

# The Tai-Kadai Languages

Edited by  
Anthony V. N. Diller  
Jerold A. Edmondson  
Yongxian Luo

ROUTLEDGE LANGUAGE FAMILY SERIES

**THE  
TAI-KADAI  
LANGUAGES**

## ROUTLEDGE LANGUAGE FAMILY SERIES

Each volume in this series contains an in-depth account of the members of some of the world's most important language families. Written by experts in each language, these accessible accounts provide detailed linguistic analysis and description. The contents are carefully structured to cover the natural system of classification: phonology, morphology, syntax, lexis, semantics, dialectology, and sociolinguistics.

Every volume contains extensive bibliographies for each language, a detailed index and tables, and maps and examples from the languages to demonstrate the linguistic features being described. The consistent format allows comparative study, not only between the languages in each volume, but also across all the volumes in the series.

*The Austronesian Languages of Asia  
and Madagascar*

Edited by Nikolaus Himmelmann &  
Sander Adelaar

*The Bantu Languages*

Edited by Derek Nurse &  
Gérard Philippson

*The Languages of the Caucasus*

Edited by Alice Harris

*The Celtic Languages*

(new edition forthcoming)  
Martin J. Ball & James Fife

*The Dravidian Languages*

Edited by Sanford B. Steever

*The Germanic Languages*

Edited by Ekkehard Konig &  
Johan van der Auwera

*The Indo-Aryan Languages*

Edited by George Cardona &  
Dhanesh K. Jain

*The Indo-European Languages*

Edited by Paolo Ramat &  
Anna Giacalone Ramat

*The Iranian Languages*

Edited by Gernot Windfuhr

*The Khoesan Languages*

Edited by Rainer Vossen

*The Mongolic Languages*

Edited by Juha Janhunan

*The Oceanic Languages*

Edited by John Lynch, Malcolm Ross  
& Terry Crowley

*The Romance Languages*

Edited by Martin Harris &  
Nigel Vincent

*The Semitic Languages*

Edited by Robert Hetzron

*The Sino-Tibetan Languages*

Edited by Graham Thurgood &  
Randy J. Lapolla

*The Slavonic Languages*

Bernard Comrie & Greville G. Corbett

*The Tai-Kadai Languages*

Edited by Anthony Diller,  
Jerold A. Edmondson and Yongxian Luo

*The Turkic Languages*

Edited by Eva Csato & Lars Johanson

*The Uralic Languages*

Edited by Daniel Abondolo

# THE TAI-KADAI LANGUAGES

Edited by

Anthony V. N. Diller,

Jerold A. Edmondson

and

Yongxian Luo

 **Routledge**  
Taylor & Francis Group  
LONDON AND NEW YORK

First published 2008  
by Routledge

Published 2014 by Routledge  
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

Simultaneously published in the USA and Canada  
by Routledge  
711 Third Avenue, New York, NY, 10017, USA

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

© 2008 Anthony V. N. Diller, Jerold A. Edmondson and Yongxian Luo

All rights reserved. No part of this book may be reprinted  
or reproduced or utilized in any form or by any electronic,  
mechanical, or other means, now known or hereafter  
invented, including photocopying and recording, or in any  
information storage or retrieval system, without permission  
in writing from the publishers.

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

*Library of Congress Cataloguing in Publication Data*  
The Tai-Kadai languages / edited by Anthony V. N. Diller,  
Jerold A. Edmondson and Yongxian Luo.  
p.cm. – (Routledge language family series)

Includes bibliographical references.

1. Tai-Kadai languages – History. 2. Tai-Kadai languages – Grammar,  
Comparative. I. Diller, Anthony Van Nostrand; II. Edmondson, Jerold A.;  
III. Luo Yongxian; IV. Series.

PL4113. T35 2008  
495.9'1–dc22  
2007046237

ISBN 13: 978-0-700-71457-5 (hbk)

---

# CONTENTS

---

<i>Illustrations</i> .....	x
<i>Abbreviations</i> .....	xiii
<b>PART 1</b> .....	<b>1</b>
<b>1. INTRODUCTION</b> .....	<b>3</b>
<i>Anthony Diller</i>	
1.1. TAI-KADAI: PROGRESS AND PROSPECTS .....	3
1.2. ACKNOWLEDGMENTS .....	3
1.3. ORGANIZATION .....	4
1.4. TERMINOLOGY AND ITS INTERPRETATIVE NUANCES .....	5
<b>2. SINO-TAI AND TAI-KADAI: ANOTHER LOOK</b> .....	<b>9</b>
<i>Yongxian Luo</i>	
2.1. INTRODUCTION .....	9
2.2. THE SINO-TAI HYPOTHESIS: A REVIEW OF DIFFERENT THEORIES .....	10
2.3. THE ROLE OF TAI IN OLD CHINESE RECONSTRUCTION .....	16
2.4. PROGRESS AND NEW FINDINGS .....	20
2.5. SUMMARY .....	23
<b>PART 2</b> .....	<b>29</b>
<b>3. RESOURCES FOR THAI LANGUAGE RESEARCH</b> .....	<b>31</b>
<i>Anthony Diller</i>	
3.1. INTRODUCTION .....	31
3.2. PHONETICS, PHONOLOGY AND ORTHOGRAPHY .....	32
3.3. SYNTAX AND SEMANTICS .....	35
3.4. SOCIOLINGUISTIC, COMPARATIVE AND HISTORICAL STUDIES .....	46
<b>4. VERBS AND MULTI-VERB CONSTRUCTIONS IN LAO</b> .....	<b>83</b>
<i>N. J. Enfield</i>	
4.1. INTRODUCTION .....	83
4.2. VERBS, VERB CLASSES, ASPECT-MODALITY MARKING .....	85
4.3. ARGUMENT STRUCTURE IN SINGLE-VERB CLAUSES .....	86
4.4. MULTI-VERB CONSTRUCTIONS .....	104
4.5. CONCLUSION .....	176

**5. SHAN AND OTHER NORTHERN TIER SOUTHEAST TAI  
LANGUAGES OF MYANMAR AND CHINA: THEMES AND  
VARIATIONS .....184**

*Jerold A. Edmondson*

5.1. BACKGROUND.....	184
5.2. LOCATIONS STUDIED.....	188
5.3. A GLIMPSE AT SHAN SETTLEMENT HISTORY.....	190
5.4. COMPARATIVE RESULTS.....	194
5.5. CONCLUSION.....	202

**6. THE TAI LANGUAGES OF ASSAM.....207**

*Stephen Morey*

6.1. THE TAI PEOPLE IN ASSAM.....	207
6.2. NOTES ON THE PHONOLOGY OF THE TAI LANGUAGES OF ASSAM.....	212
6.3. NOTES ON THE MORPHOSYNTAX OF THE TAI LANGUAGES OF ASSAM.....	219
6.4. THE DEVELOPMENT OF BOUND TAM MARKERS IN TAI AITON.....	222
6.5. CHANGED TONE IN PHAKE.....	234
6.6. BI-DIALECTALISM IN THE TAI KHAMYANG SPEECH COMMUNITY.....	240

**7. THE LUE LANGUAGE.....254**

*John F. Hartmann*

7.1. LUE AND NEIGHBOURING DIALECTS.....	254
7.2. PHONOLOGY.....	263
7.3. VARIATION.....	288

**8. THE TAI DIALECTS OF NGHỆ AN, VIETNAM (TAY DAENG, TAY YO,  
TAY MUONG).....298**

*Michel Ferlus*

8.1. INTRODUCTION.....	298
8.2. LES THAI ET LES DIALECTES THAI DU NGHỆ AN.....	299
8.3. GÉNÉRALITÉS PHONÉTIQUES.....	300
8.4. TAY DAENG (OU TAY THANH).....	301
8.5. TAY YO.....	304
8.6. TAY MUONG.....	309
8.7. LE PROTO THAI ET LES DIALECTES.....	314

**9. ZHUANG.....317**

*Yongxian Luo*

9.1. INTRODUCTION.....	317
9.2. PHONOLOGY.....	321
9.3. WORD CLASSES AND BASIC SYNTACTIC PATTERNS.....	324
9.4. GRAMMATICAL PROCESSES.....	342
9.5. SIMPLE MAIN CLAUSES.....	347
9.6. COMPLEX CONSTRUCTIONS.....	366
9.7. ACKNOWLEDGEMENTS.....	376

<b>10. BOUYEI PHONOLOGY .....</b>	<b>378</b>
<i>Wil C. Snyder</i>	
10.1. INTRODUCTION .....	378
10.2. PHONOLOGY OF WANGMO BOUYEI .....	378
10.3. COMPARISON WITH OTHER DIALECTS .....	381
<b>11. SAEK REVISITED.....</b>	<b>389</b>
<i>Wilaiwan Khanittanan</i>	
<b>PART 3.....</b>	<b>393</b>
<b>12. FOUR-WORD ELABORATE EXPRESSIONS IN YUNNAN TAI LUE: A PAN-TAI CULTURAL TRAIT? .....</b>	<b>395</b>
<i>Amara Prasithrathsint</i>	
12.1. INTRODUCTION .....	395
12.2. EXAMPLES OF FOUR-WORD EXPRESSIONS IN TAI LUE .....	397
12.3. THE STRUCTURE OF THE FOUR-WORD EXPRESSIONS IN TAI LUE .....	400
12.4. THE GRAMMATICAL STATUS OF FOUR-WORD EXPRESSIONS .....	401
12.5. CONCLUSION .....	402
<b>13. TAI AESTHETICS .....</b>	<b>404</b>
<i>Thomas John Hudak</i>	
13.1. LINGUISTIC TECHNIQUES FOR THE CREATION OF <i>phay<sup>l</sup>rɔʔ<sup>h</sup></i> .....	404
13.2. <i>phay<sup>l</sup>rɔʔ<sup>h</sup></i> AND POETIC CONSTRAINTS .....	407
13.3. CONCLUSION .....	412
<b>14. THE OLD ZHUANG SCRIPT .....</b>	<b>415</b>
<i>David Holm</i>	
14.1. METHODOLOGY .....	418
14.2. A SAMPLE TEXT .....	418
14.3. RELATIVE FREQUENCY OF CHARACTER TYPES AND READINGS .....	422
14.4. PHONETIC READINGS .....	423
14.5. REGIONAL VARIATION .....	425
14.6. ZHUANG CHARACTERS .....	426
14.7. CONCLUSION .....	426
<b>PART 4.....</b>	<b>429</b>
<b>15. AN ETYMOLOGICAL SPECULATION ON THE SEQUENTIAL INDICATOR <i>kɔʔ</i> IN THAI NARRATIVE .....</b>	<b>431</b>
<i>Somsonge Burusphat</i>	
15.1. INTRODUCTION .....	431
15.2. SEQUENTIALITY AS A SIGNIFICANT FEATURE OF NARRATIVE DISCOURSE .....	431
15.3. COMPARISON OF THE SEQUENTIAL INDICATORS IN SOUTHEAST ASIAN	



LANGUAGES .....	434
15.4. THE SEQUENTIAL INDICATOR <i>kɔɔ</i> <sup>3</sup> AS AN AREAL FEATURE .....	437
15.5. HISTORICAL IMPLICATIONS .....	438
15.6. CONCLUSION .....	442
<b>16. A HISTORICAL STUDY OF /thii/ IN THAI .....</b>	<b>445</b>
<i>Pranee Kullavanijaya</i>	
16.1. THE AIM OF THE STUDY .....	445
16.2. SOURCES OF DATA .....	445
16.3. /thii/ IN MODERN THAI .....	445
16.4. A HISTORICAL DEVELOPMENT OF /thii/ .....	451
16.5. A COMPARISON OF A GRAMMATICALIZATION PATHWAY OF /thii/ AND A HISTORICAL DEVELOPMENT OF /thii/ .....	461
16.6. THE EQUIVALENTS OF /thii/ IN OTHER TAI LANGUAGES .....	465
16.7. CONCLUSION .....	466
<b>17. BIPOLAR DISTRIBUTION OF A WORD AND GRAMMATICALIZA- TION IN THAI: A DISCOURSE PERSPECTIVE .....</b>	<b>468</b>
<i>Shoichi Iwasaki</i>	
17.1. INTRODUCTION .....	468
17.2. 'BIPOLAR DISTRIBUTION OF A WORD' AND THE THEORETICAL FRAMEWORK .....	468
17.3. DATA .....	470
17.4. THE CASE OF <i>hây</i> .....	471
17.5. THE CASE OF <i>léew</i> .....	478
17.6. CONCLUSION .....	482
<b>18. DIRECTIONAL VERBS AS SUCCESS MARKERS IN THAI: ANOTHER GRAMMATICALIZATION PATH .....</b>	<b>484</b>
<i>Kingkarn Thepkanjana and Satoshi Uehara</i>	
18.1. INTRODUCTION .....	484
18.2. MOTIVATION FOR USING THE SIX DIRECTIONAL VERBS AS SUCCESS MARKERS .....	486
18.3. GRAMMATICALIZATION OF DIRECTIONAL VERBS INTO SUCCESS MARKERS .....	492
18.4. CONCLUSION .....	505
<b>PART 5 .....</b>	<b>507</b>
<b>19. KAM .....</b>	<b>509</b>
<i>Tongyin Yang and Jerold A. Edmondson</i>	
19.1. INTRODUCTION .....	509
19.2. PHONOLOGY .....	511
19.3. SYNTAX .....	516
19.4. DISCOURSE PATTERNS AS REVEALED IN NARRATIVE .....	557
19.5. SPEECH COMMUNITIES AND SOCIOLINGUISTIC ISSUES .....	578

<b>20. SUI.....</b>	<b>585</b>
<i>James Wei and Jerold A. Edmondson</i>	
20.1. INTRODUCTION .....	585
20.2. RHYMES.....	588
20.3. TONES.....	590
20.4. LANGUAGE VARIATION .....	590
20.5. WORD CATEGORIES .....	592
20.6. SYNTAX.....	593
20.7. VARIATION IN WORD FORMS.....	593
20.8. RECORDING SYSTEM.....	594
20.9. CONCLUSION .....	594
<b>21. CHADONG, A NEWLY-DISCOVERED KAM-SUI LANGUAGE IN NORTHERN GUANGXI.....</b>	<b>596</b>
<i>Li Jinfang</i>	
21.1. THE SOUNDS OF CHADONG .....	598
21.2. LEXICON .....	600
21.3. GRAMMAR .....	603
21.4. WORD ORDER .....	611
<b>PART 6.....</b>	<b>621</b>
<b>22. THE HLAI LANGUAGE.....</b>	<b>623</b>
<i>Weera Ostapirat</i>	
22.1. INTRODUCTION .....	623
22.2. HLAI PHONOLOGY AND LEXICONS.....	624
22.3. HLAI VARIETIES .....	631
22.4. OTHER LINGUISTICS FEATURES .....	645
<b>23. KRA OR KADAI LANGUAGES .....</b>	<b>653</b>
<i>Jerold A. Edmondson</i>	
23.1. INTRODUCTION .....	653
23.2. KRA OR KADAI AND THE KAM-SUI BRANCH.....	654
23.3. GELAO.....	656
23.4. LACHI .....	658
23.5. BUYANG AND EN .....	660
23.6. QABIAO .....	662
23.7. LAHA .....	663
23.8. COMPARISON OF THE SYNTACTIC FEATURES.....	665
<i>Appendix IX: Language Maps .....</i>	<i>673</i>
<i>Index .....</i>	<i>678</i>

# ILLUSTRATIONS

## FIGURES

<b>Figure 4.3.4.1-1:</b> Constituents of the Lao clause, in order .....	97
<b>Figure 4.4.2-1:</b> Elements of the Lao verb phrase, in order .....	115
<b>Figure 4.4.5.2-1:</b> Slots in the ‘manner-path-direction construction’ .....	130
<b>Figure 5.1-1:</b> Gedney diagram of Tai tones .....	187
<b>Figure 5.4.1-1:</b> Northern Shan tone values as a function of Proto-Tai tone categories and initial consonants .....	194
<b>Figure 5.4.1-2:</b> Khamti tone values as a function of Proto-Tai tone categories and initial consonants .....	194
<b>Figure 5.4.1-3:</b> Southern Shan tone values as a function of Proto-Tai tone categories and initial consonants .....	195
<b>Figure 5.4.1-4:</b> Phake tone values as a function of Proto-Tai tone categories and initial consonants .....	196
<b>Figure 5.4.2-1:</b> Taunggyi Shan <i>Sĕa<sup>al</sup></i> ‘basket’ vs <i>sa<sup>al</sup></i> ‘angry’ .....	197
<b>Figure 5.4.3-1:</b> Deplosivization of /ʔd/ → /l/, /ʔb/ → /m/, /ʔbl/ → /l/, and variably /n/ → /l/ .....	199
<b>Figure 5.4.4-1:</b> Change of /kh/ to /x/ in Shan by region .....	200
<b>Figure 5.4.6-1:</b> Fortition of /f/ to /ph/ .....	201
<b>Figure 5.4.7-1:</b> Loss of secondary articulation .....	201
<b>Figure 5.4.8-1:</b> Vowel diphthongization changes .....	202
<b>Figure 5.4.9-1:</b> Pattern of palatalizations .....	202
<b>Figure 8.1-2:</b> La branche des langues Thai-Yaay (Tai) de la famille Tai-Kadai .....	299
<b>Figure 8.4.3-1:</b> Tons du Tay Thanh (Con Cuông, Nghệ An) .....	302
<b>Figure 8.4.3-2:</b> Tons du Tay Daeng (Mai Châu Thượng, Hòa Bình) .....	302
<b>Figure 8.4.3-3:</b> Tons du Tay Daeng (Mai Châu Hạ, Hòa Bình) .....	302
<b>Figure 8.5.3-1:</b> Tons du Tay Yo (Quỳ Châu, Nghệ An) .....	305
<b>Figure 8.6.3-1:</b> Tons du Tay Muong (Tương Dương, Nghệ An) .....	310
<b>Figure 17.4-1:</b> Functions of <i>hây</i> .....	471
<b>Figure 17.5-1:</b> Overview of Grammaticalization Path for <i>léew</i> .....	479
<b>Figure 18.3.2-1:</b> Grammaticalization path of directional verbs in Thai .....	502
<b>Figure 19.4.3.3.1-1:</b> Grouping of Kam SFPs .....	569
<b>Figure 19.4.3.3.2-1:</b> Illustration of Syntactic Positions of SFPs .....	570
<b>Figure 19.4.3.3.2.2-1:</b> SFPs in COMP position .....	572
<b>Figure 23.2-1:</b> Lala Noong Lay waveform and pitch track <i>tu<sup>33</sup>ma<sup>41?</sup></i> ‘CLF-horse’ .....	655

## MAPS

<b>Map 7.1-1:</b> Showing areas of greatest Lue concentration.....	261
<b>Map 7.1-2:</b> Showing the localities in which the Tai languages of British India are spoken.....	262
<b>Map 7.1-3:</b> Showing three major Lue speech areas previously studied by Gedney (1964) and Li (1964).....	263
<b>MAP 7.2-1:</b> Showing the isoglosses of change from Proto-Tai *d to modern t and th respectively.....	280
<b>Map 7.2-2:</b> Showing vowel length distinctions between north and south.....	286
<b>Map 7.3-1:</b> Three major dialect areas within Sipsongpanna.....	289
<b>Map 7.3-2:</b> Points in Sipsongpanna where short-long vowels are distinctive among the young, middle-aged and the old.....	291
<b>Map 9.1.2-1:</b> Zhuang Speaking Area.....	318
<b>Map 15.4-1:</b> Geographical distribution of the sequential indicator in the Southeast Asian area.....	438
<b>Map Ap-1:</b> Distribution of Tai.....	675
<b>Map Ap-2:</b> Distribution of Kam-Tai.....	676
<b>Map Ap-3:</b> Distribution of Tai-Kadai.....	677
<b>Map Ap-4:</b> Distribution of Kam-Sui.....	678
<b>Map Ap-5:</b> Distribution of Gelao, Buyang and other Kadai.....	679

## TABLES

<b>Table 2.2.2.1-1:</b> Chinese–Tai Tonal Correspondences.....	11
<b>Table 4.1-1:</b> Some V1-V2 sequences with different underlying structures.....	84
<b>Table 4.2-1:</b> Five Lao verb categories based on logical/aspectual distinctions.....	85
<b>Table 4.3.2-1:</b> Grammatical distinctions between three one-place constructions.....	89
<b>Table 4.3.2-2:</b> Grammatical distinctions between five two-place constructions.....	91
<b>Table 4.4.1.3-1:</b> Three types of V1-V2 combination, by yes-answer behaviour.....	109
<b>Table 4.4.6.2.5-1:</b> Three types of resultative construction, according to semantic structure of V1.....	140
<b>Table 4.5.2-1:</b> Grammatical properties distinguishing a range of Lao V1-V2 structures.....	178
<b>Table 5.2-1:</b> Speakers and locations studied.....	188
<b>Table 5.4.2-1:</b> A comparison of sound change in Shan locations.....	197
<b>Table 6.2.1-1:</b> Consonant phonemes in phake, khamti and khamyang.....	212
<b>Table 6.2.1-2:</b> Realigned Aiton phoneme chart.....	213
<b>Table 6.2.1-3:</b> Possible consonant phonemes in Ahom.....	213
<b>Table 6.2.2-1:</b> Vowel phonemes in Phake.....	214
<b>Table 6.2.2.1-1:</b> Vowel phonemes in Aiton.....	215
<b>Table 6.2.2.1-2:</b> Examples of proto */i/ and */e/ recorded by Banchob (MS) and (1987) and compared with the proto forms after Li (1977).....	215
<b>Table 6.2.2.2-1:</b> Possible vowel phonemes in Ahom.....	216
<b>Table 6.2.3.1-1:</b> Tones in Phake.....	216
<b>Table 6.2.3.2-1:</b> Tones in Khamyang.....	217
<b>Table 6.2.3.3-1:</b> Tones produced by bidya thoumoung for Aiton.....	218
<b>Table 6.2.3.3-2:</b> Aiton tone box.....	218
<b>Table 6.3.2-1:</b> General rules of syntax for Khamti (Needham 1894: 81).....	220
<b>Table 6.4.3-1:</b> Tam markers in Aiton texts.....	225
<b>Table 6.4.4.3-1:</b> sentence Analysis in Nang Wimala'S historical story.....	228
<b>Table 6.4.4.4-1:</b> Analysis of sentences in Ong Cham'S <i>story of the Crow and the Fox</i> .....	230

<b>Table 6.5.1-1:</b> Comparison of questioning tone and high falling tone in Phake.....	235
<b>Table 6.5.1-2:</b> The questioning tone in Phake.....	236
<b>Table 6.5.3-1:</b> Imperative markers in Phake and Aiton.....	240
<b>Table 6.6-1:</b> Comparison of Khamyang initial/b/, /l/ and /m/.....	242
<b>Table 6.6-2:</b> Vowel reduction/centralization in Khamyang.....	243
<b>Table 6.6-3:</b> Comparison of Khamyang tones.....	243
<b>Table 9.2-1:</b> Northern Zhuang /r/ and their Southern Zhuang counterparts.....	321
<b>Table 9.2-2:</b> Some variations in rhymes between Northern and Southern Zhuang.....	322
<b>Table 9.2-3:</b> Comparison of some Wenshan Zhuang and Wuming Zhuang finals.....	322
<b>Table 9.2-4:</b> Tone contrast in six Zhuang dialects.....	322
<b>Table 9.2.1.1-1:</b> Fengshan Zhuang consonants.....	323
<b>Table 9.2.1.2-1:</b> Fengshan Zhuang vowels.....	323
<b>Table 9.2.1.4-1:</b> Tones in Fengshan.....	324
<b>Table 9.3.5.1-1:</b> Personal pronouns in Fengshan Zhuang.....	327
<b>Table 9.3.5.2-1:</b> Interrogative and indefinite pronouns.....	328
<b>Table 9.3.5.3-1:</b> Reflexives.....	329
<b>Table 9.3.6.1.1-1:</b> Cardinals.....	329
<b>Table 9.3.6.1.1-2:</b> ‘Eleven’ to ‘nineteen’.....	330
<b>Table 9.3.6.1.1-3:</b> ‘Twenty’ to ‘ninety’.....	330
<b>Table 9.3.6.1.4-1:</b> Names of the month.....	331
<b>Table 9.3.7.3-1:</b> Locative words.....	336
<b>Table 11-1:</b> Comparison of Saek phonemes as spoken by one speaker in 1976 and another in 2006.....	390
<b>Table 14-1:</b> Zhuang Readings of 𑜋 <i>er</i> <sup>2</sup> (Source: <i>Gu Zhuangzi zidian</i> ).....	416
<b>Table 14.3-1:</b> Typology of Graphs and Reading in Zhuang.....	423
<b>Table 16.4.4-1:</b> Occurrences of nine functions of /thi/ in the four periods.....	461
<b>Table 19.2.1-1:</b> Kam Initial Phonemes.....	512
<b>Table 19.2.1-2:</b> Examples of Kam syllable initial phonemes.....	512
<b>Table 19.2.3-1:</b> Kam tones.....	514
<b>Table 19.4.2.1-1:</b> The Six-part Anatomy of the Kam Narratives.....	559
<b>Table 19.4.3.2-1:</b> Indexing of the Kam SFPs.....	564
<b>Table 19.4.3.2-2:</b> SFP clusters without the PERF Marker.....	566
<b>Table 22.2.1.1-1:</b> Baoding and Tongshi Tonal Split System.....	625
<b>Table 22.2.1.1-2:</b> Tonal correspondences between Baoding and Tongshi.....	625
<b>Table 22.2.2-1:</b> Hlai lexical traits of Kra-Dai origin.....	630
<b>Table 22.2.2-2:</b> Lexical traits of Hlai languages.....	630
<b>Table 22.3-1:</b> Hlai varieties.....	631
<b>Table 22.3.2.3-1:</b> Hlai Tone C comparisons.....	636
<b>Table 22.3.3-1:</b> Eastern Hlai dialect tone split.....	637
<b>Table 22.3.5.1-1:</b> Cun tones in historical perspectives.....	641
<b>Table 22.4.2-1:</b> Orthography of Hlai tones.....	648

