone' are not directly cognate even though they may come from related proto-allofams. By the same token, Taraon pia[53]kɹau[35] and Kaman tɕi[55]khɹuŋ[55] (< PTB *(m-)krəw 'dove', STC #118) are not cognate with PT *kɯ 'dove/pigeon' (< PTB *(m-)kəw 'pigeon' STC #495; note that PT normally kept the PTB *kr- cluster), for they are derived from related but distinct PTB etyma. Of course, such subtle distinctions are not always possible with languages the sound laws of which are not yet well-known.

53 The Garo data are taken mainly from Burling 1983. Supplementary forms, marked by #-, are added from Momin: no date. Transcription of Garo is based on the '*combining*' (i.e. non-final) form, which is etymologically more basic (Burling 1983:69–70). Garo-Tani cognate determination is greatly facilitated by the etymological tables in Burling 1983, where the PTB etyma of many Garo roots are provided.

54 Initial efforts have been made to inspect the sound laws of Taraon, but a full-scale comparative study of Taraon and its close kin Idu has not been attempted.

55 WT 'ang-gu is more common in Central Tibetan. In Khams Tibetan, mug-gu is used instead. The normal Classical Tibetan word is phug-ron. While PT *kɯ is clearly a reflex of PTB (*m-)kəw 'pigeon' (STC #495), WT 'ang-gu shows an unexpected voiced initial *g-* (although WT -u regularly reflects PTB *-əw).

56 Consider for example PT *lɯk, Lepcha lyăk. cf. PTB *lay 'exchange' (STC #283). The PT and Lepcha forms may be related rather to Mon-Khmer, cf. Proto-Wa-Lawa *ʔloh (Diffloth 1980), Kammu (Yuan dialect) lɛɛk 'exchange' (Lindell 1974:200). The PT and Lepcha words for 'excrement' may also be of Mon-Khmer origin (Forrest 1962). The considerable Mon-Khmer contact vocabulary in Tant languages will be explored in a separate paper.

57 Unfortunately, the Kuki-Chin-Naga and Kiranti-Tibetan-Kanauri links are not considered in Bodman 1988. Lepcha certainly seems to have as many good lexical comparisons with Mikir and Ao Naga as with Tani, on Bauman 1976's evidence.

58 There are two major subgroups within Tani: Western and Eastern (Sun 1993: chapter III). As may be expected, more parallels exist between Hrusish and *Western* Tani. For example, the Western Tani root *nam 'house' (as against Eastern Tani *kjum) is obviously related to Hrusish, cf. Dhammai nen, Bangru nɛ:[55], Hruso ñe 'house'.

59 The Taraon and Kaman forms for the following items are judged to be cognate: 'bear n.', 'bird', 'blood', 'brain' (?), 'borrow', 'burn' (?), 'child/son', 'cloud', 'day', 'die', 'dog', 'dove' (?), 'dream', 'eat', 'eight', 'extinguished', 'fat/stout', 'fat n.', 'excrement', 'fire', 'fireplace', 'fish', 'float' (?), 'flower' (?), 'four', 'full', 'gall', 'guts', 'head', 'horse', 'kidney', 'kill', 'knife', 'leech', 'lick', 'listen/hear', 'melt', 'moon', 'mortar', 'name', 'neck', 'otter', 'penis' (?), 'pig', 'poison', 'ripe', 'river', 'road', 'round', 'seed', 'sharp-edged', 'smoke n.', 'stone', 'tail', 'thick', 'thin', 'thou', 'three', 'tiger', 'tongue', 'village', 'vomit', 'water', 'weave', 'wet', 'wing', and 'wood'.

60 The development to stops is not uniquely shared by Tani and Digarish, however. Matisoff 1978b:11 reports, for instance, that PTB *ts- and *dz- went respectively to th- and t- in Mpi, a southern Loloish language of Thailand. Cf. also the Queyu (Qiangic) word for 'eat' kə[35]tɐ[53] (ZMYYC).

61 For usage, consider the illustrative sentences below:

Bokar OY (Ouyang 1985: 71)

š: lamto a:to-*joŋ*-da
this road far-*more*-declarative
'This road is farther.'

Taraon (Sun et al. 1980:219)

tɐe[55] xaŋ[35]-doŋ[31]go[31] lau[55]dzoŋ[55] pɹa[55]-*joŋ*[35]
s/he I-than learn good-*more*
'S/he learns better than I do.'

62 There is an interesting look-alike in Tai: Proto-Tai *ʔjaa[A1] prohibitive; negative imperative' (F.K. Li 1977:181). [Ed.]

63 Observe the example below, taken from Sastry 1984:189 (tone marks omitted):

hã [hɨbaŋ bo-ya jyinaŋ]REL Ø-dõ kitab haŋ-de
I forest go-impf cousin Ø-obj book give-impf
'I give the book to (my) cousin who goes to the forest.'

64 Probable cognates are bolded; uncertain cognates are both boldfaced and italicized, to be taken account of separately in the cognacy calculation.

65 Many of the 'angry' forms here are compounds with a first element meaning 'mind'; e.g. PT *haŋ-, WB cit-, and Lepeha sak- (which looks deceptively like the main PT 'angry' root *fak).

66 For Taraon ta [31]ɱm[55] cf. the more transparent form ta:-hom in Chakravarty 1963.

67 The Dhammai form mu-khuʔ exemplifies a regular sound change PTB *-a > Dhammai -u, cf. also bu-ŋu 'five': tnuʔ 'eat', lu 'month/moon', zu 'son'; thu 'tooth'.

68 Sino-Tibetan languages generally do not lexicalize directionality of the loaning transaction, thus 'borrow' and 'lend' are usually expressed by identical roots. Instead, many Tibeto-Burman languages make a different distinction based on the nature of the loaned object; thus 'borrow/lend something that must itself be returned' and 'borrow/lend something that can be returned in kind' involve distinct roots, e.g. Tibetan g.yar vs. sky; Burmese hŋâ vs. khyê; Kaman a[31]ŋat[55] vs. lu[53] in the table. This contrast has not been detected in any Tani language.

69 For the ZMYYC Kaman form ŋin[53], cf. Boro 1979 ŋit; Weidert 1987:478 ŋit 'day'.

70 Note the secondary -k coda in the Taraon form kuaɯ[53] (for -ɯ < -k, cf. Chakravarty 1963 kua; Sastry 1984 kwág).

71 The Taraon word for ear' is literally kɹu[53] 'head' + naŋ[35] 'leaf'.

72 The Lepeha form is literally mí 'fire' + mak 'die'. Lepcha mak 'die' is unlikely to be cognate with PT *mit 'extinguished'.

73 The ká- 'hand' element in the Lepeha form seems unlikely to be cognate with PT *ke(ŋ) 'finger'.

74 The Dhammai form is also glossed 'heart'.

75 For the phonologically reduced Taraon form ɱ[55], cf. Chakravarty et al. 1963 um 'hair (on body)'.

76 For (ZMYYC) Kaman ɹau[53], cf. also Boro 1979 rok; Weidert 1987:479 râuk 'arm'. The r- initial of these Kaman forms is perplexing, especially since Kaman apparently maintains the PTB contrast between *l- (e.g. lăuŋ 'stone' < PTB r - luŋ; lap[53] 'leaf < PTB *lap) and *r - (e.g. ɹam[35] 'otter' < PTB *s-ram: ɹuɯl[35] 'snake' < PTB *b-ru:l).

77 In both Taraon and Kaman, several existential verbs are distinguished: Taraon i[55] and Kaman tɐau[53] occur with animate subjects. Taroan aŋ[55] and Kaman kam[35] with inanimate ones, a third Kaman existential verb tun[55] applies only to abstract qualities (Sun et al. 1980). A different type of semantic differentiation of existential verbs is reported in Apatani A, based apparently on posture of the predicated subjects, but comparative data from other Tani languages is not sufficient for deciding whether this distinction should be pushed back to the PT level. The different Tibetan existential verbs reflect rather the *pragmatic* distinction of degrees of knowledge integration: yod for fully assimilated knowledge and 'dug for new, unassimilated knowledge (DeLancey 1989).

78 Taraon ma^{31}so^{53} is undoubtedly cognate with PT *fak, both reflecting PTB *m-sak 'itch' (STC # 465). For the equation PT *-ak <-> Taraon -o, cf. also PT *rjak, Taraon lio^{53} 'lick'; PT *jak, Taraon jo^{53} 'fox-tail millet'.

79 In the sense of have knowledge of.

80 This PT root is quite unique in Tibeto-Burman. The only extra-Tani cognate known to us so far is Tshangla ŋar 'laugh'.

81 In languages that distinguish 'listen' from 'hear', forms for both meanings are given (in that order), separated by a semicolon. In Tani, the same root occurs for both meanings; the punctual, nonvolitional sense 'see' is expressed by adding to the root a resultative verbal particle-poŋ. This is true of such other pairs as 'listen' vs. 'hear'; 'search' vs. 'find'. The Garo form means 'hear'.

82 In languages that distinguish 'look' and 'see', both forms are given (in that order) separated by a semicolon.

83 This is not considered cognate with PT *-kin, because the regular reflex of the PTB medial vowel *-i- seems to be -ă- (i.e. short -a-) in Kaman (but *-i- or *-ɯ- in PT); e.g. săŋ^{35} 'tree' < PTB *siŋ; a ^{31}măŋ 'name' < PTB *r-miŋ; măn^{53} < măt < PTB *mit 'extinguished'; ntshăn 'claw' < PTB *m-(t)sin.

84 The -ŋ in the ZMYYC Kaman form a^{31}muŋ^{35} seems secondary; cf. Weidert 1987:358 ʔmùk; Boro 1978 a-muk. both keeping the original -k coda; the latter Kaman forms are cognate with PLB *myok^{L} (Matisoff 1972 #133) < PTB *mruk STC:112.

85 The Dhammai form go could not be cognate with PT *gam because the expected Dhammai equation to PT (and PTB) *-am is -en; e.g. Dhammai lem-baŋ (< len-) PT *lam 'road'; Dhammai nen, Western Tani *nam 'house'; Dhammai ñen, PT *nam 'smell v.'; cf. also Dhammai sen < PTB *śam 'iron' (STC #228).

86 For the Taraon form pa^{31}ŋŋ^{55}, cf. Chakravarty et al. 1963 pa:-haŋ.

87 The Taraon and Kaman words are composed respectively of 'child' + 'protect' and 'child' + 'nest'. As for the Lepcha forms, kap-pǔńis literally 'covering, that which covers'; while 'ayeń-tyól is 'child' + 'accompany'.

88 Cf. the Chakravarty et al. 1963 tha:ik for Taraon and Boro 1979 tŏuk for Kaman, both retaining the -k coda.

89 More precisely 'cooked rice'. For the Kaman form ɕat^{53}, cf. Weidert 1987:479 má-syât 'boiled rice' (root = syâ 'eat' plus nominaltzing dental suffix -t).

90 The Taraon form pla^{35} seems to come from earlier *plaŋ (cf. Midu prã 'salt') and therefore phonetically quite distant from PT *lo.

91 The Taraon form o^{53}- is judged to be cognate with PT *-ap. For the equation PT -ap <-> Taraon -o, cf. also PT *krap, Taraon khro 'weep'.

92 The resemblance between Dhammai ji to PT *jup is misleading, for the Dhammai form could originate from a nasal-final rhyme, cf. Bangru dʐẽ^{33}; Hruso ǰum 'sleep'. The Lepcha compound is literally mik 'eye'+ krap 'hang down'.

93 The Garo word means 'saliva'; from ku 'mouth' + ci 'water'.

94 The Taraon form ɕau^{55} seems to come from a checked syllable, cf. Chakravarty et al. 1963 shyeb 'sweet'.

95 WT khrap occurs only in the phrase khrap-khrap 'weeper, cry-baby'. The normal 'weep' meaning has been taken over by the ngu root.

96 WT 'dab-ma (< N + lap) is a direct cognate of PT *lap. The dental stop initial is transparently caused by the homorganic nasal prefix N- (represented orthographically by the *achung*). For more evidence of the effects of achung, cf. 'dom (< N + lom) 'fathom' < PTB *la(:)m (STC p.71); 'do (< N + lo)~ zlo 'say, repeat'; this view is also strongly supported by the identical delateralizing effect of the m- nasal prefix, cf. WT mda (< m + *la); PTB *mla ~ bla 'arrow' (STC fn. 313). For a different interpretation of the provenance of this WT form (owing perhaps to a different view on the phonetic nature of WT achung). cf. Matisoff 1985a:443–4 as well as STC: 122–3; fn. 338, 339.

97 In WT, the root -ning 'year' occurs only in compounds, such as na-ning 'last year'.

References

Anonymous. 1989. *Luobazu shehui lishi diaocha* ('Investigations on the society and history of the Luoba nationality"'. Vol. II. Lhasa: Tibetan People's Press.

Anonymous. 1991. *Zangmianyu Yuyin he Cihui* ('Tibeto-Burman phonology and lexicon"'. Beijing: Chinese Academy of Social Sciences Press. ("ZMYYC")

Bauman, James. 1976. "An issue in the subgrouping of the Tibeto-Burman languages: Lepcha and Mikir." Paper presented at the 9th International Conference on Sino-Tibetan Languages and Linguistics. Copenhagen, Oct. 22–24, 1976.

Benedict, Paul K. 1972. *Sino-Tibetan: a Conspectus.* Contributing editor: James A. Matisoff. Cambridge: Cambridge University Press.

———. 1976. "Sino-Tibetan: another look." *Journal of the American Oriental Society* 96.2:167–97.

Bodman, Nicolas C. 1988. "On the place of Lepcha in Tibeto-Burman: a lexical comparison." *Linguistics of the Tibeto-Burman Area* 11.1:1–26.

Bor, N. L. 1938. "Yano Dafla grammar and vocabulary". *Journal of the Royal Asiatic Society of Bengal* 4:217–81.

Boro, A. 1978. *Mÿu Dictionary.* Shillong: Research Department. Arunachal Pradesh Administration.

Bradley, David. 1993. "The subgrouping of Proto-Tibeto-Burman." Paper presented at the 26th ICSTLL. Osaka, Japan, September 13–17.

Burling, Robbins. 1959. "Proto-Bodo." *Language* 35.3:433–53.

———. 1971. "The historical place of Jinghpaw in Tibeto-Burman." Papers on Tibeto-Burman historical and comparative linguistics from the Second Annual Meeting on Sino-Tibetan Reconstruction. (= OPWSTBL Vol II). ed. by F. K. Lehman, 1–54. Urbana: University of Illinois.

———. 1983. "The Sal languages." *Linguistics of the Tibeto-Burman Area* 7.2:1–31.

———. 1992. "Garo as a minimal tone language." *Linguistics of the Tibeto-Burman Area* 15.1:1–12.

Campbell, G. (ed.) 1874. *Specimens of languages of India.* Calcutta: Bengal Secretariat Press.

Chakravarty, L. N., et al. 1963. *A Dictionary of the Taraon language.* Shillong: Philology Section, Research Department, North-East Frontier Agency.

Chhangte, Thangi. 1990. "Nishi grammar sampler." Paper presented at the 23rd ICSTLL. University of Texas, Arlington. October 5–7.

———. 1992. "Phonology of some Nishi (Dafla) dialects." Paper presented at the 25th ICSTLL. UC Berkeley. California, October 14–18.

Dai, Qingxia. 1989. "Guanyu woguo Zangmian yuzu xishu fenlei wenti" ('On the issue of genetically classifying the Tibeto-Burman languages of China'). *Yunnan Minzu Xuebao* 3:82–92.

DeLancey, Scott C. 1989. "New vs. assimilated knowledge as a semantic and grammatical category." Paper presented at the Winter LSA Meeting, Washington. D. C., December 30th, 1989.

———. 1991. "Sino-Tibetan languages." *Oxford Encyclopedia of Linguistics,* Vol. 4, ed. by William Bright, 445–9. New York: Oxford University Press.

Diffloth, Gèrard. 1980. "Proto-Waic." *Linguistics of the Tibeto-Burman Area* 5.2:1–182.

Diffloth, Gèrard and Norman Zide. 1991. "Austro-Asiatic languages." *International Encyclopedia of Linguistics,* ed. by William Bright, 137–42. New York: Oxford University Press.

Egerod, Søren. 1974. "Sino-Tibetan languages." *Encyclopedia Britannica* 16:796–806.

Forrest, R. A. D. 1962. "The linguistic position of Róng (Lepcha)." *Journal of the American Oriental Society* 82:331–5.
French, Walter T. 1983. *Northern Naga: a Tibeto-Burman mesolanguage.* New York: City University of New York dissertation.
Hale, Austin, ed. 1982. *Research on Tibeto-Burman Languages.* Berlin, New York, Amsterdam: Mouton.
Hodgson, B. H. 1847. *Essay the first on the Kocch, Bodo, and Dhimal tribes in three parts.* Calcutta: J. Thomas Baptist Mission Press.
Huang, Bufan. 1991. "Qiangyu zhi" ('The Qiangic branch'), *Hanzangyu gailun* ('Survey of Sino-Tibetan'), ed. by Ma Xueliang et al., 208–369. Beijing: Beijing University Press.
Konow, Sten. 1909. "North Assam group." *Linguistic Survey of India,* ed. by G. A. Grierson, Vol 3, pt. 1: 568–72. Calcutta: Superintendent of Government Printing.
LaPolla Randy J. 1992. "Anti-ergative marking in Tibeto-Burman." *Linguistics of the Tibeto-Burman Area* 15.1–1-9.
Li, Fang Kuei. 1977. *A Handbook of Comparative Tai.* Honolulu: University of Hawaii Press.
Lindell, Kristina. 1974. "A vocabulary of the Yuan dialect of the Kammu language." *Acta Orientalia* 36:191–207.
Lorrain, J. H. 1907. *A Dictionary of the Abor-Miri Language.* Shillong: Eastern Bengal and Assam Secretariat Printing Office.
Mainwaring, G. B. (Revised and completed by Albert Grünwedel). 1979. *Dictionary of the Lepcha language* (Revised and completed edition). Kathmandu: Ratna Pustak Bhandar.
Marrison, Geoffrey E. 1988. "The Adi-Dafla group of languages of North-East India: a sketch." *Prosodic analysis and Asian linguistics,* to honour R. K. Sprigg. ed. by David Bradley, Eugènie J. A. Henderson and Martine Mazaudon, 20–22. Canberra: Australian National University.
Matisoff, James A. 1972. *The Loloish Tonal Split Revisited.* (Center for South and Southeast Asia Studies Research Monograph No. 7). Berkeley: Center for South and Southeast Asia Studies, UC Berkeley.
———. 1974. "The tones of Jinghpaw and Lolo-Burmese: common origin vs. independent development." *Acta Linguistica Hafniensia* 15.2:153–212.
———. 1978a. *Variational Semantics in Tibeto-Burman: The 'organic' approach to linguistic comparison.* OPWSTBL Vol. VI, ed. by F. K. Lehman. Philadelphia: Institute for the Study of Human Issues.
———. 1978b. "Mpi and Lolo-Burmese microlinguistics." Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
———. 1985a. "Out on a limb: *arm, hand,* and *wing* in Sino-Tibetan." *Linguistics of the Sino-Tibetan Area: the State of the Art:* papers presented to Paul K, Benedict on his 71th birthday, ed. by Graham Thurgood, James A Matisoff, and David Bradley, 421–50. Canberra: Australian National University.
———. 1985b. "God and the Sino-Tibetan copula, with some good news concerning selected Tibeto-Burman rhymes." *Journal of Asian and African Studies* (Tokyo Foreign Language University) 29:1–81.
———. 1991. "Sino-Tibetan linguistics: Present state and future prospects." *Annual Review of Anthropology* 20:469–504.
Momin, K. W. No date. *English-Achikku Dictionary.* No place of publication.
Nagano, Yasuhiko, 1984. *A Historical Study of the rGyarong Verb System.* Tokyo: Seishido.

Nichols, Johanna. 1986. "Head-marking and dependent-marking grammar." *Language* 62:56–119.
Nishida. Tatsuo. 1979. "Robago no keitō" ('The system of the Luoba language'). *Gengo* 7:70–77.
———. 1984. "Review of Hale 1982." *Tōkyō Gakuhō* 65.3–4.
Ouyang, Jueya. 1985. *Luobazu yuyan jianzhi: Bengni-Boga'eryu* ('Brief description of a language of the Luoba nationality: the Bengni-Bokar language'). Beijing: Minzu Chubanshe.
Qu Ai-tang. 1984. "Jiarongyu gaikuang" ('Brief description of rGyarong'). *Minzu Yuwen* 2: 67–80.
Sanyal, Chandra Charu. 1955. "The Ṭoṭos, a Sub-Himalayan tribe." *Journal of the Asiatic Society-Science* 21:97–126.
Sastry, G. Devi Prasada. 1984. *(Taraon) Mishmi grammar.* Mysore: Central Institute of Indian Languages.
Shafer, Robert. 1947. "Hruso." *Bulletin of the School of Oriental and African Studies* 12:184–96.
———. 1950. "Classification of some languages of the Himalayas." *Journal of the Bihar Research Society* 34:192–214.
———. 1955. "Classification of the Sino-Tibetan languages." *Word* 11:94–111.
Simon. I. M. 1979. *Miji Language Guide.* Shillong: Philological Section. Directorate of Research. Government of Arunachal Pradesh.
Sun, Hongkai. 1982. "Qiangyu zhishu wenti chutan" ('Preliminary explorations on the linguistic classification of Qiangic'). *Minzu yuwen yanjiu wenji* ('Collection of papers on minority language studies'), 189–224. Xining: Qinghai Minzu Chubanshe.
———. 1983. "Liujiang liuyu de minzu yuyan ji qi xishu fenlei" ('On the minority languages of the Six-River Region and their genetic classification'). *Minzu Xuebao* ('Yunnan Minzu Chubanshe') 3: 99–273.
———. 1988. "Shilun zhongguo jingnei Zangmianyu de puxi fenlei" ('On the genetic classification of the Tibeto-Burman languages of China'). *Languages and History in East Asia: Festschrift for Tatsuo Nishida on the occasion of his 60th birthday,* edited by Akihiro Sato, 61–73.
Sun, Hongkai et al. 1980. *Menba, Luoba, Dengren de yuyan* ('Languages of the Menba, Luoba, and Deng peoples'). Beijing: Zhongguo Shehui Kexue Chubanshe.
Sun, Jackson T.-S. 1993. *A Historical-Comparative Study of the Tani (Mirish) Branch in Tibeto-Burman.* Berkeley: University of California dissertation.
———. 1994. "Linguistic characteristics of the Tani (Mirish) branch in Tibeto-Burman." *Bulletin of the Institute of History and Philology* 65.1:175–220.
Thomas, David and Robert K. Headley, Jr. 1970. "More on Mon-Khmer subgroupings." *Lingua* 25:398–418.
Thurgood, Graham, 1985. "The 'Rung' languages: notes on their protomorphosyntax and subgrouping." *Acta Orientalia* 46: 79–99.
———. 1986. "A brief comment on Das Gupta's works." *Linguistics of the Tibeto-Burman Area* 9.2:91–6.
Weidert, Alfons. 1987. *Tibeto-Burman Tonology.* Amsterdam/Philadelphia: John Benjamins.
Wolfenden, Stuart N. 1936. "Notes on the Jyârong dialect of eastern Tibet." *Toung-Pao ser.* 2.32:167–204.

Part 2

SINO-TIBETAN HISTORICAL RECONSTRUCTION

13

THE NUMBER "A HUNDRED" IN SINO-TIBETAN

J. Przyluski and G. H. Luce

Source: *Bulletin of the School of Oriental Studies* 6, 3, 1931, 667–8.

In their *Notes d'Etymologie Taï,* published in 1926 in the *Journal of the Siam Society,* vol. xx, pt. i, MM. J. Burnay and G. Cædès have compared the various Taï words meaning "a hundred". Ahom *pāk,* Shan $pāk_1$, Khamti *pāk*[1], White Taï *pāk*[1], Thô *pāk*[1], Nùng *pāk*[1], Dioi *pā*[1]—all go back to a form **pāk,* which is very close to the sixth century Chinese (*pvk*).[1] MM. Burnay and Cædès draw this just conclusion (I translate): "As for the basic form on which Ahom *pāk,* etc., rest, it seems impossible—in view of its wide extension in Taï, and, in addition, the exact correspondence of the tones—not to assign it to the original Taï language or, at least, to the period of Taï union; it seems also impossible to separate it from Old Chinese *pak.* It remains to determine if we have here a borrowing by original Taï from Chinese, or a form common alike to Taï and to Chinese: this question remains untouched."

The next step, it seems, should be to compare, with Chinese and Taï, some forms at least of Tibeto-Burman.

Side by side with classical Tibetan *brgya,* we have Balti *rgyā,* Purik *rgiā,* Ladakhi *rgya.* The other Tibetan dialects have *gya.*[2]

In Burmese, on the other hand, we have twelfth century *ryā,*[3] modern *rā* (pronounced *yā*).[4]

We see that the final guttural, which is conserved both in Chinese and in Taï, has disappeared in Old Burmese and in the Tibetan dialects. As for the initial labial, which appears as a surd *p* in Chinese and in Taï, it reappears in classical Tibetan as a sonant, but is absent in Old Burmese and in the Tibetan dialects. The medial group, so complex in the classical Tibetan *-rgya,* becomes *ryā* in Old Burmese, and is reduced to a single vowel in Chinese and in Taï.

Various Southern Chin dialects still keep a trace of the initial labial: in Yawdwin it is a surd as in Chinese *-pra;* Chinbok has *p'ya.* We can compare also Gyarung *paryé* and Mikir *p'áró.*[5]

In the Northern Chin dialects the medial group appears to be contracted into *zā, jā, jhā,* or reduced merely to *ya.* In the Kuki dialects the same medial group gives Pūrūm *riyāh;* Hirōi, Lamgāng *arja;* Rāngkhōl and Langrong *rajā.* Compare also Pānkhū *rajā* (Central Chin sub-group).[6]

Finally, in two Aka forms cited by Hodson[7]: *phogwa* and *purrua,* a vocalic element is inserted between the initial labial and the medial group. These forms are particularly instructive because, classical Tibetan *brgya* being practically unpronounceable, we must probably assume the existence of an old vowel after the initial. Compare Gyarung *paryé* and Mikir *p'áró.*

In the light of all these indications it seems possible to reconstruct for original Sino-Tibetan some such form as **pargyak.*[8] The final and the initial are well conserved in Chinese and in Taï, while the Tibeto-Burman languages preserve better, in general, the medial element.

We are thus led to suppose, at the base of some modern words, a complex of at least two syllables. Neither Taï nor Chinese permit us to guess it. It is thus apparent that the mere comparison of Chinese and Taï does not carry us very far back.

Notes

1 Karlgren, *Analytic Dictionary of the Chinese Language*, s.v. *pai*, Nos. 685, 686.
2 Grierson, *Linguistic Survey of India*, vol. iii, pt. i.
3 *Epigraphia Birmanica*, vol. i, pt. i, p. 23 (Myazedi Inser., Pillar A, 1.2).
4 We can hardly question the common origin of the Tibeto-Burman and Taï-Chinese forms, in view of the closely similar series for the number "eight", which is in classical Tibetan *brgyad*; in eleventh-twelfth century Burmese *het, yhat, hyat,* or *rhac*; in sixth century Chinese *p^{w}at*, in Siamese from the thirteenth century *pèt.*
5 Houghton, *Essay on the Language of the Southern Chins*, p. 86, s.v. *p'yá.*
6 Duroiselle, *Ep. Birm.*, vol. i, pt. i, p. 27.
7 *JRAS.* 1913, "Note on the Numeral Systems of the Tibeto-Burman dialects," p. 331 ff. Cf. *Linguistic Survey of India*, vol. iii, pt. i, p. 622.
8 Or **parugyak*; for in view of such forms as Mikir *p'áró*, Aka *phogwa, purrua,* E. Dafla *lüg,* Chulikata Mishmi *malū,* it still seems doubtful if Siamese *ròy,* Laotian and Black Taï *hòy,* do not themselves go back to the same common origin as *pak.*

14

CONCERNING THE VARIATION OF FINAL CONSONANTS IN THE WORD FAMILIES OF TIBETAN, KACHIN, AND CHINESE

Stuart N. Wolfenden

Source: *Journal of the Royal Asiatic Society of Great Britain and Ireland* 4, 1937, 625–55.

The purpose of the present somewhat desultory notes may be said to be twofold: firstly, to emphasize the necessity of comparing the word stock of one Indo-Chinese language with that of another by word families only,[1] secondly, to make a preliminary investigation into certain variations of a particular type within such families, as there are here certain anomalies of which very careful note will have to be taken in any comparative work along these lines.

The importance of the word family rather than the single word in comparisons from language to language has, indeed, previously been brought forward by others,[2] but until quite recently its real significance has been only just glimpsed, no definite examples in illustration having been adduced.

To emphasize this particular matter at the outset we may here consider the following case.

