

THE PHONOLOGY OF PROTO-TAI

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Pittayawat Pittayaporn

August 2009

© 2009 Pittayawat Pittayaporn

THE PHONOLOGY OF PROTO-TAI

Pittayawat Pittayaporn, Ph. D.

Cornell University 2009

Proto-Tai is the ancestor of the Tai languages of Mainland Southeast Asia. Modern Tai languages share many structural similarities and phonological innovations, but reconstructing the phonology requires a thorough understanding of the convergent trends of the Southeast Asian linguistic area, as well as a theoretical foundation in order to distinguish inherited traits from universal tendencies, chance, diffusion, or parallel development.

This dissertation presents a new reconstruction of Proto-Tai phonology, based on a systematic application of the Comparative Method and an appreciation of the force of contact. It also incorporates a large amount of dialect data that have become available only recently. In contrast to the generally accepted assumption that Proto-Tai was monosyllabic, this thesis claims that Proto-Tai was a sesquisyllabic language that allowed both sesquisyllabic and monosyllabic prosodic words.

In the proposed reconstruction, it is argued that Proto-Tai had three contrastive phonation types and six places of articulation. It had plain voiceless, implosive, and voiced stops, but lacked the aspirated stop series (central to previous reconstructions). As for place of articulation, Proto-Tai had a distinctive uvular series, in addition to the labial, alveolar, palatal, velar, and glottal series typically reconstructed. In the onset, these consonants can combine to form tautosyllabic clusters or sequisyllabic structures. Regarding the rime, PT had seven vowel qualities that contrasted in height, backness, and rounding. A vowel length contrast also existed for each quality. Palatal

and lateral consonants also occurred in the coda in addition to the final consonants generally assumed. Furthermore, Proto-Tai was a tonal language whose four tonal categories *ABCD contrasted both in terms of pitch and voice quality.

Many of these Proto-Tai traits are not attested in modern Tai languages. The current reconstruction of Proto-Tai phonology is thus a demonstration of the power of the Comparative Method as well as the role that phonological theory can play in reconstruction. This thesis presents a picture of the history of Tai languages as characterized by divergent changes overridden by waves of convergent trends that transformed Proto-Tai into a network of typologically homogenous dialects that differ markedly from their parent. This analysis offers a comprehensive account of the transformation of the Proto-Tai phonology into modern systems.

BIOGRAPHICAL SKETCH

Pittayawat Pittayaporn, known among friends as Joe, attended Sarawittaya High School, Bangkok, before entering the Faculty of Arts at Chulalongkorn University, in 1998. He graduated with a Bachelor of Arts in Spanish (1st class honor) in 2002.

*pācerācariyā honti guṇuttarānusāsakā
paññāvuḍḍhikare te te dinnovāde namāmihaṃ*

Teachers and teachers upon teachers are superior instructors because of their merits.

I bow to those who provide instruction; who inspire knowledge and prosperity.

ACKNOWLEDGMENTS

Writing this dissertation was a long process full of mysteries and discoveries, as well as smiles, distresses, laughter. This dissertation represents moments shared with so many people whom this research brought into my life, and for whom I am deeply grateful. Above all, I am greatly indebted to my beloved family—my father Sanan, my mother Issayaporn, my brother Norrasate, and my aunt Manee—for their love and sacrifices made me who I am. I would like to thank them for their support and incredible patience throughout my life.

I offer this dissertation as a token of my utmost respect and deepest gratitude toward all my teachers. It is my greatest honor and privilege to be a student of my thesis advisor Professor John Whitman. He is a careful, thorough, and inspiring scholar, as well as a caring, attentive, and generous mentor. Words cannot describe how much I value those late evening meetings, those lunch discussions, those letters of recommendations, and above all the friendship he has offered. I am greatly indebted to Professor Pranee Kullavanijaya, who introduced me Tai historical linguistics, and has been a teacher to me in the truest sense of the word. I would like to express my appreciation for Professor Theraphan Luangthongkum, who imparted to me her incredible knowledge on Southeast Asian Linguistics, and who inspired me to be a good field linguist. I am tremendously indebted to Professor MR. Kalaya Tingsabadh, who has always been a respected teacher and benefactor to me. Without her kindness, my life would be very different. This dissertation is also an homage to Fang-kuei Li, William Gedney, André-Georges Haudricourt, Paul Benedict, and other scholars before me whose invaluable contributions made this dissertation possible.

I am indebted to all my language consultants who generously and patiently worked with me: Pan and Jai Kaewkhamsaen of Waritchapum; Dao Van Vinh, and

Hoang Thi Du of Sapa; Hoang Thi Cu of Bao Yen and family; Prof. Hoang Van Ma; Li Jingsen, You Zhouhui of Shangsi and their families; and Chen Yufang and Chen Yujiao of Jinxiu. My experience in the field was a memorable one. I was especially touched by the generosity and kindness of Ajarn Pan Kaewkhamseen, Hoang Thi Cu, Zhouhui, and their families for their warm welcome to their home. In addition, I thank the Department of Linguistics, Vietnam National University, Hanoi; the Office of Culture and Information, Lao Cai Province; and the School of Foreign Studies, Guangxi University for Nationalities, without whose assistance my fieldwork would not have been possible. I am also very grateful to my friend Marc Brunelle who was there at the beginning and helped bring my field research in Vietnam to light.

I would like to express my appreciation to my thesis committee members Prof. Abigail Cohn, Prof. Michael Weiss, and Prof. Draga Zec for their insightful comments, valuable suggestions, and helpful advice. Moreover, this dissertation would not be possible without generous help from so many people at various stages of my Ph.D. experience: Prof. John Wolff, Prof. James Matisoff, Prof. Michel Ferlus, Prof. Laurent Sagart, Prof. Maneepin Phromsuthirak, Prof. Prapod Assavavirulhakarn, Prof. Thak Chaloeontiarana, Prof. Kingkarn Thepkanjana, Prof. Somsongse Burusphat, Prof. Mitsuaki Endo, Prof. Jerrold Edmondson, Prof. Anthony Diller, Prof. Tran Tri Doi, Prof. Zhang Youjun, Prof. Fang Yin, Prof. Bo Wenzhe, Prof. Huang Xiulian, Dr. Tran Huu Son, Dr. Phichet Saiphan, Dr. Alexis Michaud, Dr. Peter Norquest, Dr. Weera Ostapirat, Komsorn Aukyapisut (P'Pop), Mongkol Tonphuban, Qin Xiuhong, You Huicai and family, Lu Sheng, Do Quang Son, Phan Luong Hung, Ryuichi Kosaka, Eric Johnson, Andrew Joseph, Sheila Haddad, Angie Tinti, Nancy Loncto, Wendy Treat, Gina Giambattista Cesari, Renee Milligan, Suwanna Charoensiri, Natcha Preechawattanasakul, Supaporn Phoprom, Nares Petcharat, Vu Huyen, Supawat Chomchan, Liu Futing, Zheng Wuyao, and numerous others.

I also extend special thanks to all my friends for their help, moral support, and, above all, their friendship—Nawaporn Inkaew, Arthid Sheravanichkul, Youththachai Vitheekol, and my friends from the Faculty of Arts for their long-lasting friendship and for being my strength in these difficult years; Woraporn Sukhumawasi, Pimnara Hirankasi, Benjarong Suwankiri, Ariya Chiralertpong, Worawat Sangthongsakaw, and Neeranat Photiyanon, for their wonderful friendship and for always being there for me; Tejaswini Deoskar, John Duong Phan, Marc Brunelle, Ivan Small, Alice Iida, Pamela Corey, Eileen Vo, Johanna Brugman, Andrew Joseph, Dan Kaufman, Andrew Johnson, Daena Funahashi, Claudine Ang, Samson Lim, Nikola Predolac, and Hye-Sook Lee for their ever-growing friendship, and for making these Ithaca years of my life marvelous ones. One thing in life that I am most grateful for is that I have friends, who are wonderful and true. It is not possible to mention all of them on this piece of paper, but they will always be in my heart.

I am indebted to Cornell's Graduate School, Cornell's East Asia Program, and the Lam Family Fund for South China Research for their generous financial support which made this dissertation, especially the field research possible. I express my appreciation to the Department of Linguistics, Cornell University for all its support. It is this department that has nurtured me intellectually in these past six years. I am indebted to Cornell's Southeast Asia Program not only for their financial support but also for being my "home" at Cornell. I express my deepest gratitude to the Anandamahidol Foundation under the Royal Patronage of His Majesty the King of Thailand for giving me the opportunity to pursue this Ph.D. study. I would like to pay my utmost respect and gratitude to His Majesty King Bhumibol Adulyadej and Her Royal Highness Princess Sirindhorn for their vision in recognizing the importance of education for the well-being of the people of Thailand.

TABLE OF CONTENTS

Bibliographical sketch	iii
Dedication	iv
Acknowledgments	v
List of figures	xiii
List of tables	xiv
List of abbreviations	xix
Chapter 1	1
1.1 Introduction.....	1
1.2 Tai languages.....	2
1.3 Conventional reconstruction of PT and pending issues.....	7
1.3.1 Li’s reconstruction.....	7
1.3.2 Outstanding problems.....	13
1.4 Methodology, data, and notation.....	16
1.4.1 Methodology.....	16
1.4.2 Data and fieldwork.....	18
1.4.2.1 Published materials.....	20
1.4.2.2 Fieldwork.....	20
1.4.2.3 Representative dialects.....	21
1.4.2.4 Tai-external data.....	21
1.4.3 Notation.....	22
1.5 Overview of dissertation.....	26
Chapter 2	28
2.1 Introduction.....	28
2.2 The prosodic word.....	30
2.2.1 The importance of the prosodic word.....	31
2.2.2 Well-formedness.....	32
2.2.2.1 Size requirements.....	32
2.2.2.2 Sonority.....	34
2.2.2.3 Foot typology.....	37
2.3 Reconstructing the PT prosodic word.....	39
2.3.1 The PT prosodic word in earlier proposals.....	40
2.3.2 Sesquisyllabicity in PT.....	43
2.4 The prosodic word in Southeast Asia.....	52
2.4.1 Monosyllabic languages.....	53
2.4.2 Disyllabic languages.....	56
2.4.3 Sesquisyllabic languages.....	59
2.5 PT as a sesquisyllabic language.....	63
2.5.1.1 Onset.....	65
2.5.1.2 Rime.....	66
2.5.1.3 Tone.....	67
2.6 Summary.....	68

Chapter 3	69
3.1 Introduction	69
3.2 Inventory of PT simple onsets	69
3.3 A richer dorsal series	72
3.4 “Voicing alternation” revisited	77
3.5 Lack of aspiration in PT	88
3.5.1 Li’s simple aspirates as post-PT innovations	89
3.5.2 Aspiration of original PT clusters	93
3.6 Reconstructions of PT simple initials	101
3.6.1 Voiceless stops	101
3.6.1.1 *p-	101
3.6.1.2 *t-	102
3.6.1.3 *c-	102
3.6.1.4 *k-	103
3.6.1.5 *q-	104
3.6.2 Implosive/glottalized	105
3.6.2.1 *ɓ-	106
3.6.2.2 *ɗ-	107
3.6.2.3 *ʄ-	108
3.6.2.4 *ʔ-	109
3.6.3 Voiced stops	110
3.6.3.1 *b-	110
3.6.3.2 *d-	111
3.6.3.3 *ɗ-	112
3.6.3.4 *g-	113
3.6.3.5 *G-	114
3.6.4 Voiceless fricatives	115
3.6.4.1 *s-	115
3.6.4.2 *ɕ-	116
3.6.4.3 *x-	117
3.6.4.4 *ɣ-	118
3.6.4.5 *h-	118
3.6.5 Voiced fricatives	119
3.6.5.1 *z-	119
3.6.5.2 *ʒ-	120
3.6.5.3 *ʝ-	121
3.6.6 Voiced nasals	122
3.6.6.1 *m-	122
3.6.6.2 *n-	123
3.6.6.3 *ɲ-	124
3.6.6.4 *ŋ-	124
3.6.7 Voiceless nasals	125
3.6.7.1 *hm-	126
3.6.7.2 *hn-	126
3.6.7.3 *hɲ-	127

3.6.7.4	*hŋ-	127
3.6.8	Voiced liquids and glides	129
3.6.8.1	*r-	129
3.6.8.2	*l-	130
3.6.8.3	*w-	131
3.6.9	Voiceless liquids and glides	133
3.6.9.1	*hr-	133
3.6.9.2	*hl-	134
3.6.9.3	*hw-	134
3.7	Summary	136
Chapter 4		137
4.1	Introduction	137
4.2	Tautosyllabic clusters	138
4.2.1	Clusters with medial *-r-	140
4.2.1.1	*pr-	140
4.2.1.2	*tr-	141
4.2.1.3	*cr-	142
4.2.1.4	*kr-	143
4.2.1.5	*qr-	144
4.2.1.6	*br-	145
4.2.1.7	*gr-	146
4.2.2	Clusters with medial *-l-	147
4.2.2.1	*pl-	147
4.2.2.2	*kl-	148
4.2.2.3	*ʃl-	149
4.2.2.4	*gl-	150
4.2.3	Clusters with medial *-w-	151
4.2.3.1	Velar + *-w-	151
4.2.3.2	Uvular + *-w-	152
4.2.3.3	Labial + *-w-	153
4.2.3.4	Coronal + *-w-	155
4.2.3.5	Three-consonant clusters *CCw-	157
4.3	Speculative *-w-	158
4.4	Sesquisyllabic clusters	162
4.4.1	Voiceless stop + voiceless stop	162
4.4.2	Voiceless obstruent + voiced stop	165
4.4.3	Voiced obstruent + voiceless stop	167
4.4.4	Voiceless stops + liquids/glides	169
4.4.5	Voiced consonant + liquid/glide	172
4.4.6	Clusters with non-initial nasals	175
4.4.7	Other clusters	177
4.5	Reduction of sesquisyllabic clusters	179
4.5.1	Processes of cluster adjustment	181
4.5.2	Monosyllabification of sesquisyllables	185
4.6	Summary	189