---

# ABBREVIATIONS

---

LTBA	Linguistics of the Tibeto-Burman Area
HCT	Handbook of Comparative Tai
ICSTLL	International Conference on Sino-Tibetan Languages and Linguistics
PAL 3	Pan Asiatic Linguistics: Proceedings of the Third International Conference
PAL 4	Pan Asiatic Linguistics: Proceedings of the Fourth International Conference
A	agent
ACHV	achievement
ADV	adverbial
AP	adjective phrase
AQ	audience question
ASEMI	Asie de Sud-Est et Monde Insulindien (Bulletin du Centre de Documentation et de Recherche, Centre National de la Recherche Scientifique, Ecole des Hautes Etudes en Sciences Sociales, Paris)
ASP	aspect
AUX	auxiliary
BA	disposal construction in Chinese
BEFEO	Bulletin de l'Ecole Française d'Extrême-Orient, Hanoi/Paris
BEI	passive construction in Chinese
BEN	benefactive
BENEF	benefactive
C	consonant
CAU	causative
CAUS	causative
CL	numeral classifier
CLA	classifier
CLF	classifier
CLNK	clause linker
CLS	classifier
CM	commenting on developments in story
COMP	complementizer
CONJ	conjunctive; conjunction
CT	class term
CV	action verb
DEM	demonstrative
Det	determiner
DI	demonstrative and interrogative

DIR	directional
DO	direct object
DV	directional verb
EMPH	emphasis/emphatic
EXPR	expressive
FEM	feminine
FOC	focus
FW	functional words
GEN	general
GI	greeting and informational exchange
HP	highlighting particle
IND	inducive
INJ	interjection
INT	intensifier
INTERJ	interjection
INTR	intransitive
IO	indirect object
IP	initial phonemes
IRR	irrealis
MAL	malfactive
MASC	masculine
MOD	modality
MSC	Modern Standard Chinese
MV	modal verb
NEG	negation; negative
NET	noninterventive
NL	non-linguistic instant responses and reactions
NOM	noun phrase conjunctive
NONPROX	non-proximal
NP	noun phrase
NSR	nominalizer
O	object
PAN	proto-Austronesian
PART	particle
PAT	patient
PCL	particle
PEN	personal experience narrative
PER	person
PERF	perfective
PFV	perfective
PFX	prefix
PL	plural
PMP	Proto-Malayo-Polynesian
POSS	possessive
POT	potential
PP	preposition phrase; pragmatic particle
PROG	progressive
PT	Proto-Tai
PURP	purposive
PV	verb-particle

---

Q	question
QP	quantifier phrase
RCP	reciprocal
RDP	reduplication
RE	reasoning or extensions
REC	reciprocal
REL	relative clause; relative marker; relativizer
RR	rest of the audience's response
S	single direct argument of intransitive clause; subject
SA	speech act
SAP	speech act particle
SFP	sentence-final particle
SG	singular
SLP	sentence linking particle
SPEC	specifier
SUB	supordinative
SUFF	suffix
SUP	suppine
SURF	surface structure
SV	stative verb
SVC	serial-verb construction
T	tone
TAM	tense, aspect and modality
TBU	tone-bearing unit
TOP	topic
TPC	topic
V	verb; vowel
VP	verb phrase
Y.SIB	younger sibling
1/2/3	1 <sup>st</sup> /2 <sup>nd</sup> /3 <sup>rd</sup> person pronoun



*This page intentionally left blank*

## **PART 1**

---

# **OVERVIEW CHAPTERS**

---

*This page intentionally left blank*

## CHAPTER ONE

# INTRODUCTION

*Anthony Diller*

### 1.1. TAI-KADAI: PROGRESS AND PROSPECTS

This volume attests to several eventful decades in Tai-Kadai research. The recent past has seen new data become available, substantial analyses completed and larger Tai-Kadai patterns coming into clearer focus. These successes are pointing the way to future projects. The editors hope that this volume not only can stand as a record of current progress but also can provide a stimulus for new field studies and generalizations. The latter may apply to the internal structure and relations of varieties within Tai-Kadai and to relationships with other languages and groupings; also to broader questions of theory and practice.

Further fieldwork is a matter of particular urgency. As elsewhere among the world's major language families, Tai-Kadai's territory is home to speakers of endangered languages. Concern is great that children of those now speaking many Tai-Kadai varieties are substantially modifying and even abandoning their parents' language. This volume documents varieties under threat, pointing to how important it is for more work of this kind to be undertaken while there is still time. Required will be well-trained linguistic researchers, motivated local speakers, appropriate institutional assistance and the acquiescence, if not active support, of governments and local officials. In the past two or three decades Tai-Kadai researchers have been fortunate in that policies of states in the region have generally moved in the direction of facilitating greater access to field sites; the profession's thanks are due to individuals and official units involved in this movement. Gratitude is also due to the funding bodies which have supported Tai-Kadai research. It is hoped that work reported in this volume may encourage others to address these challenges, gain necessary support and pursue further field studies in this critical area.

### 1.2. ACKNOWLEDGMENTS

Many people and institutions have contributed to the completion of this volume. Gratitude has already been expressed for cooperative support from officials and their units in the Tai-Kadai region; included here in China are the Central University for Nationalities, Beijing, and the Guangxi University for Nationalities. In Thailand, the National Research Council has been helpful. Of course, we owe sincere thanks to the many local language assistants, speakers of Tai-Kadai varieties, whose contributions are the essential basis of analyses presented here.

Themes recurring in this volume have often been points of discussion in a series of stimulating international conferences held in Bangkok. Chulalongkorn, Thammasat and Mahidol Universities have hosted these and thanks are due to Thai colleagues and their linguistics departments for contributing their professional and organizational expertise. The success of these events has helped to bring Tai-Kadai studies to maturity.

Funding agencies have been thanked but double thanks are due. Grants from the Australian Research Council (ARC) have been crucial in bringing this volume together and specifically in enabling work behind [chapters 2, 3, 4 and 9](#). The Research Centre for Linguistic Typology (RCLT) at La Trobe University, Melbourne, provided a most appropriate site for early editorial work. The Centre's inspiring directors also made possible international travel for initial publication arrangements. RCLT provided specific

support for work behind [chapters 1](#) and [6](#). Other tasks have been completed with assistance from the Department of Linguistics of the University of Sydney and final editing has relied heavily on the good offices of the Asia Institute of the University of Melbourne.

Contributors are thanked for their patience with production delays and for their efficiency in responding to requests. The same thanks are extended to staff at Routledge-Curzon.

Production has encountered challenges and many have offered help and advice. Special thanks are due Mr Mike Tianqiao Lu, Ms Haiqing Yu and Mr Phichit Roinil. Further acknowledgements appear in individual chapters.

Finally, in gratitude, the writer would like to call to mind inspiring teachers and other senior colleagues who have shared their many insights, among them: Helen Funnell, James W. Poultney, Mary R. Haas, Robert B. Jones, John U. Wulff, James Gair, David K. Wyatt, Fang-Kuei Li, Paul K. Benedict and William J. Gedney.

### 1.3. ORGANIZATION

In planning selections for this volume, the editors have kept in mind maximum service to the profession and to the likely readership. They have allocated much space to less-described Tai-Kadai varieties, several treated here in detail for the first time. On the other hand, general and comparative themes are included as well.

[Chapter 2](#) following is by Yongxian Luo and can serve as an introduction to the current state of research, debate and speculation as to Tai-Kadai's ultimate linkages. It also takes a particular stand on the Sino-Tai hypothesis and illustrates how argumentation in the field has been progressing. The chapter should be provocative and we can expect future debate.

[Part 2](#) presents overviews and resources relating to the Tai languages. Thai is the best-described language in the family with many overviews readily available. Rather than offering yet another overview, in [chapter 3](#) the writer selects bibliographic resources and introduces them with commentary, including some impression of how research has developed. Lao is the other national language in the Tai-Kadai grouping but is underrepresented in published research, especially in English. In [chapter 4](#) N. J. Enfield provides a substantial account of Lao verbs and verbal constructions. The extent to which the Lao features he analyzes are found throughout Tai-Kadai suggests a worthwhile direction for future research. [Chapter 5](#) by Jerold A. Edmondson is a survey of Shan and other northern tier Southwestern Tai languages of Myanmar and China. In [chapter 6](#) Stephen Morey describes Tai languages of Assam and includes treatment of their little-known writing system. As for other Southwestern Tai languages, Lue is described in [chapter 7](#) by John F. Hartmann and the Tai dialects of Nghe An, Vietnam, by Michel Ferlus, in [chapter 8](#). Zhuang, a grouping that includes many millions of speakers, spans the Central and Northern branches of Tai and is described in [chapter 9](#) by Yongxian Luo. In [chapter 10](#) phonology of the Northern-branch language Bouyei is outlined by Wil Snyder. [Chapter 11](#), by Wilaiwan Khanittanan, updates the author's earlier work on Saek, a displaced Northern-branch language with archaic features.

[Part 3](#) is given to some special Tai features often overlooked in standard treatments but of undoubted linguistic interest. [Chapter 12](#) by Amara Prasithrathsint considers four-word expressions, taking Tai Lue as a basis, but suggesting a Tai-wide distribution for the phenomena described. Thomas John Hudak in [chapter 13](#) draws on several Tai languages to produce a linguist's guide to Tai aesthetics. In [chapter 14](#) David Holm gives readers an illuminating exegesis of what may well be the first Tai writing system, the Old Zhuang script.

[Part 4](#) is a more tightly-focused set, with four chapters dealing with aspects of diachronic change and grammaticalization, topics of particular relevance as linguistic theory confronts languages of the Tai-Kadai type. In [chapter 15](#) Somsonge Burusphat traces a sequential indicator appearing in a range of Southeast Asian languages and speculates as to etymological connections. Pranee Kullavanijaya in [chapter 16](#) gives a diachronic account of how a Thai (and Tai) functional operator has evolved. More synchronic approaches to grammaticalization are taken up in [chapter 17](#) by Shoichi Iwasaki, analyzing bipolar distribution, and in [chapter](#)

18, by Kingarn Thepkanjana and Satoshi Uehara, analyzing directional verbs as success markers.

Part 5 is given to the Kam-Sui languages and provides valuable information on little-described varieties. In one of the more substantial chapters (19) of the volume, Tongyin Yang and Jerold A. Edmondson present a groundbreaking and comprehensive account of Kam (Dong), the numerically dominant member of the Kam-Sui group. This is followed in chapter 20 by a sketch of Sui by James Wei and Jerold A. Edmondson. Chapter 21, by Jinfang Li, is the first description of Chadong available for international readers.

Part 6 treats other little-described varieties, several endangered, but critical for understanding the deeper constituency of Tai-Kadai as a whole and for considering its ultimate relationships. The Hlai (or Li) languages of Hainan are described by Weera Ostapirat in chapter 22. The final chapter (23) by Jerold A. Edmondson describes the Kra group (or Geyang; also Kadai in a sense mentioned in the following section).

#### 1.4. TERMINOLOGY AND ITS INTERPRETATIVE NUANCES

Terminology used by those in Tai-Kadai research still shows flux: usage is not yet entirely uniform and complete professional consensus remains elusive. In some cases terminological variation is simply a matter of alternative names, but in other cases particular claims are presupposed or signaled when specific labels are selected. Specific languages may have multiple names, such as Black Tai, alternating with Tai Dam, Tai Noir, etc., and when transplanted in Thailand, as Lao Song Dam or Lao Song. Also, a single label, such as *Dai*, may subsume what linguistic criteria or speakers themselves might take to be separate languages. Finally some labels, e.g. Red Thai, have been subject of contention (see chapter 8). The editors have not enforced standards in this regard, so readers should keep in mind potential flexibility in usage across chapters. This fluidity applies to transcription systems as well.

By now it is usual for the spelling *Thai* to cover inclusively normative Standard Thai and related colloquial Central Thai varieties, including the more vernacular form sometimes referred to as Bangkok Thai. The spelling *Thai* is also used in regional or local variety names within Thailand's borders such as Southern Thai, Suphanburi Thai, etc. Northern Thai is often called Lanna or Kammueang; Northeastern Thai may be referred to as Isan or local Lao.

*Thai* in this generally-accepted sense then belongs to the larger *Tai* family of languages, so spelled, as defined by general professional consensus. Most in the field would accept at least as a point of departure Li's (1977) classification of Tai languages, which are spoken over a wide somewhat 'Y-shaped' area stretching from southern China through northern Southeast Asia and on to the northwest into Assam in northeastern India. The reconstructed parent language at this level is *Proto-Tai*. Readers need to keep in mind that *Northern Tai* and *Central Tai* designate major branches in Li's *Tai* family tree, not specific varieties. Languages of these branches are spoken in China and Vietnam. These terms are not to be confused with *Northern Thai* and *Central Thai*, which are regional variety names in Thailand belonging to Li's *Southwestern Tai* branch.

*Tai* and its Chinese pinyin equivalent *Dai* also occur in the names of specific varieties, such as *Black Tai*, *Tai-Aiton*, etc. Note also the million-strong *Tày* language of Vietnam, in Li's *Central Tai* branch. This has sometimes been referred to as *Tho*, but that name is now understood to be objectionable. In Chinese usage, *Dai* denotes a nationality unit of Yunnan. This combines speakers of what is elsewhere known as *Lue*, *Lü*, or *Tai-Lue*, along with the markedly different *Dehong* or *Tai-Dehong* (also know as *Chinese Shan* or perhaps *Nuea*, literally 'northern' or 'above', although that term may also refer to a different variety in the general area). In linguistic practice, *Dai* has been used in the sense of *Lue* alone, i.e. in this case excluding Dehong.

One should note that the *Thai/Tai* distinction sketched above has not always been made in way indicated and that this nomenclature pairing is not satisfactory to all. Some scholars, including Benedict (1975), have used *Thai* to refer to a wider (*Tai*) grouping and one sees

designations like *Proto-Thai* and *Austro-Thai* in earlier works. In the institutional context in Thailand, and occasionally elsewhere, sometimes *Tai* (and its corresponding Thai-script spelling, without a final -y symbol) is used to indicate varieties in the language family **not** spoken in Thailand or spoken there only as the result of recent immigration. In this usage *Thai* would not then be considered a *Tai* language.

On the other hand, Gedney, Li and others have preferred to call the standard language of Thailand *Siamese* rather than *Thai*, perhaps to reduce potential *Thai/Tai* confusion, especially among English speakers not comfortable with making a non-English initial unaspirated voiceless initial sound for *Tai*, which in any event might sound artificial or arcane to outsiders.

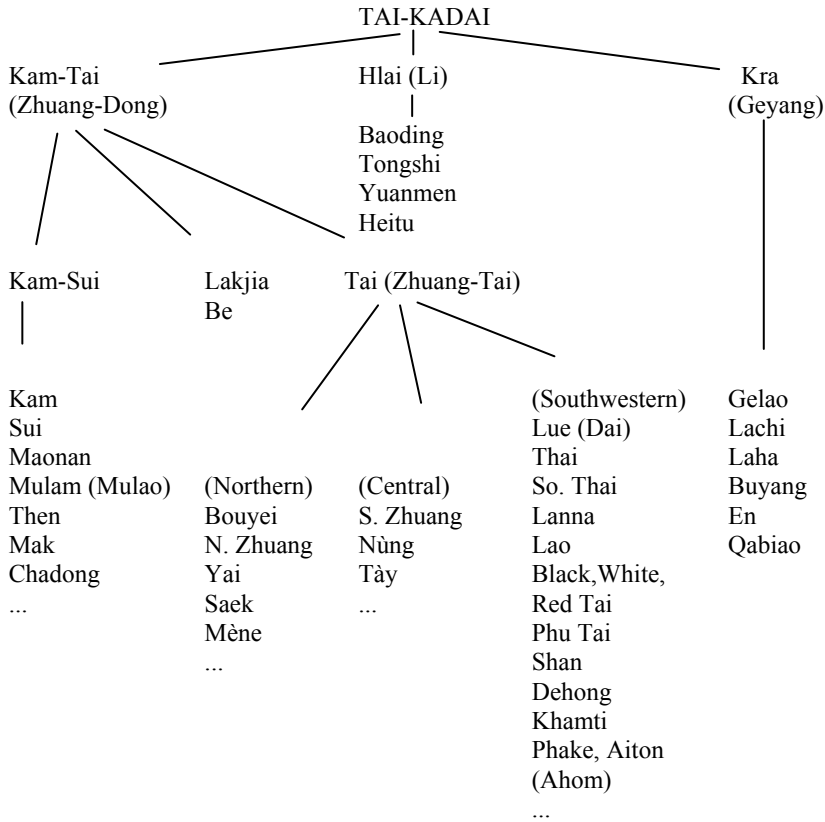
Moreover, Lao scholars I have met are not pleased with Lao being regarded as a *Tai* language. For some, Thai should instead be considered a member of the Lao language family. One or more Ancient Chinese characters for 'Lao' may be cited in support of this alternative appellation. Enfield (2002) discusses the critical issues here with clarity.

*Tai* in the sense Li has used it is now so widely accepted as to be the norm. The *Tai* language family in turn is taken as a major sub-branch of *Kam-Tai* within the greater grouping *Tai-Kadai*. Finally, whether *Tai-Kadai* is to be regarded as an isolate or whether it might be placed in some more ambitious macro-family with a designation like *Sino-Tai* or *Austro-Tai* is still a matter of great interest and current professional debate, as papers in this volume will attest.

*Kadai* as a term raises problems of its own. At least one Thai scholar has objected to *Kadai* (Ostapirat 2000, 2004), noting that it seems to mean 'ladder' in Thai. Another possibility, equally inappropriate, is that given a voiceless rendition of *Kadai*'s second syllable, i.e. to sound something like *Tai*, and given the right vowel-length and tone, *Kadai* could be mistaken for the Thai or Lao noun 'rabbit'. Furthermore, as though to complete a tortoise-and-hare scenario, it seems that *Kra*, Ostapirat's suggested term to replace *Kadai*, in Thai actually sounds like a noun referring to a large maritime turtle, *Chelonia imbricata*. However, in spite of a slow start, Ostapirat's tortoise entry *Kra* may yet overtake the *Kadai* hare and win the terminological race. If so, would *Tai-Kadai* need to become *Tai-Kra* (or could we perhaps settle for something less inventive like *Tai-Gelao*)?

The original rationale for Benedict's (1942, 1975) invention of *Kadai* was apparently to group languages of the Hlai (Li) type on Hainan Island together with the Gelao-Laha type in mainland China and Vietnam, but it would now appear that the case is rather weak for recognizing that particular subgrouping as a unified branch (by whatever name). Reinterpretation of *Kadai* to avoid the subgroup claim is one reasonable approach. *Kadai* in a new sense may be restricted to languages of the Gelao-Laha group (i.e. *Kra* or Geyang); an opposite possibility is to use it to denote the entire (Tai-Kadai) grouping.

Readers interested in ramifications of these terminology issues and in associated subgrouping proposals, debates and related diagrams might consult Edmondson and Solnit (1988, 1997); Thurgood (1994); Luo (1997, 2007); Diller (2000); Ostapirat (2000, 2004); Matisoff (2001); Edmondson (2007). These sources can be complemented with chapters in this volume. For convenience, one tentative diagram for reference follows but it should not be considered the final picture.



## REFERENCES

- Benedict, Paul K. (1942) 'Thai, Kadai and Indonesian: A New Alignment in Southeastern Asia', *American Anthropologist* 44: 576-601.
- (1975) *Austro-Thai Language and Culture with a Glossary of Roots*, New Haven: HRAF Press.
- Diller, Anthony (2000) 'The Tai Language Family and the Comparative Method', in Khunying Suriya Ratanakul, et al. (eds) *Proceedings of the International Conference on Tai Studies, July 29-31 1998*, Institute of Language and Culture for Rural Development, Mahidol University, 1-32.
- Edmondson, Jerold A. (2007) 'The Power of Language over the Past: Tai Settlement and Tai Linguistics in Southern China and Northern Vietnam', in Harris, Jimmy G., Somsonge Burusphat, and James E. Harris (eds) *Studies in Tai and Southeast Asian Linguistics*, Bangkok: Ekphimthai Ltd. 39-63.
- Edmondson, Jerold A. and David B. Solnit (eds) (1988) *Comparative Kadai: Linguistic Studies beyond Tai*, Summer Institute of Linguistics and University of Texas at Arlington.
- (eds) (1997) *Comparative Kadai: the Tai Branch*. Summer Institute of Linguistics and University of Texas at Arlington.
- Enfield, N. J. (2002) 'How to Define 'Lao', 'Thai' and 'Isan' Language? A View from Linguistic Science', *Tai Culture* 7.1: 62-67.
- Li Fang-Kuei (1977) *A Handbook of Comparative Tai*, Oceanic Linguistics Special Publications No. 15, Honolulu: University of Hawaii Press.



- Luo Yongxian (1997) *The Subgroup Structure of the Tai Languages: A Historical-Comparative Study*, Berkeley, California: Journal of Chinese Linguistics Monograph Series No. 12.
- (2007) ‘Sino-Tai Words for “Cut”’, in Harris, Jimmy G., Somsong Burusphat, and James E. Harris (eds) *Studies in Tai and Southeast Asian Linguistics*, Bangkok: Ekphimthai Ltd., 155-181.
- Matisoff, James A. (2001) ‘Genetic Versus Contact Relationship: Prosodic Diffusibility in South-East Asian Languages’, in Alexandra Y. Aikhenvald and R.M.W. Dixon (eds) *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*, Oxford University Press, 291-327.
- Ostapirat, Weera (2000) *Proto-Kra*, Berkeley, California: Linguistics of the Tibeto-Burman Area 23.1 (entire volume).
- (2004) ‘Kra-dai and Austronesian: Notes on Phonological Correspondences and Vocabulary Distribution’, in Laurent Sagart, Roger Blench and Alicia Sanchez-Mazas (eds) *The Peopling of East Asia: Putting together Archaeology, Linguistics and Genetics*, London: Routledge-Curzon, 107-131.
- Thurgood Graham (1994) ‘Tai-Kadai and Austronesian: The Nature of the Historical Relationship’, *Oceanic Linguistics* 33: 2.

## CHAPTER TWO

# SINO-TAI AND TAI-KADAI: ANOTHER LOOK<sup>1</sup>

*Yongxian Luo*

### 2.1. INTRODUCTION

A key issue in Sino-Tibetan studies is the historical link between Chinese and Tai — whether or not they are genetically related. While there is no question about the status of Tai as a distinct language family (Li 1977), the nature of Sino-Tai relationship is still much debated. Opinions can be divided into three camps: (1) Sino-Tai (De Lacouperie, Maspero, Wulff, Li, Haudricourt, Shafer, Denlinger, Xing, Zhengzhang); (2) Austro-Tai (Schlegel, Benedict);<sup>2</sup> (3) Sino-Tibetan-Austronesian (Sagaart, Reid, Xing). The Sino-Tai hypothesis assumes the membership of Tai under Sino-Tibetan while the Austro-Tai theory argues for a genetic relationship between Tai and Austronesian. The Sino-Tibetan-Austronesian hypothesis proposes a larger phylum that includes Chinese, Tai-Kadai, Miao-Yao, and Austronesian. In addition, there are still others who believe in none of these ideas (Gedney 1976, Thurgood 1994).

Western researchers now generally accept the Austro-Tai theory, but many scholars in China hold the traditional view of Sino-Tai alliance. Some of them consider the Sino-Tai connection to be the result of language shift (Dai and Fu 1995, Luo Meizhen 1992, 1994).

Scholars in the field generally postulate the time depth of Tai to be around 2500 years, and place the homeland of the Tai people somewhere in south China's Guizhou-Guangxi- area (Gedney 1965, Matisoff 1991, Wyatt 1984). Archaeological finds seem to lend support to these arguments.<sup>3</sup> Yet no coherent and systematic account of the history of the Tai people is readily available in early Chinese sources. This is not surprising, as south and southwest China was strange to the early Chinese settlers in the Central Plain. From the fragmented historical records that we can piece together, south China is home to the Bai Yue 百越/百粵 (Hundred Yue), a group of indigenous non-Chinese peoples who inhabited the vast area along the coast as far north as Shandong and south down to the Yangtze basin and west to as far as present-day Sichuan province.

Chinese historical sources indicate that when Qin Shihuang conquered Lingnan in 224 BC, the vast area was inhabited by non-Chinese speaking peoples, some of whom must have been the ancestors of Tai speakers. After Qin and Han dynasties, more systematic accounts were kept of these non-Chinese speaking populations. Barlow (2000a, 2000b) has synthesized these accounts, with an emphasis on the Zhuang people in Guangxi. Several independent studies on this topic have also appeared (Wang 1993, 1998, Xing 1984).

---

1 An earlier version of this paper was presented at a seminar at the Research Centre for Linguistic Typology of La Trobe University in November 2005. The author wishes to thank the participants for useful feedback. This study is supported by a research grant from the Australian Research Council (DP209445).

2 Matisoff (2001: 297) holds a sceptical view about this and other high-order groupings, while in the same paper (p. 316) he is treating it as a supergroup when talking about the Chinese tonal influence on Tai-Kadai, Hmong-Mien and Vietnamese.

3 For a survey on Yueh Neolithic, see Meacham 1983: 147-148.

The linguistic evidence seems to suggest that the Bai Yue were made up of at least two subgroups: one which spoke Austroasiatic-related languages and another that spoke Tai and Hmong-Mien-related languages (Norman and Mei 1976, Li Jingzhong 1995, Pan 1995, La Polla 2001). This has led to the assumption that the Kam-Tai people were generally a subset of the Bai Yue, a view that is gaining increasing acceptance.