In Tibetan we have the word family *ạbyed-pa*, P. and Imp. *ṗye, ṗyed, ṗyes*, F. *dbye* to open (vb. tr.), to separate, to keep apart, to divide, to distinguish, to classify, to pick out, to choose, to select, *ṗyed* half, *dbyad* an instrument to open the mouth by force, *dbyen-pa* difference, dissension, discord, schism, *ạbyer-ba*, P. and Imp. *byer* to disperse (as in flight), to scatter (vb. intr.), to flee in different directions, *dbye-r* (=*dbye-ru* in *dbyer-med* lit. devoid of difference) difference, distinction, *ạbye-ba*, P. and Imp. *bye* to open (vb. intr.), to divide, to separate, *dbye-ba* parting, partition, division, distinction, section, part, class, species, kind, while in Kachin we have *šă-byet* to separate into bundles, to make up into lots, *kă-čyan, lă-čyan* to divide, to deal out, to allot, to apportion, *gin-čyan* to divide, to exchange, *čyen* half, *a-čyen* band, shred, strip, *lă-ǰen* to divide equally, to apportion, to distribute, *mă-ǰen* (a cutting:) a clearing, *ǰe* to tear apart, to rend asunder,

to separate, to cut off, (*mă-ǰen ǰe* to make (lit. to cut) a clearing (*mă-ǰen*)), *ṗye* to allot, to assign, to parcel out.

There is a temptation here to equate Tibetan *ṗyed* half, with Kachin *čyen* half, and, indeed, if we were working with single words this is probably just what we should do. Following this, we should probably look for other examples of this supposed *-d*: *-n* equivalence in finals. But this would be completely wrong, and the results, on such an assumption, entirely without value, for what has happened here is that Tibetan has taken one member of the family to specialize into the meaning of "half" while Kachin has selected an entirely different member for a similar purpose. Tibetan *ṗyed* evidently belongs on the Kachin side most closely with *šă-byet*, while Kachin *čyen* must be regarded as having its nearest Tibetan relative in *dbyen-pa*.[3] But in dealing with single words we should probably never have arrived at this conclusion. Meaning must not be overemphasized at the expense of form. We can only get the correct perspective from the family as a whole, not from the individual members of the family.

From this instance alone, then—and there are hundreds of others of exactly similar type—we have clearly set before us the fact that before we can safely do anything in the way of equating single words with one another we must first be able to see as a whole the families to which the words in the proposed equation belong. Only after we have been enabled thus to understand, at least in part, how their members have been moving around within their own families, specializing their senses and changing their forms, can we ever really first propose with any safety anything in the way of equations between single members. In the sequel we shall adduce a certain number of representative families, but before entering upon this phase it may be well first to consider certain well-known facts setting limits to their assembly.

In the first place, as first clearly stated, I believe, by Simon,[4] Tibetan word families remain very true to their own single type of final, viz. guttural, dental, or labial, as the case may be, and it is impossible—except in certain peculiar cases, each of which requires its own special treatment—to establish anything in the nature of families with mixed finals.[5] This is a particularly valuable guide in instances where words without finals have obvious relatives with them, such, for instance, as *ą́kra-ba* and *mkraṅ-ba* or *ą́kraṅ-ba* hard, *bka* and *skad* speech, *sna* nose and *snam-pa* or *snom-pa*[6] to smell, for when once an open syllable form can be shown to have a definite relative in some particular class of final, say the guttural, there can be no turning back, and we must look with extreme suspicion, to say the least, upon any form which would lead us over also into a second, and unrelated, class of final. When, for instance, *ągyu-ba*, P. *ągyus* to move quickly, *ą́kyu-ba*, P. *ą́kyus* to run, *dkyu-ba* to run a race, lead us by means of *rgyug-pa*, P. *brgyugs*, F. *brgyug*, to run, to hasten, and *ą́kyug-pa*, P. *ḱyug* to run, to dart along rapidly (as fish), into the guttural final type, we cannot possibly include *ą́kyam-pa* to run about, to wander, and its relatives in final *-m*, even if the vowel alternation *u* ~ *a* ~ *o* were possible, which, in Tibetan, it almost certainly is not. This, in fact, is one principle which we must with considerable certainty rigidly enforce.

It must naturally be understood here that we are not speaking of anything but the older and more stable Tibeto-Burman languages, for, in some late dialects and in spoken Burmese, the finals have been converted into others of unrelated classes in an enormous number of cases. To see, then, into this problem with a greater feeling of certainty we shall be obliged, when extending our field beyond Tibetan itself, to select some other language or languages differing from it in sufficient degree to give us an idea of the direction in which development has proceeded, but yet which does not depart from the original stability of its finals by classes. For present purposes we will here take Kachin which stands sufficiently close to Tibetan to give us a great many families in common, while at the same time standing in relation to it in the matter of final consonants very much as does Archaic Chinese in the families put together by Karlgren. Of this we shall discuss the possible significance later; taking here the following selection of word families[7] common to Tibetan and Kachin.

Guttural finals

(1)[8] T. *ągug(s)-pa*, P. *bgug*, F. *dgug*, Imp. *ḱug* to bend, to make crooked, *kug*, *kug-kug* crooked, a hook, *ḱug*, *ḱugs* corner, concave angle, nook, creek, bay, gulf, cove, *ḱug-ma* pouch, small bag, *gug-ge-ba* bent.

K. *guk* to be bent, to be curved, *dă-guk* curve-horned, *ṅuk* to bow, to incline the head, *ṫiṅ-kuṅ* bent, to be bent, *gu* to be bent, to be curved, *a-gu* bent, humped, *lă-gu* to hang down, to droop, *ṅu* to bow, to incline the head, *kă-ṅu* to be bent, to be bowed down, to hang the head.

(2)[9] T. *ągog-pa*, P. *bkog*, Imp. *ḱog* to peel, to tear off, to remove by force, to pull off, *kog-pa* (vb. neut.) to peel, to splinter off, to scale off, to fall off in chunks, *kog-pa*, *skog-pa*, *skogs-pa* peel, rind, bark, shell, *gog-pa* to crumble off, to scale off (as plaster), *gog-po* dilapidated, ruinous, *ko-ba* hide, skin, leather.

K. *gå* to peel, to skin, to flay.

(3)[10] T. *kyog*, *kyog-po* crooked, bent, winding, *kyag-kyog* curved, crooked, *ḱyog-po*, *ą́ḱyog-po* crooked, curved, bent, *skyog-pa* to turn (as the neck), *skyogs* scoop, ladle, drinking cup, bowl, goblet, *gyog-pa* curved, crooked.

K. *ḱyåk* (crooked surface:) to be sunken, to be concave, to be depressed, *lă-ḱyåk* hole, depression, concavity, *gyåk* to be concave, to be cup-shaped, *lă-gyåk* shallow hole, depression, to be full of concave holes (as a road), *lă-gyå* to be contracted, to be crooked (as a paralysed limb).

(4) T. *bkrag* brightness, lustre, beautiful appearance or colour, *bkra-ba* glory, beautiful, fine, well, of good appearance.

K. *ḱrak* to be good, *grak* to be beautiful, fine, or good, beautiful, fine, well.

(5)[11] T. *ṅag* speech, talk, word, *mṅag-pa*, P. *mṅags* to commission, to charge, to delegate, *sṅag(s)-pa*, P. *bsṅags*, F. *bsṅag*, Imp. *sṅog* to praise, to commend, to extol, to recommend, *sṅags* incantation, *sṅo-ba*, P. *bsṅos*, F. *bsṅo*, Imp. *sṅos* to bless, to pronounce a benediction upon, to dedicate, *bsṅo-ba* a blessing, a benediction, *ṅa-ro* loud voice, cry.

K. *ṅa* to say, to speak, to tell, to relate.

(6)[12] T. *gčog-pa*, P. *bčag*, Imp. *čog(s)* to break, to smash, to crack, to split, *ąčag-pa*, P. *čag(s)* to be broken, to be smashed, to be cut off, *čogs-pa* to be broken, *ąčag-pa*, *ąčeg-pa*, P. *bšags*, F. *bšag*, Imp. *šog* to cut, to split, to cleave, *ąčog* (cutting off:) wall, partition, *čag-čag*, *čag-pa* ("breaking" the lips:) smacking the lips (in eating), *ąǰog-pa*, P. *(b)žogs*, F. *gžog*, Imp. *žog* to cut, to chop, to hew, to chip, *lčag* rod, switch, whip, stick, strike, cut, blow, hit, *ąťsog-pa*, P. *btsags*, F. *btsog*, Imp. *ťsog* to hew, to chop, to cut, to pierce, to cudgel; (also *mťsog-pa*) to find fault with, to blame, to censure, to tease, *mťsog-ma* (cleft:) fontanel, *gšog-pa*, *bšog-pa*, *gšag-pa*, *gšeg-pa*, *ąčegs-pa*, P. *gšags*, *bšags*, F. *gšag*, *bšag*, Imp. *gšog* to cleave, to split, to break open, to break through, to rend, to tear, *bžag-pa* to tear, to wear (of clothes), to burst, to crack, to split (vb. intr.), *šag-ma* (broken up:) gravel, pebbles, small stones.

K. *ťåk* to be broken, to be snapped in two, *kă-ťåk* to break (vb. tr.), to smash, *ťek* to snap, to break, *kă-ťek* ("break" the fingers:) to snap the fingers, *a-ťek*, *lă-ťek*, id.

(7)[13] T. *ądzugs-pa*, *zug-pa*, P. *btsugs*, *zugs*, F. *gzugs*, Imp. *zug(s)* to insert young plants into the ground, to plant, to erect (as a pillar by setting it into the ground), to prick, to stick into, to thrust into, to pierce, to penetrate, to bore, to sting, *ąťsugs-pa*, P. *ťsugs* to go into, to enter, to bore into, to take root in, *mdzug-gu* (also *mdzub-mo*) (inset:) finger, toe, claw, *ąǰug-pa*, P. *bčug*, F. *gžug*, Imp. *čug* to put into, to insert, to inject, *ąǰug-pa*, to be combined with, to have added to, to go in, to walk in, to enter.

K. *čyåk* to pierce, to be pierced, to be run through (as with a spear), *a-čyå* to prod, to goad, to thrust, to prick.

(8)[14] T. *gtsaṅ-ba* to be clean, to be pure, cleanness, purity, clean, pure, *gtsaṅ* clean, pure, *ąťsaṅ-ba*, P. *saṅs* to make clean, to remove impurities, to take away, to remove, *ťsaṅs* purified, clean, pure, *saṅ-ba*, P. *(b)saṅs*, F. *(b)saṅ* to clear away, to remove dirt, to cleanse, *seṅ-po*, *bseṅ-po* clean, white.

K. *seṅ* to be clean, to be pure, to clean, to clear away, clear space, a clearing, *tseṅ* to remove, to clear away, *tsån* to remove, to clear away, to put away, to drive away, to banish.

(9)[15] T. *gtug-pa*, P. *gtugs* to reach, to touch, to come up to, to overtake, to meet with, to join, *btug-pa* id., *ťug-pa* to reach, to arrive at, to come to, to meet, to light upon; > col.: to touch, to hit, to strike against, *ťug* until, to, *rdug-pa*, P. *brdugs*, F. *brdug* to strike against, to stumble over, *rduṅ-ba*, P. *brduṅs*, F. *brduṅ*, Imp. *(b)rduṅ(s)* to beat, to strike, to cudgel, to hammer, to pound, to thrash, *bduṅ-ba* id.

K. *ťuk* to hit against, to come in contact with, to come close up to, to reach, to attain, to arrive at or in, until, as far as, up to, *kă-ťuk* to hit against, to collide with, *ťu* to pound, to hammer, to pulverize, *a-ťu* to hit, to strike, to kick against.

(10)[16] T. *ťogs-pa* to strike, to stumble, to run against, *rdeg(s)-pa*, P. *(b)rdegs*, F. *brdeg*, Imp. *(b)rdeg(s)* to beat, to strike, to knock, to kick, *mťo-ba* a hammer.

K. *ťåk* to touch lightly, *kă-ťåk* id., *a-ťåk* to touch, to strike lightly, to tap.

(11)[17] T. *ąťegs-pa* to lift up, to raise up, (mounted:) to set out on a journey, *ťeg-pa* to lift, to raise, to hold up, to support, vehicle, carriage, riding animal, *ądeg(s)-pa*, P. *bteg(s)*, F. *gdeg*, Imp. *ťeg* to lift, to raise, to elevate, to support, *ťog* (uppermost part:) top, roof, *ťog-ma* summit, upper end, uppermost place, origin, source, beginning, *tog* top, top ornament, *ltag* above, over, *ltag-pa* upper place, upper part, back of the neck (i.e. top of the back), back ("top") of a knife.

K. *ťåk* to brace, to prop up, *ťek* to rise, to get up.

(12)[18] T. *ṗuṅ-po* a heap, a pile, a mass, *spuṅ-ba*, P. and Imp. *spuṅs* to heap up, to pile up, to accumulate, to amass, *dpuṅ* host, a great number, troops, army, *buṅs* mass, heap, bulk.

K. *ṗuṅ* gathering, assembly, congregation, flock, herd, *ṗåṅ* to gather, to congregate, to assemble, *ǰă-ṗåṅ* to collect, to cause to congregate, *n-ṗåṅ* a clump, a cluster, *mă-ṗåṅ* a clump, a cluster, *båṅ* to confer with, to consult, to hold council, *ṗå* to mix, to mingle, *să-ṗå* to consult, to confer, to hold council, *bå* to pack, to get one's goods together.

(13) T. *ąbyug-pa*, P. and Imp. *byugs* to stroke, to pat, to tap lightly.

K. *a-ṗuk* to tap, to pat, *gŭm-ṗuk* to tap, to pat.

(14)[19] T. *ąbug(s)-pa*, *ąbig(s)-pa*, P. *ṗug*, *ṗigs*, F. *dbug*, *dbig*, Imp. *ṗug*, *ṗig(s)* to sting, to pierce, to bore, to make a hole, to break into, to break open, *sbug-pa* id., *ṗug-pa*, *ṗig-pa* = *ąbug(s)-pa*, *ąbig(s)-pa*, *ṗug-pa* cave, cavern, *ṗug(s)* innermost part, inmost apartment, *sbug(s)* (more frequently *sbubs*) hollow, cavity, excavation, recess, interior space, *sbug-po* hollow, *ąbug* awl, chisel, punch, *bug-pa*, *bu-ga* hole, orifice, aperture, *bu-gu* hole, *sbu-gu* hollow, tubular cavity (as in plant stems).

K. *påk* hole, cavity, *kă-påk* to dig out, to hollow out.

(15)[20] T. *ąṗog-pa*, P. *ṗog* to hit, to strike, to touch, to meet.

K. *ṗåk* to hit with a thud, to whack.

(16)[21] T. *smag* dark, darkness, *mog-pa* dark-coloured, *rmoṅ-ba*, P. *rmoṅs* to be obscured, *rmoṅ-ba* obscurity, obscured.

K. *måṅ* to rise in a cloud (as dust or smoke), *kă-måṅ*, *gă-måṅ* id., *šă-måṅ* to stir up dust, to spread defamatory reports, *lă-må* to be black, to be dark, to be shaded, *n-må* to be stained, to be soiled.

(17)[21] T. *rmugs-pa* dense fog, languor, languid, *smug-pa* fog, *smug-po* dark bay, cherry brown, purple brown, *mug-pa* overcast, troubled (of the mind: *yid mug-pa* to despair), *rmu-ba* dullness, heaviness, fog.

K. *muk* to be sullen, to be sulky, to be sour-tempered, *muṅ* to be sullen, to be sulky, to be cloudy, to be overcast, to be dull, *mu* to be cloudy, to be overcast.

(18)[22] T. *mig*, old form, *dmyig* eye.
K. *myi*, *a-myi* eye.

(19) T. *miṅ*, old form *myiṅ* name.
K. *myiṅ* to name, *a-myiṅ* a name.

Labial finals

(20)[23] T. *ągebs-pa*, P. *bkab*, F. *dgab*, Imp. *ǩob* to cover, to spread over, to put on, to protect, *sgab-pa* id., *gab-pa* to hide, to conceal oneself, *ągab-pa* (cover one's actions:) to be cautious, to take care, *ąǩeb-pa*, P. *ǩebs* to cover, to spread over, *ǩebs*, *ǩyebs* covering, coverlet, *skyob-pa*, P. *(b)skyabs*, F. *bskyab*, Imp. *skyob(s)* to protect, to defend, to preserve, to save, *skyabs* protection, defence, *skyobs* help, assistance.

K. *gap* to cover, (put up a covering:) to build a house, to pitch a tent, *dă-gap* to cover (especially with something wide and flat), *mă-gap* a cover, a lid, to cover, *lă-gap* (coverer:) one who brings up the rear of a Kachin army, *tiṅ-gap* a cover, a lid, *kåp*, *a-kåp* (covering:) crust, rind, shell, *mă-kåp*, *mă-ṅåp* to cover, to shield, to defend, defender, defence, protection, *ǩiṅ-kåp* a sheath (usually of bamboo), *gup* to be covered, to be doubled over, *gup*, *n-gup* (coverer:) mouth, *ǩa-gup* a hat (*ǩa-gup gup* to wear a hat), *kă-gup*, *lă-gup* a hat, *dă-gup* to cover, to envelop, *mă-gup* (all-embracing:) every, all, *n-kup* (cover:) to turn over, to fold over, *šă-gup* to fold over, to double over, *kă-yåp* to wrap up, to cover, to encase, *yåp* to be wrapped around, to be wound around.

(21) T. *skum-pa*, P. *bskums*, F. *bskum*, Imp. *skum(s)* to contract, to draw in, to crook, to bend, *kum-pa*, *kum-po*, *kum-kum* crooked, shrivelled, *ąǩum-pa*, P. *ǩums* to shrink, to be contracted, *ǩums-pa* crooked.

K. *gum* to bow, to bend forward, to make obeisance, *n-gum* to bend forward, to be flat on the face, *n-ṅum* to bend forward, *dă-gum* bent, curved, concave, arched, *šă-gum* bent, curved, concave, arched, to lie prone, to lie on the face, *mă-gum* the ridge or comb of a house, *mă-kum* a ridge, *n-ṅup* to bend forward, *dă-gup* to bow low, to sit or kneel with the face towards the ground, to make obeisance.

(22)[24] T. *bgom-pa*, P. *bgams* to walk, to step, to stride, *gom-pa* a pace, a step.
K. *n-gam* a step, a notch (of a ladder), *lă-ǩam* to step, a step, a pace, *ǩåm* to walk, to travel.

(23) T. *ągrib-pa* to grow less, to decrease, to diminish, to grow dim, to become dark, *sgrib-pa*, P. *bsgribs*, F. *bsgrib*, Imp. *sgrib(s)* to darken, to obscure, *sgrib-pa* darkened, dark, obscured, (mental darkness:) sin, *grib* shade, shadow, filth, defilement, contamination, *srib-pa* to grow dark, to become dusk, *srib(s)* darkness, gloom, night, shady side, north side, *rab-rib* (> col. *hrab-hrib*) (lessened visibility:) mist, dimness.

K. *krip* to diminish, to subside, to become less, to die down (as a fire), *šă-krip* (take down:) to humble, to punish.

(24) T. *sgrim-pa*, P. *bsgrims*, F. *bsgrim*, Imp. *sgrim(s)* to hold fast, to force together, to twist together, *ǩrims* moral law, custom, duty, precept, rule (i.e. restrictions or prohibitions restraining one's actions).

K. *ǩrim* to act in unison, all at once, with one accord, *grim* to act in unison, *ǩrip* to act in unison, *a-ǩrip* unison, concord, unity, harmony, agreement, *lă-ǩrip* to keep step.

(25) T. *ągrum-pa* to pinch off, to nip off, to prune, to lop off, to clip, *dkrum* broken, *grum-pa* lame, crippled, *hrum-pa* to break, to smash, *ąǩrums* (torn:) carcass, game torn by wild beasts, *skrum*, *srum* (carved off:) meat (as food).

K. *krum* to trim, to prune, to lop off.

(26) T. *ągrem(s)-pa*, P. *bkram*, F. *dgram*, Imp. *ǩroms* to put in order, to lay out, to spread out together, to place together (as articles for sale), *ǩrom* market-place (where objects are laid out together for sale), crowd, assemblage, gathering.

K. *ǩram* to spread out (as a trailing plant on the ground), to become bushy (as trees), *ǩram* (the spreader:) hamadryad, *ǰă-ǩram* to widen, to spread, to attain full size, *ǩrem* to be side by side, to be in line, *lă-ǩrem* to come up beside, to edge up to, *n-ǩrem* side of the body, edge or back of a book, *ǩrep* to be in line, to be in a row.

(27) T. *ǩrab-ǩrab* a weeper, one given to tears.
K. *ǩrap* to weep, to cry.

(28) T. *ăǰum-pa*, P. *bčum*, F. *gžum*, Imp. *čum* to shrink, to contract, to draw in, *ąčum(s)-pa* to shrink, to contract.
K. *čyum* to be puckered up (as the lips), to be drawn up, to be contracted.

(29)[25] T. *ądzom(s)-pa* to come together, to meet.
K. *čyåm* to act together, to act in unison, *ǰåm* to join forces, to co-operate, together, in unison, in company, *kă-ǰåm* to gather round, to crowd round (a common centre), *čyå* to be joined, to be united, to be fitted together, to be related (as by family ties), *šă-čyå* to fit together, to adjust one thing to another, *šiṅ-čyå* to tie together, to bind up together, *mă-čyå* (meeting place:) socket, vulva, *ǰå* to be collected together, to be massed together, *mă-ǰå* to gather together, to tie together, top-knot, *n-ǰå* top-knot.

(30) T. *ńams-pa* injured, damaged, spoiled, impaired, stained.
K. *nyam* to be decayed, *čyă-nyam* decayed, crumbling, to be decayed, *nyåm* to be decayed, *čyă-nyåm* weak, failing, tottering, *nyåp* to be decayed, to be crumbling, to be broken down.

(31)[26] T. *sdeb-pa*, P. *bsdebs*, F. *bsdeb*, Imp. *sdebs* to join, to unite, to put together, to mix, *ldeb-pa* to bend back, to double down, *deb-ma* poultice, application, *lteb-pa* to double down, to bend back, *ťebs* (put together:) series, succession, *ąťeb* (added member:) surplus, extra, supernumerary, *ltab-pa*, P. *bltabs*, F. *bltab*, Imp. *ltob* to fold, gather, or lay together, *ądab-ma* flat board, wing, petal, leaf, fan, flag, *ądabs* side, surface.

K. *ťep* to be close together, to be near, *kă-ťep* to be close together, to come close up to, *ǰă-ťep* to bring close together, *a-ťep* (bring the hands together:) to clap, *tep*, *mă-tep* to pinch, to be squeezed together, to be

fastened together, *dep* to be close, to be crowded together, *šiṅ-dep* to be or act in unison, co-operation, *ṭap* layer, stratum, lamina, *kă-ṭap* to add, to superimpose, again and again, repeatedly.

(32) T. *gtam(s)-pa* full, *ltam(s)-pa*, P. *bltams*, F. *bltam* to be full, *ltem-pa* full, overflowing, *ṭam-pa* full, complete, *ṭem-pa* to be full, to be complete, to be finished, *ạṭems-pa* to be sufficient, to suffice.

K. *tåm* to have completed, to have finished, *šă-tåm* to finish, to complete.

(33) T. *snabs* mucus, discharge from the nose.

K.[27] *nep* mucus, discharge from the nose, *a-nep* id., *nyep* to be soft, mucus, *nyap* to be soft and paste-like, *a-nyap* soft, sticky, viscous, *mă-nyap* soft, moist, spongy, *nya* to be soft, to be pliable, *šă-nya* to soften, to make soft, *čyă-nya* soft, flabby.

(34)[28] T. *snub-pa*, P. *bsnubs*, F. *bsnub*, Imp. *snub(s)* to cause to perish, to suppress, to annul, to destroy, to abolish, to abrogate, *nub-pa* to fall gradually, to sink, to decay, to decline (as religion), to set (of sun and moon), *nub-mo* evening, *nub* the west, evening.

K. *nip* to cast a shadow, to be overcast, *šă-nip*, *šiṅ-nip*, *čyă-nyip* shade, shadow.

(35) T. *nom-pa*, P. *noms* to seize, to grasp, to lay hold of, *snom-pa*, P. *bsnams*, F. *bsnam* to take, to seize, to grasp, to take up.

K. *nåm* to be enclosed, to be housed, to be comprised in, to be held or contained in, *kă-nåm*, *gă-nåm* to gather, to amass, to collect, to hoard.

(36) T. *ạṗam-pa*, P. *ṗam* to be beaten, to be conquered, to be overcome, to be deprived of power to act (as *bdud* demons).

K. *ṗam* to be numb, to be benumbed, to be without the power of feeling, *kă-ṗam*, *gă-ṗam*, id.

(37)[29] T. *ạṗrab* a fluttering movement (as in *ạṗrab byed-pa* to flutter (of a wounded bird)).

K. *ṗrap* to flutter, to flicker, to flash, *n-ṗrap* lightning, *kă-ṗrap* to blink (as the eyes), to flutter (as wings), to move back and forth (as animals their ears), *kă-prap*, *prap* to flash (as lightning or a mirror).

From this selection of families in guttural and labial finals we see that the allowable alternations include only *-g* (> K. *-k*) ~ *-ṅ* ~ -O, and *-b* (> K. *-p*) ~ *-m* ~ -O, a fact which we shall have to bear well in mind when dealing with the families in dental finals, where the range is considerably wider, as is exemplified in the following cases.

Dental finals

(38)[30] T. *mgur* (bending part:) neck, throat, *dgur*, *rgur*, *sgur* crooked, *mgul* neck, throat, *dgu-ba* to bend, to make crooked, bow, inflection, bent, stooping.

K. *kun*, *tiṅ-kun* to be bent, to be curved, *tiṅ-kun* to bend, to be pliable, to wriggle, to twist, *šiṅ-kun* (assume a bending position:) to crouch down, to prowl, *mă-kun* to crouch down, *ku* to be bent, to be curved.

(39)[31] T. *rkod-pa*, P. *(b)rkos*, F. *brko*, Imp. *rkos* to dig, to dig out, to hoe, to engrave, *rkon-pa*, *skon-pa* (scooped out:) basket, *rko-pa* = *rkod-pa*.

K. *gåt* to be scooped out, *dă-gåt* to scoop out, to ladle out, *šă-gåt* to scoop up with the hand, *lă-gåt* to scoop, a scoop, a small shovel, *n-gåt* tray, shallow bamboo basket, *mă-gån* to scoop up, to collect into a heap, *n-gån* bamboo spade (used for digging graves), *gån* to be sunken, to be concave, *diṅ-gån* to undulate (as waves).

(40) T.[32] *rkun-ma* thief, theft, *rku-ba*, P. *(b)rkus*, F. *brku*, Imp. *rkus* to steal, to rob.

K. *ǩut* to rob, *lă-gut* thief, robber, *lă-gu* to steal.

(41) T. *skar-ma* star, constellation.

K. *šă-gan* star, meteor.

(42)[33] T. *skad-pa* to say, to tell, to relate, to name, to call, *skad* voice, speech, language, talk, words, *bka* word, speech, *sgo-ba*, P. *bsgo* to say, to bid, to order, *sko-ba*, P. *(b)skos* F. *bsko*, Imp. *skos* to appoint, to nominate, to name, to commission.

K. *ga* word, speech, language, *a-ga* word, instruction, command, order, *šă-ga* to call, to summon.

(43)[34] T. *ạǩyil-ba* to be twisted, to be winding, to be spiral, *skyil-ba*, P. and F. *bskyil* to bend, to make crooked, (encircling:) to pen up, to shut up or in, to dam up (also > to retain, to detain).

K. *kyit* to girdle, to gird, to tie round, *šiṅ-kyit* belt, girdle, sash, *mă-kyit* to tie, to make a knot, *gyit* to tie up, to bind together, *a-kyin* to roll into a ball, to make round, *kyin* a package, a bundle, *kă-ǩyin* to gather (lit. roll) into a heap, *gyin* to roll into ball form, to make pellet-shaped, *šiṅ-gyin* to roll together, pellet, shot, *šă-gyin* ("belt" in:) to tighten a belt, to shorten a strap, *n-gyin*, *kum-gyin* (roll-shaped:) cucumber, *yin* to be turned around, to encompass, to make a circuit, *kă-yin* to turn around, to rotate, *diṅ-yin* to be dizzy, to be giddy, *lă-yin* a four-armed reel for reeling yarn, *gyi* to be curved, to be crinkled, *mă-gyi* to be curled, to be spiral, to be kinked, *tiṅ-gyi* id., *n-gyi* (circular form:) a picture of the sun.

(44)[35] T. *bgrad-pa* to scratch, to scrape, *hrad-pa* to scratch, *ạbrad-pa*, *ạdrad-pa*, P. *brad*, Imp. *brod* to scratch, to scrape, to gnaw, to nibble at, *sbrad-pa*, *dbrad-pa* = *ạbrad-pa*.

K. *gret* to graze (as a bullet), *a-gret* to scratch (as a thorn), to graze (as a bullet), *diṅ-gret* to touch, to rub against, to graze, *kret* to gnaw, gnawing, *mă-kret* to gnaw, *ǩret* to rasp, to grate, *a-ǩret* to gnaw, *mă-kret* (scratching:) to draw a line, to rule, to strike a match, a ruler, *pret* to rasp, to grate, *n-ṗrat* to strike a match, to strike fire (as with flint and steel), *bret* rasping noise (as when tearing silk), *mă-ret* to scratch, to lacerate, *rat* to scratch, to wound, to lacerate, to cut, *a-ṗre*, *a-ṗri* to scratch, to dig (as with claws or fingers).

(45)[36] T. *ạǩrud-pa*, P. *bkrus*, F. *bkru* to wash (as *gos* clothes), to bathe (as *ka-lag* face and hands), *ǩrus* bath, washing, ablution, *ạǩru-ba* = *ạǩrud-pa*.

K. *ǩrut* to wash (as clothing, or *buṅ* (< *puṅ*) the head).

(46) T. *sred-pa* to desire, to wish for, desire, wish, *ąkren-pa* to wish, to desire, to long for.