Chapter 5	190
5.1 Introduction	190
5.2 Inventory of PT rimes.....	191
5.3 PT vowels: quality and quantity	194
5.4 Solving Gedney’s Puzzle.....	198
5.5 Enriching the codas	208
5.6 Reconstructions of PT rimes	213
5.6.1 Rimes with high vowels.....	213
5.6.1.1 Rimes with *i and *i:	214
5.6.1.2 Rimes with *u and *u:	216
5.6.1.3 Rimes with *u and *u:.....	220
5.6.2 Rimes with mid vowels.....	222
5.6.2.1 Rimes with *e and *e:.....	223
5.6.2.2 Rimes with *ɤ and *ɤ:	226
5.6.2.3 Rimes with *o and *o:.....	228
5.6.3 Rimes with low vowel *a and *a:	231
5.6.4 Rimes with diphthongs	234
5.7 Summary.....	236
Chapter 6	238
6.1 Introduction	238
6.2 3+1: The PT tonal system.....	239
6.3 Directionality of tone change	254
6.3.1 Segment-tone interaction	255
6.3.2 Contextual variation.....	256
6.3.2.1 Contour reduction.....	257
6.3.2.2 Peak sliding	257
6.3.3 Perceptual maximization.....	258
6.4 Tonal reflexes of PT in daughter varieties	259
6.5 Reconstructing the contrasts.....	267
6.5.1 Comparing *ABCD in daughter languages	267
6.5.2 Vowel duration	271
6.5.3 Pitch height	274
6.5.4 Pitch contour	275
6.5.5 Voice quality.....	277
6.5.6 PT as a “pitch-plus voice quality” tone language	282
6.6 From PT tones to modern reflexes	282
6.7 Summary.....	285
Chapter 7	286
7.1 Introduction	286
7.2 Comparison of PT reconstructions	286
7.3 Tai subgrouping.....	297
7.4 PT phonology in Kra-Dai context	304
7.5 Convergence of Tai varieties.....	308
7.6 Conclusion.....	311

Appendix A	313
Appendix B	323
Appendix C	378
References	391

LIST OF FIGURES

Figure 1-1	The Kra-Dai family tree according to Edmondson and Solnit (Edmondson & Solnit 1997: 2)	5
Figure 1-2	Proposals for subgroup structure of Tai	6
Figure 1-3	Tai dialects included in this study (see Appendix A).....	19
Figure 2-1	Sonority scale (based on Clements 1990: 292)	34
Figure 2-2	Foot typology according to Hayes (1995)	38
Figure 2-3	Phonological representation of PT prosodic word	65
Figure 3-1	Development of dorsal onsets in White Thai, Kapong, and Yay	75
Figure 3-2	Development of dorsal onsets in White Thai, Debao, and Yay	77
Figure 4-1	Weakening of *p.t- and *Ç.d-	182
Figure 4-2	Implosivization of *t.n- and *Ç.d-	183
Figure 4-3	Metathesis of *r.t-, and *p.q-.....	184
Figure 4-4	Assimilation of *p.t-.....	184
Figure 4-5	Simplification of *t.h-	185
Figure 4-6	Monosyllabification of *k.r- and *p.r-	186
Figure 4-7	Monosyllabification of *k.t- and *t.h-.....	187
Figure 4-8	Monosyllabification of *p.t-, *Ç.d-, and *Ç.t-.....	188
Figure 5-1	Proposed solution for “Gedney’s Puzzle”	206
Figure 6-1	Binary register split and neutralization of phonation type	248
Figure 6-2	Schematization of the mechanisms of tone change	256
Figure 7-1	Tai subgrouping according to Li (1960).....	297
Figure 7-2	Tentative subgrouping of Tai dialects	298

LIST OF TABLES

Table 1-1	PT consonant inventory according to Li (1977).....	8
Table 1-2	PT consonant clusters according to Li (1977).....	9
Table 1-3	PT inventory of simple vowels according to Li (1977).....	10
Table 1-4	PT inventory diphthongs according to Li (1977).....	11
Table 1-5	PT rimes with final glides according to Li (1977).....	12
Table 1-6	Chart showing the laryngeal features of the initials at the time of tone split and the original PT tone categories.....	25
Table 2-1:	Reconstruction of certain etyma under different assumptions about the PT prosodic word.....	43
Table 2-2:	Correspondences involving PT *d-.....	44
Table 2-3:	Correspondences with *q-.....	45
Table 2-4:	Correspondences involving *br- and *j-.....	47
Table 2-5:	Correspondences involving labial clusters.....	48
Table 2-6	Permissible onsets in Dehong Tai (based on Luo 1999: xiii).....	54
Table 2-7:	Complex onsets in Nhaheun (edited from Ferlus 1998: ii).....	55
Table 2-8	Complex onsets in Northern Ronglai.....	58
Table 2-9:	Onsets and syllables in Kammu (from Pittayaporn 2005c).....	61
Table 2-10	Canonical shape of PT prosodic word.....	64
Table 3-1	PT consonant inventory.....	70
Table 3-2	Voiceless dorsal correspondence sets.....	73
Table 3-3	Voiced dorsal correspondence sets.....	76
Table 3-4	Examples of voicing alternation.....	78
Table 3-5	Voicing alternation items that show additional irregularities.....	80
Table 3-6	Voicing alternation items that have clear Chinese origins.....	82
Table 3-7	Voicing alternation items with possible Chinese origins.....	84
Table 3-8	Distribution patterns regarding original voicing of some voicing-alternation items.....	85
Table 3-9	Voicing alternation items that may have been part of PT their Kra-Dai cognates.....	87
Table 3-10	Forms with Li's simple aspirated onsets not found in NT.....	90
Table 3-11	Forms with Li's simple aspirated onsets that are found in all the three groups.....	92
Table 3-12	PT clusters with medial */-r-/.....	94
Table 3-13	Examples of PT sesquisyllabic clusters that are reflected with aspiration.....	96
Table 3-14	Clusters consisting of a stop plus a liquid in PSWT.....	98
Table 3-15	Schematization of the development of PSWT stop system from PT.....	99
Table 3-16	Forms showing aspirated initials in Chinese but unaspirated in Tai.....	100
Table 3-17	Etyma with PT *p-.....	102
Table 3-18	Etyma with PT *t-.....	102
Table 3-19	Etyma with PT *c-.....	103
Table 3-20	Etyma with PT *k-.....	104

Table 3-21	Etyma with PT *q-	105
Table 3-22	Etyma with PT *ǵ-	106
Table 3-23	Etyma with PT *ǵ-	107
Table 3-24	Etyma with PT *ʔj-	108
Table 3-25	Some etyma that Li reconstructs with *j-	109
Table 3-26	Etyma with PT *ʔ-	110
Table 3-27	Etyma with PT *b-	111
Table 3-28	Etyma with PT *d-	112
Table 3-29	Etyma with PT *j-	113
Table 3-30	Etyma with PT *g-	114
Table 3-31	Etyma with PT *ǵ-	115
Table 3-32	Etyma with PT *s-	116
Table 3-33	Etyma with PT *x-	117
Table 3-34	Etyma with PT *χ-	118
Table 3-35	Etyma with PT *h-	119
Table 3-36	Etyma with PT *z-	120
Table 3-37	Etyma with PT *z-	121
Table 3-38	Etyma with PT *ɣ-	122
Table 3-39	Etyma with PT *m-	123
Table 3-40	Etyma with PT *n-	123
Table 3-41	Etyma with PT *ɲ-	124
Table 3-42	Etyma with PT *ŋ-	125
Table 3-43	Etyma with PT * ^h m-	126
Table 3-44	Etyma with PT * ^h n-	127
Table 3-45	Etyma with PT * ^h ɲ-	127
Table 3-46	Etyma with PT * ^h ŋ-	128
Table 3-47	Etyma with PT *r-	130
Table 3-48	Etyma with PT *l-	130
Table 3-49	Etyma with PT *w-	132
Table 3-50	Some etyma that Li reconstructs with *w-	132
Table 3-51	Etyma with PT * ^h r-	134
Table 3-52	Etyma with PT * ^h l-	134
Table 3-53	Etyma with PT * ^h w-	135
Table 4-1	CC- clusters in PT	139
Table 4-2	Etyma with PT *pr-	141
Table 4-3	Etyma with PT *tr-	142
Table 4-4	Etyma with PT *cr-	143
Table 4-5	Etyma with PT *kr-	144
Table 4-6	Etyma with PT *qr-	145
Table 4-7	Etyma with PT *br-	146
Table 4-8	Etyma with PT *gr-	146
Table 4-9	Etyma with PT *pl-	148
Table 4-10	Etyma with PT *kl-	149
Table 4-11	Etyma with PT *ǵl-	150
Table 4-12	Etyma with PT *kw-, *xw-, *ɣw- and *ŋw- respectively	152

Table 4-13	Etyma with PT *qw-, and *gw- respectively.....	153
Table 4-14	Etyma with PT *mw-.....	154
Table 4-15	Etyma with PT *bw-, *pw-, and * ^h mw-.....	155
Table 4-16	Etyma with PT *rw-.....	156
Table 4-17	Etyma with PT *tw-, *nw-, * ^h rw-, and * ^h lw- respectively.....	157
Table 4-18	Etyma with PT *krw-, *klw-, *q ^r w-, *trw-, and *crw- respectively ...	158
Table 4-19	Etyma with PT *wu:.....	159
Table 4-20	Etyma with PT *-wi:-.....	160
Table 4-21	Etyma with PT *we:, *wɛ:, and *wo:.....	161
Table 4-22	Etyma with PT *p.t-.....	163
Table 4-23	Etyma with PT *k.t-.....	164
Table 4-24	Etyma with PT *p.q- and *q.p-.....	164
Table 4-25	Etyma with PT *C̣.b-.....	165
Table 4-26	Etyma with PT *C̣.d-.....	166
Table 4-27	Etyma with PT *C̣.t-.....	168
Table 4-28	Etyma with PT *C̣.k-, and *C̣.q- respectively.....	168
Table 4-29	Etyma with PT *k.r-.....	169
Table 4-30	Etyma with PT *p.r-.....	170
Table 4-31	Etyma with PT *C̣.w-.....	172
Table 4-32	Etyma with PT *m.l-.....	173
Table 4-33	Etyma with PT *C̣.r-.....	173
Table 4-34	Etyma with PT *C̣.l-.....	174
Table 4-35	Etyma with PT *t.n-.....	176
Table 4-36	Etyma with PT *C̣.n-.....	177
Table 4-37	Monosyllabification in Tsat (Sagart 1993).....	180
Table 5-1	Inventory of PT simple vowels.....	192
Table 5-2	Inventory of PT diphthongs.....	192
Table 5-3	Inventory of PT codas.....	193
Table 5-4	PT vowel system according to Li (1977) and Sarawit (1973).....	195
Table 5-5	Distribution of PSWT non-high vowels.....	197
Table 5-6	Items that unequivocally go back to PT */i:/, */u:/, and */u:/, i.e. the ‘year’ set.....	199
Table 5-7	Items that unequivocally go back to PT */aj/, */au/, and */aw/, i.e. the ‘egg’ set.....	200
Table 5-8	The ‘fire’ set of Gedney’s Puzzle correspondences.....	201
Table 5-9	The ‘long’ set of Gedney’s Puzzle correspondences.....	202
Table 5-10	Comparison of solutions to “Gedney’s Puzzle”.....	203
Table 5-11	Types of languages allowed by Li’s and Sarawit’s account.....	204
Table 5-12	Tai varieties that do not belong to type 2 or type 3.....	205
Table 5-13	Etyma with irregular reflexes usually discussed under “Gedney’s Puzzle”.....	207
Table 5-14	Shared etyma with final /-l/ in Saek and Kra.....	209
Table 5-15	Etyma with PT */en/ and */-el/.....	210
Table 5-16	Etyma with final */-c/.....	211
Table 5-17	Final *-c in Tai and Be.....	213

Table 5-18	Etyma with PT *i:#	214
Table 5-19	Etyma with PT *iC and *i:C	215
Table 5-20	Etyma with PT *u:#	216
Table 5-21	Etyma with PT *u:n, *u:n, and *u:l respectively	217
Table 5-22	Etyma with PT *u:m, *u:m, *u:ŋ, and *u:ŋ respectively	218
Table 5-23	Etyma with PT *u:p, *u:t, and *u:k respectively	219
Table 5-24	Etyma with PT *wu:#	220
Table 5-25	Etyma with PT *u:#	220
Table 5-26	Etyma with PT *uC and *u:C respectively	221
Table 5-27	Etyma with PT *uk	222
Table 5-28	Etyma with PT *e:#	223
Table 5-29	Etyma with PT *eC	224
Table 5-30	Etyma with PT *et, *en, and *el respectively	225
Table 5-31	Etyma with PT *e:C	225
Table 5-32	Etyma with PT *ɣC	226
Table 5-33	Etyma with PT *ɣ:C	227
Table 5-34	Etyma with PT *ɣ:# and *ɣ:l	228
Table 5-35	Etyma with PT *o:#	229
Table 5-36	Etyma with PT *oC	230
Table 5-37	Etyma with PT *ok, and *oŋ respectively	230
Table 5-38	Etyma with PT *o:C	231
Table 5-39	Etyma with PT *a:#	232
Table 5-40	Etyma with PT *aC and *a:C respectively	233
Table 5-41	Etyma with PT *iə, *uə, and *uə respectively	234
Table 5-42	Etyma with PT *iəw	235
Table 5-43	Etyma with PT *au	236
Table 6-1	Correspondences between Tai and Chinese tones	240
Table 6-2	Phonological factors commonly found to condition tonal splits and mergers (modified from Gedney 1972a)	243
Table 6-3	Pattern of tonal splits and mergers in White Tai	244
Table 6-4	Pattern of tonal splits and mergers in Western Nung	244
Table 6-5	Pattern of tonal splits and mergers in Southern Shan	245
Table 6-6	Pattern of tonal splits and mergers in Southern Thai	246
Table 6-7	Pattern of tonal splits and mergers in Po-ai	246
Table 6-8	Pattern of tonal splits and mergers in Cao Bang	247
Table 6-9	Western Nung tonal inventory	250
Table 6-10	Pattern of tonal splits and mergers in Cao Bang	260
Table 6-11	Pattern of tonal splits and mergers in Leiping	261
Table 6-12	Pattern of tonal splits and mergers in Siamese	262
Table 6-13	Pattern of tonal splits and mergers in Black Tai	263
Table 6-14	Pattern of tonal splits and mergers in Wuming	264
Table 6-15	Pattern of tonal splits and mergers in Liujiang	265
Table 6-16	Pattern of tonal splits and mergers in Lianshan	266
Table 6-17	PT *A in individual varieties	268
Table 6-18	PT *B in individual varieties	268