The aim of this chapter is to present an update overview of the issue of Sino-Tai connection. It first offers a critical assessment of the different views of the Sino-Tai link, followed by discussions of current progress. More importantly, this study will reassess Li Fang-Kuei's 1976 position paper on Sino-Tai. Some of the issues raised in that paper will be addressed, such as shared morphological features and grammatical processes. Special attention will be paid to Li's comparative method and Benedict's teleo-reconstruction. Some new findings will be presented. The intention is to shed new light on the issue of the nature of historical relationship between Chinese and Tai.

## 2.2. THE SINO-TAI HYPOTHESIS: A REVIEW OF DIFFERENT THEORIES

### 2.2.1. Shared typological features between Chinese and Tai

Whatever their ultimate relationship may turn out to be, Chinese and Tai have the following features in common:

- 1) They share the same number of phonemic tone classes, commonly referred to as A (level), B (departing), C (rising), and D (those ending in final -p, -t, -k); regular correspondence can be established between Chinese and Tai for each of these tonal classes;
- 2) They share four classes of initials based on type of articulation, i.e. voiceless, half voiceless, voiced, half-voiced;
- 3) They share correlations between initials and tones, i.e. different initial consonants go with different tones, and the tones in turn affect the voicing of the initial consonants;
- 4) They share similar rhyme groups (Chinese 摄 *she*);
- 5) They share a significant number of lexical items, including a considerable amount of 'core vocabulary';<sup>4</sup>
- 6) They share many typological features in various aspects of grammar, including word order, phrase structure, semantic space and syntactic constructions.<sup>5</sup>

Some of the above features have been used by Denlinger (1989) as 'formal proof' for a Sino-Tai genetic relationship. Particularly important is lexical sharing, which supplies invaluable information for Old Chinese reconstruction. For example, Tai reflexes have been used in the reconstruction of initial clusters for Old Chinese (see discussion in §2.3 and §2.4 below). While some of the above Sino-Tai common traits appear to display areal features (see Matisoff 2001), others are unique to Chinese and Tai. As we shall see below, there is still further evidence that points to a deep Sino-Tai historical connection.

### 2.2.2. The traditional view on the Sino-Tai link and competing theories

#### 2.2.2.1. *De Lacouperie and Wulff*

Before the publication of Paul K Benedict's 1942 article, 'Tai, Kadai, and Indonesian: a

<sup>4</sup> This observation is different from the views held by Benedict (1975) and others (Ostapirat 2005).

<sup>5</sup> These include a number of grammaticalised items and functional operators, such as case markers such as 要 *yào* - Tai *rau* A1 (instrumental, manner); tense and aspect markers 了 *liǎo* (perfective) Tai *yù* 欲 - Tai *yaak* (irrealis), 住/著 *zhù* -Tai *yu* (progressive); 过 *guò* - Tai *kwa* B1 (past, experiential), and many negative words (Luo 1997: 98-99), among others.

new alignment in Southeast Asia', it was taken for granted that Tai was a member of the Sino-Tibetan stock. The traditional view of Sino-Tai alliance was based on the fact that Tai shares with Chinese many typological similarities, both in phonology and grammar, along with a substantial number of lexical items.

De Lacouperie (1883) did pioneering work on Sino-Tai genetic relationship. He put forward over 30 cognate words for what he called the Tai-Shan and Kunlunic (Sino-Tibetan) stock. His thesis set a corner stone for the traditional theory on the Sino-Tai.

Wulff (1934) undertook a systematic study on Sino-Tai connection. This impressive work, using the comparative method, proposed a possible genetic relationship between Chinese and Tai on the basis of over 600 carefully worked-out cognate sets. But Wulff's findings have been interpreted in different ways by scholars. Many researchers, including the present writer, would view them as evidence for a Sino-Tai link, while others, for example, Egerod (1976) and Benedict (1942, 1975), would consider them as evidence for language contact rather than genetic relatedness.

Perhaps the most significant finding of Wulff's work was his discovery of tonal correspondence between Chinese and Tai. Both Chinese and Tai have three tonal categories in open syllables, conveniently labelled as A, B, and C. Regular correspondence can be established between Chinese and Tai for these tone classes, as in [Table 2.2.2.1-1](#):

**TABLE 2.2.2.1-1: CHINESE-TAI TONAL CORRESPONDENCES**

Chinese tones	Tai tone categories	Thai tones
Ping (p'ing) [level] (unmarked)	A	(unmarked)
Qu (ch'ü) [departing] (marked by -h)	B	mai eek
Shang [rising] (marked by -x)	C	mei thoo
Ru (ju) [entering] (marked by -p -t -k)	D	(unmarked)

Each of these categories is further divided into two allotones, one for high register (A1, B1, C1) and the other for low register (A2, B2, C2). A similar distinction can be observed in syllables with final stop ending (D1, D2). This scheme has been adopted by scholars in the field for comparative work. Xing (1962) expanded these lines of arguments.

#### 2.2.2.2. *Li's position paper on Sino-Tai*

Li made his first pronouncement on the Sino-Tai genetic relationship in 1938, but a full position paper did not appear until 1976. With a sample of over 120 Sino-Tai cognates, Li presented some of the most convincing arguments for a genetic link between Chinese and Tai. Included in his database are sets of regular sound correspondences, and a significant number of core vocabulary items, such as body-part terms (blood, head, eye, arm, neck, shin, skin, flesh, leg), terms for nature and environment (sun, fog, hot), kinship terms (father, mother), everyday words (cut, chop, dig, split, soak, cooked/ripe), among others. These were analyzed and discussed in the context of genetic relationship. Tonal irregularities were considered to be an important feature between Chinese and Tai. In addition, partial correspondences were also proposed as cognate candidates. Thus, a link can be made between Chinese 鼓 'drum' (OC \*kag, MC kɔ, and MSC gu<sup>3</sup>)<sup>6</sup> and Tai \*klɔŋ, and between Chinese 肚 'stomach' (OC \*dag, MC tɔ, MSC du<sup>4</sup>) and Tai \*dɔŋ C2, and between Chinese 補 'to mend' (OC \*pag, MC pɔ, MSC bu<sup>3</sup>, Proto Tai \*foŋ A1).

It is important to note that Li was employing the standard comparative method in mining Sino-Tai cognates. His methodology is in stark contrast to Benedict's tele-reconstruction (see

6 Old Chinese reconstructions are from Li (1971 [1980]); Middle Chinese reconstruction is from Pulleyblank (1991) and Proto-Tai reconstruction is from Li (1977).

2.3 below). Li's identification of cognacy between Chinese and Tai was based on his profound knowledge of the phonological systems of Chinese and Tai. He was well aware of the differences between the tonal, isolating Chinese/Tai languages and the inflectional, non-tonal Indo-European languages. His working criteria are best exemplified in his *Studies on Archaic Chinese Phonology* (1971) and *A Handbook of Comparative Tai* (1977). His methods have served as a model for Chinese and Tai linguistics, with many followers, Chinese and western.

Li's position remains influential among Chinese linguistic circles. His paper has raised many important questions, some of which still await to be answered, a point we shall return to shortly in §2.3 and §2.4.

#### 2.2.2.3. Haudricourt and Manomaivibool

Haudricourt's 1974 work appeared as Chapter 29 in the five-part *Introduction to Sino-Tibetan* (1966-1974), edited by Robert Shafer who unequivocally grouped Tai under Sino-Tibetan. With data from 14 Tai dialects, including Siamese, Lao, Shan, Black Tai, White Tai, Tho, Nung, Po-ai, and Wuming, along with data from Sui, Mak, Be and Li, this work represented one of the most serious attempts at the reconstruction of Proto Daique (Daic). Sound correspondences have been carefully worked out for tones, initials, and finals, using the standard method of comparative linguistics. Frequent reference was made to Chinese, with several dozen good Sino-Tai cognates proposed, indicating the author's agreement with the editor's view on the genetic affiliation of Tai with Sino-Tibetan. Unfortunately, the work is riddled by too many typographical errors, due to the negligence of the type setter/copy editor.

Manomaivibool (1975) is yet another full-scale study on the Sino-Tai link along the lines of Wulff. This impressive work has postulated over six hundred correspondence sets between Chinese and Thai, some of which were also found in Wulff and Haudricourt. But unlike Wulff and Haudricourt, Manomaivibool examined Thai correspondences by comparing them with Karlgren's and Li's reconstructions of Old Chinese and Middle Chinese. The results were revealing: Thai reflexes are found to map in part with Old Chinese (the first half of the first millennium BC) pronunciation and in part with Middle Chinese (6<sup>th</sup>-11<sup>th</sup> AD, typically reflected in *Qieyun*) pronunciation. The findings have important ramifications for the reconstruction of Old Chinese, Middle Chinese, and Proto-Tai.

#### 2.2.2.4. Xing Gongwan: *Handbook of Comparative Sino-Tai*

Xing's *Handbook of Comparative Sino-Tai* (1999) is an impressive work on the topic. With over 900 proposed cognates between Chinese and Tai, the book has significantly expanded the Sino-Tai lexicon along the lines of Li (1976). The overwhelming majority of the cognates are found in Li's HCT, indicating that over 70% of Proto-Tai lexicon has Chinese links. A significant number of the shared cognates are proposed for the first time, such as Chinese 霁 (Old Chinese *\*phjən* 'rain', Proto-Tai *\*fon* A1; Chinese 盲 (OC *\*phat*) 'blind', PT *\*ʔot* D1; Chinese 拇 (OC *\*mæg*) 'hand', PT *\*mue*; Chinese 齧 (OC *\*gət*) 'to bite, gnaw', PT *\*kat* D1. Also worth noting is the fact that Tai often retains earlier meaning of words or forms which are found in Old Chinese but which are now out of use in modern Chinese, eg. Chinese 元 (OC *\*ɣwjan*) 'day', PT *\*van* A2; Chinese 食厄 (OC *\*ɣek*) 'hungry', PT *\*ɣaak* D1; Chinese 肝 (OC *\*kan*) 'night, evening', PT *\*ɣjan* A2. For Xing, the existence of words of this kind is testimony to a Sino-Tai genetic relationship.

One is struck by the regular patterns of phonological correspondences Xing has set up for Chinese and Tai. Apart from the comparative method — albeit sometimes a bit loosely applied, Xing also adopts a semantic approach. For him, a deep semantic match can better reveal a deep historical connection. He has summarized three kinds of parallel correspondences between Chinese and Tai: (a) allophones, (b) synonyms, and (c) word families (Xing 1999: 499; see §2.4.3 for more discussion). For instance, Chinese 浸 (OC *\*tsjəm*) 'soak', Dai Ya *tɕum*<sup>B1</sup> (PT *\*čum*<sup>B1</sup>), Chinese 湛 'flood' (OC *\*drəm*), Dai Ya *thum*<sup>A2</sup>

(PT), Chinese 燻 ‘boil, cook in water’ (OC *\*dəm*), Dai Ya *tum<sup>CI</sup>* (PT *\*tum<sup>CI</sup>*), Chinese 潜 ‘sink, submerge’ (OC *\*dzjiəm*), Dehong *tsum<sup>AI</sup>* (PT *čem<sup>AI</sup>*) (pp.184-185). Examples of this kind show that Chinese and Tai share extensively in morphological processes, a feature that is less likely to be borrowed.

While Xing’s method of semantic differentiation has many merits in it, he probably has stretched it a bit too far in places. For this reason, he has been criticized by a number of scholars (Ting 2000, Nie 2002, Mei 2003). Despite this, this work will remain an important source of reference for scholars in the field.

#### 2.2.2.5. *Recent works*

Very recently, Gong (2002) built on Manomaivibool’s work by reexamining different layers of Sino-Thai cognates — those that correspond to Old Chinese and those that correspond to Middle Chinese. His findings have extended Manomaivibool’s thesis. Ting (2002) reevaluated Li’s 1976 position, citing supporting evidence from recent progress in Chinese historical linguistics, as did Pan (2002), who considered Sino-Tai link in a wider context. Lan (2003), too, was concerned with different strata of Sino-Tai words. Unlike Gong and Manomaivibool, Lan’s focus is on the sound correspondences between Chinese and the Zhuang dialects in South China’s Guangxi province, which the largest group of Tai speakers in China inhabit. By comparing Tai forms with Old Chinese and Middle Chinese reconstructions by earlier scholars, he was able to draw a rough picture of the linguistic interaction between Chinese and the Zhuang at different stages in history. Zeng (2004) was tackling a similar issue of the genetic relationship between Chinese and Sui — a member of the Kam-Sui group with which Tai is closely related. Like Manomaivibool, Gong and Lan, she was looking at two types of related words between Chinese and Sui, words that reveal a deeper historical connection and words that point to late contact. The results led her to the belief that Chinese and Sui are genetically related.

At the time of writing, the present writer (forthcoming) has identified some 400 extra putative Sino-Tai cognate words that are not included in Manomaivibool’s work. He (Luo 1997b, 1998) has also discovered patterns of regular sound correspondences between Chinese and Tai in the phonetic series of sibilants and liquids as well as sibilants and velars, which lend further support to the *xiesheng* contact in Old Chinese as also revealed in some of the characters. The Tai correspondences lend strong support to the idea of existence of a sibilant complex in the proto language. In several papers (Luo 2003, 2004, 2005, 2006), he also looked at Sino-Tai sharing in morphological processes. All this was overlooked in the past. Some of his main findings will be discussed below in §2.3 and §2.4.

#### 2.2.3. **Benedict’s Austro-Tai theory and his method of tele-reconstruction**

Benedict’s *Austro-Thai Studies* (1975) was a reaffirmation and expansion of his 1942 paper, which was probably inspired by Schlegel who proposed to separate Tai from Sino-Tibetan. In both works, Benedict vehemently argued for a realignment of Tai-Kadai, first placing it with Indonesian (1942), and later linking it to a larger phylum — Austro-Tai. In his latest theory (1990), Japanese was added to the super stock, forming an even larger alliance.

The Austro-Tai hypothesis has fundamentally challenged the traditional view of the Sino-Tai genetic link. For this reason, it has sparked controversies, and generated heated discussion (see Volume 6 [1976] of *Computational Analysis of Asian and African Languages*). Authorities have different views on the validity of Benedict’s claims. For example, Egerod (1976) supported and accepted Benedict’s reconstructions while Haudricourt (1976) was somewhat lukewarm in his reaction. On the other hand, Benedict’s approach was strongly criticized by Gedney (1976), who cited evidence from Tai to show that on many occasions Benedict’s reconstructions were inadequate. Even Benedict’s ardent supporter, Matisoff, had some reservations about such ‘megalo-comparison’ (Matisoff 1990: 115) and Benedict himself cheerfully admitted himself to be a ‘lumper’ (Benedict 1990: 169).

One of Benedict's arguments was that the typological similarities between Chinese and Tai are the result of contact rather than genetic inheritance. Yet it must be pointed out that Benedict's thesis was constructed on the basis of tele-reconstruction, which was 'characterized by loose resemblances, semantic leaps and the making up of maximal earlier forms to account for cognate relationship without working out sound correspondences through step-by-step comparison' (Diller 1998: 22). Controversial as it was, Benedict's work opened up new horizons in the field. Its significance was still being evaluated and appreciated (Thurgood 1985, 1990, Reid 1988).

#### 2.2.4. Sino-Tibetan-Austronesian and Kadai-Austronesian

##### 2.2.4.1. Sagart's Sino-Austronesian hypothesis

As mentioned in §2.1, Laurent Sagart was a strong proponent of the Sino-Austronesian hypothesis. In a paper published in 1993, he proposed several dozen putative cognates between Chinese and Austronesian, claiming that the two languages are genetically related. Since Tai was viewed as a member of Sino-Tibetan, naturally it belonged to this Sino-Austronesian stock.

In a recent paper Sagart (2005b) has changed his earlier position to connect Tai with Austronesian rather than with Sino-Tibetan, although he still maintains that Sino-Tibetan and Austronesian are genetically related. In his current view, Tai-Kadai is a subgroup of Austronesian. What makes him change his mind is that 'one set of words suffices to show that at least some vocabulary is genuinely shared by Tai-Kadai and AN, not as a result of chance' (2005b: 177). Sagart's set of Kadai-Austronesian cognates is made up of three lexical items (ibid: 178):

	PAN	PMP	Tai	Lakkia
die	maCay	matay	ta:i <sup>1</sup>	plei <sup>1</sup>
eye	maCa	mata	ta <sup>1</sup>	pla <sup>1</sup>
bird	—	manuk	nok <sup>8</sup>	mlok <sup>7</sup>

Furthermore, Sagart provides a piece of supporting evidence that there is a reversal in the singular and plural number with the second person pronouns between Proto-Austronesian and Proto-Tai-Kadai:

	PAN	PKT
2SG	-Su	*məu
2pl	-mu	*sou

On the basis of evidence of this kind, Sagart concludes that 'Tai-Kadai has its origin in an early AN language called here 'AAK'. AAK was a daughter language of PAN, and a close relative of PMP.' (ibid: 180)

For his Sino-Tibetan-Austronesian phylum, Sagart offers 61 putative cognate sets between Proto-Austronesian, Old Chinese and Tibeto-Burman. Included in the sets are items such as 'bone', 'elbow', 'breast', 'foot', 'head', 'cloud, cloudy', 'sunlight', 'water', 'cave', 'dig', 'hold in the mouth', 'lick', among others (2005a: 164). In addition, 14 cultural lexical items are also compared (ibid: 165), such as 'paddy', 'chicken', and 'crossbow'.

##### 2.2.4.2. Kadai-Austronesian

In an impressive paper, Ostapirat (2005) has made what would appear to a convincing case for a genetic link between Kra-Dai (his term for Tai-Kadai) and Austronesian by

gathering new data from the Tai-Kadai languages in China and Vietnam and comparing with Austronesian. 50 carefully selected putative cognate words are cited for Tai-Kadai, of which 20 are on Yahontov's 35 basic word list, and 42 on Swadesh's 100 basic word list. Of these 50 Tai-Kadai core lexical items, 26 are found to have Austronesian connection. Among the 26 possible cognate items, 9 are not on the Swadesh's list.

Ostapirat's proposed Tai-Kadai-Austronesian etyma are listed below.

Gloss	PAN	PKd
eye	maCa	m-tja A
fart	qe(n)tut	C-tot D
hand	(qa)lima	mja A
leg	paqa	C-ka A
shoulder	qabaRa	*m-ba B
bear (n.)	Cumay	C-me A
louse (head)	kuCu	C-tu A
sesame	leŋa	l-ŋa A
moon	bulaN	m-djan A
black	tidem	hl/dəm A
eat	kaen	kan A
grandmother	aya	ja C
I	aku	ku A
excrement	Caqi	kai C
grease	SimaR	mal
head	qulu	krai B
nose	ijun	teŋ
tooth	nipen	l-pən
bird	majuk	ŋok D
leaf	(ʔabag)	Hlai betu
fire	Sapuy	pui A
water	daNum	ʔun C
live, raw	qulip	(k-)dɛp D
child	aNak	lak D
this	i-ni	ʔ-ni C/B
you	kamu	mə A/B

Having established the lexical correspondence, Ostapirat goes on to discuss Kadai tones in Austronesian roots. Following the theories of tonogenesis by Haudricourt and others, he conjectures that Kadai tones might have some correlations with Austronesian syllable types.

A comparison between Sagart's list and Ostapirat's shows that the correspondences belong to different sets of vocabulary items.<sup>7</sup> A question may now be asked: What is the nature of these lexical items? What can be said about the status of such lexical items? To what extent do they reflect the real picture of genetic relationship? How should we interpret them? In the sections below we shall look at more evidence from the Tai side.

<sup>7</sup> Sagart (1993) also postulates over 200 possible cognate words between Chinese and Austronesian, including body part terms like 'palm of the hand, sole of the foot' PAN *\*Da(m)pa*, Chinese 附 [GSR 102] OC *\*phag*, terms for nature and environments; 'open expanse of land or water' PAN *\*bawaŋ*, Chinese 潢 [GSR 707e] *\*gwaŋ* 'lake, pool'; abstract concepts like 'to oppose' PAN *\*baŋkal*, Chinese 奸 [GSR 139cd] *\*kan* or *\*k-r-an* 'treacherous; disobey; violate', among others. Some of Ostapirat's items are also in Sagart's list, eg. 'black', 'sesame', 'to eat'.



### 2.3. THE ROLE OF TAI IN OLD CHINESE RECONSTRUCTION

Whatever the ultimate relationship between Tai and Chinese may eventually turn out to be, researchers all agree that Tai plays an important role in Old Chinese reconstruction. Scholars in the field are constantly citing Tai materials as supporting evidence (Nishida 1960, Haudricourt 1956, 1974, Li 1970, Benedict 1976, Bodman 1980, Baxter 1992). Indeed, the significance of Tai in OC reconstruction cannot be overestimated, as shall be seen below.

#### 2.3.1. Initial clusters in OC as retained in Tai

It is generally believed that a number of initial clusters should be reconstructed for Old Chinese (Karlgren, Li, Gong, Baxter). The argument behind this is that in Chinese orthography, there are a significant number of *xiesheng* characters where alternations of stop initials (p, b, t, d, k, g) and liquid l- are most common. However, almost no modern Chinese dialect has provided any diachronic link. Much to the comparativists' delight, Tai supplies valuable information.

One of most frequently-cited examples is the word for 'indigo' in Tai. The Chinese cognate of is word, now pronounced with a liquid initial /l/ as 藍 *laan*<sup>4</sup> in modern Chinese, is orthographically associated with a phonetic series 藍, with a velar /k/. One would expect to find a velar cluster of some sort (\*kl/r- or \*gl/r-) in the modern dialects. Yet no modern Chinese dialects are found to retain such an initial. However, many Tai dialects have *graam*<sup>42</sup> for this item for which Li has proposed a proto cluster \*gr- for Tai, thus attesting the existence of the velar cluster in the proto language.

Another example is the reconstruction of labial clusters \*pl-, \*bl- for Old Chinese. Chinese orthography also suggests that words like 變 'change' (OC, MC MSC biàn) must have derived from an initial cluster involving a labial stop /p/ and a liquid /l/, as the character has the phonetic element 蠻 'luan', pointing to a liquid initial. Karlgren (1957: 67) has proposed \*plian (GSR 178o) for 變. Again no modern Chinese dialects are found to preserve initial \*pl-. Much to the comparativist's delight, many Tai dialects have reflexes containing initial /pl-/ for which Li has reconstructed a labial cluster \*pl- for Proto-Tai. The Tai form is undoubtedly related to Chinese.

Norman and Mei (1971) observe that in the Chinese Min dialect, there exist a group of words for which a proto \*s- cluster may be proposed. Significantly, Tai again supplies supporting evidence, which points to a deep connection between Chinese and Tai, for which Luo (1997a, b) has proposed a sibilant complex for Sino-Tai (see §2.4.1).

#### 2.3.2. Retention of bilabial initials which have developed into labial-dentals in Modern Chinese

In their admirable studies of Old Chinese phonology, the great Qing scholars have made some important discoveries about the phonological system of Old Chinese. One such discovery was that labial-dental initials were lacking in Old Chinese. Reflexes in some modern Chinese dialects such as Cantonese provide supporting evidence. Significantly, a set of cognate words in Tai with labial initials are found to correspond to Modern Chinese labial-dental initials, supplying further evidence for hypothesis of the absence of labial-dental initials in OC.

Gloss	Chrt	Pinyin	GSR	OC	MC	Thai	DH	LM	Yay	FS	HCT
father	父	fù	102a	pjagh	puǎʔ	phəwB2	po B2	poo B2	po B2	po B2	*bo B2
cooking pot	釜	fǔ	102f	bjagʷ	buǎʔ	məw C1	mo C1	moo C1	mo C1	mo C1	*hmo C1

bee, beeswax	蜂	fēng	1197s	phjung	phuawŋ	phĩŋC1	phun C1	Nung phuŋ C1	–	–	(*phĩŋ C1)
skin	膚	fū	69g	pjag	puǎ	pliakD1L	pək D1	puuk D1L	pyaak D1L	pyaak D1L	*plaak D1L
lungs	肺	fèi	501g	phjadh	puajh	pɔɔt D1L	pɔt D1L	pyt D1	put D1S	put D1S	*pɔt D1L
fat	肥	fèi	582a	bjəd	puaj	phii A2	pi A2	pi A2	pi A2	pi A2	*bi A2
divide	分	fēn	471a	pjiən	pun	pan A1	pan A1	pan A1	pan A1	pan A1	*pan A1

The development of labial sounds into labial-dentals was a feature of late Middle Chinese. The retention of labial initials in Tai indicates a deep historical connection between the two languages.

### 2.3.3. The reconstruction of final -g for OC: Some supporting evidence

A key feature in Li's Old Chinese reconstruction is that of the final \*-g for a group of rhymes that are represented with open-syllables in Middle Chinese. In Li's system, we find codas like \*-ag, \*-ig, \*-əg, and \*-ug, with no open syllables in Old Chinese except the \*-ar (see below). This has been criticized by a number of scholars (Wang 1985) for being typologically implausible and unusual. However, Li must have some reasons to believe that codas of this kind must have existed in early history. The following examples were cited by him (1976: 41):

Gloss	Chrt	Pinyin	GSR	OC	MC	Thai	DH	LM	Yay	FS	HCT
fog, mist	雾	wù	1109t	mjəgwh	muǎh	mək D1	mək D1	mookD1	mookD1	mookD1	*hmok D1
hat, cap	帽	mào	1062	məgwh	mawh	muak D1L	WT mök D1	muuk D1L	–	–	*muak D1L
skin, bark	膚	fū	69g	pliag	puǎ	pliak D1L	pək D1	puuk D1L	pyaak D1L	pjaakD1L	*pliak D1L
forehead	顙	lú	69p	blag	lɔ	phaak D1L	phaak D1L	phaak D1L	pyaak D1L	pjaakD1	phlaak D1L

The first and the second examples are in the traditional Chinese 候部 and 幽部 respectively, and the last two examples in the 鱼部. In our data, many more examples of this type are found in Tai, lending support to Li's reconstructions.