K. *mă-rit* to desire, to long for, to hanker after, *mă-rin* to be covetous, to be avaricious, to be greedy.

(47) T. *ągril-ba*, P. *gril* to be twisted, to be wrapped round, to be rounded, to be turned, *sgril-ba*, P. and F. *bsgril* (vb. tr.) to wind round, to wrap round, to wrap up, to wind up, *ąkril-ba* (vb. intr.) to wind round, to coil, to embrace, to clasp, *gril* a roll, *hril-po* round, globular, *ądril-ba*, P. *dril* to be turned round, to be rolled round, to be twisted, to wrap up, *ril-ba*, *ril-po*, *ril-mo* round, globular, cylindrical, *dkri-ba* to wind, to wind up, *ąkri-ba* to wind, to roll, to twist, *sri-ba* to wind round, to wrap round.

K. *rit* to twist, to wind, to twine, *tă-rin* to roll, to be rolled, *kri* to revolve, to spin, *mă-kri* a braid, *kri* to braid, *šiṅ-ri*, *sum-ri* a cord, a rope, *ri* thread, string, cord.

(48) T. *blud* release, ransom, *blus-ma* anything ransomed, *blu-ba*, P. *blus* to ransom.

K. *lăt* to get loose, to escape, to become free, *šă-lăt* to liberate, to set free, *n-lăt*, *šaṅ-lăt* freedman, escaped prisoner, *mă-lă* to loosen, to become loose.

(49) T. *ṅur-ba* to grunt (as pigs or yaks), *sṅur-ba* to snore.

K. *ṅut* to grunt (as a pig), *ṅun* to growl, to grumble, to murmur, *a-ṅun* grumbling, growling.

(50)[37] T. *ątsud-pa*, P. *tsud* to be put into, to go into, to enter, *čud-pa* id., *ądzud-pa*, P. *btsud*, *zud*, Imp. *tsud* to put into, to lay in, *ąǰud-pa* id., *ądzu-ba*, P. *ądzus*, to enter, to go in.

K. *ǰut* to be pierced, *mă-ǰut*, *num-ǰut* to pierce, to thrust in, *šă-ǰut* to pierce, to thrust through, *ǰu* to prick, to be pricked, thorn, bramble, spike, *a-ǰu* thorn, bramble, spike, to prick.

(51) T. *ąǰun-pa*, *gčun-pa*, P. *bčun*, F. *gžun* to subdue, to make tame, to make soft, *žun-pa* melted, *ąčun-pa* to be tamed, to be subdued, to be made to yield, *ąǰu-ba*, *žu-ba*, *bžu-ba*, P. *bžus*, F. *bžu* to melt, to digest, *ąǰu-ba* digestion.

K. *tun* to dissolve, to melt, *šă-tun* to liquefy, to melt.

(52)[38] T. *mtsan* grandchild, nephew, *btsas-ma*, *rtsas-ma* (brought forth :) harvest, wages, pay, *btsa-ba*, P. *btsas* to bring forth, to bear, to give birth to, *tsa-bo* grandson, *tsa-mo* grand-daughter, niece.

K. *ša*, *a-ša* child, son, daughter, nephew, niece, *kă-ša* child, young of animals, *mă-ša* human being, man.

(53)[39] T. *rṅed-pa*, P. *brṅed*, *brṅes*, F. *brṅed* to get, to obtain, to meet with, *sṅen-pa*, *bsṅen-pa* to come near, to go near, to approach, *gṅen*, *ṅen* relative, kinsman, *gṅer-ba* to apply oneself to, to take pains with, to procure, to acquire, *ṅe-r* (= *ṅe-bar*) near, *ṅe-ba* to be near, to approach, *sṅe-ba*, P. *bsṅes*, F. *bsṅe*, Imp. *sṅe* to lean against, to lie down on.

K. *ni* to be near, *a-ni* nearness, proximity, to come near, to approach, *šă-ni* to bring near, to put in proximity, *nyē*, *ṇya* to meet.

(54) T. *gsal-ba* to be clear, to be bright, to be distinct, clear, bright, pure, *sal-le-ba* clear, bright, brilliant, *sel-ba*, P. and F. *bsal*, Imp. *sol* to cleanse, to remove impurities, to clear.

K. *san* to be clear, to be transparent, to be pure, *a-san* clear, clean, pure, *ǰă-san* to clear, to purify, *tsan* to clear, to cleanse.

(55)[40] T. *gzan-pa* to eat, to devour, to gnaw, *bzan* food of animals, pasture, pasturage, *zan* pap, porridge, fodder, an eater, *ąǰan-ba* to swallow, to devour, *zas* food, nourishment, *za-ba*, *bza-ba*, P. *zos*, *bzas*, F. *bza*, Imp. *zo*, *zos* to eat, food, meat, victuals.

K. *šat*, *a-šat* (food:) boiled rice, *n-šat* food supply, *mă-šat* food basket (figurative name), *šan*, *a-šan* flesh, meat, *ša* to eat.

(56)[41] T. *sdud-pa*, P. *bsdus*, F. *bsdu*, Imp. *sdud*, *bsdu* to put together, to join, to unite (others, e.g. *kyo-šug-tu* as husband and wife), to marry, to compress, to condense, *sdud* (pressing together:) folds of a garment, *mdud* knot, bow, *dud-pa* to tie, to knot, *ądun-ma* council, association, assembly, meeting, *mdun* (meeting one:) fore-part, front side, *mdun-ma* wife, *ądus-pa* assembly, gathering, meeting, *bsdus-pa*, *ądus-pa* to consist of, to be made up of, *ądu-ba*, P. *ądus* to assemble, to come together, to meet, to join one another (e.g. *kyo-šug-tu* as husband and wife), to get married, to be pressed or crowded together, *ądu-ba* assembly, gathering, meeting.

K. *tut* to be joined, to be bound together, to be united, *mă-tut* to connect, to join, to link together, *kă-tut* to meet, to run up against, (urge to meet:) to press on, to push on (towards a goal), *dun* to tie together, to tether, to be connected, to be united, to join, to adhere to.

(57)[42] T. *sdod-pa*, P. and F. *bsdad* to sit, to stay, to remain, to abide, to halt, to come to a stop.

K. *kă-tăt* to stumble and fall, *mă-tăt* to stub the toe or foot, to kick against an obstruction, *kă-dăn* to stumble, to stop abruptly, (of *ga* speech:) to stutter, to stammer, *kă-tă*, *gă-tă* = *kă-tăt*.

(58)[43] T. *mṫon-po* high, elevated, *mṫos* high, elevated, *mṫo-ba* to be high, height, high, elevated.

K. *kă-ṫan* to bound up (as a ball), to leap (as a frog), *ṫan* to raise, to put up (as a ladder), *ṫa* to be above, to be high, to rise, *lă-ṫa* upper, *kă-ṫa* above, overhead, *mă-tă* pinnacle, summit, high, elevated, *tă* (raised:) to transport, to carry, *tsă* to be high, to be tall, to be lofty, *mă-tsă* upper regions, celestial heights, *šă-tsă* to heighten, to raise, to elevate.

(59)[44] T. *ądrud-pa*, *ąbrud-pa*, P. and Imp. *drud* to rub, to file, to rasp, to scour, to polish, to smooth, to plane, to drag, to pull along, *bgrud-pa*, P. *bgrus*, F. *bgru* to husk, to shell.

K. *rut* to rub, *a-rut* to rub, to abrade, to erase, *mă-rut* a grater (*mă-rut rut* to grate), *mă-krut* to gnaw, *krut* to snatch away, to pull away forcibly, whetstone, *n-krut* whetstone, *krit* to grind (as *wa* the teeth), *a-rit* to rub off, to pull off (as bark from a tree).

(60)[45] T. *ạdred-pa* to slide, to slip, to glide, *ạdren-pa*, P. *dran(s)*, F. *dran*, Imp. *dron(s)* to draw, to drag, to pull, to tear out, to press out, to squeeze out (as pus), to conduct (as water), to lead, to guide, to fetch, to transport.

K. *ret* to be snatched up, *kă-rạt*; *gă-rạt* to draw, to drag, to pull, to haul, *să-rạt* to drag, to scrape (as the feet), *krạn* to be dragged, to be pulled, to be led along, to be conducted, *gă-re* to tear away, to snatch away, to pluck away from.

(61) T. *gnas-pa* to stay, to remain, to dwell, to live at, *gnas* place, spot, abode, dwelling place.

K. *nat* to be fixed, to be held firmly, *mă-nat* to grasp firmly, to hold tightly, *ǰă-nat* for ever, always, constantly, *nan* to stay, to remain, to tarry, to be permanent, *mă-nan* always.

(62) T. *ạbar-ba* to open (of flowers), to bloom, to blossom. K. *pan*, *ban*, *nam-pan* a flower, *lă-pan* (Ḳauri), id.

(63) T. *ạbar-ba* to burn (intr.), to catch fire, to be ignited, *sbar-ba*, *sbor-ba*, P. and F. *sbar* to set fire to, to kindle, to light.

K. *lă-wạt* cooking place, *wan*, *a-wan* fire, *wạn* (Ḳauri), id.

(64)[46] T. *ạbud-pa*, P. *bus*, *ṗu(s)*, F. *dbu*, Imp. *ṗu(s)* to blow (intr. and tr.), to remove by blowing (as chaff), *sbud-pa* bellows (*sbud ạbud-pa* to blow bellows), *bud* (blown by wind:) a cloud of dust, *spun-pa*, *sbun-pa* (blown by wind:) chaff, husks, *sbur-ma* = *sbun-pa*, *ṗu* a puff of breath, *pu-ṫse* husks of barley, bran.

K. *wut*, *kă-wut*, *gă-wut* to blow, to puff, to blow upon (as a fire), *n-bun* (blown by wind:) dust, *bun* to scatter dust, to sprinkle liquids, *wu* dust, fine ash (*wan wu* fine ashes from a fire (*wan*)), *n-ṗu* dust, *a-ṗu* to scatter dust, to remove dust (as by shaking a garment).

(65)[47] T. *pus-mo* knee.

K. *ṗut* knee, to kneel, *lă-ṗut* knee, *ṗun* to cover (as with a blanket), to put on (as a coat), *bu* to put in place, to place upon (as on a shelf).

(66) T. *ạṗar-ba* to bound up, to leap up, to fly up (as sparks), *spar-ba*, *spor-ba*, P. and F. *spar* to lift up, to raise, *ạṗyar-ba*, Imp. *ạṗyor*, *ṗyor* to raise, to lift up, to hoist, *ạčar-ba*, P. *šar* to rise, to appear, to become visible, (of the sun:) to shine forth, *šar* the east.

K. *n-ṗat* (bring up:) to vomit, *ṗan* to rise, to be raised (as dust by the wind), *kă-ṗan* to be in motion, to be astir (as a crowd).

(67) T. *spel-ba* to augment, to increase, to add to, to multiply, to put together, to spread, to propagate, *ạṗel-ba*, P. *ṗel* (vb. neut.) to increase, to grow, to become larger, to improve, to grow better, *dpal* glory, splendour, magnificence, abundance, wealth, welfare, *ṗal-pa* (widely distributed:) usual, common.

K. *ǰă-ṗat* (enlarge the mind:) to instruct, to teach, *ǰă-ṗan* id., *šă-pan* to bring up (as a child), to rear, *pan* to be enlarged, to grow, to mature, *ṗan* to be mature, to be developed, to form, to create, to cause to be, *ba* to be big, to be great, to be large, *mă-ba* (great man:) chief, ruler, *kă-ba* big, great, large, *šă-pa* (vb. tr.) to extend, to spread out, to expand, to enlarge, *n-ba* great, big and ferocious.

(68)[48] T. *ạbyid-pa*, P. *byid*, *ṗyid* to disappear, to pass away, to make an exit, *dpyid* (time of budding forth:) Spring, *ạbyin-pa*, P. and Imp. *ṗyuṅ*, F. *dbyuṅ* to cause to come forth, to take out, to remove, to draw out, to pull out, to tear out, to produce, to bring to light, to send out, to emit, to shed (*mči-ma ạbyin-pa* to shed tears), to draw (as *krag* blood), to drive out, to turn out, to expel, to throw away, to liberate, to release, *sbyin-pa*, P. and Imp. *byin* to give to, to bestow upon, to hand over to, to deliver over to, *ạbyuṅ-ba*, P. and Imp. *byuṅ* to come out, to emerge, to go to, to proceed to, *ṗyin-pa* to go forth, to proceed, to advance, to come to, to reach, *dpyis* end, conclusion (*dpyis ṗyin-pa* to reach the end), *ạṗyi-ba*, *ạṗyid-pa* to wipe away (*mči-ma ạṗyi-ba* to wipe away tears), to blot out, to pull out, to tear out (as *rlig* the testicles), to remove, *ạbyi-ba*, P. *byi*, *ṗyi*, *ṗyis* to be wiped off, to be blotted out, to be effaced.

K. *ṗyet*, *kă-ṗyet* to vanish, to disappear, *šă-byet* to extort, to take away, to levy (as a fine), *šă-pyen* to throw out, to eject, to drive out, to expel, *pyen* to remove, to strip off.

(69)[49] T. *gšid* funeral, *gšin-po* a dead man, one deceased, *ạči-ba*, *ši-ba*, P. *ši* to die, to disappear (as a flame), to cease, death.

K. *si* to die, to expire, *a-si* death, dead, *čyă-si* dead, a dead person.

(70)[50] T. *gčid-pa*, *gči-ba*, P. *gčis*, F. *gči*, Imp. *gčis* to make water, to urinate (*gčin gčid-pa*, *gčin gči-ba* id.), *gčin* urine.

K. *ǰit* urine, *ǰin* vagina, female private parts, *ǰi* to make water, to urinate (*ǰit ǰi* id.).

(71)[51] T. *mčin-pa* liver.

K. *sin*, *a-sin*, *mă-sin* liver.

(72) T. *ṗyen*, *ạṗyen* wind, flatulence.

K. *ṗyet*, *a-ṗyet* wind, flatulence.

(73)[52] T. *ạbyed-pa*, P. and Imp. *ṗye*, *ṗyed*, *ṗyes*, F. *dbye* (vb. tr.) to separate, to keep apart, to open, to divide, to distinguish, to pick out, to choose, to select, to classify, *dpyad* an instrument for opening the mouth by force, *ṗyed* (divided:) half, *dbyen-pa* difference, dissension, discord, schism, *ạbyer-ba*, P. and Imp. *byer* (vb. intr.) to disperse, to scatter, to flee in different directions, *dbye-r* (= *dbye-ru* in *dbyer-med*, lit. devoid of difference) difference, distinction, *ạbye-ba*, P. and Imp. *bye* (vb. intr.) to open, to divide, to separate, *dbye-ba* parting, partition, division, distinction, section, part, class, species, kind.

K. *šă-byet* to divide into lots, to do up into separate bundles, *kă-čyan*, *lă-čyan* to divide, to deal out, to allot, to apportion, *gin-čyan* to divide, to exchange, *čyen* (divided:) half, *a-čyen* band, shred, strip, *lă-ǰen* to divide equally, to apportion, to distribute, *mă-ǰen* a clearing, *ǰe* to tear, to rend asunder, to divide, to separate (as combatants), to cut down (as brush)

(*mă-ǰenǰe* to make (lit. to clear) a clearing (*mă-ǰen*)), *ṗye* to allot, to assign, to parcel out.

(74) T. *byed-pa*, P. *byas*, F. *bya*, Imp. *byos* to make, to do, to cause, to fabricate, *spyod-pa*, *spyad-pa*, P. *spyad* to do, to act, to accomplish, *ạčos-pa*, P. *bčos*, *ạčos*, F. *bčo*, Imp. *čos* to make, to construct, to manufacture, to build, *ạča-ba*, P. *bčas*, *ạčas*, F. *bča*, Imp. *čos* to make, to prepare, to construct.

K. *čyen* to do (obsolete word), *šă-čyen* to do, to accomplish, to perform, *lă-čyen* work, labour.

(75) T. *ạbral-ba*, P. *bral*, Imp. *brol* to be separated from, to be deprived of, to be parted from, to be bereft of, *brel-ba* to be destitute of, to be without, to be in need of, to be poor, *ạṗral-ba*, P. *ṗral*, F. *dbral*, Imp. *ṗrol* (vb. tr.) to separate, to leave, to take away from.

K. *ran* to be apart, to be separated, to be divided, *mă-ran* to separate, to push away, to send away, *šă-ran* to place apart, to put down separately, *pă-ran* to separate, to sort out, *ṗă-ran* to judge, to decide, *kă-ran*, *gă-ran* to divide, to distribute, to apportion, *ra* to be parted, to be separated.

(76) T. *bris* picture, drawing, representation, *bris-ma* written book, *ris* figure, form, design, *ri-mo* figure, picture, painting, drawing, mark, *ạbri-ba*, P. and Imp. *bris* to write, to draw, to design.

K. *rit* to mark a boundary, to trace a boundary line, *ǰă-rit* boundary line, border, *a-rit* dividing line, *mă-ri* to mark, to rule, to make a line, *tsă-ri* a scribe.

(77) T. *rman-pa* wounded, *dmas-pa* wounded, *rma-ba*, P. *rmas* to wound, *rma* a wound.

K.[53] *n-ma* wound, cut, laceration, scar.

(78) T. *rmod-pa*, P. and F. *rmos* (cut the ground:) to plough, *rmed-pa* to plough and sow, *rmon-pa* ploughing, *rmo-ba* = *rmod-pa*.

K. *mạt*, *mạn* to cut, to slice, to shave, to castrate.

(79)[54] T. *mun-pa* obscurity, darkness, obscure, dark, *dmunpa* darkened, obscured (as *blo* the mind), *rmun-po* dull, heavy, stupid, *rmus-pa* dull, heavy, peevish, listless, foggy, gloomy, dark, *rmu-ba* dullness, heaviness, fog.

K. *mut* to be blue, *a-mut* blue, *čyă-mut* blue, (faded:) shabby, dull, dusky, *mă-mut* bluish, dark (as clouds), *dă-mun* grey.

(80) T. *rud* a fallen or falling mass (of *ka-ba* snow: a snowslip, of *ču* water: a deluge, of *sa* earth: a landslide).

K. *šă-rut* a landslide (*šă-rut ru* to cave in, to slide down,) *gum-rut* to slide down, to slip down, *niṅ-rut* to be broken, (fallen in:) ravine, gap, landslide, *ziṅ-rut* pit, pitfall, *rut* to pour out, to spill, *run* to pull down, to tear down, to demolish, to fall down, *šă-run* to demolish, to tear down, to destroy, to remove, to strip off, *ru* to fall, to tumble, to pour down.

The most striking thing that we notice here is that while in Tibetan the guttural and labial final families have only two consonantal finals (*-g* and *-ṅ*; *-b* and *-m*), the dental class has no less than five (*-d*, *-n*, *-r*, *-l*, and *-s*), a fact which puts the dental

families at a distinct advantage over the other two. This is, in itself, sufficiently peculiar to suggest a special inquiry into the dental finals to see if they have not undergone in Tibetan some special expansion from an original state where they also possessed only two types, which, by analogy with the other classes, we might suppose were *-d* and *-n*, and which actually exist in this form (*-t* and *-n*) in Kachin.

In this inquiry it seems to me that something may perhaps be deduced from the behaviour of the suffixes *-s* and *-d* with the so-called Perfect roots of verbs. Their occurrence, including the obsolete usages of *-d* with final *-n*, *-r*, and *-l*, may be summarized as follows:—

Suffix.	Guttural finals.	Labial finals.	Dental finals.
-s	(1) *-g*	(1) *-b*	
	(2) *-ṅ*	(2) *-m*	
-d			(2) *-n*, *-r*, *-l*

To begin with, the use of the same suffix (*-d*) after differing finals naturally follows the close relationship in sound between them, and it is, indeed, quite conceivable that *-n*, *-r*, and *-l* may represent simply varieties of one and the same sound, which, in fact, is supported by various developments occurring later in various parts of the Tibeto-Burman field, such as the interchange of *-n* and *-r* and *-n* and *-l* between Kachin and Tibetan,[55] and the similar alternation of *-n* and *-l* in Manipuri,[56] and of Tibetan *-n* with Manipuri *-l*,[57] and so on. Between Chinese and Tibetan Simon[58] has already proposed a number of *-n* = *-l* equations, on some of which we shall comment a little later.

This, then, would reduce our dental finals to three: *-n*, *-r*, and *-l* in one group, with *-d* and *-s* still standing apart.

Now it is a fact, the significance of which is considerable, that as far back as our knowledge of Tibetan goes, neither *-d* nor *-s* can be attached to *-d* verbs in the "Perfect" tense, but that very frequently *-d* falls out and *-s* occurs in its place, giving the type: *ạbud-pa*, P. *bus* to blow, *byed-pa*, P. *byas* to do, and indeed one receives the distinct impression that here *-s* of the "Perfect" is for older *-ds*, an idea to which support is also lent by the alternation of *-d* and *-s* in the Perfect of one and the same verb, as in *ṗyed* or *ṗyes* (P. of *ạbyed-pa* to open), *brńed* or *brńes* (P. of *rńed-pa* to find), which is a feature of frequent occurrence seeming to indicate older forms **ṗyeds*, **brńeds*, and so on.

I am thus led to restate an idea which I first put forward[59] on general grounds only, and which was afterwards advanced more definitely by Simon,[60] that *-s* as a final in Tibetan in a great number of cases presupposes older *-ds*, which is then the parallel in the dental series of *-gs* and *-bs* among the guttural and labial finals, and in what follows we shall somewhat expand this idea, which indeed finds considerable support when the word families of Tibetan, Kachin, and Chinese are examined together.

Let us now, then, fill in our scheme as follows:—

Suffix.	Guttural finals.	Labial finals.	Dental finals.
-s	(1) *-g*	(1) *-b*	(1) *-d*
	(2) *-ṅ*	(2) *-m*	
-d˙			(2) *-n, -r, -l*

We then have our two groups corresponding to the situation in the guttural and labial types.

But we are left, then, if this scheme is complete, without any "final" *-s* in Tibetan. And this, I believe, can be shown with a very high degree of probability to be the case, and that the occurrence of *-s* as a "final" in Tibetan can be ascribed to phenomena which we have already mentioned and which we shall consider in greater detail almost immediately below.

Before, however, passing to this phase of the question, we may profitably turn aside for a moment to inquire how the situation in practice agrees with the proposals so far made.

We have seen already that Kachin has only *-t* and *-n* as dental finals as against Tibetan *-d, -n, -r, -l,* and *-s*. Does, then, Kachin represent an older position in this regard than Tibetan? *A priori* this does not seem very probable, and yet it is certainly significant that its equipment is just that which we have suggested may have been original to Tibetan before its supposed expansion of the dental series.

Taking first Tibetan *-r* and *-l* we have among the individual families already listed the following representative cases:—

No. 49.	T. *-r*;	K. *-n* ~ *-t*
No. 63.	T. *-r*;	K. *-n* ~ *-t*
No. 66.	T. *-r*;	K. *-n* ~ *-t*
No. 38.	T. *-r* ~ *-l* ~ -O;	K. *-n* ~ -O
No. 62.	T. *-r*;	K. *-n*
No. 41.	T. *-r*;	K. *-n*
No. 43.	T. *-l*;	K. *-n* ~ *-t* ~ -O
No. 47.	T. *-l* ~ -O;	K. *-n* ~ *-t* ~ -O
No. 67.	T. *-l*;	K. *-n* ~ *-t* ~ -O
No. 75.	T. *-l*;	K. *-n* ~ -O
No. 54.	T. *-l*;	K. *-n*

which seem, indeed, to give support to the idea that *-r* and *-l* on the Tibetan side represent varieties of *-n*.[61] Phonetically, of course, the close relationship requires no comment, and we have already adduced other concrete instances of the same interchange elsewhere in the Tibeto-Burman family.

We then come again to the question of Tibetan *-s*. Among the word families already listed we have the following cases:—

No. 61.	T. *-s*;	K. *-t* ~ *-n*
No. 65.	T. *-s*;	K. *-t* ~ *-n* ~ -O
No. 55.	T. *-s* ~ *-n* ~ -O;	K. *-t* ~ *-n* ~ -O
No. 79.	T. *-s* ~ *-n* ~ -O;	K. *-t* ~ *-n*
No. 76.	T. *-s* ~ -O;	K. *-t* ~ -O

The evidence of cases like these seems, indeed, very straight-forward, and what appears to be an equation *-s*: *-t* in the type T. *pus-mo*: K. *put* (No. 65) seems, indeed, to find quite definite support in a number of similar instances which are of particularly frequent occurrence between Tibetan and Lepcha. One thus finds, for instance:—

T. *zas*	L. (*a-*)*zót* pasturage	(K. *šat* (food:) boiled rice)
T. *rus-pa*	L. *a-hrăt* bone	(K. *n-rut*)
T. *pus-mo*	L. *tŭk-pat* knee	(K. *put*)
T. *gñis*	L. *nat* two	
T. *gros*	L. *krut* advice	
T. *bkres*	L. *krít* hunger	
T. *dgos-pa*	L. *gat* to desire	

But the same kind of thing can also be set up between Tibetan and Chinese, as, for instance, in:—

T. *mk̆as-pa*	Ch. 黠 *ġat*	K. *ket*
shrewd, wise	wily, cunning	to be wily
T. *šes-pa*	Ch. 悉 *si̯ĕt*	
to know, to understand	to know throughly	
T. *bris*	Ch. 筆 *pli̯ĕt*	K. *a-rit*
picture, drawing	stroke in writing	dividing line

and I am strongly of the opinion that all these *-t* forms in Lepcha, Chinese, and Kachin, presuppose a *-d* somewhere in their Tibetan relatives, and that, indeed, the Tibetan words in question almost certainly at one time ended in *-ds*.

To illustrate the consequences of this possibility, I believe, in fact, that we are justified in advancing equations of the following type:—

(1) T. *pus-mo* (<**puds*) knee, Ch. 市 *pi̯wət* knee cover, K. *put* knee; etc. (See K.WF., H. 117, etc., and Family No. 65, sup.)
(2) T. *bris* (<**brids*) picture, drawing, representation, Ch. 筆 *pli̯ĕt* stylus, writing, stroke in writing, K. *a-rit* dividing line, boundary mark; etc. (See Family No. 76, sup.)
(3) T. *rmus-pa* (<**rmuds*) dull, listless, Ch. 忽 *χmwət* dull, stupid, K. *mă-mut* bluish, dark; etc. (See K.WF., H. 105, etc., and Family No. 79, *sup.*)
(4) T. *ḍus-pa* (<**ḍuds*) gathering, assembly, meeting (cf. *sdud-pa* vb. tr. to join), Ch. 隊 *d̑wəd* group of soldiers, regiment, K. *tut* to be joined together, etc. (See. K.WF., F. 151, etc., and Family No. 56, *sup.*)

And we do not then have the *-s*: *-d*, *-t* equations that appear on the surface, but a much more probable T. *-d*, Ch. *-d*, *-t*, K. *-t*.

In this view, then, a considerable number of equations involving the so-called "Perfect" root of Tibetan verbs may be expected, of which the type is again exemplified in e.g.: T. *(b)rkos* (<**brkods*) excavated, the so-called "Perfect form of *rko-ba*, *rkod-pa* to dig, Ch. 穴 *ɣi̯wet*, or 窟 *kwət* hole, pit, cave (<'excavated' place?)."

But what, then, is the final implication of all this? Clearly (1) That it is highly probable that Chinese, as well as Kachin, never possessed any final *-s* of any type, either true final or suffix,[62] and (2) that the following is a fairly definite series of equations which we may expect between Tibetan and Chinese:

Tibetan	Chinese
-d	*-d*, *-t*
-s <**-ds*	*-d*, *-t*
-n	*-n*, *-r*, *-l*(?)
-r	*-r*, *-n*, *-l*(?)
-l	*-r*, *-n*, *-l*(?)

In other words, the inescapable conclusion seems to be that there was no *-s* in pre-Archaic Chinese at all, and that archaic *-r* can be descended only from pre-Archaic *-r* or *-l* or some variety of *-n*, and I shall be obliged to withdraw a proposal I made on a former occasion[63] that the alternations *-s* ~ *-n* and *-s* ~ *-d* of Tibetan and *-r* ~ *-n* and *-r* ~ *-d*, *-r* ~ *-t* of Chinese support the descent of Archaic Chinese *-r* from *-s* in an undetermined number of cases. On the Chinese side there

is, of course, no certainty that any spread of the dental finals ever occurred in the pre-Archaic language producing a series anywhere nearly paralleling Tibetan with its *-d*, *-n*, *-r*, *-l*, and *-s*, and I thus cannot follow Simon [64] in his suggestion that, for instance, 卷 *ki̯wan* [65] a roll, a scroll, has changed an older *-l* into *-n* on the evidence of Tibetan *gril* a roll. There is small doubt to my mind that Chinese *-n* is always older than Tibetan *-l*. In this particular equation of Simon's I also seriously doubt whether 卷 *ki̯wan* and its relatives (v. K.WF. E. 1–31) are related to Tibetan *gril* and the other members of its family (v. No. 47, sup.) at all.[66] In the case of Kachin the situation is better known. Not only does the language lack a final or suffixed *-s*, but no spread of the original dental finals *-d* and *-n* can be traced.

I cannot help feeling that this elision theory for Tibetan *-d* explains a very great deal. Even where Tibetan verbs regularly end in *-s* throughout I believe there is ground for supposing the disappearance of *-d* before it. The type is :—

ɑgas-pa, P. *gas* to be split, to be cleft, to be cracked, to be burst asunder.
ɑges-pa, P. *bkas*, F. *dgas*, Imp. *ǩos* to split, to cleave, to divide, to cut open.