Table 6-19	PT *C in individual varieties	269
Table 6-20	Phonetic characteristics of PT tones	271
Table 6-21	Connection of *D to *B and *C	272
Table 6-22	Siamese tones correlated with average vowel durations in msec (Abramson 1962: 107)	273
Table 6-23	Pattern of tonal splits and mergers in Nung Fan Slihng	280
Table 7-1	PT consonants and their correspondents in previous reconstructions ..	288
Table 7-2	PT consonants and their correspondents in previous reconstructions ..	290
Table 7-3	Examples of PT sesquisyllabic clusters and their correspondents in previous reconstructions	291
Table 7-4	PT rimes and their correspondents in previous reconstructions	293
Table 7-5	Phonological criteria for subgrouping of Tai languages	299
Table 7-6	Proto-Tai and Sui reflexes of PKD back unrounded vowel.	306

LIST OF ABBREVIATIONS

CT = Central Tai

LC = Lungchow

LMC = Late Middle Chinese

MC = (Early) Middle Chinese

NT = Northern Tai

O = obstruent

OC = Old Chinese

PAN = Proto-Austronesian

PIE = Proto-Indo-European

PKD = Proto-Kra-Dai

PSWT = Proto-Southwestern Tai

PT = Proto-Tai

S = sonorant

SI = Siamese

SEA = Southeast Asia

SSP = Sonority Sequencing Principle

SWT = Southwestern Tai

Y = Yay

-f = showing irregularity in the coda

-i = showing irregularity in the onset

-t = showing irregularity in the tone

-v = showing irregularity in the vowel

CHAPTER 1

BACKGROUND

1.1 Introduction

The hardest task in comparative linguistics is to distinguish between types of similarities: universal tendencies, chance, diffusion, genetic retention, parallel development (Aikhenvald & Dixon 2001). Only when these different possible types of similarities are considered can we understand the role of common ancestry. The magnitude of this difficulty is reflected in the view that a similarity can be explained as shared inheritance only when other possibilities have been exhausted (Dench 2001). Successful reconstructions of proto-languages must therefore overcome this challenge. The reconstruction of Proto-Tai (PT), the ancestor of the Tai languages of Mainland Southeast Asia (SEA) is not an exception. Modern Tai languages share a great deal of structural similarities and identifiable phonological innovations. Yet, trying to reconstruct the phonology of PT has proven more challenging than it first seems. This is precisely because Tai languages have continually undergone the different waves of convergent trends that swept through the Mainland Southeast Asian linguistic area.

In particular, various researchers have put forward their own reconstructed phonology of PT but these proposals disagree on major issues due in part to assumptions that they hold, be it explicitly or implicitly. These assumptions arose precisely from different interpretations of shared similarities among modern Tai languages by different researchers. Unfortunately, some of these assumptions have become axioms in the study of the history of Tai languages, and have thus kept those issues unsolved. With this in mind, this dissertation presents a new phonological reconstruction of PT constrained by explicit and motivated hypotheses. Essentially, it incorporates dialect data that have become increasingly available and discusses

explicitly the interpretation of these data. As suggested by the title, this dissertation takes as its task the reconstruction of the PT phonological system rather than detailed analyses of individual etyma, or a complete lexicon of PT.

The widely accepted PT reconstruction is due to Li (1977). This reconstruction has been the basis of good scholarship on the history of the Tai languages but like other essays on PT reconstruction (Ferlus 1990; Haudricourt 1948; Liang & Zhang 1996; Luo 1997; Nishida 1954; Sarawit 1973), it lacks sufficient consideration of the issues mentioned above. This study goes beyond Li and other historical research on Tai languages in that it takes into serious consideration the structure and the phonotactics of the PT prosodic word. In particular, I claim that PT was a sesquisyllabic language that allowed both sesquisyllabic and monosyllabic prosodic words. The PT prosodic word consisted of three main components: the onset, the rime, and the tone. The rime is, in turn, subdivided into the vocalic nucleus and the coda. While I show that the PT vocalism was much simpler and symmetrical than conventionally reconstructed, I claim that the onset and the coda systems are richer than previously assumed. Furthermore, I propose a concrete albeit preliminary reconstruction of the PT tonal categories.

1.2 Tai languages

Tai is a technical term that is used to refer to various related ethnolinguistic groups in SEA, although only some Tai-speaking groups call themselves “Tai” or its variants. Though pronounced identically in English, “Tai” is used to refer to the group as a whole or to any language belonging to the family, while “Thai” is preserved for the dominant language of Central Thailand, also known as “Siamese”. In SEA five language families have been identified: Sino-Tibetan, Austroasiatic, Austronesian, Hmong-Mien, and Kra-Dai (formerly known as Tai-Kadai). The Tai languages, a sub-

branch of Kra-Dai, will be the subject of this dissertation. The hypothetical parent language from which all Tai languages develop is called Proto-Tai (PT). It has been speculated that PT was spoken in the area around the Guangxi-Vietnam border areas (Diller 2000; Gedney 1995). Nowadays Tai languages are currently spoken in China, Vietnam, Laos, Cambodia, Thailand, Malaysia, Myanmar, and India. The best known members of this group are Thai, Lao, and Shan, of which the first two are the official languages of Thailand and Laos respectively. Some other better-known ethnonyms used to refer to Tai-speaking groups include Nung, Thu, Zhuang, Bouyei, Giay, and Saek (Edmondson & Solnit 1997). These Tai languages occupy the geographical center of the Southeast Asian mainland, and have been active adopters and spreaders of many areal traits with languages to their north, e.g. Chinese, Tibeto-Burman, and Hmong-Mien, as well as further south, e.g. Austroasiatic and Austronesian (Edmondson & Solnit 1988)

Typologically, modern Tai languages share a number of phonological characteristics (Diller 2000; Gedney 1989b). All Tai languages are basically monosyllabic and tonal. A given modern language can best be described in terms of the distinction made between various parts of the syllables (Gedney 1989b: 13). Each syllable can be divided into onset, a vocalic nucleus, and an optional final consonant. With respect to the onset, most Tai languages have a fairly rich inventory which includes simple onsets and clusters of the stop-glide and/or stop-liquid types. As regards the vowel system, most languages make a three-way distinction in height (high, mid, and low) and a three-way distinction in tongue position and rounding (front, back rounded, and back unrounded). Some languages also show length contrast throughout their vowel systems, while others only contrast short /a/ and long /a:/. Final consonants are very limited in modern Tai languages. Most languages have stops and nasal finals in the labial, dental, and velar places of articulation. Some also have an

additional final glottal stop. Tones in Tai languages always show distinctions in both pitch height and contour. The number of tones in each language ranges from 4-7. In many better-studied varieties, the distributions of the tones are conditioned by the syllable structure; fewer tones can occur in checked syllables, i.e. syllables closed by an obstruent. However, as Diller (2000: 13) correctly points out, these characteristics are found widely throughout SEA.

These Tai languages form a branch of the Kra-Dai¹ language family, formerly known as Tai-Kadai. According to the subgroup structure proposed in Edmondson and Solnit (Edmondson & Solnit 1997: 2) shown in Figure 1-1², the closest relative of Tai according to this classification is the Be language (临高 *Lín'gāo*) of Hainan. The majority of Kra-Dai languages outside of Tai are spoken in the extreme South of China. The Kam-Sui group comprises Kam (侗 *Dòng*), Sui (水 *Shǔi*), Maonan (毛南 *Máonán*), Mulam (仫佬 *Mùlǎo*), Ai-cham (锦 *Jīn*), Then (佯僮 *Yánghuáng*) and Mak (莫 *Mò*). Closely related to Kam-Sui is Lakkja (拉珈 *Lājiā*) and Biao (标 *Biāo*), which form their own subgroups. The Hlai group (黎 *Lǐ*) consists of closely related dialects spoken by indigenous population of Hainan Island. Lastly, the Kra group consists of languages spoken by smaller populations, including Gelao (仡佬 *Gēlǎo*), Lachi (拉基 *Lājī*), Buyang (布央 *Bùyāng*), Pubiao (普标 *Pǔbiāo*), Yerong (耶容 *Yēróng*), and Laha. These smaller languages are scattered in Guizhou, Yunnan, and Guangxi Provinces of China, except for Laha which is spoken in Vietnam.

¹ This label is due to Ostapirat (2000)'s seminal work on the reconstruction of Proto-Kra.

² See Ostapirat (2005: 108) for another possible subgroup structure.

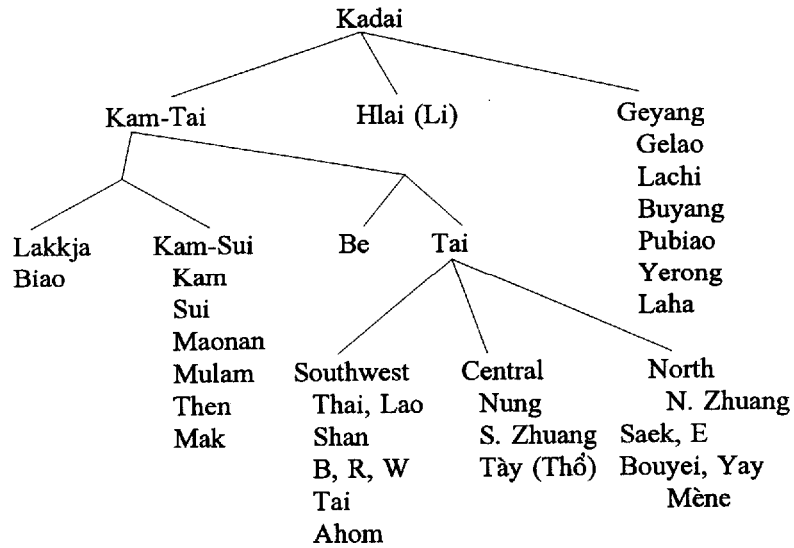


Figure 1-1 The Kra-Dai family tree according to Edmondson and Solnit (Edmondson & Solnit 1997: 2)

Although membership of Tai in the Kra-Dai family is not-controversial, the relationship of Tai to other language families is still open for investigation. Kra-Dai has been grouped with all the other families in SEA, including Hmong-Mien (Kosaka 2002) and Austroasiatic (de Lacouperie 1886), but the most widely accepted views used to be that Tai belongs with Chinese in the Sino-Tibetan family (e.g. Conrady 1896; Grierson 1903; Li 1976; Luo 1997; Manomaivibool 1975; Schmidt 1926; Wulff 1934). However, the “Sino-Tai” view is becoming superseded by the competing view that links Kra-Dai with Austronesian (Benedict 1942, 1975, 1997b; Ostapirat 2005; Sagart 2004, 2005; Schlegel 1902; Wulff 1942). This view is best represented by the “Austro-Tai” hypothesis proposed by Benedict (1942; 1975), who believes that Kra-Dai, Austronesian, and Hmong-Mien form one linguistic stock. Although Benedict’s linguistic evidence and methodology has been severely criticized, his proposal to connect Kra-Dai with Austronesian merits close attention and has found support in more recent works by Ostapirat (2005) and Sagart (2004; 2005). Although this study

does not directly address the issue of the affinity of Kra-Dai, a non-trivial link between Tai and Austronesian will become progressively clearer all through the dissertation.

Within the Tai branch, the conventional classification is that proposed by Li (1960; 1977). It divides Tai varieties into three primary subgroups: Southwestern Tai (SWT), Central Tai (CT), and Northern Tai (NT). The ancestral languages of SWT and NT have been proposed by various researchers (e.g. Brown 1965; Jonsson 1991; Kosaka 1992; Li 1977; Pittayaporn 2008b, to appear-b). However, other alternative subgrouping schemas have also been proposed (Chamberlain 1975; Haudricourt 1956; Luo 1997). Common among many of the alternative view is the recognition of similarities between Li’s SWT and CT as reflected by Haudricourt (1956)’s explicit claims that the two comprise a sister branches of NT³. Li’s and Haudricourt’s classification schemas are given in Figure 1-2.

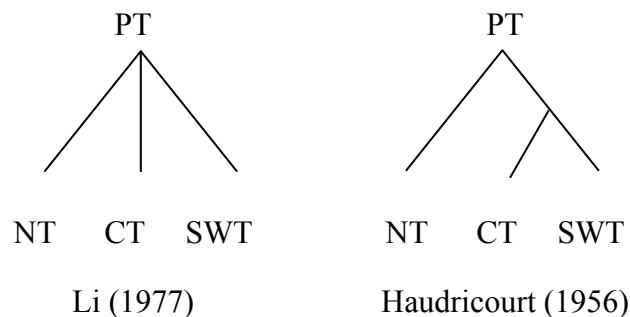


Figure 1-2 Proposals for subgroup structure of Tai

Note that, despite his bipartite subgroup structure, Haudricourt (1956) claimed that a few understudied dialects occupy an intermediate position between SWT/CT and NT. The result of this dissertation will show quite clearly that among Li’s three

³ Haudricourt (1956) uses “Thai” to refer restrictively to the subgroup consisting of Li’s SWT and CT and “Dioi” to Li’s NT.

subgroups SWT is the only valid one (see §7.3). However, Li's subgroup labels are useful and convenient labels in describing the distribution of linguistic traits shared by modern Tai varieties, and help relate the current research to the existing body of literature on Tai languages. Therefore, in this dissertation I use the terms SWT, CT, and NT as generic labels to refer grossly to clusters of languages that share a certain set of developments. In particular, I define the three clusters using the same criteria that Li (1960) uses in his proposal but without adopting his claim that they constitute genealogical subgroups of Tai. The only genealogical subgroup that I assume is SWT, which I have shown elsewhere to be a valid one (Pittayaporn 2008b). The issue of Tai subgrouping will be discussed in Chapter 7.

1.3 Conventional reconstruction of PT and pending issues

Despite a number of competing reconstructions of PT, the most widely adopted is, beyond doubt, the system proposed by Li Fang-Kuei, especially in his *Handbook of Comparative Tai* (Li 1977). Not only does this monumental work embody the advances in the study of Tai languages from the 40's through the 70's, it has also been a basis for Tai linguistics from its first appearance to the present. Even though alternative reconstructions do exist (Ferlus 1990; Haudricourt 1948; Liang & Zhang 1996; Luo 1997; Nishida 1954, 1955; Sarawit 1973), it is not amiss to refer to Li's construction as the conventional reconstruction of PT. In this section, I first outline the conventional reconstruction of PT and then proceed to discuss open issues in the reconstruction of PT phonology.