Gloss	Chrt	Pinyin	GSR	OC	MC	Thai	DH	LM	Yay	FS	HCT
place, region	處	chù	85a	thjag	tɕʰiə <sup>h</sup>	thiāk D2L	–	–	tuak D2L	tuak D2L	*d̥iək D2L
child	子	zǐ	964a	tsjəg	tsi/tsi'	tuuk D2L	luk D2	lok D2	luuk D2	luuk D2S	*l̥iuk D2L
base, foundation	基	jī	952g	kəg	ki/ki	kok D1S	kok D1S	kuk	kok D1S	kək D1S	*kok D1S
carry on shoulder	負	fù	1000a	bjag	buw'	bɛək D1L	mek D1	meek D1L	–	–	*ʔbɛək D1L

small	小	xiǎo	1149a	siagw	siaw'	lek D2S	lik D2S	–	–	leək	*dlek D2S
swaddling clothes	襌	bǎo	1057g	pəgwɣ	paw'	–	Mok D1L	Mok D1S	–	buk D1S	*ʔbuok D1L <sup>8</sup>
mortar	臼	jiù	1067a	gwjiəg wh	guw'	Khrok D2S	xok D2s	lok D2S	cok D2S	tsok D2K	*grok D2S

These examples show that Li's proposed reconstructions are not groundless speculations.

#### 2.3.4. Unique correspondences: The case of 'five', 'goose', 'gill', 'six'

In his 1976 paper, Li raised a question which has not been seriously addressed until quite recently (Ting 2002). Li's question has to do with a unique sound correspondence between Chinese and Tai in words like 'five', 'six', 'goose', and 'fish gill'. These lexical items take the velar nasal \*ŋ- in OC, where Tai correspondences show a laryngeal fricative \*h: 'goose' OC \*ŋaan, PT \*haan B1, 'five' OC \*ŋa C, PT \*ha C1, and 'gill' Chinese \*ŋək, PT \*hŋiək.<sup>9</sup> To date no Chinese dialects have been found to display correspondences of this kind. Similarly, 'six' is reconstructed as \*ljok in OC with a liquid initial. In Tai it shows a laryngeal fricative in Thai as *hok* D1, and a liquid *lək* in Po-ai, reconstructed as \*xrok D1 in HCT. These types of sound correspondences are not found in loan words. They are cited by Li as solid evidence in support of Sino-Tai link.

#### 2.3.5. Shared innovation: Finals \*-l and \*-r

Among the unsettling questions in Old Chinese reconstruction, the reconstruction of finals \*-l and \*-r is perhaps the most controversial. If these sounds once existed in history, one may expect to find some traces among the modern dialects. But so far, no reflexes among the numerous modern Chinese dialects are found to exhibit traces of these elements.

Two issues are involved here. The first has to do with the development of the traditional *ge* rhymes 歌部 which was reconstructed as \*-ar by Li Fang-Kuei (1971). Karlgren also reconstructed part of the *ge* rhymes as \*-ar in GSR. The second issue has to do with the reconstruction of final \*-r for Old Chinese.

##### 2.3.5.1. The *ge* rhymes and \*-l

Li's reconstruction of \*-ar for the *ge* rhymes is accepted by Gong (1993) and Pan (2007), who slightly revised it into \*-al. Baxter (1993) proposed \*-ei for this rhyme on the basis of reflexes from a number of modern dialects as well as Sino-Vietnamese and Sino-Korean forms. Baxter's reconstruction is questioned by Pan (2000, 2007), who believes that it fails to account for the final -i since there were already rhymes that had been reconstructed as such for Old Chinese. For Pan, Li's reconstruction has more explanatory power in that it can elegantly explain -i as lenition of \*-l.

If Pan's hypothesis of final \*-l lenition is correct, Tai supplies additional evidence. In Tai, a significant number of cognate words with final -i correspond to the traditional 歌 rhyme for Old Chinese. The following examples illustrate:

<sup>8</sup> Not included in HCT.

<sup>9</sup> The majority of Tai dialects take the laryngeal /h/ for this item, except Siamese and Po-ai, which side with Chinese in taking /ŋ-/.

Gloss	Chrt	Pinyin	GSR	OC	MC	Thai	DH	LM	Yay	FS	HCT
hang down, droop	垂	chuí	31a	tjuar	dzwið dzwi	yɔɔi A1	–	–	ruay A1	looi A1	*jroi A1
beat, hammer	捶	chuí	31i	tjuar	tɕwið, tɕwi	tɕi B1	tɕi B1	–	toy B1	toi B1	*toi B1
spittle, saliva	唾	tuò	31m	thuar	twa <sup>h</sup>	laai A2	laai A2	laai A2	nay A2	naai A2	*mlaai A2
bowl	盥	zhui	31	thuar	drwiä/ drwi	thuɔi C1	–	thuuy C1	tuay C2	tooi C2	*thuai C1
naked, bare	裸	luò	351h	luar'	lwa'	pluɔai A1	poi A1	Lao pluɔai A1	Saek puɔai A1	pjoi A1	*pluɔai A1 <sup>10</sup>
snail, shellfish	羸	luó	14b	luar	lwa	hɔi A1	hɔi A1	hoy A1	θay A1	θai A1	*srɔi A1 <sup>11</sup>
long and narrow mountain	陁	duò	11d	duarx	dwa', dwa <sup>h</sup>	dɔɔi A1	lɔi A1	nooy A2	doy A1	dɔi A1	*ʔdl/rɔi A1

These are not confined to the traditional -ar 歌 rhyme. Luo (2006a, 2006b) offers over a dozen more Tai examples that are proposed to be related to several other rhymes in Old Chinese. The forms invariably take final -i in Tai, which cannot be treated as chance occurrence.

#### 2.3.5.2. OC final \*-r and Saek -l

From *xiesheng* contact and internal evidence, along with early Sanskrit and Sino-Japanese pronunciations, Pan (2007) postulates a final \*-r for Old Chinese, which has developed into -n in the modern dialects.

It is well-known that Saek, a displaced Northern Tai language now spoken in Nakhon Phanom in Thailand, has preserved an archaic final -l. It is worth noting that this final -l is found to correspond to final -n in other Tai dialects. Final -l also occurs in Laha, a member of the Geyang group within Kadai. Characteristically, a significant number of words taking final -l in Saek are found to correspond to the traditional Chinese 元部 (-an) and 文部 (-ən) rhymes.

Saek words corresponding to the traditional Chinese \*-an rhymes include:

Saek	Thai	Fengshan	HCT	Gloss	Chrt	GSR	OC	MC
saal <sup>1</sup>	saan A1	saan A1	*san A1	fine grain	粦	154b	tsan	ts <sup>h</sup> an <sup>h</sup>
khial <sup>4</sup>	ɣɔɔn A1	–	*hɣɔɔn A1	cock's comb	冠	160a	kwanh	kwan <sup>h</sup>
tlɛl <sup>1</sup>	kron A1	tsɛn A1	*krɣuɔn A1	snore	鼾	139	han	xan
yual <sup>5</sup>	yuuɔn B2	jian B2	*juɔn B2	pass, extend	延	203a	rjan	jian
nɔɔl <sup>2</sup>	nɔɔn A1	noon A1	*hnɔn A1	worm	蠕	208c	nwan'	nwan'
sɛl <sup>3</sup>	sen C1	θɛn C1	(*sen C1)	tread, CLF.	线	155r	sjanh	sian <sup>h</sup>
sɛl <sup>6</sup>	san B1	θɛn B1	*san B1	tremble	颤	148s	tjanh	tɕian <sup>h</sup>

<sup>10</sup> Not included in HCT. Reconstructed by Luo (1997: 244).

<sup>11</sup> Not found in HCT. Reconstructed by Luo (1997: 105).

The following are Saek correspondences in OC \*-ən rhymes.

Saek	Thai	Fengshan	HCT	Gloss	Chrt	GSR	OC	MC
yal <sup>4</sup>	khanA1	hanA2	*ɣan A2	dike	垠	416h	ŋjən	ŋin
bul <sup>1</sup>	ʔbin	ʔbin	*ʔbin A1	fly, v.	𪛗	471e	pjən	p <sup>h</sup> un
vul <sup>4</sup>	fuun A2	fun A2	*vuən A2	burn	焚	474a	bjən	bun
mul <sup>4</sup>	Lao munB2	mun B2	(*muən B2)	powder	粉	471d	pjən'	pun'
sɔɔl <sup>1</sup>	sɔɔnA1	θoonA1	*sɔn A1	teach	訓	422d	hwjənh	xun <sup>h</sup>

Wider connections can be sought with Tibeto-Burman (Matisoff 2003: 383ff) for final \*-r for Proto Sino-Tibetan.

## 2.4. PROGRESS AND NEW FINDINGS

### 2.4.1. The sibilant complex

Sibilant clusters have been reconstructed for Old Chinese and Sino-Tibetan by scholars working in the field. In Li's reconstruction of Proto-Tai, consonant clusters have been reconstructed for labials (\*pl/r, \*phl/r, \*vl/r, etc.), dentals (\*tl/r, \*dl/r, \*thl/r, etc.) and velars (\*kl/r, \*khl/r, \*xr, etc.), but no sibilant clusters are proposed by Li.

There is strong evidence that suggests sibilant clusters may have existed in Proto-Tai. A number of consonant clusters have been proposed by Luo (1998) as additions to Li's PT phonemic inventory (\*sl-, \*sr-, \*zl-, etc.). More importantly, regular correspondences can be set up between Chinese and Tai for sibilants and liquids, with several dozen correspondence sets. The sibilant complex points to a deep historical connection between the two languages. Parallel development has been reported from the northern Min dialect of Chinese which displays traces of sibilant clusters (Mei and Norman 1971).

Apart from sibilant + liquid clusters, a parallel sibilant + velar cluster (\*sk-, \*zg-, etc.) is also found for Chinese and Tai. Luo (1997b) has presented evidence from the Tai side supporting his argument. Some of his Chinese comparisons were used by Li and Benedict in their reconstruction of Old Chinese and Sino-Tibetan, such as 'smell', 'needle', and 'feces'.

### 2.4.2. The issue of basic vocabulary

One of Benedict's main arguments to keep Tai apart from Sino-Tibetan is that in the area of core vocabulary, the two languages have little in common. The shared lexical items, he argues, are mainly trade terms, numerals and the like, which seem more likely to be loan words. Our data shows that this is not the case. A look at the Proto-Tai lexicon reveals that Tai shares with Chinese quite a large number of basic lexical items, including a sizeable number of body-part terms. Luo (2000) has postulated nearly 70 putative cognates between Chinese and Tai in this lexical field by synthesizing previous works by Wulff, Li and Xing, significantly expanding their inventory. For example:

Gloss	Chrt	PY	GSR	OC	MC	Siamese	Yay	HCT
head	首	shǒu	1102a	skhjəgwɣ	ɕuw'	khlaw C1	caw C1	*klau C1
ear	耳	ěr	981a	njəg'	ɲi/ɲi	hu A1	rua A2	*xrua
neck, throat	喉	hóu	113f	ɣug	ɣəw	khɔɔ A2	ho A2	*ɣɔ A2
bone	骼	gē	766c'	krak	kaijk, kɛ:jk	duuk D1	dok D1	*ʔdl/ruok D1
flesh	肉	ròu	1033a	njakwh	ɲuwk	niā C2	no B2	*niā C/B2

One may feel rather reluctant to say that words of this kind are entirely loans, as they are rather stable and are generally resistant to borrowing. If such words were excluded from the Proto-Tai lexicon, more than one third of the Proto-Tai etyma would have to go.

In addition to body-part terms, quite a significant number of basic words are found in early Chinese sources, with good correspondences in Tai. The following examples illustrate:

Gloss	Chrt	PY	GSR	OC	MC	Siamese	Yay	PT
bird	雉	luò	766g	lak	lak	nok D2	nok D2	*nlok D2
black	黠	dàn	658n	tam <sup>h</sup>	təm <sup>h</sup> / tam <sup>h</sup>	dam A1	dam A1	*ʔdam A1
white	皓	hào	1139h	gəkwh	ɣawʔ	khaaw A1	haaw A1	*xaau A1
red	赭	chēng	834m	thrjng	tr <sup>h</sup> iajŋ	ʔdeɛŋ A1	ʔdiŋ A1	*ʔdiɛŋ A1
dark purple	紺	gàn	606k	kam	kəmʔ / kamʔ	klam B1	cam B1	*klam B1
axe	斤	jīn	443a	kjən	kin	khwaan A1	vaan A1	*khwaan A1

An interesting thing about these words is that none of them are used in modern standard Chinese in their sense except in certain compounds like *hào yuè* ‘the bright moon’, found primarily in literary language, indicating that they are residues of early usage. Although the same concepts are now represented in different forms in modern Chinese (particularly the colour terms), the connections between the forms are recognizable. Some of these are obviously the original forms of the concepts designated, such as ‘axe’, the graphic for which, 斤 *jīn*, apparently a drawing of an axe, was found in oracle bone inscriptions, indicating its antiquity. This graphic serves as the semantic part of current form, 斧 *fǔ*, a form of much later appearance with a different phonetic shape.

Luo (2002, 2004, 2005, 2007, in preparation) has presented an array of basic vocabulary items from everyday words, which has expanded previous works.

#### 2.4.3. Morphological and orthographical evidence

Although consonant and vowel variations as a mechanism of morphological processes are a well-known phenomenon in Chinese and Sino-Tibetan (Karlgren 1933 Matisoff 1978, 1985), comparative work is scarce between Chinese and Tai.<sup>12</sup> A pioneering work was undertaken by Li in 1978 in which he dealt with a parallel morphological processes between Tai and Chinese in the word group ‘bent/crooked’ and ‘dig’, where he made the following remarks:

The genetic relation of Tai and Chinese needs serious study. The general impression that the phonological structure of these two languages is similar may be considered as an areal phenomenon, but is not necessarily so. The resemblance in vocabulary may be due to loans from one language to the other, but it is hard to prove. *The grammatical process of alternating initials may turn out to be an important factor to decide the relationship of Tai and Chinese.* (Emphasis added)

Li Fang-Kuei (1978: 406)

<sup>12</sup> As an exception, Dong et al. (1984) uses the notion of word families to argue for a genetic relationship between Chinese and Kam-Tai.

Drawing attention to the complexities of the issue of the Sino-Tai link, Li was calling for an in-depth investigation into the phonological and lexical structures of the two languages, proposing morphological processes as a criterion for determining the nature of Sino-Tai historical relationship. By discussing two cognate sets across Tai dialects and comparing them with Chinese, Li has presented evidence for a deep historical connection between these two languages.

In several recent papers (Luo 2002, 2003, 2004, 2005), the present writer has presented further evidence of sharing of morphological processes between Chinese and Tai. Some examples are given below for illustration.

#### 2.4.3.1. ‘Soak’ ~ ‘ooze’ ~ ‘wet’ ~ ‘sink’ ~ ‘submerge’

This group of words are typically represented by sibilant/dental + V + m.

Gloss	Chrt	PY	GSR	OC	MC	Thai	DH	LM	Yay	FS	HCT
to soak	浸	jìn	661m	tsjəm <sup>h</sup>	tsim <sup>h</sup>	čum C1/B1	tsum B1	sam C1	chum C1	sum B1	*čum B1
moisten, wet	沾	zhān	618c	tjam	triam	chum B2	yam A2	–	chom A1/B1	sum C2	*jum C2
dip into	沁	qìn	–	tshjəm <sup>h</sup>	ts <sup>h</sup> im <sup>h</sup>	čim C1/B1	tsam C1	cam C1	–	sam C1	*čiam C1
sink	沉	chén	656b-d	djəm	drim	čom A1	tsəm A1	cam A1	cham B1	sam A1	*čam A1
submerge	潛	qián	660n	dzjəm	dziam	dam A1	lam A1	nam A1	dam A1	?dam A1	*?dam A1

As the above examples show, alternations of historically voiceless vs. voiced initials and tones occur between transitive and intransitive verbs (‘soak’ vs. ‘moisten’; ‘dip into’ vs. ‘sink’, ‘submerge’) for both Chinese and Tai, suggesting the existence of an active morphological process in the proto language.

#### 2.4.3.2. ‘Chop’, ‘cut down’, ‘cut open’, ‘separate, divide’, ‘slash’

This word family prototypically takes a dental + V + t/n for both Chinese and Tai.

Gloss	Chrt	Pinyin	GSR	OC	MC	Thai	DH	LM	Yay	FS	HCT
cut fine	剽	tuán	231c	duan	duan	tɔɔn A1	tɔn A1	toon A1	toon A1	toon A1	*tɔn A1
length section	段	duàn	172a	duanh	duanh <sup>h</sup>	thɔɔn B2	tɔn B2	toon B2	ton B2	ton B2	*dɔn B2
cut off	斷	duàn	170a	duanh	duan	tɔɔn C1	tɔn A2	tan C1	ton C1	ton C1	*ton C1
prune	刪	duān	168e	tuan	tuan	–	taan A1	WT taan A1	taan A1	taan A1	*taan A1
short	短	duǎn	169a	tuanx	tuan <sup>ʔ</sup>	–	–	tin C1	tin C1	tin C1	*tin C1
cut off	剝	duò	295g	truat	twat	tat D1S	tat D1S	tat D1S	tat D1S	tat D1S	*tat D1

Similarly, we see alternations of initials and tones between verbs and nouns, and between action verbs (cut fine, section, cut off) and adjectival verbs (pruned, short). The variations in finals in Tai (ɔ ~ o ~ a ~ i) perhaps point to the effects of lexical diffusion.

#### 2.4.3.3. ‘Wide’ ~ ‘Vast’ ~ ‘Open’

This group of words takes a (labio-)velar initial plus final -aŋ for both Chinese and Tai.

Gloss	Chrt	PY	GSR	OC	MC	Thai	DH	LM	Yay	FS	HCT
wide	廣	guǎng	707h	kwangx	kwaŋ	kwaan C1	kwaan C1	kwaan C1	kuaan B1	kuaan B1	*kwaan C1
lie athwart	橫	héng	707m	gwang	ɣwaŋ	khwaan A1	xwaan A1	vaan B2	vaan A1	vaan A1	*khwaan A1
expanse of water	潢	huáng	707e	gwang	ɣwaŋ	waŋ A2	waŋ A2	waŋ A2	vaŋ A2	vaŋ A2	*waŋ A2
far apart	曠	kuàng	707o	khwang <sup>n</sup>	khwaŋ <sup>n</sup>	haan B1	haan B1	laan B1	luan B1	luan B1 <sup>13</sup>	(*xraan B1)
wide	廣	guǎng	707h	kwangx	kwaŋ	kwaan C1	kwaan C1	kwaan C1	kuaan B1	kuaan B1	*kwaan C1
lie athwart	橫	héng	707m	gwang	ɣwaŋ	khwaan A1	xwaan A1	vaan B2	vaan A1	vaan A1	*khwaan A1

The above examples again show regular correspondences in initials and tones between Chinese and Tai. More significantly still, orthography also plays an important role for the Chinese forms which are in the same *xiesheng* series sharing the phonetic element, 黃, which typically takes a labio-velar initial and a final -(a)ŋ. Examples of this kind supply solid evidence for shared morphological derivations between Chinese and Tai. Indeed such structural sharing cannot be easily explained as the results of contact or chance occurrence. If Li Fang-Kuei is right about the postulating morphological processes as an important criterion for genetic relationship, then the above examples supply strong evidence towards understanding the nature of the historical relationship between Chinese and Tai.

## 2.5. SUMMARY

Tai represents one aspect of the vast historical drama in ST. It is significant to note that the number of shared items between Chinese and Tai is far greater than between Tai and other languages in the surrounding regions such as Miao-Yao and Tibeto-Burman, indicating a close link between the two languages.

Tai and Chinese have been intermingling for centuries. Whatever the ultimate relationship between Chinese and Tai may turn out to be, Tai reflexes provide an invaluable source of information for the reconstruction of Old Chinese. They complement the vast sources of Chinese dialect data now available. They throw new light on our understanding of the linguistic situation in southern China and the surrounding areas.

Throughout this chapter, no attempts have been made to distinguish between loans from cognate words. Surely, no language is immune to loans. While some of the above-discussed comparanda will eventually prove to be loans, it would be unwise to attribute all of them to loan contact. It is sober to realize that it is often exceedingly difficult to distinguish early loans from inherited items. As Li has pointed out: 'We have as yet... no criteria to judge what are loans and what are not.' (1976: 48) A mechanism to separate loans from inherited words is yet to be worked out. More empirical work needs to be carried out before this issue can be resolved.

The following findings can be summarized for this chapter:

- (1) There is a significantly large number of shared vocabulary items between Chinese and Tai. Contrary to the claims by some earlier scholars, Chinese and Tai share extensively in basic vocabulary. Many of the shared vocabulary items are included in the Proto-Tai lexicon, which cannot be regarded entirely as loans;

13 The Yay form means 'village (larger than *baan*)' (Gedney 1991: 198); the Fengshan form means 'open space in a village; courtyard'.



- (2) Derivational morphology is a common feature between Chinese and Tai. It lends strong support to the assumption of a deep historical connection between Chinese and Tai;
- (3) The linguistic evidence presented above seems to suggest that the Sino-Tai relationship is a lot deeper than previously thought;
- (4) A number of the lexical items in Tai appear to have look-alikes in Austronesian, the nature of which remains to be further examined;
- (5) Despite the progress made, the nature of the historical relationship between Chinese and Kam-Tai still remains open.

## REFERENCES

- Aikhenvald, Alexandra Y. and Dixon, R. M. W. (eds) (2001) *Areal Diffusion and Genetic Inheritance*, Oxford University Press.
- Barlow, Geoffrey G. (2000a) *The Zhuang: Ethnogenesis*. Available on line at: <http://mcel.pacificu.edu/as/resources/zhuang/contents.html> (accessed 4 December 2006).
- Barlow, Geoffrey G. (2000b, updated 2005) *The Zhuang: A Longitudinal Study of Their History and Their Culture*, available at: <http://mcel.pacificu.edu/as/resources/zhuang/index.html> (accessed 4 December 2006).
- Baxter, William (1992) *Handbook of Old Chinese Phonology*, Berlin/New York: Mouton.
- Benedict, Paul K. (1942) 'Thai, Kadai, and Indonesian: A new alignment in Southeast Asia' *American Anthropologists* 44: 576-601.
- (1972) *Sino-Tibetan: A conspectus*, Cambridge: Cambridge University Press.
- (1975) *Austro-Thai Language and Culture*, New Haven: HRAF Press.
- (1976) 'Sino-Tibetan: Another Look', *Journal of the American Oriental Society* 96.2: 167-197.
- (1990) *Japanese/Austro-Thai*, Ann-Arbor: Karoma.
- Bodman, Nicholas (1980) 'Proto-Chinese and Sino-Tibetan: Data towards establishing the nature of the relationship', in Coetsem, Frans Van and Waugh, Linda R. (eds.) *Contributions to Historical Linguistics: Issues and Materials*, Leiden: E.J.Brill, 35-199.
- Conrady, August (1896) Eine indochinesische causativ-denominativ-Bildung und ihr Zusammenhang mit den Tonaccenten, Leipzig: Otto Harrassovitz.
- Dai, Qingxia, and Fu Ailan (1995) 关于汉藏语系语言的分类问题 [On the classification of Sino-Tibetan languages]. Seminar talk given at the meeting of the Chinese Society for Minority Languages on 'The Theories and Methodologies on Sino-Tibetan Studies'. Also available on line at: <http://www.wai-yu.com/read.php?tid=3474> (Accessed 10 March, 2007). (In Chinese).
- De Lacouperie, Terrien (1883) *Chinese and Siamese*. Academy, August.
- (1887) 'The Languages of China before the Chinese', *Transaction of the Philological Society*. (Reprinted 1969).
- Denglinger, Paul B. (1989) 'The Chinese-Tai linguistic relationship: a formal proof', *Monumenta Serica* 38: 167-171.
- Diller, A.V.N. (1998) 'The Tai language family and the comparative method', *Proceedings of the International Conference on Tai Studies*, 1-32, Bangkok: Mahidol University.
- Diller, A.V.N. (2003) 'Austroasiatic strata in Thai', in Henning Anderson (ed.) *Language Contacts in Prehistory: Studies in Stratigraphy*, Amsterdam: John Benjamins, 159-176.
- Dong, Weiguang, Cao Guangqu and Yan Xuequn (1984) 'Hanyu he dong tai yu de qinshu guanxi [The genetic relationship between Chinese and Kam-Tai]', *CAAAL* 22: 105-121.
- Egerod, Soren (1976) 'Benedict's Austro-Thai hypothesis and the traditional views on Sino-Tai relationship', *CAAAL* 6: 51-63.
- Gedney, William J. (1965) 'Review of J. Marvin Brown: *From ancient Thai to Modern Dialects*', *Social Science Review* (Bangkok) 32: 107-112.