Here Chinese relatives in *-d* and *-t* appear to be to hand in 害 *ɣâd* (to cut :) to injure, 膾 *kwâd* to cut meat to pieces, to mince, 劊 *kwâd* cut off, 刲 *kwâd*, *ki̯əd* to cut, to wound, 劇 *ki̯wăd* sharp, to cut, to wound, 栔, 契 *ǩiad* to cut, a notch, 剖 *kât* to cut, 戛 *ket* lance, 刮 *kwăt* cut off, scrape off, 鉞 *ɣi̯wăt* halberd, 鍥 *ǩiat* sickle, to cut, 刷 *ɣi̯wət* to incise, and we also have in Kachin *ǩåt* to cut, to hew, to sever, to cleave, *a-ǩåt* to make a semicircular incision with a dah, *lă-ǩåt* to cut obliquely, and finally, I believe, Tibetan itself supports *-s* < *-ds* here with : *ɑgyed-pa*, P. *bgyes*, F. *bkye* to divide, to scatter, to disperse, to dismiss, to distribute, the transitive verb form corresponding to the intransitive *ɑgye-ba*, P. *gyes* to be divided, to be separated, to come apart.

The origin of this class of verb is, indeed, perhaps to be traced to the habit which Tibetan has of using its so-called "Perfect" roots as secondary Present forms. In this way, for instance, *rgyas-pa* is used as a Present, though properly the Perfect of *rgya-ba* to increase, *ɑgyes-pa* is employed as a Present though really the Perfect of *ɑgye-ba* to disperse, giving, in practice *rgya-ba* or *rgyas-pa*, P. *rgyas*, and *ɑgye-ba*, or *ɑgyes-pa*, P. *ɑgyes*. There is only one more step for the verb then, to take before the "Perfect" form has entirely usurped the field and the verb appears with *-s* throughout, exactly as it does in the case of *ɑgas-pa* and *ɑges-pa* above.

It thus transpires that we get a somewhat different view of many Tibetan word families, and especially among those with dental finals is it evident that only a critical examination of all the material which they offer with, where possible, a comparison with the related families in other languages, can give us any sure idea of where the equations from word to word should be set up.

The same principle, indeed, applies to words which are apparently completely isolated in their own languages. A case in point is Tibetan *ǩa* mouth. Only from

the Kachin family *ǩan* to be pushed back, to be separated, *mă-ǩan* to push open, to force apart, *sŭm-ǩan* to be open, to gap, *mă-ǩa* to be open (as a door), to open (as the mouth), *jă-ǩa* to part, to separate, *ǩa* to be remote, to be at a distance, to lose (as a tooth), *ka* to be parted, to be separated, to gap, to be open, *čyiṅ-ka* door, gate, *lă-ka* to crack open, to chap, *n-ka* door, does it appear that Tibetan *ka* was originally "the opener". However reasonable the unsupported supposition [67] might have appeared, one could never from Tibetan alone have supported it with a series of similar nature.

Finally, also, as an additional example of how very essential the word stock of one language may be to the correct evaluation of that of another, we may take the following case.

The writer has previously[68] suggested that in Tibetan there was formerly, with some probability, a word family of the type *-n* ~ -O with the general sense of "to bend". Chinese alone gives one this idea, but within the Tibeto-Burman family additional proof is supplied by Kachin,[69] where we have *kun, tiṅ-kun* to be bent, to be curved, *šiṅ-kun* to bend (as a stick), to be pliable, to wriggle, *šiṅ-kun, mă-kun* (assume a bending position:) to crouch down, *ku* to be bent, to be curved.

Erratum

note 29, line 1. For *"ąpra-va"* read *"ąpra-ba"*.

Notes

1 This has already been very clearly stated by Karlgren, "Word Families in Chinese," *Bull. of the Mus. of Far Eastern Antiquities*, No. 5, Stockholm, 1934, p. 9, quoted hereinafter as K.WF.

2 See, for instance, Dragunov, *OLZ.*, 1931, Sp., 1088–1089.

3 For the Chinese connections here see K.WF.; H. 26–44, and *JRAS.*, 1936, pp. 415–416.

4 "Tibetisch-Chinesische Wortgleichungen," *Mit. d Seminars f. Orient. Spr.*, Bd. xxxii (1929), Berlin, 1930, Abt. 1, p. 162. This work I shall hereinafter quote as "Wortgl."

5 Assimilation is sometimes clearly the cause of an aberrant final. A well-known example is *mdzub-mo* thumb, beside *mdzug-gu*, in a family with guttural finals (cf. "Language", *JLSA.*, iv (1928), p. 278 (No. 4)). There can be but small doubt here that *mdzub* owes its final *-b* to the influence of the following suffix. At other times associative interference has probably been at work, but the forces involved are obscure and each case will need special treatment by itself.

6 With these, Simon (Wortgl., p. 162) would include *snabs* mucus, which gives us a good example of how assemblies uncontrolled by families in related languages may go astray, even under the best of judgment. *snabs*, we shall see below (No. 33) belongs with another family entirely, but from Tibetan alone this would never be apparent.

7 The entries, as far as possible, follow the order *-g, -ṅ*, -O; *-b, -m*, -O; *-d, -n, -r, -l, -s*, -O. The assemblies must not be regarded as necessarily correct in all particulars. In some cases there are other possible alignments which I have indicated in the footnotes.

8 Cf. K.WF., A. 266–275.

9 Cf. K.WF., A. 35–7, 330–1.

10 Cf. K.WF., A. 266–275.

11 Cf. K.WF., A. 27–8?

12 Cf. K.WF., B. 150–189.

13 Related to No. 50 with dental finals. Cf. K.WF., B. 20–5.

14 Cf. K.WF., B. 295–301.

15 Cf. K.WF., B. 112–128, 569–571. Cf. No. 10 inf.

16 Cf. K.WF., B. 112–128, 569–571, and cf. No. 9 sup.

17 Cf. K.WF., B. 64–84.

18 Cf. K.WF., D. 188.

19 Cf. K.WF., D. 62–71?

20 Cf. K.WF., D. 89–94.

21 Cf. K.WF., D. 1–23, 24–7.

22 Cf. K.WF., D. 60–1.

23 Cf. K.WF., I, 1–20.

24 The underlying sense here may be one of "succession" or "series" (of events or actions). Cf. Lepcha *góm, a-góm* series, chain, train, continuity, whence: *toṅ góm* a step, a stride (lit. foot (*toṅ*) step), *a-dyaṅ góm* a step, a pace (lit. foot (*a-dyaṅ*) step), *kă-góm* a stride.

25 Cf. K.WF., K. 79–84.

26 Cf. K.WF., K. 16–19.

27 On the Kachin side it is possible that *nya, šă-nya*, and *čyă-nya* belong with T. *mṅen-pa* flexible, pliant, supple, soft, as their final affinities are uncertain. Cf. Burmese *ṅan.*

28 Cf. K.WF., K. 87–90?

29 Also (?): *ąpra-va*, P. *ąpras* to kick, to jerk (as the legs); *pra-ba*, id.; *ąpras-pa* stroke, blow, kick.

30 Cf. K.WF., E. 153–162, and *JRAS.*, 1936, pp. 404–5.

31 Cf. K.WF., E. 114–123.

32 Also (?): *gud* separation, solitude, seclusion (> CT. loss, damage), *gun* loss, damage.

33 Cf. K.WF., E. 138–144.

34 See *JRAS.*, 1936, p. 408. Jäschke (Dic., p. 358) notes *ąpyil-ba* to wind, to twist, as a form of *ąkyil-ba*. Note, however, K. *pyin* to go round, to encircle, windlass, capstan.

35 Cf. K.WF., G. 24–34?

36 Cf. K.WF., E. 238–246.

37 Related to No. 7 with guttural finals.

38 Cf. K.WF., F. 46–7.

39 Cf. K.WF., G. 40–5.

40 Cf. K.WF., F. 100–102. It is possible that the Tibetan and Kachin families here assembled are distinct and that the only relative that Kachin has in Tibetan is *ša* flesh, meat. The position remains in doubt.

41 Cf. K.WF., F. 150–3.

42 Cf. K.WF., F. 243–6.

43 Cf. K.WF., F. 63–7. It is possible that K. *tså, mă-tså*, and *šă-tså* do not belong here, but rather with T. *gtso-bo* highest in perfection, most excellent (of its kind).

44 Cf. K.WF., G. 24–34, and No. 60 below.

45 Cf. No. 59 above.

46 Cf. K.WF., H. 54–7.

47 Cf. K.WF., H. 111–134.

48 Cf. K.WF., F. 247–250, H. 141–2. This and the next three families are very confusing. They all appear to belong together at base, as is indicated by the various collocations of the type *mči-ma ąpyi-ba*, etc., where the families are used together as though one. Probably the form *ąpyin-pa* liver, of the Berlin Gzer-Myig (v. Francke, *Asia Major*, vol. i, p. 287) is something more than a homophone, and indicates recognized relationship (present in the scribe's mind) between this word and *ąbyin-pa* to tear out, of the present family. See footnote to No. 71.

49 Cf. K.WF., F. 48–49. See note to No. 68.
50 Cf. K.WF., F. 147–9 ?. See note to No. 68. The basic idea here seems to be that of "ejection" (ejected matter). *mči-ma* (welling forth:) tears, may belong here. Cf. the Chinese family quoted.
51 See note to No. 68. Possibly the basic meaning here is "that which is torn out" (pointing to very early hepatoscopy in Tibet ?).
52 Cf. K.WF., H. 26–44.
53 Note also K. *mya* to be torn, to be ragged, *a-mya* to tear, to lacerate, to maul, *čyă-mya* torn, ragged. These, however, may belong with T. *dmyal-ba* to cut up.
54 Cf. K.WF., H 94–110.
55 See above, Nos. 38, 41, 49, 62, 63, 66, and 43, 47, 54, 67, 75.
56 As e.g. in *čen-ba* ~ *čel-ba* to run, *han* ~ *hal* the causative infix, *ƙanba* ~ *ƙal-ba* to think, *na-ton* ~ *na-tol* nose, *pun-ba* ~ *pul-ba* to tie, to bind, *sa-gon* ~ *sa-gol* (Beng. *ghōrā*?) horse, etc. This *-n* ~ *-l* and also an *-r* ~ *-l* alternation likewise come out strongly in the Bodo and related languages of Assam, notably in Garo, Båṛå, and Kachari: G. *pān* ~ *bol* tree, Mikir *iṅ-čin*, G. *šil* iron, G., K. *săn*, G. *sāl* sun, G., K. *šur*, G. *šil* iron, K. *bi-bār*, G. *pārr*, *bi-bal* flower, K. *bār*, G. *lam-pār*, *bāl* air, and others. Note also Tibetan *ṗar* exchange, barter, G., K. *ṗāl* to sell.
57 As e.g. T. *rgan-pa*, M. *a-hal* old, T. *mťon-po* high, M. *ma-tol* summit, etc.
58 Wortgl. pp. 183–4.
59 *Outlines of Tibeto-Burman Linguistic Morphology*, London, 1929, p. 19, n. 1. This will be abbreviated hereafter as *"Outlines"*.
60 Wortgl., p. 185.
61 It should be noted here that it must have been at a very early time indeed when Tibetan first placed the basic roots of what are now its word families in the final *-n*, and *-r*, and *-l* classes. This is shown by the very great comparative rarity of alternations between *-n* and *-r* in the same family (for examples, see Nos. 53 and 64 in the preceding pages, and *JRAS.*, 1936, p. 415, No. 22, and p. 406), and the almost total absence, so far as I am aware, of *-n* ~ *-l* and *-r* ~ *-l* alternations of similar type (for *-r* ~ *-l* see No. 38, sup.). In other words this move must have taken place while the root was still single and alone, and free of prefixes, which were only added to it in mobile fashion as required (cf. *Outlines*, p. 53). Though at other times families in *-n* and *-r*, *-n* and *-l*, or *-r* and *-l* may approach each other fairly closely in both form and meaning, I do not believe that any such cases involve alternations between these finals. Such families seem to me to be altogether too evidently distinct (v. *JRAS.*, 1936, pp. 407–408, and (for a contrary view), K.WF., p. 36).
62 By this I, of course, mean *-s* standing alone as final, or following a vowel as a suffix. We have no knowledge as to the existence of a *-ds* or *-ts* combination in Chinese at some very early time. It is, of course, just possible that a suffixed *-s* of this nature is at the bottom of the *-d*: *-t* differentiation in Archaic Chinese.
63 *JRAS.*, 1936, p. 402.
64 Wortgl., p. 183, and No. 312.
65 K.WF., E. 8.
66 Cf. also e.g. Simon No. 320, T. *ąṗral* Ch. 分 *ṗi̯wən* to divide (K.WF., H. 38). The Chinese family here, I believe, belongs with an entirely different one in Tibetan (v. *JRAS.*, 1936, p. 415, No. 22) where a Tibetan *-n* Ch. *-n* equation is present. For the Tibetan and Kachin relatives of T. *ąṗral-ba* see above No. 75.
67 From Tibetan alone one, of course, gets a fairly clear impression of the origin of this word from the fact of its carrying the sense of "opening", "gap", "vacancy", in such combinations as *ƙa ągebs-pa* to shut an opening, *ƙa skoň-ba* to fill a vacancy, etc., but the family as still extant in Kachin is lacking.
68 *JRAS.*, 1936, pp. 404–5.
69 See No. 38 sup.

15

VARIABLE FINALS IN PROTO-SINO-TIBETAN*

Randy J. LaPolla

Source: *The Bulletin of the Institute of History and Philology* 65, 1, 1994, 131–73.

In many Tibeto-Burman languages we find that there are a number of forms that are clearly related though differ in one segment. In some cases these variations may be due to regular or common alternations, such as in Tibetan, where you have dental suffixes that can nominalize a verb (e.g. rkun-po 'thief', from rku 'steal'). In other cases we cannot find any morphological reason for the variation, even though the variation may involve the same segments, as in Tibetan bka, skad 'speech'. When we reconstruct the Proto-Tibeto-Burman provenience of these cognates, we sometimes have no way of knowing which form is older, so we must reconstruct two forms that are clearly related, that are what James A. Matisoff has dubbed 'allofams'. On the Chinese side of Sino-Tibetan we find similar alternations among cognate forms, as in 亡 *mjaŋ, 無 *mjag 'negative/not have'; 往 *gwjaŋ, 𠂇 *gwjag 'go'.

This paper concentrates on variable finals, and argues that just as we find a certain amount of both rule-governed and non-rule governed variation in modern languages, in reconstructing Proto-Sino-Tibetan we should recognize the possibility of these types of variation. Second, the variation we find in PST and its immediate daughters is not as symmetrical and orderly as has been assumed. Third, the causes of the variation are complex and multifarious. Fourth, reconstructing a complex, typologically unlikely system to 'explain' the variation, such as the voiced stop finals reconstructed for Old Chinese, may also prevent us from attempting to find out the real causes of the variation. Fifth, the concept of word families is an important one, but we should not be unnecessarily constrained in our search for cognate sets by artifacts of our reconstructed system.

1. Variation in Sino-Tibetan

In working with Tibeto-Burman languages, we find that within each of the languages of the family there are a number of forms that are clearly related though differ in one segment, as in the following examples from Tibetan:

rku 'steal', rkun-po 'thief'
bka, skad 'speech'
nye 'near', nyen 'relative'
gči-ba, gčid 'to urinate', gčin 'urine'
fibye-ba (intr.), fibyed-pa (trans.) 'open, separate'
ŋi-ma, ŋin-mo 'day'
dro-ba 'to be hot', dron-mo 'hot', drod 'heat'.
ŋu-mo 'weep', ŋud-mo 'a sob'

In some cases these variations may be due to regular or common alternations, such as in Tibetan, where you have dental suffixes that can nominalize a verb, as in rkun-po 'thief', from rku 'steal', and ŋud-mo 'a sob' from ŋu-mo, 'weep', or they can have a causative function, as in fibye-ba (intr.), fibyed-pa (trans.) 'open, separate' (cf. Benedict 1972:100, 1991). In that case it will not affect our reconstruction of the Proto-Tibeto-Burman (PTB) form of, for example, 'steal', though if we find the same derivational process in other TB languages, then we might want to reconstruct that morpheme (and the morphological process) to the proto-language. In other cases we cannot find any morphological reason for the variation, even though the variation may involve the same segments, as in Tibetan ŋi-ma, ŋin-mo 'sun'; ka, skad 'speech'; and Dulong muʔ55 (< *muk; LaPolla 1987), rɯ31 mɯt55 'cloud'.

We find similar types of groupings on the Chinese side of Sino-Tibetan as well. These groups of related items are known as 'word families', following Karlgren's famous article (1933; see also Karlgren 1956). Karlgren, and later Wolfenden (1937), argued that in doing cross-language comparative work it is these word families that we should compare, not individual lexical items. Both Karlgren and Wolfenden felt that there were certain regularities to, or restrictions on, the type of variation within each word family, for example a restriction on the point of articulation of the finals such that all the variant forms of one word family would involve the same point of articulation. Wolfenden (1936, 1937) classified each of the forms he presented from Tibetan as to whether they were in the 'velar series', the 'dental series' or the 'labial series' of variation. He did not suggest a historical reason for this type of restriction on the variation.

Because of the recognition of these word families, in doing the comparative work necessary for reconstructing PTB we often need to recognize the same types of variation among languages or dialects in the family, as we often find forms that seem to be cognate in all but one segment, either the initial, the vowel or the final. If it is a case where the variation cannot be seen to be morphological, then we

have to see if it is a matter of one language being aberrant, as in the case of some of the -k and -t finals of Maru and the -n ~ -ŋ variation due to the causative infix of Lepcha,[1] or of a large number of languages being split (possibly along genetic lines) between having one form or the other, as in the case of 'dream', where all languages in Sino-Tibetan having cognate forms except for the Lolo-Burmese languages descend from *r-maŋ, with a velar nasal final, while the Lolo-Burmese forms descend from *r-mak, with a velar stop final. We might want to say in this case it is due to idiosyncratic phonetic change at the Proto-Lolo-Burmese level. The tendency in reconstruction work is still to attempt to reconstruct a single proto-form for the variant forms, though most cases are not as neatly distributed as the case of 'dream', and in these cases, when we reconstruct the PTB form we have no way of knowing which form is older, so we must reconstruct two or more alternate forms which represent the possible variations within the word family. Matisoff (1978:17ff) has dubbed these proto-variants 'allofams' (forms within the same word family, the term being based on analogy with 'allomorph' and 'allophone'), and he marks them with the symbol '≍' (from > and <, as we do not know which way the relationship goes). Among the most common of the alternations we find is variation in the vowels of closed syllables (e.g. -i- ≍ -u- ≍ -a-), variation between pure vowel and dipthong (e.g. a ≍ ay) (see for example Matisoff 1985), and also variation between stop and nasal final or stop and open final.[2] In this paper I will be concentrating on variable finals. Here are some examples from TB of the type of allofams I will be talking about:[3]

*ka ≍ kat	'speech'	*m-si ≍ m-sit	'comb'
*la ≍ lap	'leaf'	*pa ≍ pan	'palm'
*pyaw ≍ pyam	'fly (v.)'	*ra ≍ rat	'cut'
*k-lok ≍ k-loŋ	'stone'	*yu(w) ≍ yuk	'descend'
*ma ≍ mat	'disappear'	*ya ≍ yan	'night'
*du ≍ dut ≍ tu ≍ tut	'join, tie, knot'		

On the Chinese side of Sino-Tibetan the question of word families is very much intertwined with the concept of rime categories (韻部 yùnbù).[4] From the study of the rhyming patterns and xié-shēng (諧聲) phenomena[5] of Old Chinese (OC), we are accustomed to thinking in terms of Chinese words belonging to certain rime categories, and these rime categories to belonging to certain groups of rime categories (類 lèi). The larger groupings are based on the observance that words belonging to a particular category sometimes rhyme with words in certain other rime categories, or the Chinese characters used to represent words belonging to a particular rime category will have the same phonetic components as words in certain other rime categories. We assume this happens because these particular rime categories have similar rimes. An example is the rime categories yú (魚), dúo (鐸), and yáng (陽), which are all said by Li Fang-kuei (1980) to have the vowel *-a plus a velar final consonant: *-g, *-k, and *-ŋ, respectively.

It is because of these 'contacts', as they are referred to, in rhyming or graphic components that Li (following Karlgren) reconstructs a final *-g for what is an open syllable in later Chinese (e.g. 魚 *ŋjag 'fish').[6] When we find words with the same vowel but different finals with the same point of articulation either rhyming with each other or sharing a phonetic component, we call this 'direct transfer' (對轉 dùi zhuǎn) or 'connected rhymes' (通韻 tōng yùn). We find examples of this kind of cross-rhyming in the Shi Jing (詩經) (from Wang 1980b):[7]

芼 *magw, 樂 *ŋrakw	(宵藥通韻)	《周南：關雎》
敦 *tən, 遺 *rjəd, 摧 *sdəd	(文微通韻)	《邶風：北門》
艾 *ŋadh, 難 *nan	(祭元通韻)	《閔予小子之什：訪落》
來 *ləg, 贈 *dzəŋh	(之蒸通韻)	《鄭風：女曰雞鳴》

It is not the case that the rhyming patterns always follow the tong yun patterns. In this case it is called 'combined rhymes' (合韻 hé yùn).[8] Here are a few examples (From Wang Li 1980b):

業 *ŋjap	作 *tsak			(盍鐸合韻)	《蕩之什：常武》
答 *təp	退 *thwədh			(緝微合韻)	《節南山之什：雨無正》
躬 *kjəŋw	天 *thin			(中眞合韻)	《文王之什：文王》[9]
林 *ljəm	冰 *pjəŋ			(侵蒸合韻)	《生蕩之什：常武》
言 ŋjan	行 graŋ			(元陽合韻)	《蕩之什：常武》
人 *njin	訓 *xwjənh	刑 *giŋ		(眞文耕合韻)	《清廟之什：烈文》
服 *bjək	熾 *thjəgh	急 *kjəp	國 *kwək		
				(國之緝合韻)	《南有嘉魚之什：六月》
贍 *tjam	相 *sjaŋ	臧 *tsaŋ	腸 *drjaŋ		狂 *gwjiaŋ
				(談陽合韻)	《蕩之什：桑柔》

We also find variation within these larger rime classes where a character will have two pronunciations differing only in the final consonant (e.g. 度 *dagh/dak, 殺 *sriadh/sriat, 告 *kəgwh/*kəkw, 莫 magh/mak (for lists of these characters see Downer 1959, Wang 1980a:213ff), or where two different characters will represent what seems to be the same word, though the reconstructed pronunciations for the two characters differ in the final consonant:

夜 *riagh	夕 *rjiak	'night' (cf. Mei 1979:120ff)
無 *mjag	亡 *mjaŋ	'no, not have'
于 *gwjag	往 *gwjaŋ	'go'
女 *nrjagx	孃 *nrjaŋ	'woman'

Following are some xie-sheng contacts involving different finals:

瞿 *kwjagh	矍 *kwjak	卒 *tsət	醉 *tsjədh
赭 *tjiagx	赤 *khrjak	弗 *pjət	費 *phjədh
逋 *pag	搏 *pak	必 *pjit	秘 *pjidh
借 *tsjiagh	昔 *sjiak	等 *təŋx/təgx	待 *dəgx
賴 *ladh	爛 *lanx	能 *nəh	態 *hnəgh
由 *rəgw	迪 *diəkw	脈 *mrik	派 *phrigh

We can also compare forms from Proto-Tibeto-Burman with forms in Old Chinese, and we come up with some interesting variations.

PTB *la ≍ lap, OC *rap (葉) 'leaf'.
PTB *ka ≍ kat 'speech', OC *kal (歌) 'sing, song'.
PTB *ba, OC *bak (薄) 'thin'.
PTB *mra ≍ mraŋ, OC *mragx (馬) 'horse'.
PTB *graŋ ≍ *grak, OC *gljaŋ (涼) 'cool, cold'.
PTB *kap, OC *gap ≍ *kabh (吒) 'to cover, cover'.
PTB *san ≍ *sat, OC *san (散) ≍ *sat (撒) 'sow, pour out, disburse'.
PTB *ŋa ≍ *ŋan, OC *ŋal (鵝) ≍ *ŋran (雁) 'goose'.
PTB *tu ≍ *tuŋ, OC *dugh (豆) ≍ *təŋ (Cf. 登) 'bean'.
PTB *na ≍ *naŋ, OC *njagx (汝) ≍ *nəgx (乃) ≍ *njəŋw (戎) '2sg pronoun'.

2. Problems of methodology

Since both sides of the family seem to exhibit the same pattern of variation, we should be able to reconstruct this pattern of variation to Proto-Sino-Tibetan, but there are two problems involved with this hypothesis. First, Wolfenden's 'rule' of Tibetan word families is the result of his chosing some words over others that do not fit his pattern. I found a counter example after looking through a Tibetan dictionary for less than two minutes: sbu-gu 'hollow, cavity; the narrow interior of anything, a tube', sbugs 'hollow stalk, a tube; hole, excavation, interior space', fibugs-pa 'to hollow out, bore', sbun-gter 'meaningless, without substance, hollow, vain', sbub-khoŋ 'a hollow ball', sbub-mo 'hollow tube', sbur-ma 'chaff, husks'. Second, not all of the items that vary within one point of articulation in TB vary within the same point of articulation in OC. For example, Wolfenden gives Written Tibetan rmu-pa 'dullness, heaviness, fog', mun-ba 'obscurity, darkness, obscure, dark'; Kachin śa 'child', WT btsa-ba 'to bear children', tsha-bo, mtsan 'grandchild, nephew'; and WT rkun-ma 'thief', rku-ba 'steal' as all being in the dental series, while their Chinese cognates are all in the velar series: *mjugh (霧), *tsəgx(子), and *khugh (寇) respectively. We can also add OC *pjag/*pragx (扶 / 把), TB *pa ≍ pan 'palm'. The opposite situation exists for OC *pjidh ≍ *pjit (畀), TB *biy ≍ biŋ 'give'. If we were to hold strictly

to the 'same series consonant' rule, we would have to say that the forms in these word families are not cognate.

The problem of which forms to select exists for anyone attempting to identify word families, or even simple cognates. Each researcher has his or her own standards of rigorousness as to what constitutes an acceptable correspondence. Karlgren and Wolfenden limited their word families to only those forms whose finals had the same place of articulation, but as Pulleyblank (1972:11, 1973:120) has argued, 'One can easily find sets of words with the same initial consonant and closely similar meanings but quite different finals that are at least as plausible as the word families collected by Karlgren . . .' Among the examples Pulleyblank gives are the following (1972:11–12, 1973:121):

尼 *njid 'near, close' 昵 *njit 'intimite, familiar; glue' 狃 *njəgwx 'be familiar with, treat with contempt' 粘 *njam 'to glue, stick to'.
累 *ljədx 'bind, wrap around' 繚 *liagw, 'bind round, wrap' 摎 *kiəgw, liəgw 'tie round, strangle' 綸 *ljən 'woof, twist a cord, cord'.
From Pulleyblank 1991:30 we can also add
呼 *xag, 喝 *xat 'shout'; 舉 *kjagx, 揭 *kjat 'lift'.

Wang Li (1980a, 1980b, 1982) accepted the concept of classes of rime categories, but unlike most scholars working on Old Chinese, did not follow Karlgren in reconstructing the finals *-b, *-d and *-g.[10] Possibly because of this he was not restricted in his search for cognate characters in Chinese (Wang 1982). He has 101 pairs of suggested cognates where the finals have different points of articulation (or would have in a system with *-g, *-d and *-b). Here are a few examples (converted to Li Fang-kuei's system of reconstruction): [11]

吾	*ŋag		言	*ŋjan	
我	*ŋarx		語	*ŋjagx	'language, speech'
卬	*ŋaŋ	'lsg pronoun'			
委	*ʔwjarx		于	*gwjag	
汙迂譆	*ʔwjag	'bent'	爰	*gwjan	(preposition)
喜	*hjəgx		弗	*pjət	
欣	*hjən	'happy, happiness'	不	*pjəg	'not, negative'
急	*kjəp		架	*krarh	
亟	*kjək	'urgent, rushed'	格	*krak	'(clothes) rack'
額	*ŋrak		剔	*thik	
顏	*ŋran	'forehead'	剃	*thidh	'cut hair'
甲	*krap		恨	*gən	
介	*krat	'armor'	憾	*gəm	'regret'

It would be difficult, given the semantic correspondences (in most of the above examples, each of the characters is defined using the other from the pair), plus the fact that all other segments of the syllable match up exaclty, it would be unwise to throw out these correspondences simply because the finals do not have the same point of articulation. Doing so would also mean we would have to say the phonetic and semantic correspondences between these sets (and many others) are purely coincidental and not due to etymological relatedness.