1.3.1 Li's reconstruction

The conventional reconstruction assumes that the development of PT etyma can be accounted for in terms of single syllables consisting of the onset, the vocalic

nucleus, the optional coda, and the tone (Edmondson & Solnit 1997; Li 1977). In this system, the onset position shows a four-way phonation-type contrast among consonants: 1) aspirated voiceless, 2) plain voiceless, 3) glottalized, and 4) plain voiced. However the phonation-type contrast is neutralized in coda position so that only plain voiceless stops and voiced nasals and glides are allowed. Table 1-1 gives the inventory of consonants of PT as conventionally reconstructed. Those sounds that can occur in the coda position are underlined.

Table 1-1 PT consonant inventory according to Li (1977)⁴

		labial	alveolar	palatal	velar	glottal
stops	aspirated	*p ^h	*t ^h	*c ^h	*k ^h	
	voiceless	*p	*t	*c	*k	*ʔ
	glottalized	*ʔb	*ʔd			
	voiced	*b	*d	*j	*g	
fricatives	voiceless		*s		*x	*h
	voiced		*z		*ɣ	
nasals	voiceless	* <u>m</u>	* <u>n</u>	* <u>ɲ</u>	(* <u>ŋ</u>)	
	voiced	* <u>m</u>	* <u>n</u>	* <u>ɲ</u>	* <u>ŋ</u>	
liquids and glides	voiceless	* <u>w</u>	* <u>r</u> * <u>l</u>	* <u>j</u>		
	voiced	* <u>w</u>	* <u>r</u> * <u>l</u>	* <u>j</u>		

These consonants can be combined into clusters of two consonants. While the initial consonants may be a stop, a fricative, or a nasal, the medial must either be a liquid or a glide. Note that Li (1977: 236-242) posits labiovelar units *kw-, *k^hw-, *gw-, *ŋw-, *xw-, and *ɣw- but states that it is not clear whether they are unitary phonemes or clusters consisting of a velar consonant followed by medial *-w-.

⁴ Li's segments are re-transcribed as discussed in §1.4.3.

Table 1-2 PT consonant clusters according to Li (1977)

		labial	alveolar	palatal	velar	glottal
aspirated stops	-r-	*p ^h l/r	*t ^h r		*k ^h r	
	-l-		*t ^h l		*k ^h l	
	-w-				*k ^h w	
voiceless stops	-r-	*pr	*tr		*kr	
	-l-	*pl	*tl		*kl	
	-w-				*kw	
glottalized	-r-	*ʔbl/r	*ʔdl/r			
	-l-					
	-w-					
voiced stops	-r-	*br	dl		*gr	
	-l-	*bl	dr		*gl	
	-w-				*gw	
voiceless fricatives	-r-	*fr-				
	-l-					
	-w-				*xw	
voiced fricatives	-r-	*vr				
	-l-	*vl				
	-w-				*ɣw	
nasals	-r-	*ml/r	*nl/r			
	-l-					
	-w-	*mw			*ŋw	

In addition to clusters with medial liquids shown above, there are also clusters with medial glides. Li's system has three glides altogether: *j, *w, and *ɥ. While *j, and *w can occur as onset, medial glide, or coda, *ɥ is allowed only medially and finally. This set of medial glides requires special consideration. In *Handbook of Comparative Tai*, Li (1977) posits only two medial consonants explicitly: *-l- and *-r-. However, the three medial glides can be uncovered in different parts of the system. The first and the main part is the vocalism. Li (1977: 298) distinguishes two

types of high vowel in diphthongs—accented *i, *u, and *u on one hand, and unaccented *i, *i, and *u on the other. For example, he reconstructs a contrast between *ie (1977: 280-281) vs. *ie (Li 1977: 269-270). Since such a fine contrast seems implausible, I re-interpret Li’s unaccented *i, *u, and *u as glides *j, *u, and *w respectively. According to this reinterpretation, the contrast between *ie and *ie is in fact a contrast between a diphthong *ie and a sequence *je. Therefore, under this interpretation the conventional reconstruction allows for three medial glides *-j-, *-u-, and *-w-, in addition to the two medial liquids *-l- and *-r-.

As for the rime, the conventional reconstruction of PT has nine simple vowels without length contrast, and a large number of complex vowels. These vowels can be optionally followed by a final consonant from the set of permissible codas shown in Table 1-1 (underlined). Among simple vowels, distinctions are made for three vowel heights, backness, and rounding as shown in Table 1-3.

Table 1-3 PT inventory of simple vowels according to Li (1977)

	front	back	
	unrounded	unrounded	rounded
high	*i	*u	*u
mid	*e	*ɣ	*o
low	*ɛ	*a	*ɔ

These simple vowels can be combined together to form diphthongs. One of the most prominent properties of Li’s inventory of complex vowels is the contrast between diphthongs with accented high vowels and those with unaccented high vowels (Li 1977: 298). However, I consider only the former as true diphthongs in the

conventional reconstruction. This is because the latter is re-interpreted as a sequence of a glide followed by a vowel. For comparison, these sequences are given in parentheses in Table 1-4.

Table 1-4 PT inventory diphthongs according to Li (1977)

		high			mid			low		
		*i	*u	*u	*e	*ɣ	*o	*ɛ	*a	*ɔ
high	*i				*ie (*je)	*iɣ	(jo)	(jɛ)		
	*u		*uu	*uu (*uɰ)	*ue	*uɣ	uo	uɛ	ua (uɰa)	uɔ (uɰɔ)
	*u				*ue	*uɣ (*wɣ)	uo (wo)		ua (wa)	uɔ (wɔ)

In addition to the diphthongs shown in Table 1-4, more diphthongs and triphthongs are included in Li's reconstruction. However, these complex vowels only end in *i, *u, or *u. These final high vowels can be re-interpreted as post-vocalic *-j, *-ɰ, and *-w respectively, parallel to Li's unaccented high vowels in the prevocalic position discussed above. Table 1-5 presents the inventory of these re-interpreted rimes with final glides in the conventional reconstruction.

Table 1-5 PT rimes with final glides according to Li (1977)

		*-j	*-ɥ	*-w
high	*i			*iw
	*ɯ	*ɯj		
	*u			
mid	*e	*ej	*eɥ	*ew
	*ɤ	*ɤj *wɤj	*ɤɥ	*ɤw *jɤw
	*o		*oɥ	*ow
low	*ɛ	*ɛj		*ɛw
	*a	*aj *ɥaj		*aw
	*ɔ	*ɔj		
diphthongs	*iɤ	*iɤj		
	*ia	*iaj		
	*uɤ	*uɤj		
	*ua	*uaj		

In addition to the onset and the rime, each PT word in the conventional reconstruction also has tone *A, *B, *C, or *D. While *A, *B, and *C occurred in syllables ending in a vowel or a nasal, or ‘non-checked’ syllables, *D only occurred in syllables with a final obstruents, or “checked syllables.” Li (1977: 25) states that *D only occurred in checked syllables, in which tonal contrast was neutralized. However, he posits *D as a separate tone because it is impossible to identify it with the one of the other three tones. Moreover, the conventional reconstruction does not provide concrete reconstruction of the abstract tonal categories *A, *B, *C, and *D.

1.3.2 *Outstanding problems*

The conventional reconstruction of PT phonology outlined above has played a major role in the historical study of Tai languages but many questions still await solutions. These problems include the unaccounted onset correspondence sets, the complexity of the vocalic system, the peculiarities of the displaced NT language Saek, spoken in northeastern Thailand, and the abstractness of tones. In this sub-section, I briefly introduce the issues and then assess them from a holistic perspective.

The first problem concerns onset correspondences. Although the conventional reconstruction nicely accounts for the majority of the correspondence sets, minor sets of data are still left as puzzles. The best-known of these sets is probably the so-called “voicing alternation,” which refers to an apparent discrepancy regarding voicing of the onsets in different dialects. This discrepancy is generally described as a disagreement between SWT and CT on one hand, and NT on the other. For a sizeable set of etyma, reflexes in SWT and CT point to earlier voiceless onsets while their NT counterparts indicate original voiced onsets (Gedney 1989a). Take the etyma ‘to arrive’ and ‘ripe, cooked’ as examples. Based on SWT and CT reflexes alone, they should be reconstructed with **t^h* and **s-* respectively. In contrast, they should be reconstructed as **d-* and **z-* respectively if only NT reflexes are taken into consideration. The conventional reconstruction fails to account for this set of data.

In addition to the “voicing alternation” just described, there are many other sets of onset correspondences that still await explanation. These data sets often involve a small subset of etyma but the correspondence seems too systematic to be disregarded as irregular. One example is **tr-* which Li reconstructs for etyma like ‘eye’ and ‘to die’. This cluster is reflected as /t-/ in the SWT language Siamese and the NT dialect of Po-ai, but as /h-/ in the CT dialect of Lungchow. At first glance this choice of reconstruction seems adequate but data from other varieties cast doubt on this

treatment of the correspondence. In particular, the CT dialect of Bao Yen from my own fieldwork regularly has /p^hj-/ for this set of etyma, indicating that the original PT onset must have contained a labial segment. The conventional reconstruction has no way to account for this correspondence.

The second problem for the conventional reconstruction is the complexity of the vowel system. As apparent in the review above, Li's reconstructed PT vocalic system contains a massive set of complex vowels with very fine phonemic contrasts. Even after re-interpreting many of his complex vowels as sequence of vowels and glides, the problem still remains equally damaging. A large part of this problem has to do with the paradigm of data sometimes referred to as "Gedney's Puzzle" (Gedney 1972b). This problematic paradigm of vowel correspondence consists of four sets of vowel correspondences involving high vowels. A number of Tai etyma unequivocally go back to PT simple vowels *i, *u, and *u in the conventional reconstruction. Similarly, another set of etyma clearly go back to Li's *ɣj, *ɣw, and *ɣu. However, in addition to these two sets, there are a number of items that seem to point to earlier *i, *u, and *u in some languages but *ɣj, *ɣw, and *ɣu in others. To complicate the picture, the reverse patterns also exist. The massive vowel system in the conventional reconstruction is in large part an attempt to account for this paradigm of data.

The third challenge for the conventional reconstruction of PT is the aberrancy of Saek, a displaced NT language spoken in northeast Thailand and central Laos. This language shows many peculiarities that cannot be reconciled within the conventional model of PT phonology (Gedney 1989c; Haudricourt 1963b). For example, there are two reflexes of the PT sound that Li reconstructs as *ʔb- in Saek. The first one is the expected /b-/ as in 'leaf' and 'to fly'. The other reflex is /v-/ as in 'shoulder', and 'thin'. This pattern is left unexplained in Li (1977) as there does not seem to have been any conditioning environment for the split of PT *ʔb-. Another example of Saek

peculiarities, is the famous /-l/, found nowhere else in Tai. While recognizing that Saek final -l might have been inherited from PT, Li does not include a final liquid in the reconstruction due to lack of data.

The last but in no way less challenging issue for the conventional reconstruction is the abstractness of PT tones. Although it is well-established that PT had four tonal categories, it is not clear what they were like. In other words, it is unknown what the contrastive characteristics of *A, *B, and *C, were and how *D related to the other three categories. The conventional reconstruction remains silent about the phonological content of these categories.

From a holistic perspective, these problems largely stem from one serious shortcoming. In particular, Li does not address explicitly the structure of the PT prosodic word and its role in the reconstruction. Since words serve as the basic domains for comparative reconstruction, the researcher's assumptions about the phonotactics of the prosodic word in the PT inevitably predetermine the reconstruction. Li's reconstruction assumes at least implicitly that PT was typologically very similar to its modern descendants. Among other things, it assumes that like all modern Tai languages PT was monosyllabic. It also assumes *a priori* that the PT vowel system did not have a quantity contrast. These are assumptions that cannot be taken for granted. Recent research (e.g. Benedict 1982; Brunelle 2008; Ferlus 1975, 1992; Thurgood 1996, 2005) has shown how language contact leads to restructuring of the canonical word shape, how monosyllabic tonal languages arise from languages that have non-tonal polysyllabic canonical roots, etc. For example, all modern Chinese varieties have fully-developed tone systems, but it is uncontroversial that their ancestor Old Chinese was non-tonal (Baxter 1992). Therefore, from a holistic perspective, the issue of the PT prosodic word must not be taken for granted in the search for a solution to these problems.

In addition, these problems result from lack of comparative data. Li's *Handbook of Comparative Tai* (Li 1977) is an honorable piece of work because it brings together lexicographical data from various dialects available during those early decades. However, as Gedney pointed out in 1967, most major Tai-speaking areas were still understudied (Gedney 1989b). Only for Thailand and southern Guizhou were dialect studies on fine geographical grids available. As evidenced by the discovery of Saek, data from previously unknown dialects may lead us to question old assumptions and provide new insights into the reconstruction of PT phonology.

1.4 Methodology, data, and notation

Given the challenges discussed above, it is imperative that assumptions about PT phonology be critically assessed. Moreover, it is of equal importance to minimize gaps in our knowledge of modern dialects. Therefore, this dissertation approaches the reconstruction of PT by using a large set of comparative data that have now become available as a basis for discussing the problematic aspects of the conventional reconstructions. In this section, I first describe the methodology and data used in this dissertation before proceeding to discuss the notation used.

1.4.1 Methodology

The method employed in this dissertation is the Comparative Method which first developed in the 19th century to study the history of Indo-European languages. However, it emphasizes three important methodological facets that are sometimes taken for granted in the reconstruction of proto-languages: 1) use of external evidence, 2) role of diffusion, and 3) subgroup assumptions.

With respect to use of external evidence, in this dissertation, cognates from Kra-Dai languages outside of Tai, shared vocabulary with other language families of

SEA, and typological considerations are used strictly as additional support for reconstructions proposed on the basis of Tai-internal evidence. For example, the reconstruction of final *-l (see §5.5) is motivated by the presence of final /-l/ in Saek and the vowel reflexes in Qinzhou. Final /-l/ preserved in cognates in the Kra language Laha only lends additional yet strong support for positing PT *-l. Similarly, PT *C̣.d- (see §2.3.2 and §4.4.2) was reconstructed solely on the basis of sound correspondences among modern Tai dialects. By citing the Proto-Austronesian form *qudip ‘alive’ and *tidəm ‘dark’, I merely use the non-trivial similarities as additional support. I do not use these Proto-Austronesian forms as evidence for the reconstruction of PT complex onsets. Nor do I claim that Tai is genetically related to Austronesian.