- (1976) 'On the Thai evidence for Austro-Thai', *CAAAL* 6: 65-86.
- (1991) *William J. Gedney's The Yay Language: Glossary, Texts and Translations*, edited by Thomas John Hudak (Michigan Papers on South and Southeast Asian Studies 38), Center for South and Southeast Asian Studies, University of Michigan.
- (1993) *William J. Gedney's The Saek Language: Glossary, Texts and Translations*, edited by Thomas John Hudak (Michigan Papers on South and Southeast Asian Studies 41), Center for South and Southeast Asian Studies, University of Michigan.
- Gong Hwang-cherng (1993) 'Cong hanzang yu de bijiao kan hanyu shanggu liuyin yunwei de nice [The reconstruction of final liquid endings in Old Chinese as viewed from Sino-Tibetan comparisons]', in *Xizang yanjiu lunwen ji* 4 [*Collected Papers on Tibetan Studies*] Vol. 4, Taipei: Tibetan Studies Commission; also in *Hanzang yu yanjiu lunwen ji* [*Collected Papers on Sino-Tibetan Studies*] (Language and Linguistics Monograph Series C2-2), Taipei: Institute of Linguistics (Preparatory Office), Academia Sinica, 2002, 49-65.
- Gong, Qunhu (2002) *Hanyu Taiyu Guanxici De Shijian Cengci* [*A Study of Sino-Thai Chronological Strata*], Shanghai: Fudan University Press.
- Goodenough, Waugh (1975) Introduction, in Benedict 1975, iv- xi.
- Haudricourt, Andre-Georges (1956) De la restitution des initiales dans les langues monosyllabiques: le problème du thai commun. *Bulletin de la Société Linguistique de Paris* 52: 307-22.
- (1974) Daic, in Shafer, Robert. (ed.) (1974) *Introduction to Sino-Tibetan*, Part V, Wiesbaden: Otto Harrasowitz, 451-526.
- Karlgren, Bernhard (1933) 'Word families in Chinese', *Bulletin of Museum of Far Eastern Antiquities* 5: 5-120.
- (1957) Grammata Serica Recensa, *Bulletin of Museum of Far Eastern Antiquities* 29.
- Keightley, David N. (ed.) (1983) *The Origins of Chinese Civilization*, Berkeley/Los Angeles/London: University of California Press.
- Lan, Qingyuan (2003) *Zhuanghan tongyuanci jiecì yanjiu* [*A Study of Cognates and Loan Words between Chinese and Zhuang*]. Beijing: Central University for Nationalities Press.
- LaPolla, Randy (2001) 'The role of migration and language contact in the development of the Sino-Tibetan Language Family', in Dixon, R.M.W and Aikhenvald, A. Y. (eds.) *Areal Diffusion and Genetic Inheritance*, Oxford: Oxford University Press, 225-54.
- Li, Fang-Kuei (1938) 'Language and dialects', *China Yearbook*. (Reprinted in *Journal of Chinese Linguistics* 1973.1: 1-13).
- (1970) 'Some tonal irregularities in the Tai languages', in Jakobson, Roman and Kawamoto, Shigeo (eds.) *Studies in General and Oriental Linguistics*, Tokyo: TEC Corporation for Language and Educational Research, 415-422.
- (1971) '上古音研究 Shanggu yin yanjiu [Studies in Archaic Chinese]', *Tsinghua Journal of Chinese Studies*, n.s., 9:1-61. (Shanghu Yinshuguan edition 1980, Beijing: The Commercial Press.)
- (1976) 'Sino-Tai', *CAAAL* 3: 39-48.
- (1977) *A Handbook of Comparative Tai*. (Oceanic Linguistics Special Publications No. 15.) Honolulu: University Press of Hawaii.
- (1978) 'Siamese khot', *Monumenta Serica* 33: 403-406.
- Li Jingzhong (1995) *语言演变论 Yuyan yanbian lun* [*The evolution of language*], Guangzhou: Huachang chubanshe. (In Chinese)
- Luo, Meizhen (1992) 'A second discussion of the genetic classification of the Kam-Tai languages', in Compton, Carol J. and John F Hartmann (eds) *Papers on Tai Languages, Linguistics and Literatures in Honor of William J. Gedney on his 77<sup>th</sup> Birthday*. (Occasional Paper No. 16, Monograph Series on Southeast Asia, Center for Southeast Asian Studies, Northern Illinois University.
- (1994) 'San lun tai yu de xishu wenti [A third discussion of the genetic classification of the Kam-Tai languages]', *Minzu Yuwen* 6: 1-11.
- Luo Yongxian (1996) 'Word families in Tai: a preliminary account', *Pan-Asiatic Linguistics: Proceedings of the Fourth International Symposium on Languages and Linguistics*, Mahidol University, Thailand, Vol. 3, 850-882. (Also in Luo 1997a, Ch.6, 183-224).

- (1997a) *The Subgroup Structure of the Tai Languages: A historical-comparative study*. (Journal of Chinese Linguistics Monograph Series 12.) Berkeley: University of California.
- (1997b) On the Tai evidence for the \*sk- hypothesis. Paper presented at the 30th International Conference on Sino-Tibetan Languages and Linguistics, August 23-28, Beijing.
- (1998) 'Evidence for a series of sibilants clusters in Tai and the Sino-Tai relationship', *Journal of Chinese Linguistics* 26.1: 75-125. (Also appeared in as Chapter 4 in Luo 1997a).
- (2000) 'From 'Head' to 'Toe': Sino-Tai lexical correspondences in body part terms,' *Journal of Chinese Linguistics* 28.1: 68-99.
- (2001) Sino-Tai 'Eat', Paper presented at the Annual Meeting of the Southeast Asian Linguistic Society, Bangkok: Institute of Language and Culture for Rural Development, Mahidol University.
- (2002) Sino-Tai 'Cut', Paper presented at the 35<sup>th</sup> International Conference on Sino-Tibetan Languages and Linguistics, Tempe: Arizona State University. Also in Jimmy Harris and Songsome Brusphat (eds) (2007) *Studies in Tai and Southeast Asia Linguistics*. Bangkok: Ekphimthai Ltd, 155-182.
- (2003) Sino-Tai 'dike', Paper presented at the 36<sup>th</sup> International Conference on Sino-Tibetan Languages and Linguistics, La Trobe University, Melbourne.
- (2004) Sino-Tai Cognates for 'Vast/Open' and 'Hollow/Empty', Paper presented at the 37<sup>th</sup> International Conference on Sino-Tibetan Languages and Linguistics, Lund: Lund University, Sweden.
- (2005) Sino-Tai 'Soak/Moist/Submerge/Sink', Paper presented at the 38<sup>th</sup> International Conference on Sino-Tibetan Languages and Linguistics, Xiamen: Xiamen University, China.
- (2006a) Sino-Tai correspondence in Old Chinese \*-a- finals. Paper presented at the 39<sup>th</sup> International Conference on Sino-Tibetan Languages and Linguistics, Seattle: University of Washington, September 15-17.
- (2006b) Shanggu hanyu \*-l he \*-r: Dongtai yu zhengju [On the Tai evidence for Old Chinese \*-l and \*-r]. Paper presented at the International Conference on East Asia Languages and Linguistics, 24-27 December, Shanghai Normal University. Also to appear in the Conference Proceedings.
- (in reparation) Tai, Sino-Tai: the nature of the relationship.
- Manomaivibool, Prapin (1975) *Sino-Thai Lexical Correspondence*, PhD dissertation, University of Washington.
- Matisoff, James A. (1978) *Variational Semantics in Tibeto-Burman: The 'Organic Approach' to Linguistic Comparison*, Philadelphia: Institute for the Study of Human Issues.
- (1990) 'On megalocomparison', *Language* 66.1: 106-120.
- (1991) Sino-Tibetan Linguistics: present states and future prospects. *Annual Review of Anthropology* 20: 469-504.
- (2001) 'Prosodic diffusionality and Southeast Asia', in Aikhenvald and Dixon (eds), 291-327.
- (2003) *Handbook of Tibeto-Burman Phonology: System and Philosophy in Sino-Tibetan Reconstruction*. Berkeley/London: University of California Press.
- Meacham, William (1983) 'Origins and development of the Yüeh coastal neolithic: a microcosm of culture change on the mainland of East Asia', in David Keightly (ed.), 147-175.
- Mei, Tsu-lin (2002) Jianli you zhongguo tese de zhongguo yuyanxue [Building Chinese Linguistics with Chinese characteristics]. *Journal of Chinese Linguistics* 30.2: 211-240.
- Mei, Tsu-lin and Jerry Norman (1971) Shilun jige minbei fangyan zhong de laimu s- sheng zi [CL- > S- in some Northern Min dialects]. *Tsinghua Journal of Chinese Studies* 1 & 2: 95-106.
- Nie, Hongyin (2002) 'Shenceng duiying xianyi [Comments on 'deep correspondence']', *Minzu Yuwen* 1: 1-7.

- Nishida, Tatsuo (1960) 'Common Tai and Archaic Chinese', *Transactions of the Kansai University Institute of Oriental and Occidental Studies* 49: 1-15.
- Norman, Jerry, and Mei Tsulin (1976) 'The Austroasiatic in South China: some lexical evidence', *Monumenta Serica* 32: 274-301.
- Ostapirat, Weera (2005) 'Kra-Dai and Austronesian', in Laurent Sagart, Roger Blench and Alicia Sanchez-Mazas (eds), 107-131.
- Pan, Wuyun (1995) '关于支持澳泰理论的若干材料 Guanyu zhichi aotai lilun de ruogan cailiao [Evidence for Austro-Tai]', in William S-Y Wang (ed.), 113-144.
- (2002) 'On the strata of cognates in Sino-Thai corresponding lexical items', in Ting, Pang-Hsin and Anne O. Yue (eds), 27-55.
- (2007) 'Shanggu hanyu yuwen \*-l yu \*-r [Consonant Endings \*-l and \*-r in Ancient Chinese]', *Minzu Yuwen* 1: 1-11.
- Ratanakul, Suriya (1990) 'Tai peoples and their languages: a preliminary observation', in Pinit Ratanakul and U Kyaw Than (eds) *Development, Modernization, and Tradition in Southeast Asia: Lessons from Thailand*. Bangkok: Mahidol University, 49-111.
- Reid, Lawrence A (1988 [1984-85]) Benedict's Austro-Tai hypothesis: An evaluation. *Asian Perspectives* 26.1: 19-34.
- Reid, Lawrence A (2001) Austric: Is it a real language family or not? Seminar talk, Centre for Research on Language Change, Australian National University.
- Sagart, Laurent (1993) 'Chinese and Austronesian: evidence for a genetic relationship', *Oceanic Linguistics* 21.2: 1-62.
- (2004) 'The higher phylogeny of Austronesian and the position of Tai-Kadai', *Oceanic Linguistics* 43.2: 411-444.
- (2005a) 'Sino-Tibetan-Austronesian', in Laurent Sagart, Roger Blench and Alicia Sanchez-Mazas (eds), 161-176.
- (2005b) 'Tai-Kadai as a subgroup of Austronesian', in Laurent Sagart, Roger Blench and Alicia Sanchez-Mazas (eds), 177-181.
- Sagart, Laurent, Roger Blench and Alicia Sanchez-Mazas. (eds) (2005) *The Peopling of East Asia*. London/New York: Routledge/Curzon.
- Thurgood, G (1985) 'Benedict's work: past and present', in Thurgood, G., James Matisoff and David Bradley (eds) *Linguistics of the Sino-Tibetan Area: the State of the Art*, (Pacific Linguistics C 87) Canberra: Australian National University, pp. 1-15.
- (1994) Tai-Kadai and Austronesian: The nature of the historical relationship. *Oceanic Linguistics* 33.2: 345-368.
- Ting, Pang-Hsin (2000) '汉藏系语言研究法的检讨 Hanzangxi yuyan yanjiufa de jiantao [A critique of methodologies in Sino-Tibetan linguistic studies]', *Zhongguo Yuwen* 6: 483-489. Also in Ting, Panghsin and Sun Hongkai (eds) *Hanzang yu tongyuan ci yanjiu: Hanzangyu jianjiu de lishi huigu [Cognate Words in Sino-Tibetan Languages]* (vol. 1). Nanning: Guangxi Minzu Chubanshe, 427-442.
- (2002) '论汉语与台语的关系 Lun hanyu yu taiyu de guanxi — Li Fang-Kuei hantai yu tongyuan lun de jiantao [The relationship between Chinese and Tai: a review of Li Fang-Kuei's theory that Chinese and Tai are genetically related]', in Ting, Pang-Hsin and Yue, Anne O. (eds), 9-25.
- Ting, Pang-Hsin and Yue, Anne O. (eds) (2002) *Essays in Chinese Historical Linguistics: Festschrift in memory of Professor Fang-Kuei Li on his centennial birthday*, (Language and Linguistics Monograph Series No. W-2), Institute of Linguistics, Academia Sinica and University of Washington.
- Wang, Li (1985) 汉语语音史 *Hanyu yuyin shi [Chinese Historical Phonology]*, Beijing: Chinese Academy of Social Science Press.
- Wang, Wenguang (1999) 中国南方民族史 *Zhongguo nanfang minzu shi [A history of minority peoples in South China]*, Beijing: Minzu chubanshe.
- Wang, William S-Y (ed.) (1995) *The Ancestry of the Chinese Language* (Journal of Chinese Linguistics Monograph Series 8), Berkeley: University of California Press.

- Wulff, Kurt (1934) *Chinesisch und Tai: Sprachvergleichende Untersuchungen [Chinese and Tai: Comparative linguistic studies]*, Copenhagen: Levin & Munksgaard.
- Wyatt, David (1984) *Thailand: A Short History*, Yale University Press.
- Xing, Gongwan (1962) ‘论调类在汉台语比较上的重要性 Lun diaolei zai hantaiyu bijiao shang de zhongyaoxing [The importance of tones in comparative studies of Chinese and Tai]’, *Zhongguo Yuwen* 1: 15-27.
- (1984) ‘汉藏系语言及其民族史前情况试析 Hanzangxi yuyan ji qi minzu shiqian qingkuang shixi [A preliminary study of the languages of the Sino-Tibetan family and their prehistory]’, *Yuyan Yanjiu* 2: 159-172.
- (1999) 汉台语比较手册 *Han tai yu bijiao shouce [A Handbook of Comparative Sino-Tai]*, Beijing: The Commercial Press.
- Zeng, Xiaoyu (2004) 汉语水语关系论 *Hanyu shuiyu guanxi lun [On the relationship between Chinese and the Sui language]*, Beijing: The Commercial Press. (In Chinese)
- Zhengzhang, Shangfang (1995) ‘汉语与亲属语同源根词及附缀成分比较上的择对问题 Hanyu yu qinshu yu tongyuan gen ci ji fuzhui chengfen bijiao shang de ze dui wenti [The issue of comparison of roots and affixes between Chinese and genetically related languages]’, in William S-Y Wang (ed.), 269-282.
- (2004) 上古汉语音系 *Shanggu hanyu yinxi [Old Chinese Phonology]*, Shanghai: Shanghai Educational Press.

## **PART 2**

---

# **TAI LANGUAGES: OVERVIEWS AND RESOURCES**

---

*This page intentionally left blank*

## CHAPTER THREE

# RESOURCES FOR THAI LANGUAGE RESEARCH

*Anthony Diller*

### 3.1. INTRODUCTION

#### 3.1.1. Scope

Thai is spoken, at least as a second variety, by well over half of the total of 80 or 90 million speakers of Tai-Kadai languages. In some respects it reflects features of the greater grouping as a whole, but in other ways it is exceptional. Thai is also by far the most thoroughly described member of the group, with accounts going back several centuries. The purpose of this chapter is to call attention to some five hundred studies of Thai grammar and other aspects of the language. This is no means a comprehensive linguistic bibliography, which would need several times as many entries. Studies are selected here because they are *representative* of ongoing research and because they are useful in providing readers with further bibliography. Some attention is also given to how Thai linguistic research and its subfields have developed historically, including how studies cited relate to broader professional background issues, which also may shift diachronically.

*Omitted below are many studies of merit*, especially those in languages other than English. In particular, books, theses and journal articles written in Thai have generally not been included even though they contain innovative and revealing linguistic research on the language. Those seriously interested in Thai and in other Tai-Kadai languages will surely need to spend time in Thai university libraries and elsewhere where these illuminating materials are accessible.

The main focus here is on Thai; only a sample of work on Tai, Kam-Tai and Tai-Kadai is included. For further references on the wider family at its different levels, see other chapters in this volume and Huffman (1986a); anthologies edited by Gething, et al. (1976); by Khunying Suriya Ratanakul, et al. (1985, 1998); by Edmondson and Solnit (1988, 1997); and works of Morev (1991); Edmondson (2007); Luo (1997, 2007); Diller (2000); Ostapirat (2000, 2004); Thurgood (2007a). Only selected earlier works of Li, Haudricourt, Gedney, and other pioneers in the field are mentioned here; see Huffman (1986a) for fuller listings.

Thai authors are cited below by family (last) name, with some reluctance. Apologies are due to those who prefer given (first) name citation. As a statistical tendency, last-name citation has been the majority practice among Thai scholars writing for an international linguistics audience, so that is followed here. Such practice is also the norm in scientific and medical writing. Beware however that this contrasts with the humanities where Thai authorities are frequently cited and indexed by given name, a format also followed by many libraries, by Huffman (1986a), and in some other reference works.

Where relevant, several works are mentioned in more than one subsection below for convenience of readers with specific interests. Apologies to readers who find this repetition tedious.

#### 3.1.2. Bibliographies and anthologies

The comprehensive indexical bibliography of Huffman (1986a) includes many works not mentioned here. Bibliographies of Kullavanijaya and Vongvipanond (1984) and of Burusphat (2002) are also of utility.

Many useful papers appear in anthologies and conference proceedings. Of great value and



convenience to Thai linguists are collections of papers presented to conferences of the Southeast Linguistics Society (SEALS) and to symposia in the Pan-Asiatic Linguistics series (PAL; see Luksaneeyanawin, et al., 1992; Premsrirat, et al., 1996) and similar symposia (Bamroongraks, et al., 1988). Other anthologies with a strong Thai grammar focus are edited by Bickner, et al. (1986); Abramson (1997); Tingsabadh and Abramson (2001); Harris, Burusphat and Harris (2007). For earlier anthologies, see Huffman (1986a). There is also treatment of Thai in sources where focus is more broadly Tai-Kadai or Southeast Asian (e.g. Ratanakul, Thomas and Premsrirat 1985).

Ongoing resources to keep in mind are the journals *Mon-Khmer Studies*, *Crossroads*, *Linguistics of the Tibeto-Burman Area*, *Journal of the Siam Society* and *Tai Culture*, as well as publications shown on the website of *Pacific Linguistics*. Journals produced in the Thai university context publish linguistic studies of quality in Thai and occasionally in English, among them: *Journal of Language and Linguistics*, *Science of Language Papers*, *Journal of Language and Culture*.

### 3.1.3. Grammars and overviews

A cogent reason to reign in representation of Thai in this volume is the recent publication by Iwasaki and Ingkaphirom (2005) of a splendid reference grammar of the language. Questions linguists may have about the language's structures and functions will be answered and discussed in this comprehensive and well-indexed volume, with its many examples taken from naturally-occurring speech.

To place Thai grammar in its wider Tai context, I am aware of only one extensive comparative grammar of the Tai languages: the Russian work of Morev (1991). This is a work of insight, fine scholarship and includes an impressive bibliography of Russian sources. Perhaps subsequent research, including that reported in this volume, will stimulate updated comparative Tai or Tai-Kadai grammars.

The earlier reference grammar of Noss (1964), a standard for decades, retains its value with excellent examples and a good index, but today's readers may need some patience in matters of terminology and orientation. Panupong (1970), developing an initiating/non-initiating distinction, presents an impressive study both of sentence-level syntax and of inter-sentence relations. Her study remains an important milestone for those pursuing structurally-defined relationships. In a rather different linguistic tradition, the Thai grammar (in Russian) of Morev, Plam and Fomicheva (1961) is an earlier landmark analysis also worthy of study. Among grammars written in Thai, my favorite remains Bandhumedha (1979), full of fresh ideas.

For the general reader more comfortable with categories of traditional grammar, a range of pedagogical grammars is available, among which Smyth (2002) is particularly informative, clear and dependable. For a more philological, historical and anthropological account of Thai, consult Anuman Rajadon (1961, 1981).

For overviews, Haas (1969b) and Li (1974) present concise encyclopedic articles of admirable compactness and lucidity. For more specialized summaries, see Gedney (1967), reviewing Thai research up to that date. More recent overviews, covering the basics of phonology and syntax, include those of Hudak (1987), Bickner (1994), and Diller (2004). Premsrirat (2006) presents a useful sociolinguistic summary.

## 3.2. PHONETICS, PHONOLOGY AND ORTHOGRAPHY

### 3.2.1. Distinctive units

The consensus of many authorities would admit the following distinctive (phonemic) units in Thai, although transcription systems vary and individual scholars might have alternative views. Prevoiced and lax stops: b, d; voiceless, unaspirated and tense stops: p, t, c, k ('tense' here perhaps implying also a simultaneous glottal and oral release; c being alveopalatal and affricated); voiceless aspirated stops: ph, th, ch, kh (ch affricated); nasals: m, n, ng [ŋ]; fricatives and aspirates: f, s, h; liquids: r, l (distinguished rather tenuously, with sociolinguistic

tendencies towards merger and overcorrection); semivowels w, y [j]. The majority, but not all, scholars in the field admit glottal stop to the consonant inventory but some opt for an analysis in which it becomes predictable; it is generally not shown in the transcription used here. Some consonants show a range of articulations, e.g. as sociolinguistically conditioned. Several initial clusters with second component -r, -l, or -w are permitted but are often simplified in less educated or less careful colloquial speech.

Only voiceless unaspirated stops -p-, -t-, -k-, glottal stop, nasals and semivowels occur in final position. In some systems, final semivowel equivalents are indicated by vocalic diphthongs: [kay<sup>11</sup>] ‘chicken’ may appear as [kaj<sup>11</sup>] or [kai<sup>11</sup>]. The final stops are unreleased, lacking the tense quality of initial p-, t-, k-. The widely-used Haas transcription represents stop finals as -b, -d, -g, but Abramson’s (1972) instrumental studies do not show voicing. Loanwords increasingly introduce finals like -s, -f, etc.

For vowels, nine come in short-long pairs: i i:, u’ u’: (high back unrounded [u]); u u:, e e:, oe, oe: (mid central or slightly back unrounded [ə] or [ɔ]), o, o:, ae, ae:, a, a:, and o’ o’: (low back unrounded [ɔ]); three diphthongs ia, u’a and ua occur as phonologically long, with short variants found in a few exclamations or in other marginal vocabulary items. Long vowels can occur finally, as can short vowels plus glottal stop.

The five tones are usually described as mid (33), low (11), falling (52), high (45, 55 or 454) and rising (24). For most speakers, the high tone includes glottal constriction, more salient when vowel is long. (Tone is indicated here in superscripts of these paired numbers, in most cases with lexical rather than phonetic values, e.g. X<sup>45</sup> represents a contrastive high tone, regardless of its real pitch characteristics; unmarked syllables have so-called neutral tone.)

The preceding phonemic inventory is shown in a romanization differing only in minor respects from the semi-official system prescribed by the Thai Royal Institute. Other systems are encountered, but most approach a one-to-one correspondence with the semi-official system sketched above, including that of Haas (1964, 1969b). Workers in the field soon become used to variation in transcription systems, often merely a matter of symbols used but occasionally indicating more substantial claims as to phonetic detail. More on the phonetic nature of these units and how they are represented follows below.

For general discussion of the Thai inventory, see Tingsabadh and Abramson (1993a, b) and the recent grammars and encyclopedic reviews mentioned above.

Of historical interest is Bradley (1911), apparently the first instrumental analysis of the five Thai phonemic tones, although tones were noted much earlier: La Loubère (1691) counted six. Abramson (1962) marks the professional dawn of modern instrumental research in Thai acoustic phonetics and work of the highest quality by Abramson and colleagues has extended over nearly half a century.

### 3.2.2. Phonological approaches and overviews

For an overview of studies of Thai phonology up to the mid 1980s, see Rischel (1984). Over several decades, work of L-Thongkum, Luksaneeyanawin, Sutadarat, and other Thai colleagues has contributed significantly to progress in Thai phonetics and phonology and those researching these topics would surely need to attend to the full range of their work. Original phonological approaches are developed in these publications, e.g. Luksaneeyawin (1992) explains ‘three-dimensional phonology’. Note also Erikson’s and Gandour’s phonological contributions. Gandour’s work, together with colleagues, spans several decades and constitutes the principal body of neurolinguistic research analyzing how Thai language ability is affected under aphasic or other degenerative neurological conditions.

Redundancy rules constrain tone by patterns of vowel length and whether or not a syllable ends in a stop (referred to as ‘closed’ or ‘open’, or in more picturesque Thai-derived terms as ‘dead’ or ‘alive’). Closed syllables occur only with low, falling and high tones. There are also some sandhi-like rules, e.g. tones shift from the values above in fast speech, with some rising tones becoming high; the long vowels in first syllables of certain compounds are shortened. Example: [kha:ng<sup>52</sup>-lang<sup>24</sup>] ‘behind’, spelling pronunciation; cp. [khang<sup>5</sup>-lang<sup>24</sup>], normal speech. Such rules and debate over some points in the preceding summary are considered in

the classic study of Henderson (1949) and in Lodge (1986); also in work of many authorities noted below.

### 3.2.3. Consonants, vowels and tones

Work of Jimmy G. Harris (e.g. 1972, 1987, 2007) in articulatory phonetics analyzes specifics of the Thai sound system. Exact places and manners of articulation, e.g. of the units transcribed here as [d], [t], [th], [c] and [ch], are established through palatography. Harris provides useful comparisons with other languages and shows how identification of Thai phonetic phenomena has been oversimplified. In an earlier study, Brown (1962, 1965) also makes challenging observations regarding Thai consonant articulations, emphasizing complex articulations and their tonal effects.

Vowels and questions of vowel length are studied from various perspectives in works by Brown, L-Thongkum, Hartmann, Roengpitya, Svastikula and Tumtavitikul. Onsuwan (2000) inspects the stop/vowel interface. Abramson (2001) considers the stability of distinctive vowel length. For pharyngealization, see Henderson (1987); for nasalization, see Beddor, et al., (1999); for states of the glottis such as prephonation and unprephonation, see work of J.G. Harris. Diachronic studies of vowel development are mentioned in 4.5.

Tone and in particular how tone interacts with other phonological elements has stimulated much research in Thai phonetics and phonology. While most authorities concur that citation forms in Thai show five lexical tones, they are also aware that citation values shift in various contexts, leading to different phonological perspectives. In an early study, Leben (1971) argued for a segmental approach to Thai tone. Yip (1982) counters this with a laryngeal tier analysis. An autosegmental approach is outlined by Hoonchamlong (1990). A general overview of Thai tonal issues is given by Erickson (1976), emphasising physiology and by Intrasai (2001), emphasising acoustics; see also Gandour (1976); Robertson (1982). Tone and vowel length are considered by Gandour (1977, 1984), and in a wider comparative context by L-Thongkum and Teeranon and Intajamornrak (2007). Tumtavitikul (1993) studies how consonant onset affects tonal parameters; see also L-Thongkum (1992). Gandour, Potisuk and Harper (1996) discuss stress and vowel length. For tonogenesis, see [section 3.4.6](#).