In terms of xiesheng contacts, Mei & Gong (1992) discuss several examples that differ in rime class, such as 豆 *dugh: 短 *duanx: 登 *təŋ and 取 *tshugx: 最 *tsuats: 叢 *dzuŋ. Pulleyblank (1991:30) also gives the following forms (which are not only phonetically related, but most likely etymologically related as well): *khjagx/h (去) 'leave, go away from', *khjag (祛) 'dispel, exorcise', *khjat (竭) 'go away'. We can also add *khjap (怯) 'cowardly, afraid'. [12]

We have at least three choices when faced with a situation such as we have in Sino-Tibetan. We can attempt to account for all possible variations (or most of them) by reconstructing a very complex proto-language using phonetic symbols (see for example Coblin 1986, where *-g is reconstructed to account for correspondences between OC *-g and TB *-k), we can use non-phonetic symbols to mark those alternate correspondences that are unresolvable (as for example when Austronesianists use *L to represent *l or * ɬ, or we can reconstruct a simple system and try to either explain the variations by some morphological or phonetic means or simply allow a certain amount of variation in our word families. This is a question of methodology. The first method is problematic because the resultant system is often typologically unrealistic (e.g. having three phonemically distinct *-r phonemes), while the second gives an incomplete and formulaic reconstruction. A cross between the two occurs in the case of the voiced finals of Old Chinese, as they are meant both to phonetically explain a particular correspondence, and to serve as symbols for unresolvable correspondences.[13] This gives us a system that not only does not satisfactorily account for the data, but also gives us a typologically very unlikely system with voiced final consonants and no open finals at all. [14] It is the third methodology I believe is the proper choice given the situation in Sino-Tibetan.

3. Possible explanations

The first thing I would like to suggest is that it is not necessary to assume that the rhyming or xiesheng contacts were anything less than true rhymes and accurate phonetic borrowings. To assume they were not (as is implied by the voiced stop final hypothesis) weakens the whole theoretical underpinings of the traditional methods of Chinese historical phonology. We must assume the creation of xiesheng characters and the use of rhymes was relatively strict.[15] That is, it is not necessary to say that when a yú bù (魚) word rhymed with a yáng bù (陽) word, that it was *-ag rhyming with *-aŋ. In these cases it was very likely *-a rhyming with *-a

or *-aŋ with *-aŋ, with the difference due to variation of the final of that character/phonetic. If we accept variation in prefixes, initials, and vowels, then accepting variation of finals should not be very problematic.

Dong Tonghe (1981:268) argues that given the variation we find in the finals, 'we cannot say that the characters with stopped finals in Middle Chinese originally had no final consonant in OC, and so could rhyme and have xiesheng contact with non-stopped characters, as if we say this then the contacts between non-stopped rimes should be chaotic; they definitely would not be this clearly separated'. He suggests the only alternative is to follow Karlgren's lead and reconstruct *-g, *-d, and *-b.

The most cogent arguments presented in favor of the voiced final consonant hypothesis are those given in Ting 1979, 1987. In Ting 1979 Chinese loans to Tai are examined (citing Li 1945), and it is shown that of the 12 earth-branch (地支) callendrical signs, one, *mjədh (未), appears in Tai dialects with a -t final, and six of the seven other items reconstructed for OC with voiced stop finals (*-g or *-gw) appear with glide finals in the Tai dialects. The seventh, *ŋagx (午), appears with an -a final in all three dialects. Ting argues that the fact that in all three Tai dialects considered OC *-g, and *-gw have regular but different reflexes is evidence that these characters had different finals in OC. That is, if these characters had simple vowel rimes with open finals (e.g. *-ə), then it would be difficult to explain the appearance of off-glides in all the Tai dialects. Just as some of the off-glides in Modern Mandarin descend from OC voiceless stop finals, Ting argues these Tai off-glides descend from OC voiced stop finals. Ting explains the change of the *-d final of OC *mjədh to Tai -t and not to a glide by reference to the fact that the *-d final rimes (脂微祭) rhymed with rusheng rimes as late as the Nan-Bei-Chao period, while the *-g and *-gw rimes gradually stopped rhyming with rusheng rimes during the Han period. Ting also points out the possibility that the difference is related to the fact that *mjədh is the only qusheng word among all of the 12 callendrical signs.

Ting then (p. 731ff, citing Li 1976) gives a number of lexical items from Siamese that are suggested to be cognate to certain Chinese items, though here the correspondences are less regular, as there are sets of OC *-ag corresponding with Thai -ɔɔŋ, OC *-ag/-ug corresponding with Thai -(a)ak/-ɔɔŋ, OC əgw/agw corresponding with Thai -uak/-ok, OC *-ad/id corresponding with Thai -ɔɔt/-et, and OC *-ag corresponding with Thai -aa. Ting takes the former sets as evidence of stop finals in Chinese, and explains the last set as due to the loss of *-g with compensatory lengthening of the vowel.

Next, Ting gives two sets of OC-Tibetan/Burmese correspondences. The first set shows some possible cognate sets where the OC form is reconstructed with a voiced stop final and the Tibetan/Burmese forms have voiceless stop finals. The second set shows possible cognates where the reconstructed OC form has a voiced stop final but the Tibetan/Burmese forms have open finals or glides. Ting argues that the sets where the Tibetan/Burmese forms have stop finals shows that at least some of the OC forms must have had consonant finals, and since the Chinese rime categories cannot be split up, then it must have been Tibetan and Burmese that have changed (p. 733).

In Ting 1987 further evidence is given to show that at least some characters had stop finals of some type. It is shown from an analysis of the cross-rhyming patterns of the different tones that there was a very strong connection between qu and rusheng in the Shijing, but that this connection weakened or changed gradually through the Western Han and Eastern Han periods to the point that in the Wei-Jin period rhyming patterns only those rimes reconstructed with dental finals showed cross-rhyming between the qu and rusheng words. There was in fact an increase in dental cross-rhymes as the velar cross-rhymes decreased (p. 62). Ting suggests that the reason why only the qusheng words, and not the ping and shang-sheng words, show this close connection with the rusheng words is that the pitch value of the qusheng must have been closer to that of the rusheng than were the other tones (p. 61, citing Dong 1954:189). The reason for the drop in velar contacts in later periods is suggested to be that *-g was lost earlier and faster than *-d (p. 63). No reason is given for the increase in dental qu-ru cross-rhymes. In the Wei-Jin period not only do the ping and shang-sheng words not rhyme with rusheng words, they also do not rhyme with qusheng words. Ting's explanation for this is that something about the pitch value of the qusheng caused stop finals to be retained while they were lost from the ping and shang-sheng words.

This is very solid philological work, and there is no reason to doubt Ting's main conclusion that the relevant lexical items had consonant finals in Old Chinese. The question then is was it a voiced stop final or a voiceless one, and do all of those words in the traditional rime categories necessarily share this consonant? Ting's answer is that it was a voiced consonant and all the words in the category traditionally thought to not have a voiceless stop final shared the same voiced stop final. This is one possibility, but not the only one. We are now all in agreement that many variations in the initals of Middle Chinese are due to different prefixes in OC (see for example Pulleyblank 1962–62, 1972, 1973a; Bodman 1980, Benedict 1987, Mei 1989, Baxter 1992). In the same way much of the variation in the finals of Middle Chinese can be explained as due to qusheng (去 聲 'departing tone') derivation (see Downer 1959, Pulleyblank 1962–62, 1972, 1973a,b, 1977–78, Mei 1980, Baxter 1992). Rather than assuming that since some words in a particular rime show contacts with rusheng words all words in the rime must have had stop finals, Pulleyblank (1977–78) and Baxter (1992) reconstruct consonant finals only for those items that actually show rusheng contacts, and reconstruct non-stop finals for those words which do not show rusheng contacts. Pulleyblank and Baxter both reconstruct voiceless (rather than voiced) stop finals in those words that show rusheng contacts, assuming that these finals were later lost due to the influence of an *-s suffix which later developed into the departing tone (and possibly a *-ʔ final that developed into the rising tone).[16] All of the evidence presented by Ting is consonant with this hypothesis, and in fact more so than the voiced stop hypothesis, as the Thai and Tibeto-Burman evidence is of a voiceless stop, not a voiced one, and it explains why *ŋagx (午) (which does not have rusheng connections and so is reconstructed with an *-n final by Baxter) does not

show evidence of a consonant final in the Tai dialects. The fact that Li's *-g and *-gw have different off-glide reflexes in the Tai dialects cannot be taken as evidence of voiced stop finals, as any system that differentiates these two rime categories (之 幽) can account for this, especially if you 幽 is reconstructed with an off-glide (e.g. əw). The open final hypothesis also explains the open *-a(a) finals in the Siamese, Tibetan, and Burmese words presented by Ting, as they are all items that do not show rusheng contacts (e.g. 五吾魚無死父余塗署稼), without having to assume the irregular loss of a voiced final in some but not other words. The rhyming patterns are also explained more satisfactorily than by making ad-hoc guesses about pitch contors, as suggested by Dong Tong-he.

What this hypothesis means is that the original tone categories of OC do not coincide completely with those of Middle Chinese. Whereas rusheng is considered a separate tone in Middle Chinese, the three 'tones' (*-Ø, *-ʔ, and *-s) of OC could appear on any type of syllable, including those with voiceless stop finals. According to Baxter (Baxter 1992:309), the *-s suffix ('post-final' in Baxter's book) then caused the loss of the voiceless stop finals in the following stages ('H' is the representation of the Middle Chinese departing tone in Baxter's system):

*-ps	>	*-ts	>	*-js	>	-jH
*-ts	>	*-js	>	-jH		
*-ks	>	*-s	>	-H		
*-wks	>	*-ws	>	-wH		

Given the fact that 90% of all rusheng contacts with non-rusheng words involve qusheng words, this hypothesis explains quite a bit of the variation of finals within Old Chinese. Given this system the contacts would also not be 'chaotic', as feared by Dong Tonghe (see above). This analysis has other strong points as well. As Baxter points out (1992:336), Karlgren originally reconstructed voiced stops both to explain the rusheng contacts and the development of the qusheng, whereas in Li Fang-kuei's system the qusheng is separate from the voiced final, so 'it appears to be a coincidence that *-ad occurs only in qusheng, or that qusheng words often have clear and obvious rusheng connections, while words in other tones can usually be connected with rusheng only indirectly'.[17] It is also not a coincidence that we cannot find TB cognates for any of the OC *-dh and *-bh words that match exacty: as these words were created by a derivational process within Chinese, we would expect to find TB cognates only for the underived forms (i.e. *-t, *-p), not the derived forms. For example, we have TB *r-mok 'to wear on head', OC *məgwh (帽) 'hat'; TB *nup ≍ *nip 'enter, sink', OC *nəbh (內) 'inside'; TB *mu:k 'foggy, dark', OC *mjugh (霧) 'fog'. The *-s (written as *-h in Li Fang-kuei's system) of OC only occasionally matches up with cognates in TB languages, as in Written Tibetan rmugs [rmuks] 'thick fog', though the nominalizing *-s we find in TB is presumably cognate to the OC *-s. In terms of phonetic

motivation for sound change, Baxter (1992:311) also mentions that the traditional view that *-b merged with *-d has no phonetic motivation (as *-p did not merge with *-t), whereas *-ps > *-ts can be explained as assimilation of the final to the suffix.

The qusheng (and possibly the shangsheng) derivation hypothesis assumed by Pulleyblank and Baxter explains quite a bit of the variation of finals in Chinese, but not all of it. One important reason for the reconstruction of a series of voiced stop finals is the supposed symetry of the contacts between different finals. But do we really find a symetrical system of variation? The evidence is that we do not. Out of the 99 tong yun rhymes marked in Wang Li's Shijing Yundu (1980b), 48 are *-Ø (Li's *-g) ~ *-k, and 15 are *-w (Li's *-gw) ~ *-k. Except for the well known shift of *-m, *-n to *-ŋ (8 and 6 tokens respectively), no other pattern shows such regulariity (i.e. all have four or less tokens). If we reconstruct the yōu (幽) and xiāo (宵) rimes as *-əw and *-aw respectively (rather than as Li's *-əgw, *-agw) and the jué (覺) and yào (藥) rimes as *-əuk and *-auk respectively (rather than as Li's *-əkw, *-akw), then the total number of tong yun rhymes where the difference is the presence or not of a final *-k is 63, or 62%.[18] This is quite significant, statistically, given the large number of tong yun possibilities.[19] In some cases this *-k may be a type of derivational morpheme, as suggested by Puleyblank (1972:13, 1973:122) as an explanation for the correspondence between the pronouns *gwək (或) 'some one', *mak (莫) 'no one', *djəkw (孰) 'which one' and possibly *krak (各) 'each' and the forms *gwjəgx (有) 'there is', *mjag (無) 'there is not', *djəd (誰) 'who', and *kjagx (舉) 'all' respectively. Karlgren himself (1933: 37) mentions that in those cases where a TB form with an open final corresponds to a *-k or *-t final in OC (e.g. 'hundred'), 'these -k and -t must be an innovation, some kind of suffix in one or several Sinitic languages but not primary and common to them all.' He does not take the obvious step and use this to explain the same type of variation within Old Chinese. Examples involving variation of final *-t would include the forms from Pulleyblank (1991:30) given in section 2 above, and the different negative particles used in OC: *pjəg (不): *pjət (弗); *mjəg (毋): *mjət (勿) (see Takashima 1988). Pulleyblank (1991) suggests that Sino-Tibetan had morphological *-n and *-t suffixes to explain the correspondences among these items and between certain other words in Chinese (such as *ŋjagx (語) ~ *ŋjan (言) 'language, speech') (cf. the *-n 'collective' suffix suggested by Benedict (1972:157ff)). If we accept the *-g final hypothesis, we have to say that the phonetic and semantic similarities of these two items (and dozens of pairs like them) are entirely coincidental, whereas if we take these *-g finals to actually be open finals, then it is a simple matter of *-t/*-n suffixation.

Some variation may also be due to a coalescence of two forms, as suggested for Tibetan by Walter Simon (1941, 1942, 1957). Simon's idea was that many of the finals in Tibetan, such as -g, -n, -l, -r, -s were from the coalescence of two syllables, the second of which originally also had lexical content, such as

-s < sa/so 'place'. We find synchronic variation in Tibetan that points to this kind of development, such as da-ra ~ dar-ba 'type of buttermilk', źa-la ~ źal 'clay', bu-ga ~ bug 'hole', lco-ga ~ lcog 'lark', nya-ga ~ nyag 'steelyard', yi-ge ~ yig 'letter', and tha-ga-pa ~ thag-pa 'to weave'. If Proto-Sino-Tibetan had a particle similar to Tibetan -ga, which Das (1902:203) says 'is sometimes used as an affixed particle of a word to complete it', then this would be at least one explanation for the large number of *-Ø ~ *-k variations. Aside from the possibility of coalescence resulting in *-k, and the examples of coalescence we are familiar with in Chinese (e.g. 諸 from 之於), coalescence might explain at least a few of the other odd finals in OC. For example, in one cognate set suggested by Wang Li (1982:435) with 何 *gar, 曷 (害) *gat, and 胡 *gag, all question particles, Wang includes 盍 *gap 'negative question ('why not') particle' which according to a commentator on the Guo Yu (國語) is from the coalescence of *gar and *pəg (何不). Changes in the pronunciation of characters caused by their use in connected speech is also suggested by Gong (Mei & Gong 1992:676) as a reason for some characters having unusual pronunciations.

Yet I am not suggesting that these are the only answers. There most probably are other explanations as well. Coblin (1976:52) mentions that in Tibetan 'each verb whose perfect, future and imperative forms end in root final -ŋ has final -n in its present root' (e.g. ɦphen, ɦphaŋ, ɦphaŋs, phoŋ/phaŋs 'throw, cast'). Modifying an idea from Shafer (1951:1028–9), he suggests that the present forms originally had a -d suffix (some forms show this suffix in older texts), and that the -n final was due to assimilation to this suffix. It may be that some such assimilatory process could explain some of the variations between homorganic stop and nasal final in Chinese as well. All these variations may be due to a combination of factors, some morphological, some phonetic. An example of the latter is the change of some PST velar finals to OC dental finals after high front vowels.[20] One type of variation may even have multiple sources (e.g. Mei (1980:439) suggests that the qusheng *-s may have had more than one source). Future research would of course be needed to sort out which process determined which variations, and if possible, what motivated the different processes, as has been done in isolating and understanding qusheng derivation (see the references mentioned above, especially Mei 1980).

The system of finals I suggest for PST, and the regular correspondences between OC, PTB, and PST, then are as follows:

PST **-Ø	>	OC *-Ø	PTB *-Ø
PST **-p	>	OC *-p	PTB *-p
PST **-t	>	OC *-t	PTB *-t
PST **-k	>	OC *-k	PTB *-k
PST **-ŋ	>	OC *-ŋ	PTB *-ŋ
PST **-w	>	OC *-w	PTB *-w
PST **-y	>	OC *-y	PTB *-y
PST **-l	>	OC *-y/-Ø	PTB *-l

PST **-r	>	OC *-y/-n	PTB *-r
PST **-s	>	OC *-t	PTB *-s

This set is similar to that proposed in Baxter 1992. Below I compare the rimes proposed in Li 1980 with those in Baxter 1992, TB forms and my proposed ST forms.

rime	*Li*	*Baxter*	*TB*	*ST*	*# of sets in Appendix*
之	-əg	-ɨ(/-ɨks)	-a	-ə(/-əks)	8
職	-ək	-ɨk	-ak	-ək	8
蒸	-əŋ	-ɨŋ	-aŋ	-əŋ	3
幽	-əgw	-u(/-uks)	-uw	-əw(/-əuks)	8
覺	-əkw	-uk	-uk	-əuk	7
中(冬)	-əŋw	-uŋ	-uŋ	-əuŋ	4
緝	-əp/-əbh	-[ɨ,u,i]p(s)	-ap/-up	-əp/-up[21]	7
侵	-əm	-ɨm	-am/-um	-əm/-um	8
微	-əd	-ɨj(/-its)	-əy/ər/ey/iy	-əy/ər/ey/iy	11
文	-ən	-ɨn/-un	-ul/-un	-ul/un	6
物	-ət	-ɨt/-ut	-ay/-at	-ət	1
歌	-ar	-aj	-a/-ay/-al	-a/-ay/-al	17
月/祭	-at/-ad	-at(/-ats)	-at	-at	9/(1)
元	-an	-an	-an/-ar	-an/-ar	17
葉	-ap/-abh	-ap(/-aps)	-ap/-ep	-ap/-ep	6
談	-am	-am	-am	-am	3
魚	-ag	-a(/-aks)	-a	-a(/-aks)	30
鐸	-ak	-ak	-ak	-ak	6
陽	-aŋ	-aŋ	-aŋ	-aŋ	6
宵	-agw	-aw	-aw/-uw	-aw/-uw	7
脂	-il	-ij(/-its)	-iy	-iy(/-its)	9
眞	-in	-in	-in/-il	-in/-il	7
佳(支)	-ig	-e(/eks)	-i	-i(/eks)	2
質	-it	-it	-it/-ik	-it/-ik	9
錫	-ik	-ek	-ik	-ik	2
耕	-iŋ	-eŋ	-iŋ	-iŋ	9
侯	-ug	-o(/-oks)	-uw	-uw(/-oks)	9
屋	-uk	-ok	-uk	-uk	5
東	-uŋ	-oŋ	-uŋ/-waŋ	-uŋ/-waŋ	3
					227

It can be seen from this comparison that a system such as Baxter's, without voiced stop finals, is closer to the independently reconstructed TB forms, and allows us to reconstruct a more phonetically and typologically plausable Sino-Tibetan system than one with voiced stop finals. [22]

4. Conclusions

There are several points I would like to make in this paper. First, just as we find a certain amount of both rule-governed and non-rule governed variation in modern languages, it is necessary to recognize the same types of variation in the proto-language we are attempting to reconstruct. Second, the variation we find in PST and its immediate daughters is not as symmetrical and orderly as has been assumed. Third, the causes of the variation are complex and multifarious. Fourth, reconstructing a complex, typologically unlikely system based on broad generalizations such as the voiced stop final hypothesis not only is unsatisfactory from the typological point of view, but also effectively ends our search for the real causes of the variation. As mentioned earlier (footnote 13), Li Fang-kuei saw the stop final hypothesis as a stopgap measure, not the final solution. Especially given how little we really know about Sino-Tibetan lexical morphology, to limit the possibilities we are willing to consider would be very unwise. Fifth, the concept of word families is an important one, but we should not be unnecessarily constrained in our search for cognate sets by artifacts of our reconstructed system or methodology.

While recognizing the existence of variation, it is also important to emphasize that in terms of methodology we can only recognize variation within the context of regularity. We must first establish solid regular correspondences to establish what is regular, and to serve as the anchor that allows us to be able to talk about variation. For example, I can feel confident that OC *rap 'leaf' and TB *la 'leaf' are cognate (even if I did not know about the *la ≍ *lap variation within TB) because the initial and the vowel correspond regularly (i.e. there are half a dozen or more parallel examples of each) and the meanings match exactly. We should not push etymologies or cognate sets where we have to explain variation of almost every segment in the forms, as for example when Benedict (1987:48) attempts to support a proposed shift in Chinese from *s-k- to *t- by comparing TB *mkha 'sky, heaven' with Chinese tian 天 'sky, heaven', which he reconstructs as *skhien/thien, giving PST *(-)ka(-n) 'with the PST "collective" plural *-n suffix (= "the heavens") (reg. vowel shift before final dental.)' We then have variation of the prefix, the initial, the vowel, and the final, all within the same set. Were each of these types of variation proposed on the basis of multiple examples of the same type of correspondence appearing in isolation (i.e. the other segments of the forms corresponding regularly), we might be able to accept the cognacy of the forms in such a set, but not only are we asked to accept this set without evidence of such regular correspondences, we are asked to accept this set as corroborating evidence for a proposed development within Chinese!

Appendix: List of suggested OC-PTB correspondences[23]

	Li	*Baxter*	*PTB*	*GLOSS*
魚部				
1.魚	*ŋjag	*ŋ(r)ja	*ŋya	'fish'
2.苦	*khagx	*khaʔ	*ka	'bitter'
3.吾／卬	*ŋag/*ŋaŋ	*ŋa/*ŋaŋ	*ŋa ≍ *ka	'lsg pronoun'
4.五	*ŋagx	*ŋaʔ	*b/l-ha	'five'
5.汝	*njagx	*njaʔ	*na (see below)	'2sg pronoun'
6.狐	*gwag	*gwa	*gwa	'fox'
7.罟	*kwag	*kwa	*gwan ≍ *kwan	'net'
8.斧	*pjagx	*p(r)jaʔ	*r-p-wa	'axe'
9.父	*bjagx	*b(r)jaʔ	*pa (=pwa)	'father'
10.笆	*prag	*pra	*g-p(w)a	'bamboo'
11.豝	*prag	*pra	*p-wak	'pig'
12.雨	*gwjagx	*w(r)jaʔ	*r-wa-ŋ	'rain'
13.于	*gwjag	*w(r)ja	*s-wa (?)	'go'
14.無	*mjag	*m(r)ja	*ma	'no, not'
15.扶／把	*pjag/*pragx	*p(r)ja/*praʔ	*pa-n	'palm'
16.訝	*ŋragh	*ŋras	*ŋra	'meet, encounter'
17.鼠／貉	*hrjagx/*g(l)ak	*h[r,l]jaʔ/*gak	*rwak	'rat, mouse'
18.武	*mjagx	*Np(r)jaʔ	*d-mak	'soldier, war'
19.馬	*mragx	*mraʔ	*mra-ŋ	'horse'
20.簏	*kjagx	*k(r)jaʔ	*kak(PLB,JAM1972:30)	'basket'
21.膊	*phak	*phak	*pak(PLB, JAM 1972:40)	'dismantle'
22.膚	*khwak	*kwhak	*kwak	'skin'

(continued)

	Li	*Baxter*	*PTB*	*GLOSS*
23. 薄	*bak	*bak	*ba	'thin'
25. 戶	*gwagx	*g(w)aʔ	*gwa ≍ m-kha	'door'
26. 雒	*glak	*C-rak	*k-rak	'fowl, bird'
27. 補	*pagx	*paʔ	*pa	'patch, mend' [24]
28. 渡	*dagh	*dak(s)	*da	'ford, cross(a river)'
29. 睹	*tagx	*taʔ	*ta	'see'
30. 夜 / 夕	*riagh/*rjiak	*(l)jAks/*z(l)jAk	*s-la ≍ g-la	'moon' (see Mei 1979)
(惡	*ʔak/ʔag	*ʔak/ʔaks	*WT ʔag	'bad, evil')

	Li	*Baxter*	*PTB*	*GLOSS*
陽部				
1. 涼	*gljaŋ	*g-rjaŋ	*graŋ ≍ *grak	'cool, cold'
2. 望	*mjaŋh	*mjaŋs	*mraŋ	'look, see'
3. 岡	*kah	*kaŋ	*kaŋ (PLB)	'mountain top'
4. 孟	*mraŋh	*mraŋs	*maŋ	'big/older brother'
5. 量	*ljaŋ	*C-rjaŋ	*g-raŋ	'measure/count' [25]
6. 梗 / 硬	*kraŋx/ŋraŋh	*kraŋʔ/ŋraŋs	*krak ≍ *kraŋ	'hard, solid, stiff'

	Li	*Baxter*	*PTB*	*GLOSS*
歌部				
1. 歌	*kar	*kaj	*ka ~ kat	'speech'
2. 沙	*srar	*sCraj	*sa ⩙ *tsa ⩙ *say	'earth, sand'
3. 鵝	*ŋar	*ŋaj	*ŋa-n (see 雁)	'goose'
4. 蜾	*kwarx	*kwajʔ ~ kojʔ	*kway	'bee, wasp'
5. 移	*rar	*ljaj	*lay	'change'
6. 播	*parh	*pajs	*bwar	'spread, sow'
7. 皤	*bar	*baj	*pwa:r	'white'
8. 疲	*bjiar	*b(r)jaj	*bar	'tired'
9. 駕	*krarh	*krajs	*s-ga	'saddle, yoke horses'
10. 唾	*thuarh	*thojs	*m-twa ⩙ s-twa	'spit, vomit, spittle'
11. 妥	*snarx	*hnajʔ	*na-r	'rest, cease motion'
12. 何	*gar	*gaj	*ga-ŋ ⩙ *ka	'what, which'
13. 鹺	*dzar	*dzaj	*tsa	'salt, salty,
14. 荷	*gar	*gaj	*s-gal ⩙ gur	'carry on back'
15. 蝸	*kwrar	*kwraj ~ kroj	*kroy	'snail'
16. 簸	*parx/h	*pajʔ/s	*pway	'husks, shavings'
17. 我	*ŋarx	*ŋajʔ	*ŋay	'lsg pronoun'

	Li	*Baxter*	*PTB*	*GLOSS*
元部				
1. 蒜	*suanh	*sons	*swa-n	'garlic'
2. 餐	*tshan	*tshan ~ *sran	*dza ⩙ *dza	'food, eat'
3. 辮	*bianx	*benʔ	*bat ⩙ *ban	'braid'
4. 雁	*ŋranh	*ŋrans	*ŋa-n	'goose'
5. 炭	*thanh	*thans	*tal ⩙ *dul	'dust, ashes, charcoal'
6. 員 / 圓	*gwjan	*wjɨn	*wal	'circle'
7. 連 / 聯	*ljan	*C-rjan	*ren	'connect'
8. 鮮	*sjan	*sjen	*sar	'fresh'
9. 乾	*kan	*kan	*kan	'dry'
10. 酸	*suan	*son	*swa:r	'sour'
11. 霰	*sianh	*skens	*ser	'sleet/hail'
12. 鐫 / 鑽	*tsjuan/*tsuan	*tsjon/*tson	*tswan	'pointed, to bore'
13. 燔 / 焚	*bjan/bjən	*bjan/bjun	*b(w)ar	'burn'
14. 斷	*duanx	*ŋtonʔ	*da:n	'cut'
15. 瓣	*brianh	*brens	*ba:r	'flower, petal'
16. 散	*sanx/h	*sanʔ/s	*san (PLB, JAM1985#40)	'sow, disburse'
17. 犬	*khwianx	*kwhi/enʔ	*(s-)kwəy	'dog'
(翻	*phin/phjian	*phin	*pyam	'fly')

	Li	*Baxter*	*PTB*	*GLOSS*
祭部				
1. 大 / 多	*dadh/tar	*lats/*taj[26]	*tay	'big'
2. 刺 / 列	*lat/ljat	*C-rat/C-rjat	*(g-)ra-t ≍ *(g-)rya-t	'cut, scrape'
3. 發	*pjat	*pjat	*-pat (PLB, JAM1972:35)	'send forth, vomit'
4. 殺	*sriat	*s(C)rjat	*sat	'kill'
5. 脫	*hluat	*hlot	*g-lwat 0 *s-lwat	'release, let loose'
6. 滅	*mjiat	*mjet	*s-mit	'destroy'
7. 八	*priat	*pret	*b-g-ryat	'eight'
8. 割	*kat	*kat	*(s-)kat	'cut'
9. 撒	*sat	*sat	*sat (PLB, JAM1985#40)	'pour out, disburse'

	Li	*Baxter*	*PTB*	*GLOSS*
葉部				
1. 葉	*rap	*ljap	*la-p	'leaf'
2. 接	*tsjap	*tsjap	*tsyap	'connect'
3. 蓋	*gap ≍ *kabh	*gap ≍ *kaps	*kap	'to cover, cover'[27]
4. 蝶	*diap	*lep	*s-lep	'butterfly'[28]
5. 躡	*ljap	*C-rjap	*rap	'tread(upon), trample'
6. 疊	*diap	*[d,l]ep	*tap	'fold'

	Li	*Baxter*	*PTB*	*GLOSS*
談部				
1. 彡	*sram	*sram	*sam ≍ *tsam	'hair'
2. 嵌	*khram	*khram	*r-kam	'precipice'
3. 談	*dam	*lam	*g-dam	'talk'
(藍	*gram	*g-ram	WT rams	'indigo, blue')