The role of diffusion is also a factor that cannot be ignored in applying the Comparative Method. Certain traits that are shared widely among Tai dialects may in fact be innovations. A good example is the case of PT uvular consonants (see §3.3). Although no attested modern Tai dialects have a distinctive uvular series, cases of velar stop /k-/ in some dialects corresponding to /h-/ in others unequivocally indicate that PT must have had uvular consonants *q-, *χ-, and *g-, which contrasted with velar *k-, *x-, *g-, and *γ-. Another example is PT final *-c (see §5.5), which is not attested in any modern Tai varieties. However, cases where Saek /-k/ corresponds to /-t/ in other varieties must go back to a final palatal *-c. This interpretation is supported by the Kra-Dai languages Be and Hlai. The lack of uvular series and final *-c must undoubtedly be explained as traits that diffused from one Tai dialect to another throughout the Tai-speaking area. These two cases are good illustrations of why one should not underestimate the effect of diffusion in applying the Comparative Method.

Furthermore, the interpretation of sound correspondences in this dissertation does not make references to any particular proposal of genealogical subgrouping of Tai languages. This is because arriving at a subgroup structure without establishing the proto-language is theoretically highly problematic. In particular, interpreting shared similarities as retentions or innovations depends entirely on the reconstructed PT language. As subgrouping arguments take shared innovations as evidence, it is not possible to assume subgroups of Tai languages without having a clear understanding of the proto-language. Although I have established elsewhere (Pittayaporn 2008b) that SWT forms a valid subgroup within Tai (and I will, for presentational purposes, occasionally refer to Proto-SWT), the status of SWT varieties as a subgroup is never used here as evidence for PT reconstruction.

1.4.2 Data and fieldwork

The data used in this dissertation cover a broad geographical area. Figure 1-3 is a map of all dialects included in this study. The varieties included in this dissertation belong to all three dialect groups: SWT, CT, and NT. They represent most of the different geographical areas where Tai languages are spoken, including Thailand, Laos, Malaysia, Myanmar, northeast India, Yunnan, Guangxi, and Guizhou. The only two areas for which availability of data does not permit close examination include Central and Eastern parts of Laos, Central Vietnam, and the interior of northern Vietnam. However, reports on varieties from these areas (Chamberlain 1984; Đoàn 1996; Haudricourt 1960; L'École Française d'Extrême-Orient 1938; Simmonds 1965) have also been taken into account.



Figure 1-3 Tai dialects included in this study (see Appendix A)

The dialect data used in this dissertation comes from both published materials and my own fieldwork. Appendix A lists all the varieties included in this study and their sources. In total, data from 68 dialects are used in the reconstruction.

1.4.2.1 Published materials

The main set of data used is from Gedney's unpublished wordlist (Gedney n.d.), which has now been published as part of the *William J. Gedney's Comparative Tai Source Book* (Hudak 2008). Another main source of lexical data comes from a comprehensive survey of Zhuang dialects of Guangxi and Yunnan 壮语方言研究 *Zhuàngyǔ Fāngyán Yánjiū* (Zhang et al. 1999). In addition to these two sources, other resources that play invaluable roles in this dissertation include published and unpublished wordlists, dictionaries, and dialect descriptions in English, Thai, and Chinese (see Appendix A). Worth special mention is the survey of Bouyei dialects 布衣語調查報告 *Bùyīyǔ Diàochá Bàoɡào* (Chinese Academy of Sciences 1959) which includes 40 localities altogether. These Bouyei dialects will be referred by their numbers in the original source, e.g. Bouyei (point 4) or Bouyei(4) refers to dialect point 4 in *Bùyīyǔ Diàochá Bàoɡào*.

1.4.2.2 Fieldwork

Equally important is my own fieldwork carried out on a number of dialects in Thailand, Vietnam, and China. Altogether I carried out fieldwork on five Tai varieties. Three of them are spoken in northern Vietnam, one of the most understudied Tai-speaking areas. During the winter of 2008, I collected data from the Tai dialects of Sapa, Bao Yen, and Cao Bang. The first two are spoken in Lao Cai province while the third is spoken in Cao Bang province. In Thailand, I carried out fieldwork on the

Kapong dialect of Phu Thai spoken in Sakon Nakhorn Province in 2005. In China, I collected data of the Shangsi dialect in southeastern Guangxi in spring of 2009. Data from these five varieties are invaluable for the task of reconstructing of PT phonology.

1.4.2.3 Representative dialects

Given the number of dialects included in this study, it is not possible to discuss and present all data in this dissertation. In the discussion, three dialects from the three clusters are consistently used: Siamese, Lungchow, and Yay. They represent SWT, CT, and NT respectively. These dialects are chosen mainly because of the richness and quality of the lexical data available. Moreover, choosing these dialects allows easy comparison between this dissertation and Li (1977). Siamese and Lungchow are used in Li (1977) as representative dialects. Yay is not included in Li (1977) but his NT representative Po-ai, is closely related to it.

1.4.2.4 Tai-external data

In addition to data from Tai varieties, data from other Kra-Dai languages as well as languages from other families are also used as additional evidence. Although this dissertation frequently makes references to the Kra-Dai language Lakkja (L-Thongkum 1992; Liu 1999; Pittayaporn 2008a), Proto-Austronesian (Wolff in press), Late Middle Chinese (Pulleyblank 1991), (Early) Middle Chinese (Schuessler 2009), Later Han Chinese (Schuessler 2007, 2009) and Old Chinese (Schuessler 2007, 2009), the reconstruction is based solely on Tai-internal evidence. As discussed above, the Tai-external data are used only as clues, or to help confirm certain choices of reconstruction. Unless otherwise indicated, forms cited in this dissertation from these languages are from the above sources. Sino-Tai etyma are cited from Li (1976) and/or Manomaivibool (1975), except where indicated.

1.4.3 Notation

Because this dissertation incorporates data from diverse sources and compares reconstructions by different researchers, it is necessary to establish notational conventions that facilitates comparisons between different Tai dialects and different studies on PT phonology. All modern dialect forms as well as reconstructed PT forms that appear in this dissertation have been re-transcribed according to the IPA standard (International Phonetic Association 1999) as detailed below.

- 1) Aspiration is transcribed as a superscripted [h] rather than a normal [h], e.g. Siamese /p^ha:^{B1}/ ‘to split’ rather than /pha:^{B1}/, and PT *^hma: ‘dog’ rather than *hma:.
- 2) Glottalization is transcribed as a superscripted [ʔ] rather than a normal [ʔ], e.g. Debao /ʔja:^{A1}/ ‘medicine’ rather than /ʔja:^{A1}/.
- 3) The palatal semivowel is transcribed as [j] rather than [y] which represents the high front rounded vowel, e.g. Siamese /ja:^{A1}/ ‘medicine’ rather than /ya:^{A1}/.
- 4) Syllable-final semivowels are treated as glides [j], [ɥ], and [w], e.g. Li’s *vɛi^A ‘fire’, *mɛu^A ‘cat’, and *hmou^B ‘new’ are retranscribed as *vɛj^A, *mɛw^D, and *hmouɥ^B respectively.
- 5) Li’s unaccented high vowels *_̣i, *_̣ɨ, and *_̣u are re-transcribed as glides *j, *ɥ, and *w respectively, e.g. Li’s *pjet^D ‘duck’, *fɿa^C ‘cloud’ and *dluom^A ‘wind’ are written as *pjɛt^D, *fɥa^C, and *dlwom^A respectively (see §1.3.1). The undermark diacritic indicates that Li considers these semivowels.

- 6) High and mid back unrounded vowels are transcribed as [u] and [ɤ] rather than [ĩ] and [ə], e.g. Lungchow /mu:^{A2}/ ‘hand’ rather than /mĩ:^{A2}/, and /fɤn^{A2}/ ‘firewood’ rather than /fən^{A2}/ . The symbol [ə] is reserved for reduced and neutral vowels only.
- 7) [tɕ] and [ç] both represent the same voiceless palatal affricate but are considered stops phonologically and re-transcribed as [c], e.g. Siamese /com^{A1}/ ‘to sink’ rather than /tɕom^{A1}/ or /çom^{A1}/ . Similarly, [dz] and [j] are re-transcribed as [j]. In addition, [ɕ] is used instead of [š], e.g. /ɕam^{A1}/ ‘to sink’ rather than /šam^{A1}/ . Two exceptions are Li’s *č- and *j-, which are retranscribed as *tɕ- and *dz- because Li (1977: 164-173) explicitly states that the two PT phonemes were affricates.
- 8) Long vowels are indicated by [:] rather than double vowels, e.g. Siamese /ja:^{A1}/ ‘medicine’ rather than /jaa^{A1}/ .

The notation of tones in modern dialects deserves special attention. It is the area where the largest number of different practices co-exists. These conventions range from using diacritics, numbers, or tone letters, to using syllable-final letters (Edmondson & Solnit 1997). To allow for ease of analysis and presentation, notations containing etymological information are widely used in historical work. A few systems of this sort are available, but all share one common characteristic: they make reference to PT tone categories *A, *B, *C, and *D and the laryngeal features of the leftmost onset consonants.

An essential method used in approaching the onset-tone interactions in the history of Tai languages is known as the ‘tone-box’ method, formalized by Gedney (1972a). It is well-established that PT had a system of three tones on non-obstruent-final syllables, or “non-checked syllables”: *A, *B, and *C. There is no tonal contrast

on obstruent-final syllables, or “checked syllables”; this type of syllables are then said to be toneless, or conventionally to have *D tone. Some time after PT had diversified into various languages, sound changes involving splits in the tonal system conditioned by the laryngeal feature of the initial consonants took place and spread through out the SWT-speaking territory (Gedney 1991; Pittayaporn 2008b). Each of the four PT tones then split into two or more tones. Some of the resulting tones then merged or further split. Checked syllables usually underwent an additional split conditioned by vowel length.

In talking about initials and tones in Tai languages, it is useful to draw a chart showing the laryngeal features of the initials at the time of tone split and the original PT tone categories. In Table 1-6, the order of the laryngeal features presented vertically, and the order of the PT tones, presented horizontally, must not be altered. A lexical item is then annotated with the name of the box that corresponds to its PT initials and tones. For example, Thai /ma:⁵/ ‘dog’ which comes from PT *^hma:^A is written as /ma:^{A1}/ to signal that the original tone was the *A tone and the original initial was voiceless. Note that *DS and *DL represent PT *D on syllables with short and long vowels respectively.

Table 1-6 Chart showing the laryngeal features of the initials at the time of tone split and the original PT tone categories.

		*A	*B	*C	*DS	*DL
1	aspirated *f-, ^h m-, p ^h -,...	A1	B1	C1	DS1	DL1
	unaspirated *p-,...	A1	B1	C1	DS1	DL1
	glottalized/implosive *ɓ-,...	A1	B1	C1	DS1	DL1
2	voiced *v-, m-, b-,...	A2	B2	C2	DS2	DL2

Although it is well established that the basic type of tone split is the so-called “register split”, which is conditioned by voicing of the initials, splits conditioned by aspiration, glottalization, and frication have also been reported for Tai languages (Gedney 1989b; L-Thongkum 1997). However, in reconstructing initial consonants, knowing what the original voicing specification of those initials was is generally sufficient. Therefore, this dissertation adopts Li (1977)’s convention, in which 1 is the assigned code for tones developed from original voiceless stops and 2 for those developed from voiced stops. When there is a need to allow for more fine-grained categorization of the proto-initials, additional notation will be introduced.

Note that the notation described above only applies to forms in Tai dialects. Reconstructed forms from languages other than Tai are presented in their original notation. Also, note the use of Chinese characters in this dissertation. While simplified characters are used generally in this dissertation, traditional characters are used in discussions regarding etymology of those characters. For example, the Tai dialect of Lungchow is written in Chinese with simplified characters 龙州 *Lóngzhōu* but the Chinese etyma for ‘needle’ and ‘pair’ are written with traditional characters 鍼 *zhēn* and 雙 *shuāng* respectively.

1.5 *Overview of dissertation*

The remainder of this dissertation consists of six chapters. The next chapter, Chapter 2, addresses the issue of the PT prosodic word while the following three, Chapters 3, 4, and 5, deal with the segmental components of the PT prosodic word: the onset and the rime. These are followed by a chapter that addresses the issue of the PT tones, Chapter 5. The last chapter then discusses the contributions and implications of this dissertation. An overview of each chapter is as follows:

Chapter 2 discusses the structure of the PT prosodic word. I first discuss the importance of the prosodic word in the reconstruction of PT. Then, I show that the monosyllabic view espoused by the conventional reconstruction fails to account for the range of comparative data and argue that PT was in fact a sesquisyllabic language. I also provide a brief typological survey of the canonical shape of the prosodic word in SEA languages. Lastly, I claim that PT was a sesquisyllabic language whose prosodic word can be either monosyllable or sesquisyllable and then outline the phonotactics of PT onsets, rimes, and tones.

Chapter 3 deals with the PT onset system focusing on simple onsets. I propose a system of PT simple onsets that departs from the conventional reconstruction in a number of major respects. First, I reconstruct a contrastive series of uvular consonants in addition to the five places of articulation in Li's system. Second, I argue that the so-called "voicing alternation" is in fact two distinct phenomena—borrowings and reduction of sesquisyllables. Last but not least, I claim that PT did not have contrastive aspiration.

Chapter 4 takes up the issue of complex onsets in PT. I first show that the monosyllabic hypothesis with only canonical clusters fails to account for the existing range of onset correspondences. I then present the inventory of reconstructible PT

complex onsets that includes two types of PT complex onsets: tautosyllabic clusters and sesquisyllabic clusters. Lastly, I show that the process through which Tai languages became monosyllabic cannot be viewed as a unified process of monosyllabification.

Chapter 5 discusses the rime system of PT. I first argue that PT had a full-fledged vowel length contrast which interacted closely with syllable structure. I account for the so-called “Gedney’s puzzle” by positing four sets of rimes that do not add any complexity to the system. Lastly, I propose that final *-l, *-c, and perhaps *-ŋ also existed in PT in addition to those codas assumed in the conventional reconstruction. Crucially, the proposed system takes seriously the effect of medial and coda consonants on the development of the vowels.

Chapter 6 proposes a concrete reconstruction of the PT tonal categories *A, *B, *C, and *D, already well-established in the conventional reconstruction. I argue that the contrast between PT tones was based not solely on pitch but also on voice quality. In particular, I posit a modal tone with a mid level contour for *A, a creaky tone with a low rising contour and a relative long vowel duration for *B, and a high falling contour ending with a glottal constriction with a relatively short vowel duration for *C. As for checked syllables, I posit a low rising contour for *D.

Finally, Chapter 7 discusses the contributions of this reconstruction. I first compare the proposed reconstruction of PT phonology with Li’s conventional reconstruction. Then, I propose a preliminary subgrouping schema of Tai languages as well as a historical scenario that explains the diversity and the convergence among modern Tai dialects.