Questions of pitch/amplitude components of tone and stress and of the interaction of tone, rhythm, sandhi and sentence intonation have been vigorously pursued for decades. A procession of representative studies includes Kroll (1956), Warotamasikkhadit (1968), Whitaker (1969), Hiranburana (1971), Gsell (1972), Noss (1972, 1975), Pantupong (1973), L-Thongkum (1978, 1984), Luksaneeyanawin (1983), Court (1985), Peyasantiwong (1986) and Wong-opasi (1994).

Musical recitation raises interesting research questions. Still of value is List's (1961) study of linguistic tone and song melody. Tumtavitikul and Promkhuntong (2007) present results of an instrumental study of how Thai classical poetry is chanted.

Closely related is how tone, vowel-length and other properties are manifested in speech-act particles, interjections and the like. Perhaps belonging here too is Cooke's (1992) discussion regarding a possible sixth tone in Thai. Do these items show phonological properties somewhat different from other lexical material? Chuenkongchoo (1956) is among the earliest studies; Peyasaniwong (1979, 1981, 1986) further develops the analysis; also Lodge (1986). For conjunctions and linker syllables, consult Bee (1975). Chaimanee's (1994) study of filled hesitation pauses breaks new ground in a related area.

### 3.2.4. Psycholinguistics and phonology; phonesthesia

Psycholinguistic investigations of tone include work of Abramson (1971) on whispered Thai and of Van Lanker and Fromkin (1978), who report different neural processing for contrastive speech tone and non-lexical musical tone. Psycholinguistic implications of the range publications of Gandour and colleagues deserve attention.

Perception studies using experimental protocols have often involved comparative as well as psycholinguistic issues. Wayland and Guion (2003) investigate tonal perception among native and non-native speakers. For consonant perception, including voicing issues, see the

early contrastive studies involving Thai and English of Melamed (1962) and Kanasut-Roengpitya (1965). Other perception studies are by Donald (1978); Carney, et al. (1988). Beach, Burnham, and Kitamura (2001) investigate bilingualism and Thai bilabial stop production and perception. Harris and Bachman (1976) study how Thai speakers perceive consonants in other Tai languages.

General longitudinal and acquisitional studies include Tuaycharoen (1977, 1979) and Imsri and Idsardi (2003). Bilingual acquisition is the topic of Sarawit (1976).

Direct sound-meaning relationships, including sound symbolism, phonesthesia and quasi-morphological or morphophonemic functions of phonological material are the topic of studies by Henderson (1965), Kam (1980) and Chamberlain (1992). These processes seem to be of more importance for Thai than for English, etc., and merit further study.

### 3.2.5. Orthography and Romanization

For the Thai writing system, Haas (1956), the first complete description in English, remains a useful resource. Danvivathana (1987) presents the system in all of its detail along with historical treatment of how letter shapes and inventories have altered over the centuries. For those simply seeking a practical introduction to Thai orthography, Iwasaki and Ingkapirom (2005) and Smyth (2002) can be recommended, and many pedagogical manuals, CD-ROMs, etc., are available. The encyclopedia entries cited above also contain overviews.

A compact summary of current symbols and basic principles, along with Lao comparisons, can also be found in Diller (1996a). This study includes a historical sound-change rationale for what may seem today like arbitrary and complex system of rules. For more on orthography and diachronic sound change, see [section 3.4.6](#).

Given that the Thai writing system is not only practically efficient but also a longstanding component in the Thai sense of national identity, alternatives have never posed a real threat (Anuman Rajadon 1961, 1981; Aeosrivongse 1984; Diller 1993, 2002). True, Kings Rama IV and Rama VI each came up with a reformed writing system and a change to romanization was tentatively considered just after political changes of 1932, yet these were never popular options. A few putative simplifications in Thai spelling decreed during the regime of the 1940's were quickly undone and forgotten (Hudak 1986).

King Rama VI (Vajiravudh, r. 1910-1925) proposed different transcription systems for native words and Indic loans in Thai, his own name, pronounced [wachi<sup>45</sup>rawut<sup>45</sup>], providing a good example of the latter type. Thai family names are often romanized this way: through transliterating etymological Sanskrit letter values rather than indicating modern sounds. Owners of Indic-component Thai names of this category have reported to me hesitancy to change romanization for a surname that was royally granted both in Thai form and also as romanized in the King's etymological-Sanskritic transcription system. Inconsistent application of this two-fold system also accounts for frequently seen toponym pairs like *Dhonburi*, *Thonburi*; *Ubol*, *Ubon*, etc. (Ronakiat 2007). For more on romanization and transcriptions, including disparagement of the Haas system and apparently phonetics in general, see Prince Dhaninivat (1970).

## 3.3. SYNTAX AND SEMANTICS

### 3.3.1. Syntactic typology

Many, but far from all, authorities in the field would concur in a general way with the following first-approximation sketch. Most (including the writer) would also offer qualifications, counter-examples, definitional quandaries and further debate, as is clear from following entries. As to commonly-cited typological parameters, Thai characteristics include basic transitive syntactic order [S + V-trans + O], in more semantic frameworks represented as AVO, and most frequent intransitive order [S + V-intrans]. [V-intrans + S] occurs also occurs in presentational or existential contexts as mentioned below. Understood noun phrases are widely unstated and construed (i.e. are subject to zero anaphora or deletion, etc., depending on analytic framework) and topicalized orders are common. Taken together, these factors give

rise to a number of alternative pragmatically- or functionally-determined surface orders in actual discourse.

Nominal modification order in Thai strongly accords with the inherited Tai pattern [nominal head] + [modifier], with possessives and relative clauses also following head nouns. Interestingly, Indic compounds have introduced a contrasting [modifier] + [nominal head] ordering. Some commercial and institutional NP names use the Indic order: compare *Chulalongkorn-mahawithayalay* (Indic compound) and *Mahawithayalay Thammasat* (as though Tai/Thai noun phrase, although components are both Indic); *Ao-Thai Gas* [a:w<sup>11</sup> thay<sup>33</sup> kae:s<sup>45</sup>], ‘Gulf-of-Thailand Gas’, Indic compounding order for the whole NP, but components are etymologically Thai and English; the ‘Gulf-of-Thailand’ subcomponent shows Thai/Tai head-modifier order. (This issue is now strictly academic, given Caltex’s acquisition.)

Prepositions precede their nominals. Many auxiliaries precede their main verb, but others follow. Positioning of lexical items with semantic functions of auxiliaries in English, or at least with translational similarities, is not clear-cut. Sometimes reinforcing correlative auxiliary components are found on both sides of a verb complex, e.g. the progressive aspect sequence [kam<sup>33</sup> lang<sup>33</sup>] X [yu:<sup>11</sup>] ‘to be in process of Xing’ (Kullavanijaya and Bisang, 2004); the deontic sequence [sa:<sup>24</sup> ma:t<sup>52</sup>] X [day<sup>52</sup>] ‘to be able to X’.

Morphosyntactically, as a tonal language with many typically isolating features, Thai retains a core of basic monosyllabic Tai words. However, in the current lexicon this inherited base is statistically overwhelmed by vocabulary from other sources, often polysyllabic with tonally attenuated or perhaps ‘neutral’ reduced syllables. Some prefixal syllables show at least a weak semantic content: [ma-] codes a set of fruit-bearing flora such as [ma-muang<sup>52</sup>] ‘mango’ and [ma-phra:w<sup>45</sup>] ‘coconut’. This is transparently compound reduction: [ma:k<sup>11</sup>], now ‘areca nut’, originally had a wider ‘fruit’ meaning. Compounding of several types is active and common; some sandhi phenomena apply to compounding. Moderate use is made of full and partial repetition. Derivational processes of several types are mentioned below but no use whatsoever is made of obligatory inflectional morphology to indicate tense, aspect, transitivity, specificity or number. These either are coded lexically, understood from context, or left vague.

Nominal word classes accepted by most scholars are common and proper nouns, including a copious supply of titles and epithets, shading into pronouns; also in the nominal class or else in classes of their own are classifiers, number words (i.e. lexical numerals), and deictics. As modifiers follow head nouns prolifically and endomorphically, rather large noun phrases can be built up. A few nouns are homonymous with verbs, mainly instrumentals as in other languages: [thay<sup>24</sup>] ‘a plow; to plow’; [prae:ng<sup>33</sup>] ‘a brush; to brush’.

Open verbal classes are more controversial. For the protracted debate regarding adjectives and/or adjectival verbs, see [section 3.3.6](#). Progressive grammaticalization is at the bottom of several definitional quandaries: this process not only moves full verbs into preverbal and postverbal auxiliary subclasses, but also into preposition-like covers marking semantic case for following nominals. These construction types impinge on wider *serial verb* patterns (3.6). Closed functional classes widely recognized include prepositions (but challenged by Warotamasikkhadit 1988; see also Indrambarya 1994), conjunctions, intensifiers and quantifiers, speech-act and polite particles and interjections. Whether or not one or more additional adverb classes might be needed depends on how generously one defines other classes or on which tests are applied.

Syntactic overviews are presented in the encyclopedic articles mentioned above, with issues developed at greater length in other sources cited. For those interested in how syntactic constructions function in actual speech, an excellent place to go first is Iwasaki and Ingkaphirom (2005). Mention should be made also of many fine theses and other studies in Thai not considered here that describe individual constructions with great insight.

### 3.3.2. History of Thai syntactic research

Historically, interest in Thai syntactic issues by outsiders is of long standing and in varied frameworks. Progress in understanding the nature of Thai can be traced from the brief but

valuable sketch and translated lexical lists of La Loubère (1691), a diplomat-trader, through to the first explicit grammar in English by James Low (1828), an East India Company trader-cartographer. He began by informing readers that Thai has ‘no grammar’ (meaning inflectional morphology?) but then went on to treat parts of speech with understanding. His book is a technical as well as linguistic milestone: it was printed using the first moveable-type Thai printing-press font, developed by Ann Judson and her missionary colleagues (Winship 1986). Unfortunately the text was replete with myriad typographic errors. Noteworthy here is Smyth (2001), who outlines the early study of Thai by outsiders and provides a useful bibliography and discussion of the early grammars.

A substantial grammar was produced by another missionary, Bishop D. J. B. Pallegoix (1850). He projected the full panoply of Latinate categories onto Thai, including even the future perfect tense. Pallegoix’s Latin-Siamese-French-English lexicon, with Thai script and a credible romanization, was an impressive *tour de force*. More importantly, as I suspect, Pallegoix’s scholarly friendship and intellectual exchanges with the Buddhist monk Prince Mongkut were influential in promoting the sentiments leading to a vision of Thai as the standard normative language such as we know it today (Diller 2001a). After Mongkut’s coronation in 1851 he was subsequently known as King Rama IV. He turned his attention to language reform and to ‘correct’ Thai, apparently the first time a monarch had pursued normative linguistic interests at that level of detail. He focused not only to lexical issues but even on syntactic minutiae normally of concern only to linguists, such as prescription of different classifier constructions (elephants and horses were to be counted directly, without an idiosyncratic classifier [tua<sup>33</sup>] ‘lit. ‘body’, since these animals were considered higher in dignity than others). Khanittanan (1987a, b) considers linguistic features in royal writing as they evolved, complemented by the more socio-historical study of Aesrivongse (1984, in Thai), whose insights are behind Diller (1993, 2001a).

Pedagogical grammars and readers, too many to enumerate here, contain insightful discussions of specific constructions and exemplify them in context as well. Especially dependable earlier sources informed by linguistics include Haas (1945); Brown (1967); Anthony, French and Warotamasakkhadit (1967); Jones, Mendiones and Reynolds (1969); Yates and Tryon (1970); Kuo (1982). Landmark syntactic analyses written by those with native-speaker intuitions and reasonably comprehensive in scope include the formidable, influential and abstruse normative grammar of Upakit-Silapasan (1939, in Thai); Warotamasakkhadit (1963), the first grammar written in a generative framework; Panupong’s (1970) inter-sentence study mentioned above; Pankhuenkhat (1978) and Bandhamedha (1979, in Thai), well-organized and full of insights.

### 3.3.3. Theoretical perspectives in syntax

Approaches to Thai syntax have included most of the familiar linguistic frameworks of the mid- to late-20<sup>th</sup>-century: traditional, structuralist, tagmemic, generative, functional and more discourse-oriented studies. Upakit-Silapasan (1939, in Thai), mentioned above, presents essentially a traditional analysis employing familiar Western Classical grammatical categories, but disguised by neo-Indic nomenclature and making use of several authentically Indic notions, especially *karaka* (semantic deep case) theory. In methodology, Noss (1964) and Panupong (1970) pursue the substitution-frame methodology standard for structural linguistics of their period and are both convincing examples of that approach. As noted above, generative work begins with Warotamasakkhadit (1963) and continues on, as that framework advances, through representative post-Aspects productions such as Bandhamedha (1976) and Surintramont (1979) on deletion, Warotamasakkhadit (1979) on topicalization, Rodman (1977), on coordinating constraints and Wongbiasaj (1980), on movement transformations. Bounding and subadjacency are covered by Panpothong (2001). For generative semantics, Stine (1968), on the instrumental case, is a good early example.

Deletion or nominal omission has frequently been a focus of theoretical treatment. As noted above, in terms of the patterns S + Vintrans and S + Vtrans + O, noun phrases S and O are often construed from context. Nominal omission applies to compounding processes as

well. In some conventional expressions, head nouns in compounds are optionally or regularly missing: [kho<sup>24</sup>-tha:n<sup>33</sup>] ‘request alms’, but also ‘beggar’ with head noun [khon<sup>33</sup>] ‘person’ missing. Note also noodle, rice dishes and other culinary productions, e.g. [phat<sup>11</sup>-thay<sup>33</sup>] ‘Thai-style fried (noodles)’; [phat<sup>11</sup>-kaphraw<sup>33</sup> mu:<sup>24</sup>] ‘(rice topped with) basil-fried pork’, possibly giving rise by analogical formation to [maek<sup>45</sup>-kaphraw<sup>33</sup> mu:<sup>24</sup>] ‘Mc(Donald’s-style bunned patty with) basil pork’. However, after true prepositions, which are few in the language, nominals resist deletion. Verb-derived coverbs sometimes also reject deletion of following nominal, sometimes not, providing a test for degree of grammaticalization. There are important constraints too, some sentence-internal and some sensitive to macro issues of discourse organization. These are considered with particular insight and cogency by Grima (1978, 1986). For taking missing nominals as empty categories as they were construed in the frameworks of the 1980’s, see Lehman and Pingkarawat (1985), Cole (1987), Kobsiriphat (1988) and Pingkarawat (1989). In the post-Government- and-Binding framework, see Hoonchamlong (1991) and Laksinaking (1991) on anaphora. Outsiders may be astonished that grammaticality judgments underpinning generative work of this type are not infrequently contested by native speakers themselves, a topic examined ethnolinguistically by Diller and Khanittanan (2002).

Numerous additional frameworks have been used. Under the institutional aegis of the University of Hawaii’s East-West Center, work making reference to case grammar and to the specific format ‘lexicase’ (see Starosta 2001), includes Kullavanijaya (1974), on word classes; Savetamalya (2001); Clark (1978), with comparisons to Vietnamese coverbs; and Clark and Prasithrathsint (1985), on synchronic lexical derivation. Deep case, that is semantic role relationships rather than those derived from surface syntax, informs the approach of Lekawatana (1970) and also of Vijchulata (1978), who also develops a stratificational analysis. Tagmemic work includes Patamapongse (1971); Phinit-Akson (1972); Punyodyana (1976). Deepadung (1989) exemplifies residential grammar. For Montague grammar, consult Godden (1981). Diller (1997) wonders whether ‘subject’ is needed in Thai syntax, referring to role and reference grammar. Schiller (1992) produces an autolexical analysis, a framework also used in Wong-Opasi’s (1994) treatment of compounding.

### 3.3.4. Compounding, nominalization and morphological processes

Compounding and questions of complex word formation are treated by Gehr (1951); Fasold (1968); Osipov (1969); Warotamasintop (1975); Vongvipanond (1992); Witayasakpan (1992); Kullavanijaya (1992); Manomaivibool (2000). L-Thongkum (1994) presents a comparative Tai-Kadai study. Vongvipanond (1992) treats doublets with components of related meaning, a type widespread in the Thai lexicon. See also Wong-Opasi (1994), noted above. The astonishingly large set of compounds relating to emotions and personal attributes with component [cay<sup>33</sup>] ‘heart, mind’ have attracted much syntactic, semantic and anthropological interest; see the comparative Southeast Asian survey of Matisoff (1986). For Thai data analysed in various ways, see Lee (1987); Diller and Juntanamalaga (1990); Moore (1992). Another common compound type has a component meaning ‘head’, which Juntanamalaga (1992) convincingly relates to kinesic tabu beliefs relating to heads and feet. Thai nominalization is similar to compounding in some respects. Prasithrathsint (1996), (1997), (2006), (2007) presents a compelling set of diachronic and comparative analyses covering this topic.

Some compounding heads referring to people such as [nak<sup>45</sup>], [cha:w<sup>33</sup>] and [phu:<sup>52</sup>] are in effect bound morphemes. [khon<sup>33</sup>], another ‘person’ word, functions as a common noun but also as a generic compounding head and as classifier. Other classifiers do so as well.

Quasi-morphological derivational processes involving vowel ablaut, tonal variation and even some consonantal interchange is sometimes referred to as elaboration and can be found throughout Southeast Asia, surveyed by Henderson (1967), Nacaskul (1976) and Williams (1991). For Thai, Haas (1964) provides many further examples. An early brief analysis of this material along with more straight-forward reduplication is given by Haas (1942). Sookgasem (1997) provides a valuable expansion of reduplication types with discussion of theoretical

ramifications. Various patterns of semi-repetition (elaboration or echo-syllables) not only account for common lexical forms but also, for many speakers, show all the signs of active derivation processes. Some patterns of vowel and tone alternation are used to supply vibrancy and emotive wit to informal spoken language. Kullavanijaya (1997) shows how some of these processes are used in intensifying. The deictic system too shows some quasi-morphological paradigmaticity, discussed in Henderson (1967). Iwasaki (2004) shows how topic-marking particles have been derived from such forms. Diller and Juntanamalaga (1988) speculate as to how the current deictic system may be the residue of diachronic reorganization.

Traditional Thai grammatical study (e.g. Upakit-silapasan 1939; Anuman Rajadhon 1981) makes much of morphological processes affecting Indic loanwords borrowed into Thai. Some changes, such as deletion of many Indic final syllables, are simply a matter of assimilation to the Thai phonological system. Other changes, such as those in morphological recombination of Indic roots to create Thai neologisms, may show prefixation, assimilation, etc., that reflect processes in the Indic donor systems. Pali-Sanskrit terminology such as [sama:t<sup>11</sup>] ‘samasa compound’, [son<sup>24</sup>thi<sup>45</sup>] ‘sandhi compound’, may be used and are even sometimes applied to compounds non-Indic in provenance. See Gedney (1947); Wan Waithayakorn (1970); Prasithrathsint (1994).

### 3.3.5. Nominal substitutes and classifiers

Pronouns and nominal substitutes immediately lead into sociolinguistics, as the forms in question are generally more sensitive to such constraints than they are to the more usual pronominal properties of number and even person. The definitional borderline between pronouns, conventionalized epithet-substitutes and ordinary common nouns has been drawn in various ways. Also, classifiers (below) are part of the story as they have anaphoric functions and show other pronominal traits as well. Early studies of note include Cooke (1965), which also compares Vietnamese and Burmese data using distinctive features; Campbell (1969), a useful comprehensive orientation; Palakornkul (1972, 1975) emphasizing sociolinguistic usage; Hatton (1973, 1978), discussing translation; Strecker (1980); Sugamoto (1989). Truwichien (1980) discusses the important topic of address avoidance with great insight, while her longer study (1985) impressively integrates relevant socio-cultural matters. Gething (1986) discusses similar issues. Hatton (1978) and Hoonchamlong (1992) elucidate differences between male- and female-speaker first-person forms and usage. This topic is further developed by Diller and Chirasombutti (2000), who suggest that Thai women are required by the prevailing linguistic system to ‘place themselves’ through self-reference selections in a more finely-determined social space than that required of male speakers. Chirasombutti (1995) provides further detail and a comparison with Japanese. Wijeyewardene’s (1968) contribution at the tabu end of this field is of great anthropological interest.

Hinds (1988) presents a lucid analysis of reflexives, including discourse-based examples. His argument sounds a note of caution for researchers quick to assume that anaphoric properties of reflexives in Thai such as [tua<sup>33</sup>], [e:ng<sup>33</sup>] and [tua<sup>33</sup>e:ng<sup>33</sup>] equate formally to English translational counterparts. Namthammachat (1975) and most theses written in generative frameworks also treat these issues.

Classifiers can claim a respectable store of professional literature. Most classifiers are derived from homonymic common nouns like [tua<sup>33</sup>] ‘body’; a few are from verbs, e.g. [phu:k<sup>11</sup>] ‘to tie’, as a verb; ‘bundle’ as a classifier for monastic manuscripts. They constitute nearly an open (sub)class, especially if combined with generics like [ya:ng<sup>11</sup>] ‘kind, sort’ and measure words, both formal, like [ki<sup>33</sup>lo:<sup>33</sup>] ‘kilometer; kilogram’, and conventional, like [kae:w<sup>52</sup>] ‘(drinking) glass’, as used in a measuring context. These types share many classifier patterns. For many common nouns, classifiers are required in counting expressions, where regular order is [head noun] + [number word] + [classifier]: [no’ng<sup>45</sup> so’ng<sup>24</sup> khon<sup>33</sup>] [younger-sibling two classifier] ‘two younger siblings’. Classifiers are also used to suggest definiteness, especially with deictics in the pattern [head noun] + [classifier] + [deictic], but other modifiers can occur in the [deictic] position if they are used to specify, especially



contrastively. The counting and definiteness patterns can be combined. In these patterns and elsewhere in the language, head nouns are frequently ‘missing’: they are construed from discourse context, giving syntacticians room to theorize as to zero anaphora, deletion, empty categories, etc.

Idiosyncratic classifiers are most strongly required for counting items that are discrete, concrete, and part of Thai cultural life. Two or three dozen are used quite commonly in informal conversation, but many more are recognized: McFarland (1944) gives a list of 82; however, it takes the Thai Royal Institute (Ratchabanditsathan 1995) many more than that and a booklet of 128 pages to prescribe the complete system. As items become more abstract or obscure they tend either to be counted with a general classifier [an<sup>33</sup>], or counted directly with no overt classifier, or else with a ‘repeater’ construction where a single lexical form occurs in the pattern [head noun] + [numeral] + [classifier] both as head noun and as classifier. A few concrete items, such as several body parts, are also counted in this manner. The system is far from rigid, showing individual and sociolinguistic variation, bureaucratic prescription as above, late acquisition (some items typically learned at school), and, as noted above, diachronic instability both as to the forms used and as to what each classifies (e.g. [tua<sup>33</sup>], originally for animates with bodies, is now on the rise).

Haas (1942, 1978) was perhaps the pioneer in their analysis along modern linguistic lines and her study remains an admirably concise overview, keeping in mind the passing of time. As a great help both to language learners and to researchers, her dictionary (1964) specifies one or more classifiers for virtually every noun, including indication of those that use the ‘repeater’ construction. Hiranburana (1978) concentrates on a hierarchical semantic classification, with Placzek (1984, 1985, 1992) providing more detailed studies of this type, with interest in perceptual salience. Plam (1974) supplies a treatment both semantic and syntactic, as do Kölver (1979) and Hundius and Kölver (1983); see also Zhang (1992). Kookiattikoon (2001) looks in depth at their syntax.

Along with pronouns, classifiers are among the Thai word classes highly subject to borrowing and diachronic change, with few forms still shared throughout the language family and, for those that are, some variation in the sets of nouns classified. Palakornkul (1976) and Deepadung (1997) substantiate recent changes or changes in progress, while Jachontov (1971) and Krupa (1978) suggest diachronic dynamics. Gandour, et al. (1984) attend to acquisition issues, as does Carpenter in a series of studies (1986, 1987, 1991, 1992). Gandour, et al., (1985) describe their dissolution in aphasic situations.

For comparative classifier studies, see [section 3.4.4](#).

### 3.3.6. Verbs, transitivity, predicate constructions and grammaticalization

Thai verbs are a robust class but should a separate class of adjective be recognized? There are good reasons to take adjectives as a subclass of verbs, following Gedney (1947), Haas (1964), Panupong (1970), Hudak (1987) and others who subcategorize these items as a type of verb, perhaps ‘adjectival verb’. Such forms do not require (or even admit, in most cases) a copula such as [pen<sup>33</sup>] or [khu:<sup>33</sup>] ‘to be’ to form complete predicates and they interact with auxiliaries, negatives and question/answer patterns mainly in the way full verbs do. But good arguments to the contrary have been advanced too, variously invoking comparative constructions, the ability to occur with certain prefixal formatives, semantic effects of repetition, etc. See Noss (1964); Prapa (1996), Smyth (2002), Iwasaki and Ingkaphirom (2005) and Thai traditional grammar. These authorities, it would seem, would need to acknowledge that complete Thai sentences could lack overt verbs. Nominalizing formatives [ka:n<sup>33</sup>-] for verbs and [khwa:m<sup>33</sup>-] have been often been used to distinguish these categories but a number of verbs with cognitive-emotive meanings like ‘understand’, ‘detest’, etc. accept both prefixals equally well. To call adjectival verbs ‘stative verbs’ has been one attack, but one might well wish to call verbs like [yu:<sup>11</sup>] ‘to be located’ stative verbs as well and these fail most of the putative adjective tests.

As though to make tightly defining verbs yet more onerous, Thai main verbs are difficult to distinguish absolutely from forms one might wish to take as auxiliaries: markers of timing,

aspect and modality. In serial coverb constructions they intrude on prepositions.