	Li	*Baxter*	*PTB*	*GLOSS*
宵部				
1. 號	*gagwh	*gaws	*gaw/*kaw	'call, yell'
2. 燒	*hrjagw(-hŋrjagw?)	*hŋjew	*tsyow	'cook, burn'
3. 嗷	*ŋagw	*ŋaw	*ŋuw	'cry'
4. 膮	*sagw	*saw	*sa:w	'fat'
5. 漂	*phjiagw	*phjew	*pyaw	'float'
6. 熬	*ŋagw	*ŋaw	*r-ŋaw	'fry, roast'
7. 梟	*kiagw	*kew	*ku	'owl'
(刀	*tagw	*taw	*s-ta	'knife')

	Li	*Baxter*	*PTB*	*GLOSS*
之部				
1. 來	*ləg	*C-rɨ(k)	*ra	'come'
2. 母	*məgx	*m(-r)o/iʔ	*ma	'mother'
3. 子	*tsjəgx	*tsjɨʔ	*tsa	'child'
4. 耳	*njəgx	*njɨʔ	*g/r-na	'ear'
5. 乃 / 汝	*nəgx/*njəŋw	*nɨʔ/*njuŋ	*na-ŋ(cf. 仍 *njəŋ)	'2sg pronoun'
6. 牛	*ŋwjəg	*ŋwjɨ	*ŋwa	'cow'
7. 織	*tjək	*tiɨk	*tak ≍ *trak	'weave'
8. 息	*sjək	*sjɨk	*sak	'breath'
9. 咳	*khəg	*khɨ	*ka:k	'cough'
10. 極	*gjək	*g(r)jɨk	*kak(PLB JAM 1972:31)	'limit, peak'
11. 蝠	*pjək	*pjɨk	*ba:k	'bat'
12. 亟	*djək	*Ljɨk	*dyak (PLB JAM 1972:30)	'really'
13. 翼	*rək	*ljɨk	*lak	'arm, wing'
14. 陟	*trjək	*trjɨk	*l-tak	'ascend'
15. 革 / 鞹	*krək ≍ *kwhak	*krɨk ≍ *kwhak	*kok ≍ *r-kwak	'skin'
16. 右	*gwjəgx	*wjɨʔ(s)	*g-ya ≍ gra	'right (side)'
(黑	*hmək	*hmɨk	*Tib smag	'black')
(友	*gwjəg	*wjɨʔ	*Tib grogs	'friend')
(植 / 直	*djək/drjək	*djɨk/drjɨk	*dzuk	'plant, erect')
(止	*krjəgx	*tjɨʔ	*kriy	'foot')

	Li	*Baxter*	*PTB*	*GLOSS*
蒸部				
1. 夢	*mjəŋ	*mjɨŋ(s)	*smaŋ ≍ *smak	'dream'
2. 蠅	*rəŋ	*jɨŋ	*b-/k-raŋ ≍ yaŋ	'fly'
3. 烝	*tjəŋ	*tjɨŋ	*taŋ	'firewood, pine, fir'

	Li	*Baxter*	*PTB*	*GLOSS*
微部				
1. 衣	*jəd	*ʔjɨj	*g-wa-t	'clothing'
2. 飛	*pjəd	*pjɨj	*byer	'fly'
3. 類	*ljədh	*C-rjut/ps	*tɛrəy	'class'
4. 尾	*mijədx	*mjɨjʔ	*r-may ≍ *mey	'tail'
5. 寐	*mjiədh	*mjits	*r-mwiy ≍ *s-mwiy	'sleep, dream'
6. 火	*hmərx	*hmɨjʔ	*s-mey	'fire'
7. 微	*mjəd	*mjɨj	*mwəy	'small'
8. 紱 / 韍	*pjət	*pjut	*put ≍ *pit	'knee, knee covers'
9. 開	*khəd	*khɨj	*ka	'open'
10. 唯	*gwjəd	*wjijʔ	*wəy(=wiy)	'copula'
11. 桑	*ljədx	*C-rjujʔ	*(s-)rwey	'cane, creeper'
12. 韋	*gwjəd	*wjɨj	*kwər	'skin, hide, leather'
(揩	*gwət	*gut	*r-ko-t	'dig')

	Li	*Baxter*	*PTB*	*GLOSS*
幽部				
1. 目	*mjəkw	*m(r)juk	*mik/*myak	'eye'
2. 腹	*pjəkw	*p(r)juk	*puk	'belly'
3. 六	*ljəkw	*C-rjuk	*d-ruk	'six'
4. 毒	*dəkw	*duk	*duk/*tuk	'poison'
5. 九	*kjəgwx	*k(w)juʔ	*d-guw/d-gaw	'nine'
6. 舅	*gjəgwx	*g(r)juʔ	*kuw	'uncle'
7. 寶	*pəgwx	*puʔ	*puw	'precious'
8. 菽	*hnjəkw(?)	*stjiwk	*s-nuk	'bean'
9. 鳩	*kjəgw	*k(r)ju	*kuw	'pigeon'
10. 牢	*g-ləgw	*C-ru	*kuok(PLB, JAM1973:31)	'pen, corral'
11. 夙	*sjəkw	*sjuk	*C-sok(PLB, JAM1972:55)	'morning, early'
12. 柔	*njəgw	*nju	*now	'soft'
13. 帽	*məgwh	*muks	*r-mok	'hat, wear on head'
14. 篤	*təkw	*tuk	*tu:k ≍ *tow	'thick'

	Li	*Baxter*	*PTB*	*GLOSS*
文部				
1. 銀	*ŋjiən	*ŋrjɨn	*ŋul	'silver'
2. 本	*pənx	*pɨnʔ	*pul	'root'
3. 糞	*pjənh	*pj[ɨ,u]ns	*pun	'dung, fertilizer'[29]
4. 鈍	*dənh	*duns	*dul	'dull'
5. 貧	*bjiən	*brjɨn	*bul	'poor'
6. 昏 / 悶	*hmən/mənh	*hmun/*mɨns	*s-mun ≍ *r-mun	'dark, dull, stupid'[30]
(洗	*siənx	*sɨnʔ	*m-s(y)il	'wash')
(蠅	*pjən	*pjɨn	*byer	'fly')

	Li	*Baxter*	*PTB*	*GLOSS*
中部				
1. 中	*trjəŋw	*k-ljuŋ	*tsyu:ŋ=tu:ŋ	'middle'
2. 弓	*kjəŋw	*kwjɨŋ	*ku:ŋ	'bow'
3. 躬	*kjəŋw	*k(r)juŋ	*guŋ	'body'
4. 蟲	*drjəŋw	*lrjuŋ	*dyuŋ	'bug'
(宮	*kjəŋw[31]	*k(r)juŋ	*kyum	'house')

	Li	*Baxter*	*PTB*	*GLOSS*
緝部				
1. 立	*gljəp	*C-rjɨup	*g-ryap	'stand'
2. 泣	*khljəp	*khrjɨp	*krap	'cry'
3. 汲	*kjəp	*g(r)jɨp	*ka:p	'draw water'
4. 捻	*niəp	*nep	*nyap	"pinch'
5. 習	*rjəp	*zlj[ɨ,u]p	*s-lap	'learn/teach'
6. 内 / 入	*nəbh/njəp	*nups/njup	*nup ≍ *nip	'enter/sink'
7. 十	*djəp	*gjip	*gip	'ten'

	Li	*Baxter*	*PTB*	*GLOSS*
侵部				
1. 熊	*gwjəm	*wjum	*d-wam	'bear'
2. 含	*gəm	*g[o,u]m	*gam	'hold in mouth'
3. 飲	*ʔjəmx	*ʔ(r)jumʔ	*am	'drink'
4. 荏	*njəmx	*njɨmʔ	*njam	'soft'
5. 煖	*rjəm	*zljum	*lum	'warm'
6. 三	*səm	*sum	*g-sum	'three'
7. 枕	*krjəmx	*Kjumʔ	*kum	'pillow'
8. 針	*krjəm	*kj[ɨ,i]m	*kap	'needle'
(林	*gljəm	*C-rjɨm	Lushai ram	'forest')

	Li	*Baxter*	*PTB*	*GLOSS*
侯部				
1. 軀	*khjug	*kh(r)jo	*s-kuw=s-kəw	'body'
2. 口	*khugx	*kh(r)oʔ	*kuw (GB)	'mouth'
3. 乳	*njugx	*njoʔ	*nuw/*nəw (DL nuŋ^{55})	'breast, milk'
4. 谷	*kuk	*kok	*grok	'ravine'
5. 寇	*khugh	*khos	*r-kuw	'steal, thief'
6. 豆	*dugh	*dos	*tu-ŋ (Cf. xiesheng)	'bean'
7. 角	*kruk	*drok	*kruw	'horn'
8. 摳	*khug	*kho	*ku	'lift, raise'
9. 務	*mjugh	*m(r)jo(k)s	*mow	'effort, work'
10. 霧	*mjugh	*m(r)jo(k)s	*muw ≍ *mu:k	'fog'
11. 局 / 曲	*gjuk/*khjuk	*ɦkh(r)jok/*kh(r)jok	*guk/*kuk	'bent'
12. 嗽	*suk	*sok	*su (w)	'cough'
13. 屬	*djuk	*djok	*dźuk (PLB)	'vulva' (see Mei 1979)
14. 樹	*djugh	*djos	*dzuk	'plant, erect'

	Li	*Baxter*	*PTB*	*GLOSS*
東部				
1. 孔	*khuŋx	*khoŋʔ	*kuŋ	'hole'
2. 洞	*duŋh	*doŋs	*dwa:ŋ	'cave, pit, hole'
3. 巷	*gruŋh	*groŋs	*g-rwa-ŋ	'village/street'

	Li	*Baxter*	*PTB*	*GLOSS*
脂部				
1. 二	*njidh	*njijs	*g-ni-s	'two'
2. 四	*sjidh	*s(p)jij/ts	*bliy	'four'
3. 死	*sjidx	*sjijʔ	*siy	'die'
4. 屎	*hrjidx	*xjijʔ	*kliy	'shit'
5. 細	*sidh	*si[j,t]s	*ts(y)iy ≍ *ziy	'small, fine'
6. 姊	*pjidx	*pjijʔ	*piy	'grandmother'
7. 邇	*njidx	*njejʔ	*ney	'near'
8. 水	*hwrjidx	*h[l]jujʔ	*lwi(y)	'water'
9. 日	*njit	*njit	*niy(=nəy)	'sun, day'
10. 汁	*tshjit	*tshjit	*tsiy	'juice, paint'
11. 血	*hwit	*hwit	*s-hwiy(=s-sywəy)	'blood'
12. 畀	*pjidh	*pjits	*biy (DL biŋ)	'give'
13. 節	*tsit	*tsik	*tsik	'joint'
14. 蝨	*srit	*srit	*s-rik=*śrik	'louse'
15. 結	*kit	*kit/k	*kik	'tie'
16. 篳	*pjit	*pjit	*pyik(JAM1970:26)	'thicket'
17. 蛭	*tjit	*dujit	*m-li:t	'leech'
18. 一	*ʔjit	*ʔjit	*it	'one'
19. 米	*mid	*mij	*may ≍ *mey	'rice'
(脂	*tjid	*kjij	*tsil	'fat')
(切	*tshit	*tshit	*tsyat	'cut')
(堅	*tsjit	*tsjik	*WT rtsig-pa	'masonry, etc.')

	Li	*Baxter*	*PTB*	*GLOSS*
眞部				
1. 眠	*min	*min	*myel	'sleep'
2. 辛	*sjin	*sjin	*m-sin	'liver'
3. 齗	*snjinx	*hjinʔ	*r-nil ≍ *s-nil	'gums'
4. 年	*nin	*nin/ŋ	*niŋ	'year, harvest'
5. 薪	*sjin	*sjin/ŋ	*siŋ ≍ *sik	'wood, tree'
6. 民	*mjin	*mjin	*r-mi	'people, person'

	Li	*Baxter*	*PTB*	*GLOSS*
耕部				
1. 頸	*ljingx	*C-reng?	*m-ling	'neck'
2. 生 / 腥	*sriŋ/siŋ	*srjeŋ	*śriŋ	'live, raw'
3. 名	*mjiŋ	*mjeŋ	*r-miŋ	'name'
4. 定	*diŋh	*deŋs	*diŋ	'certain'
5. 盈	*riŋ	*ljiŋ	*bliŋ	'full'
6. 平	*bjiŋ	*brjeŋ	*pleŋ	'flat'
7. 脛	*giŋ	*geŋ/kh-ljeŋ(?)	*r-k(y)aŋ	'leg/shank'
8. 赬	*hljiŋ	*hljeŋ	*kyaŋ	'red'
9. 狌	*sriŋ	*srjeŋ	*sre-ŋ	'weasel'
(清	*tshjiŋ	*tshjeŋ	*tsyaŋ ≍ *syah	'clean, clear, pure')[32]

	Li	*Baxter*	*PTB*	*GLOSS*
佳部				
1. 滴	*tik	*tek	*tki ≍ *tsak	'drip, drop'
2. 隻	*tjik	*tjek	*g-tyik	'one'
3. 縊	*?igh	*?jeks	*?ik	'strangle'

Notes

* An earlier version of this paper was presented at the 25th International Conference on Sino-Tibetan Languages and Linguistics, Oct. 14–18, 1992, U.C. Berkeley. I would like to thank all those who gave me comments on early drafts of this paper, especially William Baxter, W. South Coblin, James A. Matisoff, Tsu-lin Mei, Edwin G. Pulleyblank, Jackson T.-S. Sun, Pang-hsin Ting, and an anonymous reviewer.

1 Maru has innovative -uk and -it appearing whereever the cognate forms in other languages would lead us to reconstruct *-uw and *-iy respectively (Burling 1966, contra Wolfenden's (1939) view that the Maru -k is original). In Lepcha (Maniwaring 1876:93) causatives are formed by infixing -y- after the initial consonant (e.g. thór 'to escape', thyór 'to cause to escape'). If the final consonant of the simplex form is -ŋ, then the corresponding final in the causative form is -n (e.g. hróŋ 'to ascend', hryón 'to cause to ascend').

2 Shafer (1951:711) uses 'morphophonic' to refer to morphophonemic alternation of vowels, and 'morphosymphonic' for the morphophonemic alternation of consonants.

3 Most of the Tibeto-Burman reconstructions I will be discussing are from the work of Paul Benedict, especially Benedict 1972, and James A. Matisoff (e.g. 1978, 1985, 1989, 1992), though some are from Coblin 1986 or are reconstructions/word families I have put together myself (see LaPolla 1987 and also the appendix to this article). As the works just mentioned cite many of the same examples, I will not mark the source of each individual example.

4 I will here use the spelling 'rime' to mean the part of the syllable excluding the initial consonant or cluster (itself simply called the 'initial'), and 'rhyme' for the usual sense of this word as the poetic use of assonance.

5 This is where two characters share the same phonetic component.

6 The reconstructed forms for Old Chinese I will be using in the body of the paper are based on the system outlined in Li 1980, including forms adapted from other sources.

7 A number of the items mentioned below (e.g. 艾，熾，退，賴) are considered ru sheng (入聲) rhymes by Wang Li, due to his hypothesis that OC ru sheng words could be divided into 'long ru' (長入) and 'short ru' (短入) tones, where the long ru became Middle Chinese qu sheng words, while the short ru remained ru sheng words, yet are considered qu sheng words in OC by Li Fang-kuei. As I am using Prof. Li's system in this paper, I have modified some of the examples taken from Wang Li's work to conform to Prof. Li's system.

8 The type of rhyme where the finals are the same but the vowels are different (known also as hé yùn or as 'side transfers' (páng zhuǎn 旁轉)) are not relevant to the present discussion and so will not be discussed here.

9 Wang Li (1980b:334) considers 躬 to be in the 侵 (*-əm) category, but Li Fang-kuei (1980:43) treats this word as being in the 中 category, and reconstructs it as *kjəŋw. As I am using Prof. Li's system in this paper, I have used his reconstruction here.

10 Wang Li was quite clear about his lack of appreciation for Karlgren's reconstruction of OC: 'In short, Karlgren's research on Middle Chinese phonology was fruitful (是有成績的), but his research on Old Chinese was not very fruitful (是沒有多大成績的)' (1980a:68).

11 This is not to say that I accept Wang Li's system of reconstruction or the cognacy of all the sets he proposed in his 1982 book, but the cognacy of the items in each of the sets given here is difficult to deny on any grounds but the difference in final consonant.

12 I have doubts about how the judgement of what is a phonetic in a particular character and what is not is made. For example, *ʔjəgh/ʔjək (意 / 憶) 'think, remember' has 音 (*ʔjəm) as part of the character. The Shuowen (說文解字) and Karlgren both treat this as a hul yl (會意) character, so *ʔjəm is not seen as a phonetic in this character, but

generally in characters with the heart radical, the rest of the character is the phonetic, and *ʔjəm is phonetic in a large number of other characters (the Shuowen includes 暗痊窨喑黯闇諳欿). Compare this with 短 *duanx, which the Shuowen says has the character 豆 *dugh as its phonetic, and 媼 *ʔəgwx which the Shuowen says has 溫 *ʔən as its phonetic. It seems then the decision as to whether *ʔjəm is or is not a phonetic in *ɔjəgh/ɔjək is not due only to the difference in final, but involves some degree of arbitrariness.

13 Li Fang-kuei (1983:401) mentions that he used *-b, *-d, and *-g 'merely as an orthographic device without going into their phonetic details. There is no Chinese dialect or Sino-Tibetan dialect, so far as I know, in which there are two series of [final] stops' (see also Li 1980:33).

14 See Baxter 1992:332ff and Pulleyblank 1992:372–375 for further typological arguments against reconstructing a system with voiced stop finals for Old Chinese.

15 Cf. Duan Yucai's statement that 'characters with the same phonetic element must be of the same rhyme group' (同聲必同部) (《六書音韻表》, 蘇州保息局本, p. 22, cited at Wang Li 1980a:60).

16 The idea of an *-s suffix to explain the origin of the departing tone goes back to Haudricourt 1954, and the idea of a glottal stop suffix to explain the origin of the rising tone goes back to Pulleyblank 1962 and Mei 1970.

17 See also Li Yifu 1984 for reasons why ji bù (祭部, Li's *-adh) and yuè bù (月部, Li's *-at) should be considered one rime. In Ting's study of the Wei-Jin period cross-rhyming patterns, the vast majority of contacts were between ji (祭部) and yuè (月部) (Ting 1987:62).

18 Because Bodman, Coblin and others see *-gw etc. as a single final rather than seeing the *-w as part of the vocalism, they give the correspondence TB *-k,OC *-kw. Not seeing the *-w/-u- as a possible part of the vocalism causes them to miss seeing the variation between *-gw and *-kw and the variation of *-Ø and *-k as the same phenomenon.

19 It is interesting to note that of the 110 suggested word families Karlgren (1933: 98–100) lists that differ in having a final consonant or not (the latter including those ending in *-g, *-d, and *-b), 57 of them, more than half, involve a velar final (40 *-k, 17 *-ŋ).

20 E.g. OC *tsit (節), PTB *tsik 'joint'; OC *srit (蝨), PTB *ś-rik = *śrik 'louse'; OC *kit (結), PTB *kik 'tie'; OC *pjit (篳) PLB *pyik 'thicket'; OC *nin (年). PTB *niŋ 'year, harvest'; OC *sjin (薪), PTB *siŋ ≍ *sik 'wood, tree'.

21 It may be that ST *-ip and *-im are reflected in OC *-əp and *-əm respectively, as suggested by Gong (1980:468), but I have not found any solid correspondences that would either support or disprove this suggestion.

22 It is not my intention to argue specifically for Baxter's system. It would also be possible to modify Li Fang-kuei's system by removing the voiced finals, much as suggestions have been made to modify it in other ways, such as recognizing the *-s suffix (Mei 1980) and having *r- for lăi (來母) initials (Gong 1990). The good points of Baxter's theory are that it not only incorporates these ideas (both of which originated with Pulleyblank), but that it is a theory worked out character by character rather than by broad generalization.

23 I have evaluated the cognate sets suggested by Benedict (1972, 1987), Bodman (1980), Coblin (1986), Gong (1980, 1990, 1991), Matisoff (1985, 1989, 1992, etc.), Yu Min (1989), and others, plus have put together some new sets. I have been very rigorous and conservative in evaluating the correspondences, including here only those forms for which I have solid PTB reconstructions and the correspondences of which seemed uncontroversial (e.g., I have generally followed the 'same series final' rule). I have excluded all those sets suggested by other authors where only a Written Tibetan form is available, though in a few cases I put likely cognates in parentheses after

the regular correspondences. This does not mean these will not turn out to be valid cognate sets, just that at present we do not have enough comparative data available to reconstruct PTB forms; it is unwise to reconstruct a PTB form based entirely on a Written Tibetan form.

24 The reconstruction of the TB form is based on WB pha, JP ká[31] pa[31], Zaiwa pho[51], Bijiang Nu pha[35], Mawo Qiang spa, Tangut pa (based on the use of *pa (巴) in transliteration), Achang pho[55], and Langsu pho[31].

25 The reconstruction of the TB form is based on WT graŋ WB khraŋ, Geman Deng krung[55], Darang Deng xa[31] rueng[35], Menba dsaŋʔ and Lahu ʔɔ[33] (the etymology for this form is given as PLB *riy in Matisoff 1990, but the etymology suggested here (<*raŋ) better fits the usual Lahu pattern of *-aŋ >-ɔ). Bokar (Bo'erga) Luoba ruŋ 'to measure' may also fit here, though the usual Bokar reflex of PTB *-aŋ is -oŋ, as in jupmoŋ 'dream' (Jackson T-S. Sun, p.c.).

26 Baxter suggests that it is the latter form, meaning 'much, many' (and 侈 *hljajʔ 'great, large') that is cognate to TB *tay, not the former, as usually assumed.

27 The reconstruction of the TB form is based on WT kha gteod 'a cover', sgab-pa 'to cover'; Dulong ta[55] kop[55] 'a cover', kap[55] 'to put a cork in a bottle'; JP ma[31] kap[31], Geman Deng ŋkhap, rGyarung ta pkap 'a cover'.

28 The reconstruction of the TB form is based on WT phye-ma-leb, Lushai pheng-phe-hlep, WB lip-pra, Naxi phe[33] le[31] 'butterfly'.

29 The reconstruction of the TB form is based on JP man[31] phun[33], Darang Deng tш[31] phш[35], Zaiwa phun[55], Langsu phun[35]. WT brun may also be related to this form.

30 This set is tentative, as the PTB form is based on only WT mun-ba 'dark', rmun-po 'dull' heavy, stupid', WB hmun 'dim, dusky, blurred'. I could not find cognates in any other languages (in the materials I had available).

31 Both 宮 and 躬 are in the 冬 rime category, which is often reconstructed with an -m final, which is then said to have changed to a velar nasal. If we accept this hypothesis, then the 'house' set is probably valid and the 'body' set is not, while if we do not accept it (i.e. assume 'body' was always a velar nasal in Chinese), then the 'body' set is valid and the 'house' set is not.

32 Benedict (1972:53) mentions that the TB forms might reflect an old *-ya- ≍ *-i- alternation. If so, this would be a solid cognate set.

References

Baxter, William H. 1992. *A handbook of Old Chinese phonology*. Berlin & New York: Mouton de Gruyter.

Benedict, Paul K. 1972. *Sino-Tibetan: a conspectus*. (Princeton-Cambridge Studies in Chinese Linguistics II, James A. Matisoff, Contributing Editor) Cambridge: Cambridge University Press.

Benedict, Paul K. 1987. Archaic Chinese initials. *Wang Li memorial volumes*, English volume, ed. by Ma Meng et al. (The Chinese Language Society of Hong Kong). Hong Kong: Joint Publishing Co.

Benedict, Paul K. 1991 Proto-Burmese-Yipho nominalizing *-t suffix. *Linguistics of the Tibeto-Burman Area* 14.2:149–154.

Bodman, Nicholas C. 1980. Proto-Chinese and Sino-Tibetan: Data towards establishing the nature of the relationship. *Contributions to historical linguistics: Issues and materials*, ed. by Frans van Coetsem & Linda R. Waugh. Leiden: E. J. Brill.

Burling, Robbins. 1966. The addition of final stops in the history of Maru (Tibeto-Burman). *Language* 42.3:581–86.

Coblin, Weldon South. 1976. Notes on Tibetan verbal morphology. *T'oung Pao* 62:45–70.

Coblin, Weldon South. 1986. *A Sinologist's handbook of Sino-Tibetan lexical comparisons*. Monumenta Serica Monograph Series, Vol. 18. Nettetal: Steyler Verlag.

Das, Sarat Chandra. 1902. *A Tibetan-English dictionary*. Calcutta: The Bengal Secretariat Book Depot.

Dong Tonghe (董同龢). 1954. 《中國語音史》。台北：中華文化出版事業委員會.

Dong Tonghe (董同龢). 1981. 《漢語音韻學》。台北：文史哲出版社.

Downer, G. B. 1959. Derivation by tone change in Classical Chinese. *Bulletin of the School of Oriental and African Studies* 22.2:258–90.

Gesang Jumian (格桑居冕) 1991. 藏文字性法與古藏語音系，《民族語文》1991.6:12–22,35.

Gong Hwang-cherng. 1980. A comparative study of the Chinese, Tibetan, and Burmese vowel systems. *Bulletin of the Institute of History and Philology* 51.3:455–490.

Gong Hwang-cherng (龔煌城). 1990. 從漢藏語的比較看上古漢語若干聲母的擬測. *A collection of essays in Tibetan studies*, Vol. 3. Taipei: Committee on Tibetan Studies.

Gong Hwang-cherng (龔煌城). 1991. 從漢藏語的比較看漢語上古音流音韻尾的擬測. ms. Taipei.

Haudricourt, André. 1954. De l'origine des tons en vietnamien. *Journal Asiatique* 242:69–82.

Karlgren, Bernard. 1933. Word families in Chinese. *Bulletin of the Museum of Far Eastern Antiquities* 5:9–120.

Karlgren, Bernard. 1956. Cognate words in the Chinese phonetic series. *Bulletin of the Museum of Far Eastern Antiquities* 28:1–18.

LaPolla, Randy J. 1987. Dulong and Proto-Tibeto-Burman. *Linguistics of the Tibeto-Burman Area* 10.1:1–42.

Li Fang-kuei. 1945. Some Old Chinese loan words in the Tai languages. *Harvard-Yenching Journal of Asiatic Studies* 8:333–342.

Li Fang-kuei. 1976. Sino-Tai. *Computational Analyses of Asian & African Languages* 3:39–48.

Li Fang-kuei (李方桂. 1980. 《上古音研究》。北京：商務印書館.

Li Fang-kuei. 1983. Archaic Chinese. *The origins of Chinese civilization*, ed. by David Keightley, 393–408. Berkeley & Los Angeles: UC Press.

Li Yifu (李毅夫). 1984. 上古韻祭月是一個還是兩個韻部。《音韻學研究》，第一輯，中國音韻學研究會編, 286–293. 北京：中華書局.

Mainwaring, George Byres. 1876. *A grammar of the Rong (Lepcha) language: as it exists in the Dorjeling and Sikim hills*. Calcutta: Printed by C. B. Lewis.

Matisoff, James A. 1970. Glottal dissimilation and the Lahu high-rising tone: A tonogenetic case-study. *Journal of the American Oriental Society* 90: 13–44.

Matisoff, James a. 1972. *The Loloish tone split revisited*. Research Monograph No. 7. Center for South and Southeast Asia Studies, University of California, Berkeley.

Matisoff, James A. 1978. *Variational semantics in Tibeto-Burman*. Philadelphia: Institute for the Study of Human Issues.

Matisoff, James A. 1985. God and the Sino-Tibetan Copula, with some good news concerning selected Tibeto-Burman rhymes. *Journal of Asian and African Studies*, Tokyo Foreign Languages University, #29.

Matisoff, James A. 1988. *The Dictionary of Lahu* (University of California publications in linguistics, 111.) Berkeley & Los Angeles: University of California Press.
Matisoff, James A. 1989. Palatal suffixes in Sino-Tibetan. Paper presented to the 22nd International Conference on Sino-Tibetan Languages and Linguistics, University of Hawaii.
Matisoff, James A. 1992. On the 'regularity' of Sino-Tibetan sound correspondences. Paper presented to the 25th International Conference on Sino-Tibetan Languages and Linguistics, U.C Berkeley.
Mei Tsu-lin. 1970. Tones and prosody in Middle Chinese and the origin of the rising tone. *Harvard Journal of Asiatic Studies* 30:86–110.
Mei Tsu-lin. 1979. Sino-Tibetan 'year', 'month', and 'vulva'. *Tsing Hua Journal of Chinese Studies* 12.117–133.
Mei Tsu-lin (梅祖麟). 1980. 四聲別意中的層次。《中國語文》1980.6:427–43.
Mei Tsu-lin. 1989. The causative and denominative functions of the *s- prefix in Old Chinese. *Proceedings of the Second International Conference on Sinology*. Taipei: Academia Sinica.
Mei Tsu-lin & Gong Hwang-cherng (梅祖麟、龔煌城). 1992. 上古音對談錄. *Chinese languages and linguistics I: Chinese dialects*, 641–719. Taipei: Institute of History and Philology, Academia Sinica.
Pulleyblank, Edwin G. 1962–63. The consonantal system of Old Chinese (2 parts). *Asia Major* (New Series) 9:58–144, 206–265.
Pulleyblank, Edwin G. 1972. Word families in Chinese: A reconsideration. *Journal of Asian and African Studies* 5:1–19.
Pulleyblank, Edwin G. 1973a. Some new hypotheses concerning word families in Chinese. *Journal of Chinese Linguistics* 1:113–114.
Pulleyblank, Edwin G. 1973b. Some further evidence regarding Old Chinese *-s and its time of disappearance. *Bulletin of the School of Oriental and African Studies* 36.2: 368–73.
Pulleyblank, Edwin G. 1977–78. The final consonants of Old Chinese. *Monumenta Serica* 33:180–206.
Pulleyblank, Edwin G. 1991. Some notes on morphology and syntax in Classical Chinese. *Chinese texts and philosophical contexts: Essays dedicated to Angus C. Graham*, ed. by Henry Rosemont, Jr. LaSalle, III.: Open Court Press.
Pulleyblank, Edwin G. 1992. How do we reconstruct Old Chinese? *Journal of the American Oriental Society* 112.3:365–382.
Shafer, Robert. 1951. Studies in the morphology of Bodic verbs. *Bulletin of the School of Oriental and African Studies* 13:702–24; 1017–31.
Simon, Walter. 1941. Certain Tibetan suffixes and their combinations. *Harvard Journal of Asiatic Studies* 5:388–89.
Simon, Walter. 1942. Tibetan daŋ, ciŋ, kyin, yin, and fiam. *Bulletin of the School of Oriental and African Studies* 10:954–975.
Simon, Walter. 1949. The range of sound alternation in Tibetan word families. *Asia Major* (New Series) 1:1–15.
Takashima, Ken-ichi. 1988. Morphology of the negatives in oracle-bone inscriptions. *Computational Analyses of Asian and African Languages* 30:113–133.
Ting Pang-hsin (丁邦新). 1979. 上古漢語的音節結構. *Bulletin of the Institute of History and Philology* 50.4:717–739.