CHAPTER 2

THE PROTO-TAI PROSODIC WORD

2.1 *Introduction*

Never seriously discussed in historical studies of Tai languages is the structure of the PT prosodic word. In reconstructing Proto-Tai, the focus has always been on reconstruction of individual segments to account for individual correspondences. This lack of attention to the organization of segments into larger units has ironically hindered progress in solving the mysteries in segmental reconstruction. This is because in all natural languages individual segments must be organized within larger domains such as syllables, and words. These prosodic domains crucially pose certain constraints on how segments are combined. Therefore, reconstruction of segments and consequently individual etyma depends largely on the researchers' implicit or explicit assumptions about possible phonological shapes of prosodic words in the proto-language.

A well-known example is the case of the *DVD constraint on the Proto-Indo-European (PIE) root, which has been used as important evidence for the Glottalic Theory (Gamkrelidze & Ivanov 1972). In the Standard Theory of PIE stops (Mayrhofer 1986), PIE phonology had three contrastive phonation types in stop consonants—voiceless, plain voiced, aspirated voiced. These stops, when combined, yield nine logically possible CVC root types. Nevertheless, it has been generally assumed that there was a constraint against CVC roots having two plain voiced stops, i.e. *deg (Iverson & Salmons 1992). The Glottalic Theory, in contrast, believes that the three contrastive series of stops in PIE was in fact (aspirated) voiceless, ejectives, and (aspirated) voiced. In support of this segmental reconstruction, the Glottalic Theory claims that the *DVD constraint is common in ejective languages. In this view

*DVD root would have in fact been *T'VT', which is not allowed in ejective languages. This putative *DVD constraint has however been challenged by various Indo-Europeanists such as Iverson and Salmons (1992), Barrack (2002), and Cooper (to appear). As the root is an obligatory part of a prosodic word, this example from Indo-European illustrates how assumptions about possible phonological shapes of the prosodic word, or some part of it, could influence the reconstruction of the segments.

Another example closer to Tai is the case of Old Chinese (OC). Karlgren (1964), Li (1980) and Baxter (1992) all assume that the OC was a monosyllabic language, while Sagart (1999), and Handel (1998) argue that OC allowed for sesquisyllabic words. The different hypotheses about the shape of the OC prosodic word have led to major differences in segmental reconstruction. An example comes from cases of etyma whose Middle Chinese (MC) onsets point to OC *t- or *th-, but whose word-family connections point to another initial consonant. Assuming that OC was strictly monosyllabic, Karlgren (1964) reconstructs *təp and *tʰiəwət for 答 *dá* 'to respond' and 出 *chū* 'to exit' respectively. In contrast, Sagart (1999: 90), who holds that sesquisyllabic words were common in OC, posits *^at-gip and *^bt-khut for the two etyma. This OC case is especially relevant not only because of the geographical proximity to Tai but also because a case can also be made for reconstructing PT either as a monosyllabic or sesquisyllabic language.

Therefore, it is crucial that possible shapes of the PT prosodic word be established before proceeding to reconstruct the onset, the rimes, and the tones. In this chapter, I first review the structure of the prosodic word as assumed by earlier proposals on PT reconstruction and show that the conventional view that PT words were all monosyllables cannot account for the comparative data. Crucially, I argue that the view that PT was a sesquisyllabic language is the best hypothesis regarding the PT prosodic word. Then, I briefly discuss the prosodic word and its typology from the

metrical and prosodic points of view (Hayes 1995; Liberman 1975; McCarthy & Prince 1986; Nespor & Vogel 1986; Ussishkin 2005). Lastly, I put forward a proposal establishing an explicit characterization of the PT prosodic word. In particular, I claim that PT prosodic words can be either monosyllables, or sesquisyllables. Each word contained exactly one iambic foot headed by a heavy syllable. The phonation-type and place contrasts in the onset were neutralized in the coda position. Lastly, each prosodic word is specified for one lexical tone.

2.2 *The prosodic word*

In the framework of Prosodic Phonology (Inkelas 1989; McCarthy & Prince 1986; Nespor & Vogel 1986; Selkirk 1981), the prosodic word is described as “the lowest constituent of the prosodic hierarchy which is constructed on the basis of mapping rules that make substantial use of non-phonological notions (Nespor & Vogel 1986: 109).” It is also characterized formally as the prosodic category that is larger than the foot but smaller than the phonological phrase. It is the phonological domain that roughly corresponds to the morphological word, and thus the smallest prosodic constituent that can stand alone (Dixon & Aikhenvald 2002: 24-25; Zec 2005)⁵. Establishing the PT prosodic word, therefore, means figuring out what possible PT words may look like phonologically. In this section, I first discuss the importance of the prosodic word in synchronic phonology. Then, I outline some well-formedness conditions relevant to the discussion of the PT prosodic word.

⁵ According to Dixon and Aikhenvald (2002: 24-25), a morphological word can stand alone only when it is also a prosodic word. For examples, lexical words and free function words in standard Serbian can stand alone and carry a pitch accent but bound function words must be cliticized to a host prosodic word (Zec 2005).

2.2.1 *The importance of the prosodic word*

The prosodic word is very important from a synchronic perspective. It is the domain in which rules and/or constraints are applied to maintain the syllables and the feet in conformity with universal or language-specific phonological principles. Finnish, for instance, requires that vowels within a prosodic word be all front or all back (van der Hulst & van de Weijer 1995: 498). Examples of permissible words are *väkkärä* ‘pinwheel’ and *pöytä* ‘table’, both of which have only front vowels. Similarly, *makkara* ‘sausage’ and *pouta* ‘fine weather’ exemplify Finnish words with back vowels. This means that words like **väkkara* or **poutä* do not exist in the language. Some other examples of rules that make references to the prosodic word include vowel harmony in Hungarian and stress assignment in Latin (Nespor & Vogel 1986).

Furthermore, the prosodic word is important in synchronic phonology because individual languages impose certain phonotactic requirements on their prosodic words. The Austroasiatic language Kammu described by Svantesson (1983) is a good example. In this language, prosodic words must contain exactly one iambic foot (Pittayaporn 2005c). In other words, a simple Kammu word must either be a heavy monosyllable, i.e. /jim/ ‘red’ and /cè:m/ ‘wet’, or a sesquisyllable, i.e. /kʔi:p/ ‘centipede’ and /k.màʔ/ ‘rain’. Words like **/bi/* or **/tu/* are not permissible because they only have one short vowel, which makes them too light for a Kammu prosodic word. Similarly, non-compound words like **/(pə)(təban)/*, or **/(sa:)(kətum)/* do not exist in Kammu because they consist of two feet⁶, exceeding the weight maximum for the Kammu prosodic word. In addition to the case of Kammu, some additional examples of phonotactic requirements imposed on the prosodic word include the ban of lax vowels at the end the prosodic word in German (Hall 1997) and the constraint in

⁶ The foot is represented here by ().

Dutch that within a prosodic word the first consonant in a heterosyllabic consonant sequence has to be more sonorous than the second (Booij 1997).

In short, the prosodic word is a prosodic domain that plays an important role in determining word-level phonotactics and other phonological phenomena within the word. Any comparative reconstruction therefore needs to take seriously the shape of the prosodic word in the proto-language. For example, if we were to reconstruct Finnish, we must therefore not posit words that contain both front and back vowels. Similarly, if we were to reconstruct Kammu, we must include only heavy monosyllabic words and sesquisyllabic words in the reconstructed lexicon.

2.2.2 Well-formedness

A well-formed prosodic word in a given language is one that simultaneously satisfies all the well-formedness requirements of that language including the well-formedness requirements imposed on the syllables (Goldsmith 1995). In this subsection, I discuss three most relevant phonological concepts: size requirements, sonority, and foot typology.

2.2.2.1 Size requirements

Numerous languages of the world exhibit restrictions on the size of possible prosodic words. Many impose restrictions on the minimal size a prosodic word can have, while fewer prohibit prosodic words that exceed a certain size. These types of restrictions can be referred to as word minimality and word maximality, respectively. These size requirements should not be thought of as primitive phonological conditions but should be characterized as effects of interactions among various prosodic and metrical constraints (de Lacy 2004; McCarthy & Prince 1986).

Between the two types of size requirements, word minimality is found more generally cross-linguistically. Although cross-linguistic surveys have shown that there is no strong correlation between the minimal word size and the minimal foot (Downing 2006:48-58, 94-100), diverse languages from all over the world require that their prosodic words be minimally one well-formed metrical foot. In SEA, a number of languages including Burmese (Green 2005), Thai (Bennett 1994, 1995), Kayah Li (Bennett 1995), Khmer (Huffman 1972), Indonesian (Cohn & McCarthy 1998) and Moken (Pittayaporn 2005a) prosodic words have been shown to contain at least one well-formed bimoraic foot.

The maximality requirement is not found as commonly as its minimality counterpart. Maximality effects have been shown to operate in Maori and perhaps in Ura (de Lacy 2004). Although unfortunately very few studies have been done on maximality, especially in Southeast Asian languages, some languages including Burmese (Green 2005), Kammu (Pittayaporn 2005c), Moken (Pittayaporn 2005a) and Indonesian (Cohn in progress) have been shown to have at least a strong dispreference for prosodic words that exceed one metrical foot.

Although literature on interaction between word minimality and word maximality is lacking, many languages seem to show effects of both resulting in a situation where the well-formed word is equal to a well-formed foot. For example, in Maori, word minimality requires that a prosodic word be at least one metrical foot while word maximality bans any prosodic words that are larger than one foot. As a result of the interaction, normal prosodic words in Maori contain one single foot and optionally one unfooted syllable. Prosodic words with two feet are allowed only under special circumstances (de Lacy 2004). In Modern Hebrew, prosodic words optimally consist of two syllables that form one metrical foot. Trisyllabic words are allowed only to realize trisyllabic affixes (Ussishkin 2005). Similarly, in Khoekhoe, lexical

words must be exactly one metrical foot (Brugman 2009) due to the joint effect of minimality and maximality. In SEA, languages that require their prosodic word to be exactly one metrical foot are very common. Some clear cases of such languages include Kammu (Pittayaporn 2005a), Moken (Pittayaporn 2005a), and, to a large extent, Burmese (Green 2005) and Indonesian (Cohn in progress).

2.2.2.2 *Sonority*

A well-formed prosodic word is one that can be divided exhaustively into one or more well-formed syllables (Booij 1997). One of the ingredients in the well-formedness of the syllables and the prosodic word is sonority. In particular, strings of segments are syllabified into syllables according to constraints on the relative sonority of the segments. These syllables are then combined to form the prosodic word. Although its definition is an issue of much debate, it is clear that sonority is linked at least partly to intensity or loudness (Clements 1990; Parker 2002). Sonority can be understood, and can be characterized in terms of hierarchy or scales. Many versions of this hierarchy have been proposed (e.g. Hooper 1976; Selkirk 1984; Steriade 1982), but in this dissertation I adopt Clements (1990)'s proposal that arranges non-syllabic segments from less sonorous to more sonorous classes as given Figure 2-1.

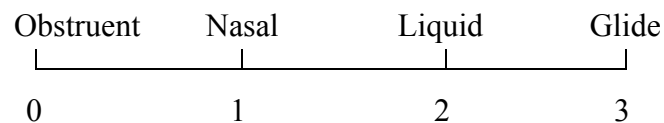


Figure 2-1 Sonority scale (based on Clements 1990: 292)

In the scale given here, obstruents are the least sonorous and can be said to have the sonority value of 0. On the opposite end of the scale, glides are the most

sonorous among non-syllabic segments and have the sonority value of 3. Nasal and liquid consonants occupy intermediate positions along the scale, having the sonority values of 1 and 2 respectively. These sonority values are very useful in describing the sonority profile of segments or string of segments.

Most relevant to the issue at hand is the relative sonority among consonants that form a complex onset in a well-formed syllable. Cross-linguistically, permissible consonant clusters are subject to an important constraint, generally known as the Sonority Sequencing Principle (SSP), which states that “in any syllable, there is a segment constituting a sonority peak that is preceded and/or followed by a sequence of segments with progressively decreasing sonority values (Selkirk 1984: 116).” In other words, in a given syllable sonority is lowest at the edges and highest at the syllabic nucleus. Because well-formed syllables are subject to SSP, as part of the syllable complex, onsets are generally expected to follow the same principle as well.

The role of SSP in well-formedness of complex onsets can be illustrated by the case of Spanish initial clusters. Spanish strictly obeys SSP so that all permissible complex onsets consist of a stop followed by a liquid or a glide. For instance, /kl-/ , /kr-/ and /kw-/ are well-formed initial clusters as in /klabe/ ‘key’, /krisis/ ‘crisis’, and /kwota/ ‘quota’. This is because /k-/ has lower sonority than /l-/ , /r-/ and /w-/ , which are closest to the syllabic nucleus. In contrast, /lk-/ , /rk-/ , /wk-/ , are not allowed because the /k-/ is less sonorous than /l-/ , /r-/ , and /w-/ violating SSP (Harris 1983; Hualde 1991).

In addition to SSP, languages may further impose a minimal sonority distance on complex onsets (Levin 1985; Selkirk 1984; Venneman 1972; Zec 2007). Such constraint requires that members of complex onsets meet a certain difference in sonority values. Recall the sonority values of different types of consonants in Figure 2-1. In Spanish, /kl-/ , /kr-/ are well-formed clusters because the sonority distance between the

obstruent /k-/ and the following liquids is exactly 2. Similarly, the difference in sonority values between the initial obstruent /k-/ and the glide /w-/ is 3, exceeding the required minimal sonority distance. On the other hand, */pn-/ and */ml-/ are banned in Spanish, even though they do not violate SSP. This is because the sonority distance within each of the consonant sequences is only at 1 respectively (Harris 1983).

Although the role of sonority in syllable well-formedness is robustly attested across languages, clusters that are not predicted by sonority-based generalizations are relatively common. The most well-known case is English clusters /sp-/, /st-/, and /sk-/ as in *spy*, *star*, and *sky* respectively. In these clusters, the /s-/ and the following stops form sonority plateaus in violation of SSP, but they are still well-formed (Halle & Vergnaud 1980; Steriade 1982). Modern Hebrew allows biconsonantal clusters that consist of an obstruent followed by another obstruent, e.g. /dg-/ as in /dgalim/ ‘flag’, /ks-/ as in /ksamim/ ‘magic’, and /pg-/ as in /pgarim/ ‘carrion’. Instead of a rising sonority, such clusters form sonority plateaus as both the first and the second consonants have the sonority value of 0.