As noted, Thai verbs show no formal marking as to transitivity, nor do they overtly indicate finite-nonfinite status. The ability of nominals to be deleted or not to occur overtly presents a moderate challenge to transitivity analysis, but most authorities concur that many lexical verbs are basically in a transitive class, a few like [hay<sup>52</sup>] ‘to give’ in a ditransitive one, many more are intransitive and still others alternate in potential transitivity status, perhaps along with other minor semantic effects. Landmark studies of verb classes include Noss (1964), Panupong (1970) and Sindhavananda (1970). More recent discussions are by Sriphen (1982) and Thepkanjana (1992), who develops the useful notion of a transitivity continuum, taking account of verb semantics. Savetamalya (1992), emphasising patient subjects, discusses transitivity using a case-grammar approach. Negation of predicates and elsewhere is considered by Kanasut-Roengpitya (1974) and Lagsanaging (1992).

Semantics and pragmatics interact in transitivity issues. Panupong (1978) and Diller (1997) wonder about how best to analyze single verb forms exhibiting alternating transitivity. Part of the question must include pragmatics: in a rather common construction type, topicalized undergoer or patient/theme object occurs without overt agentive subject: O-undergoer (S-agent) V-trans. Does this merge in a gradient way with S-theme V-intrans? Perhaps related to this problem is a small class of high-frequency verbs such as [mi:<sup>33</sup>] ‘to have; there is/are’ and [koet:<sup>11</sup>] ‘to be born; to happen’. Here a single form appears to have both a transitive use and also to occur intransitively in a pattern of V + S when showing existential or presentative meanings. Sookgasem (1992) clarifies the issues, also considered in the lexibase system by Indrambarya (1996).

This leads to the question of marked passive or pseudo-passive constructions, well-researched topic of long-standing interest (Bergen 1875). Prasithrathsint (1988) documents substantial diachronic change in a range of passive-like constructions. A common issue in the current language involves a verb [thu:k<sup>11</sup>] ‘to come in contact with, touch’. This is widely recognized with a shifted, generalized meaning more like ‘to undergo, suffer’ and with the function of an adversative passive: [mae:w<sup>33</sup> thu:k<sup>11</sup> ma:<sup>24</sup> kat<sup>11</sup>] [cat undergo dog bite] ‘the cat was bitten by the dog’. In this construction, the form [thu:k<sup>11</sup>] retains nearly all of the formal syntactic collocational properties of a main verb, if not quite retaining normal verb semantics. This has led some authorities to resist labelling the construction as ‘passive’ *per se*, as argued forcefully and entertainingly by Noss (1972b), holding that using ‘passive’ here would be too much of an imposition of Western grammatical conceptions. (Is it churlish to observe that Noss registers no similar problem with ‘subject’ and other Western category labels, which his grammar liberally utilizes?)

In any event, as Khanittanan (1979), Prasithrathsint (1988) and others have maintained, the construction has been widely used to translate English and other Western-language passives. Perhaps partly because of resulting translation genres, usage of the [thu:k<sup>11</sup>] construction is assuredly becoming used in less adversative contexts, especially in Thai technical discourse (‘the metal was dissolved in acid’) and among middle-class speakers (‘I was invited by him to the party’). This recalls Prasithrathsint’s (1988) demonstration that earlier Thai ‘passive’ constructions have been moderately changeable. A syntactically similar, but less common, form [do:n<sup>33</sup>] ‘to be hit by’ is more stable in its adversative semantics. The (pseudo-)passive problem is treated structurally and functionally in several doctoral theses and in briefer works, among them: Filbeck (1973a); Lekawatana (1975); Thanyarat (1983); Wongbisaj (1979b) in a generative framework and Savetamalaya (2001) in a lexibase one. Morev (1996) investigates these matters from the perspective of diathesis, which includes consideration of causatives and other transitivity-shifting issues. Gero (1977) and Gsell (1979) are concerned with a similar range of issues.

Serial verb constructions, directionals and causatives have been the focus of much syntactic research. The preceding discussion indicates that the pseudo-passive markers mentioned above have many properties of verbs; as such they may be implicated in the wider phenomenon of serial verb constructions. Definitions have varied but many take this type of construction to consist of two or more verbs or verbal predicates strung together without overt

marking of coordination or subordination. In many instances, or in all if so defined, at least one nominal argument is shared, such as subject/agent or object/patient. Usually at least one such nominal is understood, i.e. is an empty category or zero anaphor. Crucial here also is the observation that along with many other serializing languages Thai lacks morphological marking distinguishing finite from non-finite. Also, to assume that the first verb in sequence were in all cases the dominant one leads to quandaries. Constructions of this type are a Southeast and East Asian areal feature, as documented by Clark (1978, 1992) and Clark and Prasithrathasint (1985) in a lexibase framework. Bisang (1996), with grammaticalization and ‘great attractors’ in mind, analyzes Thai examples in this wider context too, as does Post (2007). Analyses reveal both broad similarities and specific differences with comparable data in nearby languages.

Needleman (1973a) is among the first studies in which a formal post-Aspects generative framework confronts the challenge of the seemingly flat multi-verb structures of Thai verb serialization, followed by Filbeck (1975), Vis (1978) and Sereechareonsatit (1984). Thepkanjana (1986) should be credited, it seems to me, with a notable advance by turning attention to specific semantic subclasses of verbs as a constraint in how longer complexes are contextually built up and strung together, a direction developed by Chuwicha (1993) and in other work. Somewhat similar in orientation, but using the lexibase apparatus with robust use of synchronic derivation, Clark and Prasithrathasint (1985) offer an analysis of verb forms showing context-sensitive differences in sense or structural properties. Working in the same basic framework, Wilawan (1992, 1993) goes on more radically to supply an argument rejecting the ‘serial verb’ characterization entirely and taking the relevant constructions to be coordinate or subordinate clauses within a system of sentence adjuncts. For more consideration of these matters, further references and a scheme based on symmetrical/asymmetrical properties of these constructions, see Diller (2006a), which also mentions properties of shared arguments, negation, modality and timing.

The verbs [pay<sup>33</sup>] ‘to go’ and [ma:<sup>33</sup>] ‘to come’ and similar verbs of motion enter into serial constructions of great frequency, mixing directional, temporal, aspectual and evaluative nuances in intriguing ways that have stimulated much study. These forms are considered in many of the sources cited above, but studies focus on them specifically as well. Gandour (1978a) associates ‘come’ and ‘go’ with deictic properties, with Treerat (1990) going on to come (!) to grips with how syntax, predicate semantics and discourse context interact. Temporal-aspectual functions of [pay<sup>33</sup>] offer a particular challenge, with interpretations of timing sometimes suggesting past completive, sometimes future continuative. [khaw<sup>52</sup>] ‘to enter’, hence inchoative, and some other directionals contribute to the topic as well. Bickner (1985) and Rangkupan (2001) analyze usage of directionals as it indicates psychological perspective in narrative and in other contexts.

The verb [hay<sup>52</sup>] ‘to give’ in serial constructions promotes a different polysemous complex involving notions like benefactive, causative and complementizer of controlled action. Areal considerations are again relevant; see comparative analyses by Hermann (1979); Pooisrakit (Poo-israkij 1995), Iwasaki and Yap (1998); Post (2007). Vichit-Vadakan (1976) notes differences in intended versus inadvertent causality in constructions with [hay<sup>52</sup>], [tham<sup>33</sup>] ‘to do’ and the compound [tham<sup>33</sup>hay<sup>52</sup>]. Other studies are by Kumlert (1976), Khamseen (1978), Indrambarya (1992), with Warotamasikkhadit (1994) offering questions as to prevailing terminology. Grima (1978), while not limited to [hay<sup>52</sup>] constructions, presents an especially cogent discussion of the associated types of zero anaphora. See also [chapter 17](#).

Temporal-aspectual information may be marked by the use of directionals to indicate temporal-aspectual nuances is mentioned above. Other marking of this type, along with modality, can usually be traced back to main-verb sources too, although such verbs may or may not still be in active usage. A problem encountered by those attempting to establish a fixed linear order for lexical forms in auxiliary constructions, e.g. Anthony (1964), Dellinger (1975), is that many of the dozen or more commonly used forms can function in different surface positions with slightly different meanings and collocational dynamics; see critique of Warotamasikkhadit (1979). Thus the form [a:t<sup>1</sup>] can function epistemically as ‘apt to, likely

to' but also deontically as 'able, capable of'. The epistemic position is more peripheral, preceding positions of irrealis marker [ca<sup>11</sup>] and negative [may<sup>52</sup>], with the deontic alternate closer to the verbal core and following positions of such formatives. The favored structuralist solution has been to expand the lexicon: Noss (1964) distinguishes two [a:t<sup>11</sup>] homonyms in different form classes.

Kimsuvan (1992) considers alternations with [yu:<sup>11</sup>] 'be located; be happening; at'. This form along with others has been investigated in detail by Kullavanijaya and Bisang (2004) in a selection-theory approach. Similar syntactic and semantic alternations apply to several other forms. [day<sup>52</sup>] 'can' (and many further possible glosses) is notorious, with functions ranging widely over the epistemic, deontic and temporal-aspectual semantic territory. Enfield (this volume) has given much attention to the counterpart of this form in Lao and much of what he describes in this volume for the Lao would apply to Thai as well. Other approaches involve linkage of semantics to various treatments of phrase structure, an especially well-designed analysis being that of Sookgasem (1990) in *Head-driven Phrase Structure Grammar*. See also studies of Kanchanawan (1978), with interest in machine translation; Eknaiyom (1979), using internal reconstruction; and theses of Kullavanijaya (1968); Scovel (1970); Boonyapatipark (1983). Howard (2000) and Koenig and Muansuwan (2000) inspect perfectivity in detail, e.g. use of the form [lae:w<sup>45</sup>], originally a verb 'to finish'.

Much of the above discussion relates in one way or another to grammaticalization paths, especially to situations where diachronically prior constructions and senses coexist in the current language along with their evolved alternates, with little or no phonological indication of difference. Enfield (2006) presents a cogent discussion of grammaticalization issues relating to Lao, in most cases with direct application to Thai as well. Not only verbs and derived auxiliaries are at issue, but also other paths like [verb – coverb – preposition], cp. [ca:k<sup>11</sup>] 'to leave (now very restricted); from' or [noun – preposition], cp. [lang<sup>24</sup>] 'back (body part noun); in back of (as though preposition or adverb)', with yet another function being classifier for houses. For more on what is essentially grammaticalization, if not explicitly so identified, see work of Clark (from 1978 onwards), Jagacinski (1991), Juntanamalaga and Diller (1992), Bisang (1996), Diller (2001b). Iwasaki (2004) considers the grammaticalization of topic-marking form [nia<sup>52</sup>]. See Part Four of this volume.

### 3.3.7. Other parts of speech and constructions

Questions are taken up in work by Kullavanijaya (1980) and Santaputra (1984). A common yes-no question type is coded by a final particle written as though [may<sup>24</sup>] but usually pronounced [may<sup>45</sup>]; the ultimate diachronic source is undoubtedly the preverbal negator [may<sup>42</sup>]. This and other final particles with various speech-act, politeness and discourse functions are analyzed in a number of studies, Cooke (1989) providing a comprehensive orientation, with other studies by Bhamoraput (1972), Peyasantiwong (1981), Kendall, Yoon and Hye-Suk (1986), Horie (Ingkaphirom) and Iwasaki (1996). Phonological reduction of these forms is studied by Peyasantiwong (1979). Bandhamedha (1979) presents a convenient grouping of families of particles and perceptive analyses of functions. For specific treatments, see Cooke (1979) for the [si<sup>45</sup>] set and Neill's (1989) narrative-based analysis of [na<sup>45</sup>] and [chay<sup>52</sup> may<sup>45</sup>]. Diller and Juntanamalaga (1992) take up the [oe:y<sup>24</sup>] set. Strings of particles can occur together. Whether such forms are subject to ordering rules is considered by Warotamasikkhadit (1975) and Prasithrathsint (1974).

Conjunctions are considered by Thomas (1979) and by Clark (1994), who presents a comparative study and argument for a topicalizing function. Jagacinski (1991) takes up complementizer [wa:<sup>52</sup>] and other complement types. For relative clauses and factitive noun clauses marked by formative [thi:<sup>52</sup>], see Eknaiyom (1971); Suktrakul (1975); Sornhiran (1978); Kuno and Wongkhamthong (1981b); Savetamalya (1996) and, for a convincing diachronic analysis, Kullavanijaya, this volume. Comrie (1996) treats clauses of this type comparatively with particular attention to Japanese. Morev (1994) describes possessive constructions. For reciprocals, Bee (1972) comes to grips with the poly-functional [kan<sup>33</sup>], variously a reciprocal, gathered-plural marker and male first-person pronominal form.

Constructions and their functions of sentences with copular forms or equatives [pen<sup>33</sup>] and [khu':<sup>33</sup>] are taken up by Kuno and Wongkhomthong (1980, 1981a). As part of the wider picture, Warotamasikkhadit (1969, 1976a) treats other idiosyncrasies of the verb [pen<sup>33</sup>] along with verbless sentences.

A postverbal nominal in a construction that might seem SVO superficially on inspection might not be a typical O-type patient semantically: it might turn out to have a locative, manner or instrumental function. Instrumental constructions are described by Warotamasikkhadit (1986). The postverbal position can also accommodate affected body-part terms, e.g. in pain expressions. These introduce the issue of how pain is represented and how this affects construction choice, as discussed by Diller (1980) and by Iwasaki (2002); also in a more comparative anthropological mode by Fabrega and Tyma (1976).

Quantification is taken up by Stine (1981). Whatever the theoretical approach, there can be little doubt that Thai quantificational phrases are prone to occur at the end of predicates, even if this means splitting an earlier noun phrase. Wongbiasaj (1979a) considers this type under quantifier floating. In a related area, Haas (1946) and Kullavanijaya (1997) give examples of a range of intensifying techniques, several of which involve final position as well.

### 3.3.8. Discourse, conversation and sociolinguistically-based studies

Not a few of the studies cited above consider units greater than single sentences, among them Panupong (1970), Grima (1986) and Hinds (1988b). Iwasaki and Ingapirom (2005), and in their other work, frequently rely on conversation-based examples and cogently take such wider contexts into account. Thomas (1988) argues that Thai grammar needs to recognize the language as 'paragraph-efficient' rather than as based on rules confined to single clauses, with Vongvipanond (1988) arguing for two types of linkage devices: 'macro-cohesive' and 'micro-cohesive'.

An early concern with discourse can be traced in the brief study of Hatton (1975). Vongvipanond (Ekniyom) (1977, 1982) was among the first to consider topicalization in thesis-length detail, along with other discourse-pragmatic phenomena. A similar focus informs Messenger's (1980) dissertation, analyzing theme in discourse. Interest in topic and topic-marking devices continues in the work of Iwasaki, e.g. (2004). Schmidt (1994) is a treatment of aspect in discourse.

1986 was an especially fruitful year for dissertations devoted to Thai discourse by native speakers. Burusphat (1986) presents a comprehensive discourse analysis based on narrative folklore, with texts presented in a convenient appendix, fully transcribed, glossed and translated. Chanawangsa (1986) studies cohesion from several perspectives, while Chodchoey (1986) uses transcribed materials to uncover strategies in oral discourse.

Conversational Analysis (CA) refers to micro-analysis of discourse along sociological lines, including special attention to pauses, repetitions, self-corrections, kinesics, etc., usually ignored in other frameworks. Moerman (1988) presents stunning examples of the insights that this style of analysis can reveal. An extra dimension of complexity in his study is the combination of Northern Thai, Lue and Central Thai in the text material, which is presented in meticulous oral transcriptions. Bilmes (1992) makes similar points. Turn-taking and speaker overlap is of interest in these studies, with Hinds (1988) providing a provocative parallel between Thai driving behavior and informal conversational interaction. While filled pauses are also of interest in CA, Chaimanee (1996) instead devotes attention to them in a comparative study of native and non-native speakers' hesitations.

In other frameworks, an early study considering interpersonal issues in discourse organization is Hatton's (1978) analysis of first-person reference in narratives. Using natural conversation, Hartmann (1993) describes communication in market-places, with the more general thesis-length study of Baron (Meepoe) (2001) analyzing reference to persons in conversation. Patrakom (1977) deals with similar issues from a more philosophical perspective, developing a hermeneutic approach to the characterization of persons. Neill (1989) explicitly develops a synthetic view of discourse as an interaction of grammar, rhetoric and background socio-cultural knowledge.

Khanittanan (1987a, 1987b, 1988b) uncovers diachronic trends in discourse genre and makes a compelling case for the evolution of a more ‘autonomous’ and abstract type of Thai, partly as the result of different communicative needs arising from modernizing social trends. This genre is less tied to traditional constraints of immediate interpersonal communication, apt to use abstract nominalizations and to make nominal material explicit. In a similar vein, differences between spoken and written discourse are clarified by Chodchoey (1988). Person (1996) turns attention to specialized genre: oral sermons of a Buddhist monk.

For Thai-Japanese comparisons, see Ruetaivan’s (1999) study on how motion events are represented in narrative discourse.

### 3.3.9. Semantics and lexical fields

Several of the syntactic studies already cited show a strong semantic focus. The distinctive feature framework is used by Cooke (1965) in his comparative study of pronominal reference. Hiranburana (1978) develops a related hierarchical-taxonomy approach to classifying Thai classifiers, with several works of Placzek (e.g. 1992) producing a more finely-tuned analysis with particular semantic attention to shape. Gething (1986) discusses the extent to which distinctive features relate to cultural matters. As mentioned above, Thepkanjana’s work (e.g. 1986) is notable for uncovering how semantics of verb classes can constrain the makeup of serial verb constructions.

Compounding has attracted semantic attention. The studies cited with regard to compounding [cay<sup>33</sup>] ‘heart, mind’ to yield emotional terminology treat semantics as well as syntax. Juntanamalaga’s (1992) study classifies the cluster of meanings associated with [hua<sup>24</sup>] ‘head’ as it occurs in compounds. Vongvipanond (1992a) presents a substantial semantic and syntactic analysis of compounds where component parts are synonyms.

Additional studies link meaning with syntactic constructions in specific semantic fields. As mentioned, Thai pain terms are described in a comparative context by Fabrega and Tyma (1976) and by Diller (1980), who points to interesting syntactic features of these constructions. Their syntactic analysis is substantially advanced by Iwasaki (2002). A similar topic, traditional Thai disease terminology, is developed in a more ethnographic study by Bamber (1987). Still relating to physiology but in a cheerier subfield, Reed (1976) produces a semantic analysis of Thai gastronomic terms.

Representative thesis-length treatments of semantic topics include Gething (1972) and Terayanont (1988). Varied approaches are attested. Gething (1968, 1972) develops a structural redundancy methodology which he applies to analysis of nominals. His later work (1975, 1979) treats the semantics of locatives and other expressions with comparative reference to Lao. Diller (1994), in a comparative volume, attempts to address issues raised in the ‘semantic primitives’ enterprise. Folk taxonomy is the focus of studies by Simmonds (1978) and by Stott (1978), describing vernacular forest nomenclature. Historical semantics is taken up by Khanittanan and Placzek (1982), who trace how inherited Tai ‘psychic’ vocabulary, such as [khwan<sup>24</sup>] ‘(roughly) psyche, spirit’, has undergone semantic readjustment with the introduction of Indic loans such as [win<sup>33</sup>ya:n<sup>33</sup>] ‘(roughly) soul’. Another historical study of anthropological interest is that of Gething (1977).

Toponyms as a reflection of culture, social organization and environment are treated in impressive studies of Hartmann (2007) and Prasitrathsint (2007b). These link names of villages and other geographical units in Thailand to the wider Tai context, along with consideration of the Southeast Asian ecosystem and socio-political factors.

Kinship: of anthropological merit and impact are studies of Thai kinship terms and their extended usages. Gething (1986a) discusses distinctive features organizing Thai kinship and occupational terms. In a more extended study, Prasitrathsint (2001) presents a componential analysis of the complex system with useful comparative references. For diachronic/comparative development, one could consult also Black Tai (Fippinger, 1971) and check Strecker’s (1980, 1984) Proto-Tai reconstructions. Formative studies are by Benedict (1943, 1945), with Chinese comparisons. Also of interest to anthropologists is the fact that Thai speakers appear to enjoy ‘disrupting’ their lexically normative kinship system in

manners exotic (to outsider anthropologists, at least): parental terms [pho':<sup>42</sup>] 'father' and [mae:<sup>42</sup>] 'mother' are used by parents (among others) in referring to or addressing their children: thus, one calls one's daughter 'mother'; but the daughter may call her mother [mae:<sup>42</sup>] too. Also, these terms may precede children's given names or nicknames in the manner of a title. There is yet more: sibling terms are the widely used in address and reference among spouses; see Haas (1969).

### 3.4. SOCIOLINGUISTIC, COMPARATIVE AND HISTORICAL STUDIES

#### 3.4.1. Sociolinguistic perspectives on Thailand

Studies cited above dealing with pronouns, kinship terms, speech-act particles and discourse are especially likely to be of interest to sociolinguists. In this subsection we note more general descriptions of the rather complex setting in Thailand important in understanding communicative functions of Central Thai in their social setting. For those requiring an overview of Thailand's sociolinguistic situation, Premrirat (2006) is recommended. The issue of *International Journal of the Sociology of Language* 186 (2007), devoted to Thai sociolinguistics, includes range of current work, some mentioned below. For those interested in the development of the subfield of Thai sociolinguistics, comparisons could be made with the sociolinguistic review of Debyasuvarn (1973) and also with an issue of the journal *Language Sciences* 10 (1988), devoted to papers on Thai sociolinguistics of that era. Roop's (1969) earlier study of language diversity remains useful for perspective.

Beebe (1975, 1976, 1981) presents an impressive set of Labovian studies correlating articulatory alternatives with variables like occupation, age and gender. There is also treatment of these issues a language-learning context by Brown (1967), who usefully distinguishes colloquial 'Bangkok Thai' from the standard language; the former variety shows cluster simplification, substitution of /l/ for /r/, etc.

Multilingualism and language hierarchy in Thailand is the theme of the impressive and informative survey of Smalley (1994), expanding his (1988) introduction. For Smalley, Thailand's local dialects and minority languages are part of a hierarchical structure with standard Central Thai at the top. Speakers tend to become bilingual or multilingual in varieties positioned over them in the hierarchical model.

Following from Smalley's observations, investigators find that local dialects of the Tai family in Thailand are by now most frequently components in speakers' bidialectal (or multilingual) competence: outside of the Central-speaking area virtually all speakers with basic public schooling have at least a functionally passive competence in Central Thai, although active competence and most daily-life communication may be in a local variety. These local varieties are not infrequently mixed with or influenced by the standard language. Chamberlain (1972) and Khanittanan (1973) describe tonal influence of the standard language on Northeastern or Lao varieties. Diller (1979) investigates how tones and segments in Southern and Central Thai are systematically combined in sociolinguistically salient ways. In particular, hybridization is described whereby Southern Thai tones are articulated for what is otherwise Central Thai. This can apply even to the standardized written form. For the north, the interplay of Kam-mueang (Northern Thai), Lue and Central Thai in their sociological context is amply elucidated and documented by Moerman (1965, 1988) in a closely-focused conversational-analysis framework. Similarly, but in a different framework, Premrirat's (2007) study of endangered languages in Thailand calls attention to this urgent line of research, recalling Bradley's (1992) discussion of the disappearance of Ungong. Morita (2003, 2007) studies the important question of assimilation and language shift among Sino-Thai speakers; her bibliographies are a useful resource for those interested in this issue. From another quarter, Chunsuvimol (1980) considers communicative networks of Thai workers in Singapore.

Dialectology studies of local and regional varieties in Thailand are too numerous to be listed here: many are catalogued and described by Tingsabadh (1984), with leading studies discussed in Smalley (1994). Many fine theses completed in Thai universities describe

individual local varieties. Brown (1962) and Hartmann (1980) treat issues of subgrouping. These works and sources mentioned therein can be consulted for Kam-mueang (Northern Thai), Southern Thai and so-called Isan Thai (essentially Lao varieties spoken in Thailand). Chantavibulya (Panupong) (1959) appears to be the first detailed linguistic description of a local dialect (Songkhla, Southern Thai). For comparative reference, a remarkable nearby Southern tonal system is acoustically and physiologically analysed by Rose (1997).

Central Thailand is far from uniform in local speech. For the central-west area, Suphanburi Thai is described by Tingsabadh (1988, 1992), who makes the intriguing and methodologically challenging observation that speakers of this variety really cannot cite their local forms in isolation. Court (1972) describes the unusual tonal system of Traat.

Early landmark comparisons with focus on tone compare varieties within Thailand, and sometimes beyond: Haas (1958); Egerod (1961); Jones (1965). Brown (1962, 1965) is the first book-length work to cover the entire country in detail. It establishes a convenient visual display for comparing tonal systems and remains a valuable dialectology resource. Many later works describe new varieties or provide further analyses, e.g. Strecker (1979); see Tingsabadh (1984).

Sociolinguistic and political commentary on Thai as a standard national language is provided by Aesrivongse (1984, in Thai), whose viewpoint is summarized and extended by Diller (1988, 1991).

### 3.4.2. Registers, polite speech and special styles

Studies of conversational Thai by Moerman (1977, 1988), Peyasanitwong (1981, 1986), Hinds (1988a), Chodchoey (1986), and by Iwasaki and Ingkaphirom (2000, 2005) elucidate features of colloquial spoken Thai that set it off from the standard written form. Differences are directly confronted by Tiancharoen (1987) and by Chodchoey (1988). Khanittanan (1988b) demonstrates how written Thai has evolved in an 'autonomous' direction not anchored in aspects of personal interaction as regularly encoded in colloquial conversation.