Ting Pang-hsin (丁邦新). 1987. 上古陰聲字具輔音韻尾說補證。《國立台灣師範大學國文學系國文學報》16:59–66.
Wang Li (王力). 1980a.《漢語史搞》。北京：中華書局.
Wang Li (王力). 1980b.《詩經韻讀》。上海：上海古籍出版社.
Wang Li (王力). 1982.《同源字典》。北京：商務印書館.
Wolfenden, Stuart N. 1936. On certain alternations between dental finals in Tibetan and Chinese. *Journal of the Royal Asiatic Society* 1936:401–416.
Wolfenden, Stuart N. 1937. Concerning the variation of final consonants in the word families of Tibetan, Kachin and Chinese. *Journal of the Royal Asiatic Society* 1936:625–655.
Wolfenden, Stuart. 1939. On the restitution of final consonants in certain word types of Burmese. *Acta Orientalia* 17:153–168.
Yu Min (俞敏). 1989. 漢藏同源字譜稿。《民族語文》1989.1:56–77,6; 1989.2: 49–64.

16

A COMPARATIVE STUDY OF THE CHINESE, TIBETAN, AND BURMESE VOWEL SYSTEMS*

Hwang-cherng Gong

Source: *The Bulletin of the Institute of History and Philology* 51, 3, 1980, 455–89.

I. Introduction

The Sino-Tibetan family comprises hundreds of languages and dialects. Among them, the most important languages having long histories in written form are Chinese, Tibetan, and Burmese. Chinese preserves literature of the first millenium B.C., and of the language in that time we already possess considerable reliable knowledge. For Tibetan there is an Inscription of A.D. 821–822, which was studied by Fang-kuei Li (1956). The earliest document for Burmese is the Myazedi inscription of A.D. 1112, studied by Nishida (1955, 1956). The purpose of this paper is to compare the vowel systems of these three literary languages and to reconstruct the vowels of their parent language.

The development of comparative Sino-Tibetan linguistics is closely connected with progress made in the field of Chinese historical linguistics. When Conrady published his *Eine Indochinesische Causativ-Denominativ-Bildung und ihr Zusammen-hang mit den Tonaccenten* in 1896, the reconstruction of Middle Chinese had not yet begun; consequently, he had to base his comparison on modern Chinese dialects. In 1916 Laufer listed 96 Chinese-Tibetan cognates in the Appendix to his article, "The Si-hia Language." In his comparison he marked most of the Chinese forms with asterisks; however, his reconstruction was made more on *ad hoc* basis than on any clearcut principle.

The first systematic reconstruction of Middle Chinese (called Ancient Chinese at first) was made by Karlgren in 1919 in his "Etudes Historiques." His work provided solid ground for comparative study, and his *Analytic Dictionary of Chinese and Sino-Japanese,* published in 1923, became an indispensable reference book for all students in the field. The road to an extensive comparative study was thus opened.

Simon's "Tibetisch-Chinesische Wortgleichungen: Ein Versuch" (1929) was the first attempt at a systematic comparative study. He gave 338 Chinese and Tibetan likely cognates, with Chinese represented by Middle Chinese forms as reconstructed by Karlgren, and compared them by their final consonants, initials, and vowels. However, as Karlgren (1931:30) pointed out in his review of Simon's work, "Every Chinese vowel seems to correspond to a whole row of different Tibetan vowels, and each of these Tibetan vowels in its turn corresponds to a long series of Chinese vowels." From such correspondences it would be difficult to reconstruct the vowels of the parent language.

However, failure in establishing sound correspondences was largely due to the circumstance that the historical study of Chinese and Tibetan had not yet been finished. Karlgren first published some research on Archaic Chinese (called Old Chinese in this paper) as early as 1923 in his *Analytic Dictionary;* and the research was later joined by Simon and Fang-kuei Li. But achievements in this field were not applied in comparative study until after 1940 when Karlgren published *Grammata Serica.* During this time there was very little progress in the study of Tibetan historical phonology, one of the few thoroughly modern linguistic approaches to the internal reconstruction of Tibetan being made by Li in 1933.

Grammata Serica replaced *Analytic Dictionary* and became the pivot in the comparative study. Most authors turned then to Archaic Chinese for comparison, with the exception of R. Shafer, who remained with Ancient Chinese.

Karlgren's Archaic Chinese system later was partially revised by subsequent study. Tung (1945) recombined the two parts respectively, of the rhyme categories *yü* 魚 and *hou* 侯, which had been split by Karlgren, and reconstructed a final consonant *-g for all the members of these two classes. Tōdō (1957) reconstructed a final consonant *-r for the *ko* 歌 rhyme category. Archaic Chinese thus appeared to be a language without open syllables, as Simon argued long ago. But on the other hand Wang Li (1957) reconstructed a whole series of open syllables for the *yin-sheng* 陰聲 part of the *yü* 魚, *hou* 侯, *chih* 之, *chih* 支, *yu* 幽, and *hsiao* 宵 rhyme categories. Opinions were divided so far as the final consonants were concerned.

In 1971 Fang-kuei Li published his "Studies on Archaic Chinese Phonology." This article, written tersely in 61 pages, integrates new developments in this field during the previous decades and contains many new solutions of his own to various problems in this reconstruction. Chinese has been regarded as a language with a complicated vowel system. In comparative Sino-Tibetan research, this has been a great obstacle. Li, starting from his basic hypothesis of "the same rhyme category, the same vowel," revising the theory of Yakhontov (1963) and Pulleyblank (1962–63) concerning the -l- medial for words of the second division, and explaining the double rhymes in the second division with *ia, arrived at a reconstruction which is in accordance with his basic hypothesis. The vowel in the *chih* 脂, *chen* 眞, *chia* 佳, and *keng* 耕 categories was reconstructed as *e by Karlgren. This reconstruction has been generally accepted since then and it seems that no one has ever wondered why there was *u and no corresponding *i. An Archaic Chinese rhyme category generally contains words of all divisions, but it happens

that the four rhyme categories mentioned above lack words of the first division. Li changed *e into *i and solved all these problems in one stroke. At the same time, the reconstruction *i solves a puzzle in our comparative study.

It is on this reconstruction that I base my study. I have examined many proposed cognate words and selected those which seem certain to me, added some of my own findings, and tried to fix the rules govering these cognate words. I do not mean to deny other kinds of correspondences, but I think a substantial number of examples should be required to establish them.

In the following examples, Chinese tones "level," "rising," and "departing" are designated by A, B, and C in order to facilitate comparison with the Burmese tone system.

II. The vowel systems of Old Chinese, Written Tibetan, and Written Burmese

According to Li (1971:24) there are four vowels: *i, *u, *ə, *a, and three vocalic clusters: *iə, *ia, *ua in Old Chinese.

Vowels	i	u	Vocalic clusters: iə, ia, ua
	ə		
	a		

In written Tibetan there are five vowels: i, u, e, o, a.

i	u
e	o
a	

The vowel system of Written Burmese needs some explanation before it can be applied in comparative study. For convenience of discussion, I cite the following list of finals given in Pulleyblank (1963:216):

	Level	Creaky	Heavy		Final Stop	
(a)	ā	a	āḥ	[a]		
	aŋ	aŋ.	aŋḥ	[iŋ]	ak	[ɛʔ]
	ań	ań.	ańḥ	[iŋ, i, ɛ]	ac	[iʔ]
	an	an.	anḥ	[aŋ]	at	[aʔ]
	am	am.	amḥ	[aŋ]	ap	[aʔ]
(i)	ī	i	īḥ	[i]		
	in	in.	inḥ	[eiŋ]	it	[eiʔ]
	im	im.	imḥ	[eiŋ]	ip	[eiʔ]
(u)	ū	u	ūḥ	[u]		
	un	un.	unḥ	[ouŋ]	ut	[ouʔ]
	um	um.	umḥ	[ouŋ]	up	[ouʔ]
(e)	e	e.	eḥ	[e]		

(ai)	ay	ai.	ai	[ɛ]	
(o)	o'	o.	o	[ɔ]	
	oŋ	oŋ.	oŋḥ	[auŋ]	ok [auʔ]
(ui)	ui	ui.	uiḥ	[o]	
	uiŋ	uiŋ.	uiŋḥ	[aiŋ]	uik [aiʔ]

"Level," "Creaky," and "Heavy" represent three different tones. In the present study they will be designated as A, B, and C, respectively. As the vowel length is correlated with the tones and has no phonemic significance, it will be omitted in my transcription. The above list shows that Written Burmese, like Written Tibetan, has five vowels. Irregularity in the distribution, however, suggests that this is not original. As we can see from the table, only a, i, and u can combine with final consonants -m/-p and -n/-t, whereas e and o cannot. The vowel e always occurs alone, whereas o occurs only in front of velar finals -ng/-k. Shafer (1941:22) posits the following shifts:

*-ing > -əing > -ań
*-ik > -əik > -ać

Parallel to this are the shifts:

*-ung > -əung > -əung (transcribed as -ong)
*-uk > -əuk > -auk (transcribed as ok) > -auʔ

By means of this postulation, the parallelism in distribution of a, i, and u is restored.

-ang(k)	*-ing(k) > ań(ć)	*-ung(k) > -ong(k)
-an(t)	-in(t)	-un(t)
-am(p)	-im(p)	-um(p)

However, the counterpart of o[ɔ] is not e[e], but ai[ɛ], as can be seen from their sound values given in square brackets. From the way they are written in the Burmese writing system and from their modern pronunciation, it can be easily inferred that *au and *ai have undergone the following shifts:

*-au > [ɔ] (transcribed as -o)
*-ai > [ɛ] (transcribed as -ai)

It seems that what we usually transcribe as o has in fact two sources: *-u- (in -ung and -uk) and *-au. The former must have already broken into -au- and coalesced with original *-au at the time the Burmese writing system came into being. The later divergent development is conditioned by the presence or absence of the final consonants -ng and -k.

**-au	> *-au	> [ɔ]	(transcribed as -o)
**-ung	> *-aung	> [aung]	(transcribed as -ong)
**-uk	> *-auk	> [auʔ]	(transcribed as -ok)

The counterpart of e[e] is ui[o], as their modern pronunciation indicates. In the Myazedi inscription, e is written iy, while ui is written with the sign for u below and the sign for i above the consouant. Blagden (1914) transcribed the sound as ui, and since then this transcription has been generally followed. As for its sound value, opinion differs. Wolfenden (1929:197) supposes it was pronounced like the Dutch colloquial "ui" in huis, buis. Nishida (1955:21–22) takes it to be[ɯ]; and since there are in the inscription words of other origin written -uy (which is later written -we in Written Burmese), he writes it ö in order to avoid confusion. The same sound is transcribed as b1 in Miller (1957: 42), and Pulleyblank interprets it as /iw/ in the article quoted above. In a comparative study, Benedict (1972) posits -ui < *-uw in the main text, but in a new footnote (No. 188, p. 57) declares *-əw to be preferable to *-uw.

In my opinion, a new proposal must take the following facts into consideration: (1) It corresponds to OC *-ug and WT -u; (2) it is written with the signs for u and i in the Myazedi orthography; and (3) the later development shows parallelism with -iy > -e. In order to account for all these, I posit the following sound shifts:

ST *-ug > Myazedi -uï > *-uw > -o
ST *-id > Myazedi -iy > -iy > -e

At the time of the Myazedi inscription, the second element of ui (our uï must have sound close to i, for it was written with that sign. A sound which goes back to *g and is similar to i might have retained the features of both. It seems therefore reasonable to assume that the sound was back (like g), high and unrounded (like i). (I write [ï], which is equal to [ɯ] of the International Phonetic Alphabet.) The shift -uï > -uw can be easily explained by assimilation.

As far as -uing and -uik are concerned, Pulleyblank quotes Shorto to the effect that words with these finals may not be native Burmese. If we exclude them from the list, we get the following system:

1. Closed syllables

(a)	ang	ak
	an	at
	am	ap
(i)	ing > añ	ik > ać
	in	it
	im	ip
(u)	ung > aung (=ong)	uk > auʔ (=ok)
	un	ut
	um	up

2. Open syllables

(a)	a	ay (=ai)	aw (=o)
(i)	i	iy (> e)	—
(u)	u	uy (> we)	uï > uw (=ui)

According to this analysis, the vowel system of Written Burmese goes back to an earlier three-vowel system.

i u
a

III. Vowel correspondences and their reconstruction

A	ST (Sino-Tibetan)	*a		OC (Old Chinese)	a	:
	WT (Written Tibetan)	a	:	WB (Written Burmese)	a	

Involved in this kind of correspondence are Chinese words in the *yü* 侯, *yang* 陽, *ko* 歌, *chi* 祭, *yuan* 元, *yeh* 葉, and *t'an* 談 categories. Difficulties arise when one bases such a comparison on the reconstructioon of Karlgren, who splits the *yü* 魚 category in two parts, one having open syllables with final -o and another having closed syllables with final -ag. The diacritical marks employed to distinguish words of different divisions complicate the matter further and lead to wrong conclusions. The use of Middle Chinese in comparative study and the choice of incorrect cognates also increases confusion. As a matter of fact, the correspondence of ST *a is the clearest one. The problem of medials and final consonants exceeds the scope of the present study and will not be discussed here.

(The number in parentheses refers to the phonetic series in Grammata Serica Recensa)

1.	OC	ngag	B 五	five (58, a)
	WT	lnga		five
	WB	nga	C	five
2.	OC	ngag	A 吾	we, my, our (58, f)
	WT	nga		I, we
	WB	nga	A	I
3.	OC	ngjag	B, C 語	speak (58, t)
	WT	ngag, dngag		speech, talk, word
4.	OC	ngjag	A 魚	fish (79, a)
	WT	nya		fish
	WB	nga	C	fish
5.	OC	khag	B 苦	bitter, suffer (49, u)
	WT	kha		bitter
	WB	kha	C	bitter

6. OC	khag C 苦		difficulty, hardship (*KYSH* 93)
WT	khag-po		difficult, hard
	dka-ba		difficult, hardship
WB	khak		difficult, hard
7. OC	pljag A 膚		skin (69, g)
WT	pags, lpags		skin, hide
8. OC	mjag A 無		not, no (103, a)
WT	ma		not
WB	ma B		not
9. OC	tsjag A, tshjiag B 且		on the point of, will soon (46, a)
WT	cha		to be about, to be on the point
WB	ca B		to begin, make a beginning or commencement
10. OC	dag C 渡		to ford (801, b)
WT	'da		to pass over
11. OC	njag A 如		if (94, g)
	njak 若		if (777, a)
WT	na		if, in case, supposing
12. OC	njag B 汝		you (94, j)
WB	nang A, B		you, your
13. OC	kjag B 擧		lift, raise (75, a)
WT	'khyog pf. khyag		to lift, lift up
14. OC	bjag B 父		father (102, a)
WT	pha		father
WB	a-pha B		father
15. OC	mjag A 巫		magician (105, a)
WT	'ba		magician, sorcerer, conjurer
16. OC	tag B 睹, 覩		see (45, c', d')
WT	lta		to look, to view
17. OC	mrag B 馬		horse (40, a)
WT	rmang		horse, steed (see Coblin 1974)
WB	mrang C		horse, pony
18. OC	prak < *priak 百		hundred (781, a)
WT	brgya < *brya		hundred (see Li 1959 p. 59)
WB	a-ra A < a-rya		hundred
19. OC	·ak 惡		bad, evil (805, h)
WT	ʔag		bad
20. OC	khrjak 赤		red (793, a)
WT	khrag		blood
WB	hrak		to be ashamed, to be shy

21. OC	phjang B 紡		spin (740, r)
WT	phang		spindle
WB	wang B		to spin
22. OC	pjang C 放		loosen, let go, banish (740, i)
WT	spong, spang, pf. spangs		to give up, to renounce
23. OC	dzang A 藏		conceal, to store (727, g')
WT	gsang		to conceal, secret, hidden
24. OC	tsang A 臧		good (772, f')
WT	bzang		good, fair, beautiful
25. OC	trjang A 張		give tension to a bow (721, h)
WT	thang		tense, tight
WB	tang C		to tighten, become tense or taut
26. OC	drjag A, C 除		eliminate, remove, to clear out (82, m)
WT	'dag		to clear, to wash away, to wipe off
27. OC	kang A 岡		a hill, ridge (697, a)
WT	sgang		a projecting hill or spur
WB	khang-A ruw C		a strip of high ground, a spur of a range of mountains or hills
28. OC	njang C 讓		cede, yield, give way (730, i)
WT	gnang		to give, grant, concede
WB	hnang C		to give, deliver over
29. OC	njang A 瀼		heavy with dew (730, f)
WT	*na*-bun		fog, thick mist
	khug-*rna*, khug-*sna*		fog, mist, haze
WB	hnang C		dew, fog, mist
30. OC	nang B 曩		in past time, formerly (730, k)
WT	*gna*-bo		ancient
31. OC	njang A 攘, 穰		expel, sacrifice to expel evil influences (730, e, g)
WB	hnang A		to drive, to drive away
32. OC	krang B 梗		strong (745, e)
	ngrang C 鞏, 硬		hard, stiff, firm
WT	mkhrang, khrang		hard, solid, firm
WB	rang B		mature, firm
33. OC	gljang A 涼		cold (755, l)
WT	grang		cold, cool
34. OC	ljang A, C 量		to measure (737, a)
WT	'grang		to number, to count
	grangs		number
WB	khrang A		to measure with a measure of capacity
35. OC	priat 八		eight (281, a)

	WT	brgyad < *bryad	eight (see Li 1959 p. 59)
36.	OC	tar B, C 癉	disease, suffering (147, e)
			wearied, disease arising from overwork
	WT	ldar	to be weary, tired
37.	OC	ljar A 籬	hedge (23, g)
	WT	ra	fence, enclosure, wall
38.	OC	nga A 鵝	goose (2, p)
		ngran C 雁	wild-goose (186, a)
	WT	ngang	goose
	WB	ngan C	goose
39.	OC	gar A 河	river (1, g)
	WT	rgal	a ford, to ford (a river)
40.	OC	gar A, B 荷	carry (1, o)
	WT	sgal	load of a beast of burden
		'gel, pf. bkal fut. dgal	to load, to lay on a burden
		khal	burden, load
	WB	ka B	to harness, attach a daught animal to a carriage
41.	OC	tjan C 顫	shivering, shaking, trembling (Analytic Dictionary p. 279)
	WT	'dar	to tremble, shudder, shiver
		sdar	trembling
42.	OC	nan A, C 難	difficulty, calamity (152, d)
	WT	mnar	to suffer, be tormented, torture
43.	OC	sjan A 鮮	fresh fish, fresh meat (209, a)
	WT	gsar	new, fresh
	WB	sa B	to make anew, do afresh
44.	OC	kan A 竿	bamboo pole, rod
	WT	mkhar	staff, stick
		'khar	staff
45.	OC	tshan C 燦	bright, splendid (154, b)
	WT	mtshar	bright, shining, of metals, fine, beautiful
46.	OC	djan A 纏	bind, wind (204, c)
	WT	star	to tie fast, to fasten to
	WB	ta A	to cling to
47.	OC	tshan A 餐	eat, food, meal (154, c)
	WT	'tshal	to eat
		'tshal-ma	breakfast
48.	OC	trjan B 展	roll over, unfold, develop (201, a)

	WT	rdal	to spread, to extend
49.	OC	kan A 干	shield, violate (139, a)
		gan C 扞, 捍	to ward off, protect, guard (139, q, i′)
	WT	'gal	violate, to counteract
	WB	ka A	a shield of any kind, to make a barrier against, ward off, debar
50.	OC	han A 鼾	to snore
	WT	hal	to pant, to snort
51.	OC	kan A 乾	dry (140, c)
		gan B 旱	drought, dry (139, s)
	WB	khan C	to be dried up, exhausted, as a liquid
52.	OC	gjan C 健	strong (249, g)
	WB	kyan C	to be well, healthy
53.	OC	pran A 健	turn round (182, a) (*KYSH* 381)
	WB	pran A < plan	to return, to repeat
54.	OC	hrjab C < hrjaps 般	generation, epoch (339, a)
		rap 葉	generation, epoch (633, d)
	WT	rabs	generation
55.	OC	kab C < kaps 蓋	to cover, conceal, a cover (642, q)
		gap 蓋	to thatch, to cover (642, q)
	WT	'gebs, pf. bkab, ft. dgab	to cover, to put on a cover
56.	OC	krap 甲	shell (629, ā)
	WT	khrab	shield, scales
57.	OC	tsjap 接	connect (635, e)
	WB	cap	to join, unite, connect
58.	OC	tam A 擔	carry on the shoulder (619, k, h)
	WB	tam C	to bear or carry on the shoulder
59.	OC	dam A 談	speak (617, l)
	WT	gtam	talk, discourse, speech
60.	OC	phjam C 泛, 汎	to float (641, b; 625, f)
		phjam C 氾	overflow, inundate (626, c)
		bjam A 氾	disperse, float about (626, c)
	WT	'byam	to flow over, to be diffused
61.	OC	grjam A 鹽	salt (609, n)
	WT	rgyam-tshwa	a kind of salt, like crystal
		lgyam-tshwa	a kind of rock-salt

B ST *i

As mentioned above, Li (1971) reconstructs *i for the words in the *chia* 佳, *keng* 耕, *chih* 脂, and *chen* 眞 categories. The vowel of these words corresponds to Tibetan i. Since there are no -im and -ip syllable types in Old Chinese, it seems reasonable to assume that ST *-im and *-ip have shifted to OC *-əm and *-əp and coalesced with the original ST *-əm and *-əp, which are reflected in Chinese words in the *ch'in* 侵 and *ch'i* 緝 categories. In Written Burmese, ST *-ing and *-in have shifted to aṅ, while ST *-ik and *-it have shifted to aḱ. In this analysis of Written Burmese, we have seen syllables like -in and -it. However, it turns out that words with these finals do not play a part in the comparison. It seems that we are dealing here with a renovation.

This irregularity in the correspondences of final consonants seems to have been caused by dialectal shifts in Chinese (*-ing > -in, *-ik > -it) on one hand, and by the morphophonemic alternation in Sino-Tibetan languages on the other.

62.	OC	tik 滴	a drop, to drop (written as 滴 in the *SW*)
	WT	thigs	a drop
		'thig	to drop, to fall in drops
		gtig(s)	to fall in drops, to drop
		btig	to drop, to let fall in drops
63.	OC	mjing A 名	name, fame (826, a)
	WT	ming < *mying	name
	WB	mań A < *ming	to be named, have a name
		hmań B < *hming	to name, give a name
		a-mań A < *a-ming	a name
64.	OC	ljing B 領	neck (823, f)
	WT	'jing < *'lying	neck, to turn or move round
	WB	laṅ A < *ling	neck, to turn around
65.	OC	tsring A 爭	strife, quarrel (811, a)
	WT	'dzing	to quarrel, contend, fight
	WB	caḱ < *tsik	war, battle
66.	OC	· jit —	one (394, a)
	WB	ʔaḱ < *ʔik	a unit, one
67.	OC	srjit 蝨	louse (506, a)
	WT	shig < *syig	louse
68.	OC	njid C 二	two (564, a)
	WT	gnyis	two
	WB	hnaḱ < *hnit	two
69.	OC	sjin A 薪	firewood (382, k)
	WT	shing < *sying	tree, wood
	WB	saḱ < *sik	wood, timber

70.	OC	sjin A 新	new, renew (382, k)
	WB	saḱ < *sik	new
71.	OC	nin A < *ning 年	year (364, a)
	WT	na-*ning*, kha-*ning*	last year
	WB	a-hnaḱ < *hnik	a year
72.	OC	njin A < *njing 仁	kind, good (388, f). cf. 佞 *ning
	WT	snying	the heart, the mind
	WB	hnaḱ < *hnik	heart
73.	OC	rin B 引	draw the bow, pull, stretch, prolong (371, a)
	WT	ring	long, high, tall
		sring	to extend, stretch, prolong
	WB	hraṅ A < *hring	to be long
74.	OC	ljit 慄	fear (403, d)
	WT	'jigs < *'lyigs	to be afraid of a thing, fear, dread
75.	OC	tsit < *tsik 節	knots or joints of bamboo (399, e)[1]
		sjit 卩, 膝	knee (401, c)
		tshit 切	cut (400, f)
	WT	tshigs	joint, knee, knot
	WB	chaḱ < *tshik	to cut in parts
		a-chaḱ < *a-tshik	a joint
76.	OC	kit 結	tie, knot
	WT	'khyig	to bind
	WB	khyaṅ A < *khing	to tie, bind, fasten by tying
77.	OC	njit 日	sun, day (404, a)
	WT	nyi-ma	the sun, day
	WB	ne A < *niy	the sun
		ne C < *niy	a day
78.	OC	tshjit 泰, 漆	vanish (401, a, b)
		tshjid C 鬃	to varnish, to lacquer
	WT	tshi	tough, viscous, sticky matter
	WB	che C < *tshiy	paint, pigment
		ce C < *tsiy	to be sticky, adhesive
79.	OC	tid B 底	bottom (590, c)
		tid A 低	to lower (590, e), low
	WT	mthil	bottom, lowest part
	WB	mre A < mliy[2]ˇ	earth, ground, soil (For the semantic connection, cf. English bottom and German Boden.)

80. OC	hwrjid B 水		water (576, a), The character is phonetic in 癸 gwjid C
WT	rtsi		all fluids of a somewhat greater consistency
WB	re A < *riy		water (For the semantic connection, cf. WT chab < *thyab water and OC tjəp 汁 juice, sap (686, f). For the phonetic correspondence between WT and WB, cf. WT rtsi, to count and WB re < *riy, to count, enumerate)
81. OC	sjid B 死		die, death (558, a)
WT	'chi < *'syi		to die
WB	se A < *siy		to die
82. OC	sjid C 四		four (518, a)
WT	bzhi < *blyi		four
WB	le C < *liy		four
83. OC	tshjit 七		seven (400, a)
WB	khu B hnać < *khu-hnit		seven
84. OC	njid B 爾		you (359, a)
WT	nyid		self, same, thou, you
85. OC	pjid C 畀		give (521, a)
WT	sbyin, pf. byin		to give, to bestow
WB	pe C < *piy		to give, to present for acceptance
86. OC	tjid C 至		arrive, come to (413, a)
WT	mchi < *mtshyi		to come, to go, to appear
WB	ce B < *tsiy		to come, arrive
87. OC	sid C 細		small, minute (1241, l)
WT	se C < *siy		small, fine, slender
88. OC	kjit 吉		luck, auspicious, good (393, a)
WT	skyid		to be happy, happiness
89. OC	dzjin C 盡		exhaust entirely (381, a)
WT	zin		to be consumed (zin-pa med-pa, endless. cf. Ch. 無盡)
90. OC	kjəp 急		hasty (681, g)
WT	grim		to haste, to hurry
91. OC	tshjəm B 寢		lie down to sleep (661, f)
WT	gzim		to fall asleep, to sleep
92. OC	tsjəm C 浸		to soak, overflow (661, m)
WT	sib		to soak in
WB	cim A		to steep, soak in liquor

C ST *u

In Old Chinese there are no syllables of the type *-un(t) and *-um(p). However, the correspondence seems to show that ST *-un(t) and *-um(p) shift to *-ən(t) and *-əm(p) and coalesce with the original -*ən(t) (the *wen* 文 and *wei* 微 categories) and *-əm(p) (the *ch'in* 侵 and *ch'i* 緝 categories), Examples Nos. 121–125.