A more extreme case is Polish, which allows long strings of segments that violate SSP to precede the vocalic nucleus, e.g. *bdzura* ‘nonsense’, *fstręć* ‘repulsion’. Interestingly, these complex onsets in Polish have been analyzed as consisting of two separate onsets (Gussmann 1992) or a degenerate syllable followed by an onset (Cho & King 2003). Sesquisyllabic structure (see §2.4.3) found widely in SEA has been similarly analyzed as comprising a degenerate syllable followed by an onset (Cho & King 2003; Pittayaporn 2005c). Because in these cases, the consonants in each sequence are not all linked to the same syllable nodes, they in fact do not form syllable clusters in the strict sense.

Kreitman (2008) surveys 62 languages with complex onsets and concludes that there is an implicational hierarchy for the occurrence of different types of consonant

clusters (S = sonorant, O = obstruent): $SO \Rightarrow SS \Rightarrow OO \Rightarrow OS$. According to this typology, OS clusters, e.g. /pr-/ , /kl-/ , /kw-/ etc. are the most basic type of clusters. The presence of OO clusters, e.g. /pg-/ , /st-/ , /kt-/ etc. implies the presence of SO clusters. Similarly, the presence of SS clusters, e.g. /ml-/ , /nl-/ , /rl-/ etc., implies the presence of OS and OO clusters. Lastly, the presence of SO clusters, e.g. /rt-/ , /ld-/ etc., implies the presence of the other three cluster types. In this dissertation, I refer to OS clusters as “canonical” or “unmarked” clusters, and refer to the other types of clusters as “marked” clusters.

2.2.2.3 Foot typology

In Prosodic Phonology (McCarthy & Prince 1986; Nespor & Vogel 1986), a prosodic word is made up of one or more metrical feet, each of which is made of one or more syllables. By transitivity, a well-formed prosodic word consists of one or more well-formed syllables. According to Hayes (1995), cross-linguistically a well-formed foot belongs to one of the following three types: syllabic trochee, moraic trochee, and iamb. This foot typology is based on quantity sensitivity and dominance. The three basic foot types can be illustrated in Figure 2-2.

	Quantity-insensitive	Quantity-sensitive
Left dominant	syllabic trochee (× .) σ σ	moraic trochee (× .) or (×) ˘ ˘ ˉ
Right dominant		iamb (. ×) or (×) ˘ σ ˉ

× = stressed syllable . = unstressed syllable

˘ = light syllable ˉ = heavy syllable

Figure 2-2 Foot typology according to Hayes (1995)

A syllabic trochee is a metrical foot with initial prominence that counts syllables regardless of their weight (Hayes 1995). Pintupi is an example of languages that have this type of foot as illustrated by the following words: /('paŋa)/ 'earth', /('ma|a)(wana)/ 'through from behind'. The former word consists of one syllabic trochee while the latter consists of two (Hayes 1995). Among Southeast Asian languages, Indonesian is an example of a language that has syllabic trochees (Cohn & McCarthy 1998).

A moraic trochee refers to a metrical foot that consists of two moras and, similar to a syllabic trochee, has initial prominence. A moraic trochee can either consist of two light syllables or one heavy syllable (Hayes 1995). Cairene Arabic is a good example of a language that have this type of foot as illustrated by the following words: /('kata)(bitu)/ 'she wrote it (m.)' and /(?ad)(wiya)(tuhu)/ 'his drugs (nom.)'. In the former each of the two feet had two light syllables while the latter consists of one

foot with one single heavy syllable followed by two feet of two light syllables (Hayes 1995). In SEA, Tondano, spoken in Sulawesi, is an example of languages with moraic trochees (Ball 2003).

Lastly, an iamb is a metrical foot with final prominence. It can either consist of two syllables with a light first syllable, or one single heavy syllable (Ball 2003). In Creek, the prosodic word consists of iambic feet as exemplified by the following words: /(ni'ha:)/ 'lard' and /(ti:)(ni:t)(ki:)/ 'thunder'. While the first word consists of one iambic foot with a light syllable followed by a heavy one, the second word consists of three feet, each of which is a heavy syllable (Hayes 1995). In SEA, numerous languages belong to this type, including Thai (Bennett 1994, 1995), Kayah Li (Bennett 1995), Burmese (Green 2005), Moken (Pittayaporn 2005a), Kammu (Pittayaporn 2005c; Svantesson 1983), and others.

2.3 *Reconstructing the PT prosodic word*

Assumptions about the canonical shape of the PT prosodic word are a central factor underlying the controversies about the reconstruction of PT. Unfortunately, they are not usually discussed and to my knowledge not explicitly compared. Although the conventional view represented by Li (1977) among others holds that PT was monosyllabic, alternative proposals have also treated PT as sesquisyllabic (Ferlus 1990), or disyllabic (Haudricourt 1975; Strecker 1983).

Advances in the study of Kra-Dai languages have shown that Proto-Kra-Dai cannot have been a strictly monosyllabic language. Evidence for such a conclusion comes from comparative studies of lesser-known relatives of Tai. Ostapirat (2000) reconstructed Proto-Kra with sesquisyllables on the basis of onset correspondence and tonal reflexes in modern languages. For example, he posits *p.la^A 'afraid' to account for Gelao /lau^{A2}/ and Paha /pja:^{A2}/, in contrast to *plat^D 'blood' for Gelao /plɔ^{D1}/ and

Paha /pɛ:^{D1}/. While reflexes of the sesquisyllabic *p.l- have tone in the second series according to the voicing of the non-initial *l-, the reflexes for the tautosyllabic cluster have tone in the first series according to the initial *p-. Edmondson and Yang (1988) present initial correspondences among Kam-Sui languages as evidence for “preconsonants” which perhaps reflect disyllabic or sesquisyllabic structure. For example, Mulam /khyət^{D1}/ corresponding to Kam /sət^{D1}/ suggests Proto-Kam-Sui *k-jət^D ‘tail’. Similarly, Ferlus (1989; 1996b) shows convincingly that Proto-Kam-Sui was a sesquisyllabic language although all modern Kam-Sui languages are monosyllabic. For example, he reconstructs *k.s- as onset of Kam-Sui etymon ‘sour’ which is now reflected as /səm^C/ in Kam, /hum^C/ in Sui, and /k^hyəm^{C1}/ in Mulam. Similarly, L-Thongkum (1992) proposes a scenario in which Lakkja nasalized vowels developed from earlier sesquisyllabic forms, e.g. /k^hũ^{A1}/ ‘pig’ from earlier *kh-mu. Further, Osaka (2007) proposes that Be /vəi^{A2}/, Siamese /pi:^{A1}/ and Saek /pi:^{A1}/ comes from Proto-Be-Tai *Cəpii ‘year’.

This growing consensus that prosodic words in Proto-Kra-Dai were not all monosyllabic is clearly at odds with the strict monosyllabicity of cognates found among modern Tai languages. This raises the issue of what the situation at the PT stage was. In other words, it is still unclear whether the Proto-Kra-Dai prosodic word had reduced to monosyllables by the PT period. In this section, I show that the monosyllabic view of PT fails to account for the range of comparative data and argue in favor of the sesquisyllabic hypothesis.

2.3.1 The PT prosodic word in earlier proposals

The conventional view of the PT sound system holds that the proto-language was strictly monosyllabic. Li does not explicitly state in the *Handbook of Comparative Tai* (Li 1977) that PT was strictly monosyllabic, but all of his reconstructed forms are

monosyllables. Other linguists working on PT reconstruction also share this view despite their disagreement about segmental reconstruction. Nishida (1954; Nishida 1955), Liang and Zhang (1996), Luo (1997), and Sarawit (1973) all posit only monosyllables in their reconstruction. Gedney, who rarely gives precise reconstructions of PT “morphemes,” always talk about PT in terms of syllables (Gedney 1972a, 1989b).

Although the motivation for the monosyllabic view has never been discussed, it is safe to say that it stems from the fact that cognates found in modern Tai varieties are almost always monosyllables (Gedney 1989b: 15). This view clearly affects segmental reconstructions proposed by different authors. For example, Li (1977) proposes PT *dl- to explain cases in which some daughter languages indicate PT *l- but others indicate PT *r-. Li’s choice is motivated by the fact that *dl- is the only canonical alveolar cluster that he has not reconstructed. Similarly, Liang and Zhang (1996) reconstruct a marked cluster *xkr- to account for etyma that Li considers to be a special development of his *khl-.

An alternative to the monosyllabic view is to reconstruct sesquisyllabic words for PT. The term “sesquisyllable”, literally “one-and-a-half syllable”, was originally coined by Matisoff (1973) to refer to words that consist of an unstressed “minor” syllable followed by a fully stressed “major” syllable. Crucially, the minor syllable is characterized as lacking a contrastive vowel, e.g. Palaung (Shorto 1960) /k.tɛ/ ‘earth,’ and /kɾ.taʔ/ ‘tongue’⁷ (see §2.4.3). Many languages in SEA have been described as showing sesquisyllabic words, e.g. Khmer (Henderson 1952), Kammu (Svantesson 1983), Palaung (Shorto 1960), Burmese (Green 2005), Lahu (Matisoff 2001) etc., but the precise formal characterization of sesquisyllables is unclear. In fact, it is likely that

⁷ Shorto transcribe these forms as /kətɛ/ and /kəɾtaʔ/ respectively but states that “a minor syllable has no vowel other than an anaptyctic one.”

the so-called sesquisyllables differ structurally from one language to another. In this dissertation, I define a sesquisyllable as a prosodic word that consist of two syllables, the first of which is unstressed and lacks a phonological vowel. Under this definition, the sesquisyllabic structure differs from tautosyllabic clusters in that the consonants in a sequence at the beginning of a sesquisyllable are not all linked to a single syllable node.

Ferlus (1990) proposes a reconstruction of the PT system of initial consonants that includes both monosyllabic and sesquisyllabic etyma⁸. This view leads him to interpret many correspondence sets differently from Li. For example, he posits *k.t- for etyma for which Li proposes *tl-. The reflexes in daughter languages, however, can be accounted for equally well by the two accounts. In addition, the sesquisyllabic view also faces the same challenge posed by the reconstruction of vowels. Since the minor syllable of a sesquisyllabic word does not have a contrastive vowel, the problematic vowel correspondences must be accounted for in the rime of the major syllables only.

Another alternative holds that PT was disyllabic. Haudricourt (1956; 1963b) proposes that disyllabic roots be reconstructed for PT to account for some onset correspondence sets (see Chapter 4). For example, he links the Tai etyma for ‘bird’ and ‘water’ to Malay /danum/ and /manuk/ to explain the fact that the medial /-n-/ is reflected in these two etyma as /n-/ in Siamese but /r-/ in Yay, a NT—/na:m^{C1}/ and /ram^{C1}/ respectively. Furthermore, Haudricourt (1975) proposes that disyllabic roots be reconstructed for PT to account for the so-called “Gedney’s Puzzle (see Chapter 5).” For instance, he proposes that PT *apui ‘fire’ went through two different paths in different subgroups: 1) *apui > *apwi > *pwai > *vai in SWT and CT but 2) *apui > *apwi > *pwi > *vi in NT. However, it is not clear whether according to this account monosyllabic words were allowed in PT. Similar to the sesquisyllabic view, this disyllabic view seems to be motivated by the need to account for segmental correspondences, rather than by

⁸ He refers to sesquisyllables as “disyllables”.

typological considerations as is the case for the monosyllabic view. Examples of reconstructed forms under different assumptions are given in Table 2-1 for comparison.

Table 2-1: Reconstruction of certain etyma under different assumptions about the PT prosodic word

	Li	Ferlus ⁹	Haudricourt ¹⁰
‘fire’	*vɛj ^A	*vɛj ^A	*apwi ^A
‘bear’	*hmwi ^A	*hmwi ^A	*tumaj ^A
‘moon’	*ʔbluən ^A	*ɬluən ^A	*bulan ^A
‘bird’	*nrok ^D	*C.nok ^D	*manuk ^D

In this section, I have reviewed the three views regarding the canonical shape of the PT prosodic word. All of them basically constrain how the segmental materials are to be reconstructed. In the following section, I will argue that the sesquisyllabic view is the best hypothesis regarding the PT prosodic word.

2.3.2 *Sesquisyllabicity in PT*

The monosyllabic view of PT that underlies the conventional reconstruction has contributed enormously to the understanding of the history of Tai languages. However, it cannot account for the whole range of comparative data, and therefore fails to provide an accurate picture of PT phonology and its development into daughter languages. In particular, many onset correspondences show traces of PT words that

⁹ Ferlus (1990) does not reconstruct PT rimes. The rimes given in this table are taken from Li (1977).

¹⁰ Except for ‘fire’, Haudricourt does not propose PT forms. Forms found in this table are provided for illustration only.

were larger than monosyllables. Specifically, there is too wide a range of onset correspondences to be accounted for by positing a monosyllabic template for the PT prosodic word. Although these data are problematic for the monosyllabic hypothesis, they can be captured quite straightforwardly by the sesquisyllabic view.

The first set of onset correspondences that is problematic for the monosyllabic view of the PT prosodic word is that in which most dialects point to PT *d- (see §3.6.2.2), except for Saek. In Table 2-2, Siamese represents SWT, Lungchow represents CT, and Yay together with Saek represent NT. Li's PT and the reconstructed Proto-Southwestern Tai (Pittayaporn 2008b, to appear-b) onsets are also provided for comparison. This correspondence cannot be accounted for if the reconstruction of PT is limited to monosyllables with unmarked clusters.