Native speakers of Thai find it easy to characterize samples from written and spoken registers in evaluative terms like [supha:p<sup>52</sup>] 'polite' or its opposite. In effect this presupposes a diglossic arrangement which includes the distinguishing of lexical pairs that can be described as 'higher' and 'lower' in terms of speech-level, although the relationship among resulting registers is far from a simple binary one. Diller (1985, 1993) suggests that both lexical and syntactic evaluations contribute to this continuum, but it should not be concluded that written Thai always shows 'high' selections and spoken Thai 'low' ones. Sometimes written Thai opts for 'lower' expressions as in journalism: Khanittanan (1994, 2007) shows how styles are manipulated for effects in different segments of news presentation; see also Srinarawat's (2007) informative work on political slang, another use of 'lower' expressions for effect.

Politeness markers and similar indicators of interpersonal dynamics are ubiquitous and especially salient in colloquial Thai. These categories are coded by various means: most obviously by address and reference forms and by final particles such as polite final particles [kha<sup>45</sup>], [kha<sup>52</sup>] (female speaker) and [khrap<sup>45</sup>] (male speaker) and the less polite [ha<sup>45</sup>], [wa<sup>45</sup>], etc., (Bandhumedha, 1979; Peyasanitwong 1981; Cooke 1989), but also by various lexical options along the high-low continuum mentioned above. This continuum generally reflects ascribed deference as well as formality. But the system is complex. Lexical selections in formal Thai cannot be taken as necessarily indicative of marking interpersonal deference, e.g. in formal sources high-status and low-status males alike may be referred to as [na:y<sup>33</sup> X]: 'Mr. X', [na:y<sup>33</sup> thak<sup>45</sup> sin<sup>24</sup>] 'Mr Thaksin' (a former Prime Minister); [na:y<sup>33</sup> dae:ng<sup>33</sup>] 'Mr Daeng' (an unemployed unskilled manual worker)'. Compare more colloquial Thai: [na:y<sup>33</sup> X] seems less deferential than [khun<sup>33</sup> X] 'Mr. X' (lit. 'honorable X'). Note that [khun<sup>33</sup> X] is uncommon in formal Thai, e.g. in official reports, etc. Even syntactic choices are implicated in this complex. Khanittanan (1988a) elucidates strategies and issues with great insight, as do Kummer (1992) and Bilmes (2002), while Deephuengton (1992) turns to disagreement strategies. Work of Moerman, Hinds and Iwasaki keep such factors in mind as well.



Thai is remarkable for special registers relating to institutions of royalty and monkhood, although the lexical makeup of these registers shares many selections with the ‘high’ or ‘polite’ [supha:p<sup>52</sup>] level mentioned above. Particularly given the impact of broadcast media, most Thai speakers are at least passively familiar with the hundreds of lexical substitutions these registers prescribe, if not always able to control them actively. Haas (1951) and Jones (1971) lay out the complexity of the system, especially as regards degrees of deference ascribed to ranks of royal persons. Gedney (1961) provides a convenient outline of these registers, while Kanasut-Roengpitya (1973) documents them in greater detail. Diller (2006b) also gives a sketch of ‘royal Thai’ calling attention to some of its derivational processes and suggesting that functionally it shows some parallels with special avoidance genres in other languages.

Poetic Thai and other literary genres have their own special vocabulary and syntax and show points of contact with the type of language used with royalty and the monkhood. Also, the formal metrical requirements of traditional poetry provide a number of clues both as to abstract phonology and as to how Thai may have developed diachronically. Linguists undertaking work of this type include Warotamaskkhadit (1968a); Bickner (1981, 1992); Chittasophon (1984); Gedney (1980, 1989b); Hartmann (1989); Hudak (1990 and this volume).

### 3.4.3. Loans, contact and bilingualism

General contact issues are introduced in traditional philological studies, see Anuman Rajadhon (1961, 1981). A linguistic orientation is taken by Warie (1973, 1979). Sometimes overlooked is the importance of foreign language education as an aspect of contact, considered by Chirasombutti (2007).

In the preceding section, lexical items of the higher diglossic registers were noted. These in general are etymologically not inherited Tai vocabulary but are loans introduced over many centuries. Perhaps the most concentrated and conspicuous loan element in Thai is vocabulary from Indic languages Pali and Sanskrit. Gedney (1947), in a work still of great value, presents a formal analysis and catalogue of many hundreds of words of this sort, mainly nouns and verbs. Historically, Indic vocabulary in Thai can be classified into two types: (i) traditional loans, accepted into Thai from the thirteenth to nineteenth centuries; and (ii) coined neologisms created in the twentieth centuries through official institutional action. There are hundreds of both types in current use and distinguishing them is often not a simple matter. Type (i) were typically introduced in pre-modern times through Khmer or Mon intermediaries, subject to assimilation processes in those languages first, or else from Pali Buddhist sources, i.e., terms originally used in religious contexts came to be used more generally. Most of these loans are transparent in provenance, but not all. Harris (2007b) convincingly shows how Mon or Khmer mediated Sanskrit *marica* ‘hot (capiscum) pepper’, borrowed into Thai with sound changes, resulting now in [phrik<sup>45</sup>].

Type (ii) loans were created by literary and technical authorities in an environment of modernization requiring lexical enrichment, but coupled with nationalistic or esthetic sensibilities disparaging direct loans from English or other Western languages. The solution was to code the Western-based concepts with neologisms constructed from Indic (mainly Sanskrit) morphemes. The neo-Indic forms created also conformed analogically to various assimilated prototypes already established through type (i) forms; for example, there was no attempt to pronounce neo-Sanskrit material the way it would be pronounced in India. Sometimes shifts have occurred in the designated semantics of type (ii) items, especially those of the 1920s or 1930s. Wan Waithayakorn (1970), one of the key figures in the coining process, and Prasithrathsint (1994) discuss the specifics of how these (ii) items were introduced. See also Court (1984).

Austroasiatic, Austronesian and Sino-Tibetan language families have contributed many vocabulary items to Thai at different diachronic periods. Relatively recent loans from Khmer, Malay and southern Chinese varieties like Taeciw are often transparent, while earlier strata feel to Thai speakers like authentic Thai vocabulary and require linguistic scholarship to

deduce provenance. Careful analysis and correct differentiation of diachronic strata from various sources of vocabulary now used in Thai can be considered still in the infancy stage. Surely however such studies are a precondition for convincing discussions of remote or ultimate Tai and Tai-Kadai relationships, be they genetic or some form of contact, if that distinction can indeed be maintained.

Contact with Khmer and other Austroasiatic languages is a topic pursued by Khanittanan (2001) and Diller (2002), who both argue for an early period of substantial Thai-Khmer bilingualism. Evidence for this claim lies in the sweep and magnitude of Thai vocabulary that is Khmer-derived (Varasarin 1984). While some occurs at the higher speech-levels, much is basic and currently in every-day usage: [camu:k<sup>11</sup>] ‘nose’; [tapho:k<sup>42</sup>] ‘hip’; [khu’<sup>33</sup>] ‘to be (equivalent)’; [doe:n<sup>33</sup>] ‘to walk’; [koe:t<sup>11</sup>] ‘be born’; [set<sup>11</sup>] ‘to finish’. More tellingly: a number of auxiliaries, conjunctions, prepositions and other grammatical formatives are clearly of Khmer or Mon-Khmer provenance, as is the disyllabic lexical pattern, seen in ‘nose’ above, now so thoroughly incorporated into Thai that it seems native.

Turning to Austronesian, Suthiwan (1992) focuses on Malay loans into Central Thai. She is able to show several different diachronic strata on the basis of tone assignment and stop devoicing. One Ayudhya-era level is associated with a Thai court adaptation of the Javanese and Malay Panji tales and many of the loans are literary in character, but other strata have toponyms, names of fruits, etc., that are in common usage. Suthiwan also considers Malay loans into Southern Thai, as does Court (1975), who includes insightful diachronic deductions and disussion. More contentious are the earlier levels still. Schlegel (1902) advanced the hypothesis, revived and enhanced (?) by Benedict (1942, 1975, etc.), that Proto-Tai vocabulary shows enough plausible Austronesian cognates to support a standard genetic (hence Austro-Tai) relationship. This topic is pursued in detail elsewhere, whether *pro*, *con* or prevaricating: (e.g., Gedney 1976; Diffloth 1977; Reid 1984; Hartmann 1986b; Matisoff (1990); Thurgood (1994); Diller 2000; Ostapirat 2004).

Another important contact domain concerns Chinese relationships, again a matter of contention. Work of Egerod (1957, 1959b), Manomaivibool (1975, 1976, 2000), Li (1976) and Luo (2000; also this volume and sources mentioned therein) suggest that there are very early, if not genetically inherited, strata of Chinese in the Tai (and Tai-Kadai) languages. Manomaivibool (1976) carefully differentiates early strata, e.g. [plu’ak<sup>11</sup>] ‘peel; bark’ versus [phiw<sup>24</sup>] ‘skin’, which she takes to be from the same Chinese lexical source 膚, but reflecting different stages. Luo shows the magnitude, basic nature, semantic spread and emphasizes regular correspondence patterns of many items involved. Thurgood (2007a), on the other hand, calls attention to an irregularly-corresponding lexical group taken by Gedney (1979) as evidence for a new set of Proto-Tai initials. Thurgood suggests rather that the items in question represent Chinese loans of various strata, hence the irregularity. This problem is perhaps indicative of more general methodological tensions attending Chinese-Tai (or Tai-Kadai) diachronic research. Whatever the nature of the ultimate Sino-Tibetan relationship, vocabulary of early eras (e.g. terms for numbers, body parts, basic verbs, etc., on at least the Proto-Tai level) is not to be confused with much more recent southern Chinese (mainly Swatow/Taewi) loans into Central Thai (Egerod 1959a). Unlike the earlier vocabulary, later items are felt by speakers of Thai to be ‘Chinese’ and sometimes even have phonological properties that effectively mark them as such (e.g., unaspirated initial stops in items with high or rising tones, precluded as a regular possibility for inherited Tai vocabulary). Such vocabulary refers especially to foods, cooking processes, business terms and other transparently Chinese cultural concepts, but also to some pronouns now commonly heard in Thai speech, at least on Bangkok streets. Of relevance here are Morita’s (2007) work on Chinese-Thai assimilation and Srinarawat’s (1988) on language use of Chinese-background speakers in Bangkok. These are interesting to contrast with Huang’s (2007) study of Zhuang-Chinese assimilation.

Portuguese, Persian, Arabic and other languages recalling Indian Ocean pre-modern commerce have brought into Thai designations for grapes, cabbage, roses, soap and other cultural items; see Harris (2007a, 2007b).

English has had a moderate to heavy impact on Thai, explored sociolinguistically by Khanittanan (1979), Nacaskul (1979) and by Chutisilp (1984). Work on the development of the so-called adversative passive marked by [thu:k<sup>11</sup>] ‘to undergo’ often notes the use (or misuse) of the construction in translating English passives (e.g. Prasithrathsint 1988; Diller 1993). Lexically, of historical interest is the fact that many English loans like [sathe:<sup>33</sup>chan<sup>52</sup>] ‘station’, commonly used in the nineteenth century, were replaced by Sanskrit neologisms in the early twentieth century, such [satha:<sup>24</sup>ni:<sup>33</sup>]. In this case, the Indic neologism even recalls the form of the English prototype as well as its etymological connections (note Proto-Indo-European *sta*: ‘stand’, appearing in both the Sanskrit and English forms). In a sense then hundreds of such Sanskrit neologisms are an attempt to represent English concepts semantically but through portraying them in neo-Indic phonological guise. This is more in keeping with early twentieth-century nationalistic and literary sensibilities than would have been the simple inclusion of barely assimilated English forms. Nonetheless, sensibilities shift and recent decades have not avoided a substantial influx of direct English borrowing (Senawong 1992). A torrent of such loans now imbues ‘pop’ culture, teen slang, sports reporting and technology. Commercial establishments like tailor shops frequently display transliterated English names. Human nicknames like ‘Nut’, ‘Golf’ and ‘Bird’ are analyzed by Nacaskul (1987), who clarifies the phonology of such loans; even dogs get English names. The humanities and law, on the other hand, tend to eschew direct English borrowing and such loans are regularly absent from proclamations, constitutions, etc. For the English-to-neoIndic process, see Wan Waithayakorn (1970); Prasithrathsint (1994). Gandour (1979) and Bickner (1986b) discuss the complicated issue of how tone is assigned to English loans. Thai-English bilingualism is studied by Cefola (1981) and by Davis and Schoknecht (1994). Psycholinguistic studies of Burnham, et al. (1992) are of relevance here as well. In reverse, Cohen (1987) provides an entertaining look at expatriate foreigners’ acquisition (or not) of Thai in Bangkok.

#### 3.4.4. Comparative studies

Several studies compare standard Central Thai with other Tai varieties of Thailand, some with focus on attitudinal factors: Northern (Lanna) Thai (Pankhuenkhat 1976; Nokaeo 1989); Yong (Davies 1979); Northeastern Thai (Palikupt 1983).

Comparisons with Khmer (Cambodian) include focus on predicates by Martini (1956, 1957); descriptions of remarkable syntactic parallelism by Nacaskul (1971) and by Huffman (1973); study of complement constructions by Poo-israkij (1995); and diachronic lexical analysis Varasarin (1984). Khanittanan (2001) and Diller (2003) develop diachronic proposals related to these comparative studies.

Comparative classifier studies are instructive in several ways. Jones (1970) presents a landmark survey of Southeast systems. It establishes areal patterning of classifier syntax, especially among Tai languages: classifiers follow their head noun as in Thai to the west and south, but precede it, as in Tai varieties like Nung or Zhuang to the east and north. (As for an isogloss, a leading Black Tai linguist, the late Cam Trong, told me that both patterns are used in his language for different purposes. In fact, most speakers of Central Thai admit some flexibility in special contexts, such as the regular [phi:<sup>42</sup>-no:<sup>ng</sup><sup>45</sup> so:<sup>ng</sup><sup>24</sup> khon<sup>33</sup>] ‘two siblings’ as compared to the more restricted [so:<sup>ng</sup><sup>24</sup> khon<sup>33</sup> phi:<sup>52</sup>-no:<sup>ng</sup><sup>45</sup>] ‘both of them, the siblings’. As for particular lexical items, work of Jacob (1965) indicates that Khmer has supplied Thai with several classifiers and vice-versa. Conklin (1981) is a comparative study encompassing Tai (including Thai) and Austronesian classifiers. In a cross-language survey of classifiers and language standardization, Barz and Diller (1985) examine evaluative and normative feelings about classifiers, found to differ across Indo-Aryan and Southeast Asian languages. For Thai especially, sociolinguistic constraints are fine-tuned and crucial in many Thai classifier selections, as established by Juntanamalaga (1988). Thai is placed in a comparative Tai context by Burusphat (2007), who presents a comprehensive treatment of animate classifiers; see also Morev (2000).

Southeast Asian language comparisons including Thai are presented by Downer (1963); Nguyen Dang Liem (1974); Clark (1978, 1992, 1994); Clark and Prasithrathsint (1985). Cooke (1965) considers pronominal reference in Thai, Burmese and Vietnamese.

Interest in Thai-Vietnamese comparison can be traced back at least to Martini (1950) and to Honey and Simmonds (1963). Masuko and Kiritani (1991) compare Indonesian and Thai consonant articulation. See also Shimizu (1989). Syntactic comparisons of Urak Lawoi, Malay and Thai are presented by Hogan (1978). Chinese and Thai treatments of zero anaphora are compared by Cole (1984). Egerod (1971) includes Thai and Chinese in a study of phonation types. Xing (1955, in Chinese) compares Chinese and Thai ‘offspring’ compounds. Substantial comparative studies involving Thai and other Asian languages seem regrettably sparse, but include the studies of Japanese and Thai formulaic expressions by Wongkhomthong (1985) and of self-reference in these languages by Chirasombutti (1995).

Thai-English comparisons often have an applied-linguistics focus. Pioneering studies include Kruatrachue (1960) on phonology and Chaiyaratana (1961) on syntax. A sample of others includes studies on pronominalization, considered by Chomaitong (1976); on definiteness, by Lamchote (1971); on aspect, by Noochoochai (1978); and on intonation, by Kanchanathat (1977). For applied studies featuring error analysis, translation problems or difficulties Thai speakers have with specific features of English, representative studies are by Ariyapitipun (1988), Meemeskul-Martin (1984), Palmer (1969), Richards (1968), Schmidt (1988), Suwattee (1981), and by Van Syoc (1964).

### 3.4.5. Historical and philological studies

Diachronic sources relating to Tai family and to Tai-Kadai are mentioned in [section 1.1](#). This and the next subsection mainly mention works with a focus specifically on (standard and/or Central) Thai, or else the language of Sukhothai; however, controversial hypotheses tracing how Central Thai has developed over the past seven centuries depend crucially on comparative-historical Tai background to provide plausible starting points for presumed Thai-internal diachronic changes. These hypotheses also refer to orthographic history. Even though some may devalue study of written sources as being institutionally ‘philology’ rather than linguistics, I see no good reasons to disregard orthographic evidence when trying to unravel Thai diachronic sound change.

See Anuman Rajadhon (1961, 1981), Hartmann (1986a), Court (1996) and Diller (1996b 2001a) for historical hypotheses as to how Indic-based orthographies and literary culture, including Thai writing, have developed and spread in the Southeast Asian context. The inventor(s) of Thai writing certainly had Khmer orthography in mind, and perhaps Mon, but significant innovations were introduced as well. These included tone-marking, horizontal rather than vertical representation of clusters, phasing out of redundant vowel-initial graphemes and the creation of new segmental symbols as needed. The latter were usually accomplished through modifications made to existing Khmer-type letters representing similar sounds; thus a new [f] symbol was an enhancement of a given [ph] letter. The impression is strong that the original impetus for Sukhothai script involved great care and attention to Tai sound-system detail along with a semiotic attempt to represent perceived phonetic closeness through graphical similarity. Result: diachronic linguists should take this orthography seriously.

Although it is clear that Thai orthography in general can be traced back to South Indic scripts, intermediate points are still professionally debated. The tradition that Thai writing originated in the reign of King Ramkhaeng of Sukhothai (r. approx. 1279-1298) has been challenged but to date no material evidence of a different origin has been adduced. Also, the historical relations of Thai and Lao scripts are somewhat contentious, although existing material evidence strongly indicates that a Sukhothai-type script had spread east to Lao-speaking areas by the early sixteenth century and constituted the prototype for standard Lao writing. Mon writing, on the other hand, was the basis for Lanna (Northern Thai) script, currently under resuscitation, also known to the east as *tham* (i.e. ‘dharma script’, as it was used for Buddhist texts). Discussion and further sources are included in Danvivathana (1987)

and in Diller (1996a). Hudak (1986) describes official simplifications in Thai spelling during the 1940's, later rescinded.

For those with interest in how lexical, syntactic and semantic aspects of the Thai language has evolved since the Sukhothai era of about seven centuries ago, a good place to start is with the commentaries and texts of Na Nagara and Griswold (1992). They present and discuss key literary sources of the Sukhothai period, although their perspectives are mainly philological and historical. A difficulty encountered in some philological work is the tendency to assume that Sukhothai texts are a direct reflection of earlier stages of Central Thai, or even coincide phonetically with modern Thai, whereas Sukhothai Thai probably stood in a less direct line with it. Note that Brown, 1962, 1965, considers Southern Thai varieties to be the more direct descendants of Sukhothai Thai. To their credit, Na Nagara and Griswold (1992) frequently cite cognate material from non-Central dialects. Weroha (1992) too provides comparisons with local varieties. Mikami (1984) summarizes other features of Thai of the early period, while Bamroograks (1987) makes significant progress in understanding Sukhothai discourse patterns. Prasithrathsint (2007), in a convincing comparative and diachronic study of nominalization, shows that at least the type with prefixal [khwa:m<sup>A2</sup>]- 'matter (of...),' was present in the language of Sukhothai, with nominalization on the increase in succeeding centuries.

For the complex diachronic development of the pronominal system, Strecker's (1984) treatment of Proto-Tai pronouns is a good place to start, with dialect studies such as Filbeck (1973b) useful to keep in mind. Studies like these indicate that the Thai system is the result of substantial diachronic shifting and innovation. Iamchinda (1992, in Thai) is a seven-hundred year survey covering the period over which written data on pronouns are available.

Jones (1971) makes accessible King Chulalongkorn's important essay on Thai titles and ranks, a study nicely complemented by Tingsabath and Prasithrathsint (1986, in Thai), who analyze the use of address terms over some two centuries.

Etymology, areal contact and comparative-historical semantic shifting are insightfully treated by Matisoff (1986, 1992), who analyzes kinship terms and analogues of Thai 'heart/mind' compounds in [cay<sup>33</sup>]. These are considered in the wider Southeast Asian context, however issues are raised that need to be kept in mind when focusing on semantic shifts more specifically in Thai. An especially noteworthy study of the latter type is that of Khanittanan and Placzek (1982), who show how the inherited item [khwan<sup>24</sup>] 'spirit, soul' underwent semantic reorganization when an Indic term [win<sup>33</sup>ya:n<sup>33</sup>] with similar meaning was introduced in Buddhist contexts. For more on etymology, see Burnay and Coedès (1920); Li (1956, 1971, 1977); Anuman Rajadon (1961, 1981); and Thai-language sources such as Na Nakhon (1973).

Turning to diachronic phonology and phonetics regarding segmentals, we can surmise that the vowel system of Central Thai is an area where Khmer comparisons are merited, but inscriptional Khmer is not without its own interpretive challenges (Jacob 1965, etc.). Tai-internal accounts of earlier vowel inventories and of how they have evolved into the current Central Thai system have led to divergent proposals, especially regarding vowel-length (Sarawit 1973; Hartmann 1976b; Li 1977; Brown 1979; Strecker 1983; Luksaneeyanawin 1992). In any event, Sukhothai orthographic vowel distinctions show affinities both with Angkorian inscriptional Khmer and also with modern Thai, even though the varieties may not be in direct linear relationship. Dhananjayananda (1997) calls attention to one difference: what is now the [e]/[e:] distinction is not marked until the seventeenth century. The same general picture perhaps applies to long-short distinctions in low front and back vowels [ae] and [o'] as well. But does this mean that the distinctions were pronounced earlier but not marked in text until later, or that new contrastive articulations originated subsequently as well? Diachronic aspects of [ay] and [aw] sequences are elucidated in work of Bickner (1992). Finally, as L-Thongkum, Teeranon and Intajamornrak (2007) have convincingly established, diachronic consideration of vowel length change and of tonogenesis, considered below, must be considered as intertwined issues.

Work on Thai syntactic shifts has been less controversial but has a potential contribution to make to typological studies. Analyses of Sukhothai discourse by Bamroograks (1987) and Prasithrathsint (2007) have been mentioned. Grammaticalization, mentioned in 3.6 above, has been an important process in Thai diachronic syntax but I am not aware of a comprehensive text-based study organizing just how all relevant changes have occurred historically. Khanittanan (1987a, 1987b, 1988b) and Prasithrathsint (1988, 1996) have produced leading work in establishing how syntactic patterns of written Thai have shifted in the past two centuries or so, whether qualitatively or quantitatively. Prasithrathsint has traced in great detail the increase of passive-like constructions and nominalizations in written sources. Khanittanan, using a succession of royal prose compositions, identifies a number of features that characterized evolving prose writing of the mid-nineteenth century: zero anaphora of understood subjects; topic-initial sentences; paratactic constructions rather than marking with overt conjunctions; lists with quantifiers in final position. These generally reflect features today associated with spoken Thai (Messenger 1980; Chodchoey 1988; Hinds 1988a).

Over time, a new style of formal written Thai has taken shape, characterized by denser nominalization, clausal embedding and other types of subordination and more overt specification of nominals, rather than leaving zero anaphors to be construed. This has coincided with increased normative interest in specifying what is ‘correct’ Thai (Diller 1993, 2001a). Lexical selections also play a role here, with higher diglossic choices and technical Indic vocabulary characteristic (Wan Waithayakorn 1970; Prasithrathsint 1994). Resulting in what Khanittanan (1988b) refers to as a more ‘autonomous’ style, discourse of this type is comparatively depersonalized and even aloof from Thai interpersonal social dynamics. ‘Autonomous’ Thai does not encode a range of interpersonal communicative factors the way the lower colloquial style typically does with particles, finely calibrated address-reference selections and other choices. This style of written Thai is currently maintained, but as mentioned in [subsection 3.4.2](#) above, written Thai genres also admit a more colloquial type of Thai as well.

Contrasts between these written styles are sometimes manipulated for special effects, such as to produce eye-catching journalistic headlines in a lower more oral style, while providing ‘serious’ content reporting in the higher more autonomous style (Khanittanan 2007). Also, some Thai authors are adroit in exploiting the difference through stark contrasts in descriptive passages versus dialogue; Chat Kopchitti would be an example. On the other hand, many educated Thai speakers (academics in particular) can, when occasion requires it, produce a spontaneous type of oral discourse showing many features of the higher written style. A methodological consequence of this syntactic vibrancy seems to be that native Thai speakers may disagree with each other as to grammaticality judgments (Diller and Khanittanan 2002). One can hope that future scholarship will contribute to more understanding of these interacting oral and written genres and of their role in the constitution of Thai linguistic competence.

### 3.4.6. Tonogenesis and its quandaries

In diachronic phonology, one area where an understanding of Tai-wide issues impinges crucially on the more parochial history of Central Thai concerns tonal development. Also, since some of the crucial evidence generally accepted for sound changes is from written texts, matters of orthographic interpretation and development become involved as well.

The Thai writing system has been subject to diachronic phonology in some ways and resistant to it in others. Methodologically, it has been taken as important tool in uncovering and analyzing sound changes relating to tone, thus meriting linguistic concern. Most authorities posit an original (or Proto-Tai) system of three tones for open syllables and limited tonal options for stopped syllables; in general, they took one of the open tones (Li 1977). A sweeping devoicing process is thought to have subsequently occurred affecting voiced obstruents, along with loss of aspiration in formerly aspirated sonorants. For Thai (and Lao), the devoiced stops further acquired aspiration. Lexical items with the newly-arising initials were kept from merger by consonant-induced allophonic tonal distinctions, which then