Examples Nos. 114–119 indicate clearly the sound shift ST *-ul > OC *-ən. In examples Nos. 116 and 117 we have OC *-jiən instead of simply *-jən. This distinction is made to account for the divergent development from Old Chinese to Middle Chinese on one hand, and the different reflexes of labials and labiodentals in Mandarin on the other (see Li 1971: 37–38). For example:

*bjən	>	bjuən	> fen 墳	tumulus
*bjiən	>	bjěn	> p'in 貧	poor

Comparative evidence seems to indicate that this distinction in Old Chinese results in the loss of preinitial d-, and that the phenomenon is essentially a kind of compensatory lengthening. For example:

ST *dbjul → O C *dbjən > bjiən 貧 poor; → W T dbul

ST *dngjul → O C *dngjən > ngjiən 銀 silver; → W T dngul; → Myazedi nguy > W B ngwe

cf. ST *pjul → O C *pjən > pjuən > fen 分 distribute; → W T phul pf. imp. of 'bul to give

93. OC	gug C 候		attend, wait upon (113, e)
WT	sgug		to wait, to await
94. OC	khug B 叩		strike, attack (110, d)
	khug C 叩, 扣		strike (110, d, e)
WB	khauk < *khuk		to khock, rap
95. OC	dug C 逗		to remain, to stay
WT	'dug		to remain, to stay, to live, to sit
96. OC	trjug C 晝		time of daylight, day (1075, a)
WT	gdugs		mid-day, noon
97. OC	mjug C 霧		fog, mist (1109, t)
WT	rmugs		a dense fog
	rmu		fog
	rmus		foggy
WB	mru A khuw C		fog, mist, haze
98. OC	khjug A 軀		body, person (122, g)

	WT	sku	body
	WB	kuwy A	an animal body
99.	OC	njug B 乳	nipple, milk (135, a)
	WT	nu-ma	breast, female breast, bosom
	WB	nuw B	the breast of a female, milk
100.	OC	khug C 寇	to rob, robber (111, a)
	WT	rku	to steal, to rob
	WB	khuw C	to steal
101.	OC	tug C 噣, 咮	beak (1224, n; 128 u)
		trjug A, C 咮	beak (128, u; 1224 n) 噣
	WT	mchu < *mthyu	lip, beak or bill of birds
102.	OC	tjug B 枓	ladle (116, b)
		tjug C 注	to conduct water (129, c)
	WT	'chu < *'thyu	to lade or scoop, to irrigate, to water
103.	OC	kuk 穀	grain, good (1226, i)
	WB	kauk < *kuk	the rice plant, rice
		kaung < *kung C	to be good
104.	OC	tjuk 燭	torch (1224, e)
	WT	dugs	to light, to kindle
	WB	tauk < *tuk	to blaze, flame, to shine
105.	OC	thjuk 觸	to butt, knock against (1224, g) to touch
	WT	thug	to touch, to hit or strike against
		gtug	to touch
106.	OC	khjuk 曲	bend, crooked (1213, a)
	WT	'gug(s)	to bend, to make crooked
		kug	crooked, a hook
	WB	kauk < *kuk	to be crooked, not straight
107.	OC	suk 欶	suck, inhale (1222, o)
	WB	sauk < *suk	to drink, to smoke
108.	OC	khung B 孔	empty (1174, a)
		khung A 空	hollow, empty, hole (1172, h)
	WT	khung	hole, pit, hollow, cavity
	WB	khaung C < *khung	to be hollow
109.	OC	thung C 痛	to be pained (1185, q)
	WT	gdung(s)	to feel pain, to be pained
110.	OC	bung A 蜂	bee, wasp (*KYSH* p. 495)
		phjung A 螽	bee, wasp (1197, s 蜂, t 螽)
	WT	bung	a humming and stinging insect, bee

111.	OC	ljung A 龍	dragon (1193, a), ponetic in 龐
			brung
	WT	'brug	dragon, thunder
112.	OC	tuan B < *tun < **tung 短	short (169, a). The character contains the phonetic 豆 dug
	WT	thung	short
	WB	taung C < *tung	short
		tuw A < *tug	short
113.	OC	tsjət 卒	finish, die (490, a)
	WT	sdud	to close, conclude, finish
114.	OC	kən B 頸	neck
	WT	'gul, mgul	neck, throat
		mgur	neck, throat
115.	OC	pjən A 分	divide, separate, distribute (471, a)
	WT	'bul	to give
		'phul	to give
116.	OC	ngjiən A < *dngjən 銀	silver (416, k)
	WT	dngul < *dngjul	silver, money
	WB	ngwe A < nguy	silver
117.	OC	bjiən A < *dbjən 貧	poor (471, o)
	WT	dbul < *dbjul	poor, poverty
118.	OC	dən C 鈍	dull (427, i)
	WT	rtul	blunt, dull
119.	OC	djən C 順	obey, submissive (462, c)
		sdjən A 馴	docile (462, f), tame
	WT	'dul	to tame, to subdue
		dul	soft, tame, gentle
		'jun < *'djun	to subdue, make tame
		'chun < *'thjun	to be tamed, subdued
120.	OC	pjəd A 飛	to fly (580, a)
		pjən A 翁	to fly, soar (471, e)
		pjən C 奮	spread the wings, fly up (473, a)
	WT	'phur	to fly
121.	OC	hmən A 昏	dusk, evening, darkness (457, k)
	WT	mun	obscurity, darkness
		dmun	darkened, obscured
		rmun	dull, heavy, stupid
	WB	hmun A	to be dim, to be dusky
122.	OC	tsən A 尊	to honour, honorable (430, a)
	WT	btsun	respectable, noble, honourable

123.	OC	njəp 入	enter (695, a) to sink, to set in the expression 日入而息
	WT	nub	to sink, to set, west
	WB	ngup	to dive, to go beneath
124.	OC	səm A 三	three (648, a)
	WT	gsum	three
	WB	sum C	three
125.	OC	khəm A 戕	to kill (651, v)
	WT	'gum	to kill, to put to death

D ST *ə OC *ə : WT a : WB a

The OC vowel *ə in Chinese words of the *chih* 之, *cheng* 蒸, *wei* 微, *wen* 文, *ch'i* 緝, and *ch'in* 侵 categories, which correspond to a in Written Tibetan and Written Burmese, go back to ST *ə. Following are examples of this correspondence:

126.	OC	njəg B 耳	ear (981, a)
	WT	rna	the ear
	WB	na C	the ear
127.	OC	tsjəg B 子	child, treat as a child (964, a)
		dzjəg C 字	to breed, to love, fondle, written character (964, n)
		dzjəg A 慈	affectionate, loving (966, j)
		dzjəg A, C 孳	copulate, breed (966, k)
	WT	tsha < *tsa	grandchild
		btsa	to bear, to bring forth
		mdza	to love, as friends or kinsmen do
	WB	ca A < *dza	to have tender regard, to feel for another, as for one's self, a letter
128.	OC	dzrjəg C 事	serve, affair (971, a)
	WT	rdzas < *dzras	thing, matter, object
	WB	a-ra A < *dzra	a thing, subject, matter
		ca A < *rdza	a thing
129.	OC	məg B 母	mother (947, a)
	WT	ma	mother
	WB	ma B	sister. Compare the similar semantic development in Albania (Jespersen p. 118)
130.	OC	ngəg C 礙 (礙)	obstruct (956, g)
	WT	'gegs-pa pf. bkag fut. dgag	to hinder, prohibit, stop, to forbid
131.	OC	mək 墨	ink, black (904, c)
		hmək 黑	black (904, a)

	WT	smag	dark, darkness
	WB	mang A	ink
		hmang A	ink
132.	OC	dzək 賊	bandit (907, a)
	WT	jag	robbing, robbery
133.	OC	tjək 織	weave (verb) (920, f)
		tjəg C < *tjəks 織	stuff made of coloured silk (noun) (920, f)
	WT	'thag < 'tag	to weave
		thags < *tags	texture, web
	WB	rak	to weave, whether cloth, a mat, or a basket
134.	OC	sjək 息	breathe, sigh, rest (925, a)
	WB	a-sak	breath, life
135.	OC	rəng A 蠅	a fly (892, a)
	WT	sbrang	fly and similar insects without a sting
	WB	yang A	the common house fly
136.	OC	tsəng A 憎	hate (894, d)
	WT	sdang	to hate
137.	OC	hnər B 安	tranquil, at ease (354, a)
		snjəd A 綏	give repose to, calm (354, g)
	WT	rnal	rest, tranquillity of mind
	WB	na C	to cease from motion or action through desire for rest
138.	OC	pjəd B 誹	slander (579, g)
		pjəd A 非	not, wrong (579, a)
	WT	phyar-kha	blame, affront, insult
		'phya-ba	to blame, censure, chide
139.	OC	bjən A 焚	to burn, destroy (474, a)
	WT	'bar	to burn, to catch fire
		sbar	to light, kindle, inflame
	WB	pa B	to shine
140.	OC	sən A 孫	grand-son, grand-daughter (434, a)
	WT	mtshan[3] < *m-san	nephew
141.	OC	mjən A 聞	hear, to be heard (441, f)
	WT	mnyan-pa, nyan-pa	to hear, to listen
142.	OC	gljəp 立	to stand (694, a)
	WT	'khrab	to strike, to stam, tread heavily
		skrab	to beat the ground with one's feet, to stamp, tread
	WB	rap < ryap	to stand; to stop, halt, remain

143.	OC	khljəp 泣	weep (694, h)
	WT	khrab-khrab	a weeper, one that sheds tears on every occasion
144.	OC	təp 答	respond to, answer (676, a)
		təb C < təps 對	respond, reply (511, a)
	WT	'debs pf. btab, fut. gtab	to answer, to explain
145.	OC	sdjəp 習	to practise, exercise (690, a)
	WT	slob pf. bslabs	to learn, to teach
146.	OC	tjəp 摺	to fold
		diəp 疊, 褶	double (690, g) (1255, a)
	WT	ltab	to fold or gather up
	WB	thap	to place one on another, to repeat
147.	OC	kjəp 汲	draw water (681, h)
	WB	khap	to dip up, draw water
148.	OC	tjəp 汁	juice, sap (686, f)
	WT	chab < *thyab	water
149.	OC	sjəm A 心	heart (663, a)
	WT	sem(s), pf. sems, bsams, fut. bsam	to think
		bsam	thought, thinking
150.	OC	njəm B 恁	think (667, q)
	WT	nyam(s)	soul, mind, thought
		snyam	to think, suppose, fancy
151.	OC	gəm A 含	hold in the mouth, put in the mouth (651, 1′)
	WT	'gam	to put, or rather throw, into the mouth

Chinese words in the *yu* 幽 category, reconstructed as -əkw and -əgw by Li, show a different correspondence from the *chih* 之 category. It seems necessary to project the reconstruction of Old Chinese back into Sino-Tibetan.

ST	OC		WT		WB	
*-əkw	-əkw	:	-ug	:	-auk < *-uk	
*-əgw	-əgw		-u(g)		-o < *-uï, -u < *-u	
*-əngw	-ən		-ung		-aung < *-ung	

152.	OC	ljəkw 六	six (1032, a)
	WT	drug	six
	WB	khrauk < *khruk	six
153.	OC	dəkw 毒	poison (1016, a)
	WT	dug, gdug	poison
	WB	tauk < *tuk	to be poisoned

154.	OC	tjəkw 粥, 鬻	rice gruel (1024, a, b)
	WT	thug < *tug	soup, broth
155.	OC	təkw 篤	firm, solid, (1019, g). The *SW* defines it as 厚 thick
		tən < təngw A 惇, 敦	solid, thick (464, n, p). The *SW* defines 惇 as 厚 thick. For the sound change, cf. No. 164 (see Gong, 1976, pp. 63–69)
	WT	'thug < *tug	thick
		mthug < mtug	
		stug(s)	thickness, density, thick, dense
	WB	thu A	to be thick, not thin
		thu B	thickness
156.	OC	sthjəgw B 手	hand (1101, a)
	WT	sug	the hand
157.	OC	hrjəgw A, C 收	collect, harvest (1103, a)
	WT	sgrug, rug	to collect, gather, pluck
158.	OC	njəgw A 揉	to make pliable (1105, b)
	WT	nyug	to besmear, to rub gently
159.	OC	kjəgw B 九	nine (992, a)
	WT	dgu	nine
	WB	kuw C	nine
160.	OC	gjəgw B 舅	maternal uncle (1067, b)
	WT	khu-bo	uncle, on the father's side
	WB	kuw A	brother
161.	OC	krjəgw A 舟	boat (1084, a). The character is the phonetic in 𣏌 gak (?). See *KYSH* p. 713
	WT	gru	boat, ferry, ship, vessel
162.	OC	trjəgw B 肘	wrist, elbow (1073, a)
	WT	gru-mo	elbow
163.	OC	njəgw A 柔	soft, mild, tender (1105, a)
		njəgw A 揉	to make pliable (1105, b)
	WT	nyug	to rub gently
	WB	nu C	soft, to be made soft by some process
164.	OC	tən < *təngw A 墩	a heap, a mound. For the sound change, cf. No. 155 (see Gong, 1976, pp. 63–69)
	WT	rdung	a small mound, hillock
	WB	taung A < *tung	a hill, mountain

IV. Origins of Tibetan -e and -o

The vowels -a-, -i-, and -u- are shared by Chinese, Tibetan, and Burmese, whereas the vowel -ə- was maintained only in Old Chinese. In the above section we have seen how these four vowels correspond in the three languages. What remains to be analyzed now are the Tibetan vowels -e- and -o-, which are not found in Old Chinese as reconstructed by Li, nor in Written Burmese as I interpret it. The origins of these two Tibetan vowels present many difficulties in comparative study.

If we regard Tibetan -e- and -o- as inherited from the parent language, we are obliged to explain how ST *-e- and *-o- developed in Old Chinese and Written Burmese. Conversely, if we regard them as secondarily developed, we are obliged to explain how they came into existence.

It seems to me that we are here dealing with Tibetan innovations. The following facts can be pointed out in support of this view.

A. Tibetan -e- and -o- in the verb paradigm

The morphology of a language often reveals traces of phonetic change, and this seems also to be the case with Tibetan. As is well known, some Tibetan verbs show the following paradigm:

1.	'gebs-pa	'to cover'	pf. bkab,	fut. dgab,	imp. khob
2.	'debs-pa	'to answer'	btab,	gtab,	thob
3.	gson-pa	'to hear'	bsan,	gsan,	gson
4.	slob-pa	'to learn'	bslabs,	bslab,	slob(s)

In the above examples, the vowel -a- occurs in perfect and in future tenses, whereas -o- occurs in the imperative. As for the present form, we have -e- in the first two verbs and -o- in the last two verbs. In these verbs, vowels show e ~ a ~ o and o ~ a ~ o alternation. Since Schiefner (1851), many writers on the morphology of Tibetan verbs have regarded the vowel -a- as original, and the vowels -e- and -o- as secondarily developed, though different writers have different interpretations as to the process of this development. Schiefner noted, that words often were written in two different ways, varying between -a- and -e- or -a- and -o-; for instance, *kag* or *skag* ("unlucky") is also written *keg* or *skeg,* and *cag* (a plural marker) is also written *cog.* He gave twelve pairs of words showing -a- ~ -e- alternation and thirty-four pairs of words showing -a- ~ -o- alternation, and interpreted the change of -a- into -e- as *Schwächung* and -a- into -o- as *Trübung.* According to him, it was originally nothing but a natural phonetic change. Later, as it became necessary to distinguish the tenses of verbs, these coexisting forms were differentiated, with the "weakened" and "muddy" forms designating the present, and "unweakened" and "unmuddy" forms designating the perfect.

Shafer (1951), who was not satisfied with this interpretation, sought an answer in modern Tibetan dialects. Basing his arguments on the imperative suffix in Murmi, Magari, and Bahing, he reconstructed an imperative suffix *-o for Old Tibetan and explained the phonetic change *-a- > -o- as due to assimilation. For the same sound shift in the present, he assumed either an infix *-o- or a suffix *-o (for instance, he posited *g-o-san > gson, *slab-o > slob). As for the sound shift *-a- > -e- in the present, he explained it either through prefix *ind' (e)- or through suffix *-se/*-es.

Nishida (1957) reconstructed a present suffix *-ɛd, on the strength of the suffix of the same function in Purik (-ĕt) and in Balti (-ed), and explained the *-a- > -e- sound shift in the present as through assimilation. So far as the sound shift *-a- > -o- in the present is concerned, Nishida made extensive use of the infix *-o- in explaining both cases (for example, *g-o-san > gson and *ɤ-o-slab > slob), and kept the suffix *-o exclusive for the imperative, citing the imperative suffix -o in Rong (Lepcha) as additional evidence.

Coblin (1976) revised Nishida's *-ɛd into *-d and *g-o-/*ɤ-o- into *g-, and systematized the whole process of sound change in a set of rules. On the basis of comparative evidence that shows Tibetan -o- corresponding to Chinese labiovelar *gw + V (see *infra* B), I would like to suggest that Coblin's *g- be revised to *gw-; and on the ground that Tibetan -o in the open syllable partially goes back to ST *-u (see *infra* C), I propose to reconstruct the imperative suffix as *-u, instead of *-o.

B. Tibetan -o- and its correspondences in chinese

It will be shown in the following examples that Tibetan -o- corresponds to Old Chinese *-wə-, *-wa-, and *-ua- (*w here being a sign for labio-velar). The correspondences clearly show that Tibetan -o- has three different sources, for if we take Tibetan -o- as the original, we cannot explain why it has three reflexes in Old Chinese. From examples Nos. 165–176, I infer that labio-velar initials caused the contiguous vowels *-ə- and *-a- to change into Tibetan -o-.

In Li's reconstruction, the vocalic cluster *-ua- occurs only in the *chi* 祭, *ko* 歌, and *yüan* 元 categories. However, Li conjectured that it might have had a wider distribution in Proto-Chinese.

1. OC *wə : WT o

165.	OC	gwjəd	C	胃	a stomach (523, a)
	WT	grod			belly, stomach
166.	OC	gwjəg	B	友	friend, associate (995, e)
	WT	grogs			friend, companion
167.	OC	gwjəd	A	遹	go against, oppose, deviate from, err (571, d)
	WT	'gol			to part, to deviate, err

168.	OC	gwjət 掘	dig out (496, s)
	WT	rkod, rko	to dig, dig-out
169.	OC	kwjəd A 歸	return (570, a)
		gwəd A 囘	revolve (542, a)
		gwjəd A, C 圍	encircle (571, g)
	WT	'khor	circle, circumference
		'khor-ba	to turn round, to go round in a circle
		Skor	circle, repetition
		skor-ba	to surround, encircle, to return
		sgor-mo	round, a circle, a globe
		skyor-ba	to repeat, enclosure, fence
170.	OC	gwjəm A 熊	a bear (674, a)
	WT	dom	the brown bear
	WB	wam A	a bear

2. OC *wa : WT o

171.	OC	kwjak 攫	seize (778, b)
	WT	'gog	to take away forcibly, to snatch
172.	OC	gwjag A 于	go to (97, a)
		gwjang B 往	go to (739, k)
	WT	'gro	to walk, to go
	WB	krwa B	to proceed, whether going or coming
173.	OC	gwjag C 芋	taro (97, o)
	WT	gro-ma	potato
174.	OC	gwjag B, C 羽	a feather (98, a)
	WT	sgro	a large feather
175.	OC	ngwjar C 僞	false, cheat (27, k)
	WT	rngod	to deceive
176.	OC	gwag B 戶	door (53, a). The character is the phonetic in 㕣, which is a *ho-k'ou* word.
	WT	sgo	door

3. OC *ua : WT o

177.	OC	dzuar B, C 坐	sit, seat (12, a)
	WT	sdod	to sit, to stay
178.	OC	djuar A 垂	hang down, fall (31, a)
	WT	'jol < *'dyol	to hang down
179.	OC	dzjuat 絕	cut off, break off (296, a)
		tsjuat 剸	cut off (transitive)

	WT	chod < *tshjod	the cutting off, to be cut off
		gcod-pa	to cut, to cut asunder
180.	OC	thuat, duat 脫	take off, escape, careless (324, m)
	WT	lhod, lod, glod	loose, relaxed
	WB	lwat	to be at liberty, free
		hlwat	to free, release, to emancipate
		kjwat < klwat	to be loosed from its proper place
		khjwat < *khlwat	to release, free, emancipate
181.	OC	ruat 悅	pleased, glad (324, o)
	WT	brod	joy, joyfulness
182.	OC	thuar 唾	spit (31, m)
	WT	to-le	to spit

C. Tibetan -e- and -o- compared in Tibetan and Burmese

1. Tibetane -e and -o in open syllables

Miller (1956) reconstructed six vowels (*a, *i, *u, *e, *o, *bl) for Tibeto-Burman, basing on the following correspondences in open syllables:

WT	i	:	WB	e	TB < *i
	e			i	< *e
	a			a	< *a
	u			o	< *u
	o			u	< *o
	u			bl	< *bl

However, for the correspondence WT -u: WB -o there is only one example:

WT	'bu	to open, to unfold, of flowers
WB	pho	to be swelled

Pulleyblank (1963:219) arranged the correspondences in the following schema:

WT	a	:	WB	a
	e			i
	i			e
	u			ui (=Miller's bl)
	o			u

According to the analysis in this study, the table can be rearranged as:

WT		WB	TB	
a	:	a	a	< *a
e		i		< *i
i		iy		< *iy
u		uï		< *uï
o		u		< *u

As already mentioned, TB *-iy goes back to ST *-id, whereas TB *-uï goes back to ST *-ug or *-əgw. Accordingly, TB *-y and *-ï can be regarded as traces of ST *-d and *-g(w), respectively.

In addition to these correspondences, there are examples of Tibetan -o in open syllables corresponding to Burmese -wa.

WT		WB	
mtho	'a span'	thwa A	'to measure with a span'
so	'tooth'	swa C	'a tooth'

2. Tibetan -o- in closed syllables

Tibetan -o- in closed syllables often coorresponds to Burmese -wa-. Following are a few examples:

WT		WB	
nor	'cattle'	nwa C	'a bull, ox, or cow'
dong	'pit'	twang C	'pit'
sbom	'thick, stout'	phwam B	'fat, plump'
dpon	'master, lord'	wan A	'government officer'
rkon	'net'	kwan A	'a casting net'
spobs	'to dare'	wam B	'to dare'

The circumstances here are the same as in the Tibetan and Chinese comparison, both pointing to a secondary origin for the Tibetan vowel -o-.

V. Conclusion

According to the present study, the shift of vowels in Old Chinese, Written Tibetan, and Written Burmese can be summarized as follows:

ST *-a- > -a-	in all three languages
ST *-i- > -i-	in all three languages, except before labial finals in OC, where it yielded *-ə-

ST *-u- > -u-	in all three languages, except before dental and labial finals in OC, where it yielded *-ə-
ST *-ə- > OC -ə-,	WT and WB -u- before labio-velar finals and -a- elsewhere

Tibetan has -e- and -o-, which are not found in Chinese or Burmese; they are treated here as Tibetan innovations. In addition to the four vowels *-a-, *-i-, *-u-, and *-ə-, there were in ST two vocalic clusters, *-ua- and *-ia-; the former yielded WT -o- and WB -wa-. The development of the latter is not clear, however, in the examples cited above (Nos. 18 and 35), ST*-ria- yielded WT -rgya- and WB -rya-. In OC there was *-iə-, but the two comparisons (Nos. 116 and 117) cited in this study show it is a Chinese innovation. The vowel system of ST is then:

Vowels:			Vocalic clusters:	
i		u	ia	ua
	ɔ			
	a			

Notes

* This paper was presented at the 11th International Conference on Sino-Tibetan Languages and Linguistics at University of Arizona, October 20–22, 1978. I wish to thank Professors Fang Kuei Li, Fa-kao Chou, Paul K. Benedict, Pang-hsin Ting, Paul Jen-kuei Li and Tsu-lin Mei for their comments on the manuscript and Professor Lynn Struve for her editorial assistance. However, I have not followed all their advice and I am alone responsible for the errors in the article.

1 This character contains the phonetic 卽 tsjək (923, a)

2 See Yoshio Nishi (1977) p. 42.

3 Wolfenden (1928:279). Thomas (1927:74; 1951; II 24, 1955:III29).

References

Benedict, Paul K. 1972. *Sino-Tibetan: A Conspectus.* Cambridge University Press.

Blagden, C. Otto. 1914. "The Transliteration of Old Burmese Inscription." *JBRS* 4: 136–139.

Bodman, Nicholas C. 1976. "Syllable Types and Yod in Sino-Tibetan. Preliminary draft for the 9th International Conference on Sino-Tibetan Languages and Linguistics." MS.

———. 1977. "Proto-Chinese and Sino-Tibetan. Data Towards Establishing the Nature of the Relationship." MS.

Chang, Kun. 1973. "Review Sino-Tibetan: A Conspectus by Paul K. Benedict." *JAS* 32: 335–337.

Chang, Betty Shefts and Kun Chang. 1976. "Chinese *s- Nasal Initials." *BIHP* 47: 587–609.

Coblin, W. South. 1974. "An Early Tibetan Word for 'Horse'." *JAOS* 94.1: 124–125.

———. 1976. "Notes on Tibetan Verbal Morphology." *TP* 62: 45–70.

Conrady, August. 1896. *Eine indochinesische Causativ-Denominativ-Bildung und ihr Zusammenhang mit den Tonaccenten.* Leipzig: Otto Harrassowitz.

Gong, Hwang-cherng. 1976. Die Rekonstruktion des Altchinesischen unter Berücksichtigung von Wortverwandtschaften. Ph. D. dissertation. Universität München.
Jäschke, H. A. 1965. A Tibetan-English Dictionary. New York: Frederick Ungar Publishing Co.
Jespersen, Otto. 1959. *Language. Its Nature, Development and Origin.* London: George Allen & Unwin LTD.
Judson, Adoniram. 1921. Burmese-English Dictionary. Rangoon: American Baptist Mission Press.
———. 1922. English and Burmese Dictionary, 8th ed. Rangoon: American Baptist Mission Press.
Karlgren, Bernhard. 1923. *Analytic Dictionary of Chinese and Sino-Japanese.* Paris: Librairie Orientaliste Paul Geuthner.
———. 1931. "Tibetan and Chinese." *TP* 28: 25–70.
———. 1957. "*Grammata Serica Recensa.*" *BMFEA* 29, pp. 1–332.
Laufer, Berthold. 1916. "The Si-hia Language. A Study in Indo-Chinese Phonology." *TP* 17: 1–126.
Li, Fang-kuei. 1933. "Certain Phonetic Influences of the Tibetan Prefixes upon the Root Initials." *BIHP* 4: 135–157.
———. 1956. "The Inscription of the Sino-Tibetan Treaty of 821–822." *TP* 44: 1–94.
———. 1959. "Tibetan Glo-ba-'dring." Studia Serica B. Karlgren dedicata, edited by S. Egerod and E. Glahn Copenhagen. E. Munksgaard 55–59.
———. 1971. "Shang-ku-yin yen-chiu." *THJCS* 9, 1–2: 1–61.
———. 1976. "Chi-ke shang-ku sheng-mu wen-t'i." *Tsung-t'ung Chiang-kung shih-shih chou-nien chi-nien lun-wen chi,* pp. 1143–1150. Taipei: Academia Sinica.
Miller, R. A. 1956. "The Tibeto-Burman Ablaut System." *Transactions of the International Conference of Orientalists in Japan* 1: 29–56.
———. 1957. "The Phonology of the Old Burmese Vowel System as seen in the Myazedi Inscription." *Transactions of the International Conference of Orientalists in Japan* 2: 39–43.
Nishi, Yoshio. 1977. "Medials in Burmese." *Kagoshima Daigaku Shigakuka Hōkoku* No. 26: 41–52.
Nishida, Tatsuo. 1955–1956. "Myazedi hibun ni okeru chūko biruma-go no kenkyū." *Kodaigaku (Palaeologia)* 4: 17–32 and 5: 22–40.
———. 1957. "Chibettogo dōshi kōzō no kenkyū." *Gengo kenkyū* 33: 21–50.
Pulleyblank, E. G. 1963. "An Interpretation of the Vowel Systems of Old Chinese and of Written Burmese." *AM* NS 10, part 2: 200–221.
Schiefner, Anton. 1851. "Tibetische Studien." *Bull. Acad. Sci. St. Pet. 8.*
Shafer, Robert. 1940–1941. "The Vocalism of Sino-Tibetan." *JAOS* 60: 302–337 and 61: 18–31.
———. 1951. "Studies in the Morphology of Bodic Verbs." *BSOAS* 13: 702–724; 1017–1031.
———. 1974. *Introduction to Sino-Tibetan.* Wiesbaden: Otto Harrassowitz.
Shen Chien-shih. 1944. *Kuang-yün Sheng-hsi* (abbreviated as KYSH in this study) Peking.
Simon, Walter. 1929. "Tibetisch-Chinesische Wortgleichungen: Ein Versuch," *MSOS* 32: 157–228.
Thomas, F. W. 1935–1955. *Tibetan Literary Texts and Documents. Concerning Chinese Turkestan,* Part I, 1935; Part II, 1951; Part III, 1955. London: Luzac & Company, LTD.
Ting, Pan-hsin. 1975. *Chinese Phonology of the Wei-Chin Period: Reconstruction of the Finals as Reflected in Poetry.* Taipei: IHP Special Publications, No. 65.

Tōdō, Akiyasu. 1957. *Chūgokugo on'inron.* Tokyo: Kōnan Sho'in.
Tung, T'ung-ho. 1944. *Shang-ku yin-yün piao-kao.* Reprinted as IHP monograph series A, No. 21, 1967.
Wang, Li. 1957–1958. *Han-yü-shih-kao.* 3 vols. Peking.
Yakhontov, S. E. 1963. "Sochetaniya Soglasnykh v Drevne-kitaĭskom Yazyke." *Trudy dvadtsat'* pyatogo mezhdu-narodnogo kongressa vostokovedov. Moskva.