Table 2-2: Correspondences involving PT *d-

		PSWT	Siamese	Lungchow	Yay	Saek	Li
1	'handle'		da:m ^{C1}	da:m ^{C1}	da:m ^{C1}	da:m ^{C1}	*?d-
	'to acquire'	*?d-	da:j ^{C1}	daj ^{C1}	daj ^{C1}	daj ^{C1}	
	'single, only one'		diəw ^{A1}	de:w ^{A1}	dew ^{A1}	diəw ^{A1}	
2	'red'		dɛ:ŋ ^{A1}	dɛ:ŋ ^{A1}	diŋ ^{A1}	ri:ŋ ^{A1}	*?dr-
	'bone'	*?d-	du:k ^{DL1}	duk ^{DL1}	dok ^{DL1}	rɔ:k ^{DL1}	
	'raw'		dip ^{DS1}	dip ^{DS1}	dip ^{DS1}	rip ^{DS1}	
3	'star'		da:w ^{A1}	da:w ^{A1}	da:w ^{A1}	tra:w ^{A1}	*?dr-
	'to transplant'	*?d-	dam ^{A1}	dam ^{A1}	dam ^{A1}	tram ^{A1}	
	'earthworm'		duən ^{A1}	dɣ:n ^{A1}	duən ^{A1}	truəl ^{A1}	

Li (1973; 1977: 107-110, 129-131) reconstructs PT *?d- for the first set and *?dr- for the second and third sets. This set of data has also been observed by Gedney

(1989c). With respect to the contrast between the second and third sets, Li (1977: 129) simply states that *ʔdr- gives /r-/ in Saek but is occasionally reflected as /tr-/. If we assume that the PT prosodic word was strictly monosyllabic with only canonical clusters, we have no way of explaining why Saek shows the contrast and why the stop in the third set is a plain voiceless instead of a glottalized stop. On the other hand, the sesquisyllabic view accounts for this paradigm by positing a coalescence of sesquisyllabic structure *t.n- into *ʔd- for set 3¹¹.

The second set of correspondences that cannot be accounted for by the monosyllabic hypothesis with unmarked clusters, shown in Table 2-3, involves the uvular consonants (see §3.3). In addition to the NT language Saek, Siamese, Lungchow, and Yay data represent SWT, CT, and NT respectively. Li's PT and PSWT reconstructions are given for comparison.

Table 2-3: Correspondences with *q-

		PSWT	Siamese	Lungchow	Yay	Saek	Li
1	'to kill'		k ^h a: ^{C1}	k ^h a: ^{C1}	ka: ^{C1}	ka: ^{C1}	
	'torn'	*q-	k ^h a:t ^{DL1}	k ^h a:t ^{DL1}	ka:t ^{DL1}	ka:t ^{DL1}	*k ^h -
	'arm'		k ^h ɛ:n ^{A1}	k ^h e:n ^{A1}	cen ^{A1}	ke:n ^{A1}	
2	'to hang'		k ^h wɛ:n ^{A1}		ven ^{A1}	ve:n ^{A1}	
	'face down'	*qw-	k ^h wam ^{C1}	k ^h um ^{C1}	ham ^{C1}	(ŋam ^{C1})	*xw-
	'soul'		k ^h wan ^{A1}	k ^h wan ^{A1}	van ^{A1}	hɔn ^{A1}	
3	'leg'		k ^h a: ^{A1}	k ^h a: ^{A1}	ka: ^{A1}	kwa: ^{A1}	
	'to sell'	*q-	k ^h a:j ^{A1}	k ^h a: ^{A1}	ka:j ^{A1}	kwa:j ^{A1}	*k ^h -

¹¹ Dempwolff and Dyen reconstructs this form as PAN *taNem 'to plant' (Wurm & Wilson 1975: 153) but it does not appear in Formosan languages.

The above table shows three correspondences that must be reconstructed with PT *q- (see §3.6.1.5)¹². The first set is quite straightforwardly accounted for by positing PT simple *q-. The uvular stop is reflected as /k^h-/ in most SWT and CT dialects including Siamese and Lungchow, and as /k-/ in NT. The second set should go back to PT *qw- although the NT reflexes at first glance seem irregular. The usual reflex of this uvular cluster is /k^hw-/ in CT and SWT dialects and /v-/ in NT. Yay /ham^{C1}/ ‘face down’ lost the medial /-w-/ early probably due to the labial quality of the final /-m/. The sequence *-wa- in ‘spirit’ thus must have become /ɔ/ in Saek before *^hw- became *w-, i.e. PT *qwan^A > *^hwan^A > hɔn^A. The third set of correspondences shows unexpected reflexes in Saek, as noted by Haudricourt (1952), Li (1965; 1977: 195), and Gedney (1989c). Their reflexes in the rest of Tai point unequivocally to *q- but Saek shows a labial element in its reflex. It does not help to reconstruct any cluster with */-w-/ because medial */-w-/ are normally kept in daughter languages, cf. Siamese /k^hwan^{A1}/ ‘spirit’.

If PT prosodic words must be monosyllabic with unmarked clusters, these two forms must be treated as an unsystematic development as assumed by Li (1977: 195). However, data from the Kam-Sui branch of the Kra-Dai family indicate that the labiality of the Saek clusters is a retention. The modern Sui forms for ‘leg’ and ‘to sell’ are /pa:¹/ and /pe:¹/ respectively (Institute of Language and Culture for Rural Development & Thailand and the Kam-Tai Institute 2003; Li 1965; Thurgood 1988). In contrast, if we assume that PT is sesquisyllabic, we are allowed to posit */p.q-/ for these two etyma. Note that Ostapirat (2005) connects PT ‘leg’ to PAN *paqa ‘thigh’. Note that the etyma in set 2 are not found in Kam-Sui.

The third paradigm that argues against the monosyllabic view involves complex clusters whose reflexes in some daughter languages indicate PT *br- (see §4.2.1.6) but others indicates PT *ɟ- (see §3.6.3.3). The data are given in Table 2-4.

¹² For arguments for reconstructing PSWT uvulars, see Pittayaporn (to appear-b).

Table 2-4: Correspondences involving *br- and *j-

		PSWT	Siamese	Lungchow	Yay	Saek	Li
1	‘to separate’		p ^h ra:k ^{DL2}	pja:k ^{DL2}	pja:k ^{DL2}	p ^h ra:k ^{DL2}	
	‘to deceive’	*br-	p ^h ra:ŋ ^{A2}	pja:ŋ ^{A2}			*br-
	‘deficient’		p ^h rɔŋ ^{A2}	pjoŋ ^{A2}	pjoŋ ^{A2}	p ^h rɔ:ŋ ^{A2}	
2	‘machete’		p ^h ra:C ²	pja:C ²	sa:C ²	t ^h a:C ²	
	‘evening meal’	*br-	p ^h raw ^{A2}	pjaw ^{A2}	saw ^{A2}		*vr-
	‘tomorrow’		p ^h ru:k ^{DL2}	pjuk ^{DL2}	sok ^{DL2}	t ^h ɔ:k ^{DL2}	
3	‘to hate’		c ^h aŋ ^{A2}	caŋ ^{A2}	saŋ ^{A2}	saŋ ^{A2}	
	‘to taste’	*j-	c ^h im ^{A2}	cim ^{A2}	sim ^{A2}		*j-
	‘to soak’		c ^h ɛ:A ²		se:A ²	se:A ²	

The reconstruction of set 1 and set 3 is straightforward. Li (1977) posits *br- (see §4.2.1.6) for set 1, and *j- (see §3.6.3.3) for set 3. However, set 2 presents a serious problem to the simple monosyllabic view. While Siamese and Lungchow reflexes point to PT *br-, Yay and Saek reflexes all point to *j-. To account for this problematic correspondence, Li tentatively reconstructs *vr- for these etyma, but it is not clear how *vr- could have developed into a palatal stop. This correspondence, however, is not a problem for the sesquisyllabic view, which allow for complex structure *jɹ.r- in PT. Under this view, *jɹ.r- simplified to *mr-, and became *br- in Siamese and Lungchow. The cluster took a different path in Yay and Saek, where it simplified to *jr, and became *j-.

The last set of sound correspondences involves labial clusters. Not only does it pose a challenge for the view that PT was a monosyllabic language that only allowed unmarked clusters, but it also shows that PT could not have been a monosyllabic

language that had marked tautosyllabic clusters, e.g. *pt-, *qs-, either. The data in this paradigm is given in Table 2-5.

Table 2-5: Correspondences involving labial clusters

		PSWT	Siamese	Lungchow	Yay	Saek	Li
1	‘vegetable’ ‘to walk’ ‘head hair’	*p ^h -	p ^h ak ^{DS1} p ^h a:j ^{C1} p ^h om ^{A1}	p ^h jak ^{DS1} p ^h ja:j ^{C1} p ^h jum ^{A1}	piək ^{DS1} pja:j ^{C1} piəm ^{A1}	p ^h rak ^{DS1} p ^h ram ^{A1}	*p ^h r-
2	‘eye’ ‘to die’ ‘grasshopper’	*t-	ta: ^{A1} ta:j ^{A1} tak ^{DS1}	ha: ^{A1} ha:j ^{A1}	ta: ^{A1} ta:j ^{A1} tak ^{DS1}	pra: ^{A1} pra:j ^{A1}	*tr-
3	‘to be exposed’ ‘to burst’ to hunt	*t-	ta:k ^{DL1} tɛ:k ^{DL1}	p ^h ja:k ^{DL1} p ^h e:k ^{DL1} p ^h jaw ^{B1}	ta:k ^{DL1} tɛ:k ^{DL1} taw ^{B1}	pra:k ^{DL1} pre:k ^{DL1}	*pr-

Assuming the view that PT only had monosyllabic prosodic words with canonical clusters, Li (1977) reconstructs the data in set 1, set 2, and set 3 as *p^hr-, *tr-, and *pr- respectively. It should be immediately clear that such a view cannot account for the correspondences in set 2 and set 3. Specifically, both of these sets must have gone back to consonant sequences that contain a labial consonant as suggested by the fact that Saek has /pr-/ for both. The labiality in the NT language Saek in set 2 cannot be viewed as an innovation because the labial stop is also attested in the CT dialect of Bao Yen, e.g. /p^hja:^{A1}/ ‘eye’, as well as in the Kam-Sui language Lakkja, e.g. /pla:^{A1}/ ‘eye’. The simple monosyllabic view clearly cannot account for the distinction between set 2 and set 3, and would have to reconstruct *pr- for both. Although one may argue for positing *fr- for set 2, such an ad hoc solution can be easily ruled out.

Crucially, it is hard to imagine how *fr-¹³ could have become *t- in SWT and CT. Moreover, it will become clear in §3.6.9.3 that *f- did not exist in PT.

In contrast, the sesquisyllabic hypothesis straightforwardly accounts for this paradigm of data. As Li's *p^hr- must in fact be reconstructed as *pr- (see discussion on lack of contrastive aspiration in PT in §3.5), Li's *tr- and *pr- can be reconstructed with *p.t- and *p.r- respectively. According to this scenario, the etyma in sets 2 and 3 must go back to two sesquisyllabic sequences-- *p.t- and *p.r-. These two PT onsets merged in some languages but are kept distinct in others. In Siamese and Yay, *p.t- and *p.r- merged into *p.t-, and later simplified to /t-/. Similarly, Saek merged *p.t- and *p.r- but now reflects the result of merger as /pr-/. In contrast, Lungchow keeps *p.t- and *p.r- distinct, and now shows /h-/ and /p^hj-/, respectively. In short, the sesquisyllabic view can successfully accounts for both labial reflexes without losing the distinction between set 2 and set 3.

Although the disyllabic view can account for this paradigm as well as the sesquisyllabic view, the monosyllabic view with marked clusters fails to capture the distinction between the three sets of data. As will become clear in §3.5, PT did not have contrastive aspiration, meaning that Li's *p^hr- was in fact *pr-. The elaborate monosyllabic view can posit *pr- for set 1, and *pt- for set 2, but it must leave set 3 unaccounted for. Because by definition all members of a true cluster must be linked to the same syllable node, a distinction between *pr- and *p.r- is thus ruled out. Essentially, the data in set 1 must be reconstructed with PT tautosyllabic *pr- contrasting with set 3 which goes back to PT sesquisyllabic *p.r- respectively. One may propose an ad hoc solution that perhaps the etyma in set 2 may go back to marked tautosyllabic cluster *fr-, *fpr-, or even *fpt-, but as pointed out earlier PT did not

¹³ Li (1977) does in fact reconstruct *fr- but he uses it to account for a different set of data.

have *f-. Therefore, it is now clear that the view that PT was strictly monosyllabic but allowed marked clusters is also not tenable.

Not only does the empirical evidence provided so far show that PT cannot have been strictly monosyllabic, it also shows that the sesquisyllabic hypothesis can effectively account for the whole range of comparative data. Unfortunately, empirical evidence that would allow us to compare the sesquisyllabic view against the disyllabic view does not exist. However, the sesquisyllabic hypothesis is much more plausible if the chronology of the shift toward monosyllabism is taken into account. More specifically, the sesquisyllabic hypothesis offers a more plausible scenario of how modern languages evolved from the PT system.

In SEA, the typical path of evolution is from disyllabism to sesquisyllabism, and then to monosyllabism (Ferlus 1996a; Matisoff 2001). This evolutionary scenario has been amply attested in the different language groups of SEA, e.g. Chamic (Haudricourt 1956; Lee 1966; Thurgood 1999), Tibeto-Burman (Matisoff 2001), and Vietic (Ferlus 1975). Assuming this evolutionary path, it goes without saying that the monosyllabic hypothesis is the most straightforward of the three. Because all attested daughter languages of PT are monosyllabic at least in the inherited portion of the lexicon, we can derive modern monosyllabicity directly from a monosyllabic proto-language. Well-known cases of non-monosyllabic Tai languages are Siamese and Lao, but it is clear that polysyllabic words in these two languages are either loanwords—mostly from Indic languages (Gedney 1965) and Khmer (Varasarin 1984)—, or results of reduction of earlier compounds (Kullavanijaya 1992b). This fact alone would suggest that the hypothesis that the PT prosodic word was strictly monosyllabic is the most likely one.

However, I have shown earlier that the monosyllabic view is insufficient as it fails to capture the range of comparative data. Therefore, the sesquisyllabic and

disyllabic view became the only remaining candidates for the characterization of the PT prosodic word. Although the two hypotheses conflict with the typological fact that all attested modern languages are monosyllabic, they still are not implausible. There are numerous cases of typological mismatches between a proto-language and the attested modern languages. Old Chinese, for example, was uncontroversially a non-tonal language (Baxter 1992; Haudricourt 1954a; Pulleyblank 1962; Sagart 1999; Schuessler 2007) but no modern Chinese dialect attests such a typological profile.

Between the two remaining hypotheses, the sesquisyllabic hypothesis offers a more chronologically plausible scenario than the disyllabic hypothesis. The sesquisyllabic view has to assume that different branches underwent monosyllabicization separately. However, it is still not impossible because it only takes one evolutionary step to reduce sesquisyllables to monosyllables, and there are many well-documented cases. In contrast, the disyllabic view would have to assume that different subgroups went all the way from disyllabicity to monosyllabicity separately without any modern language that would attest intermediate, perhaps sesquisyllabic, steps. Therefore, the chronological plausibility renders the disyllabic hypothesis only marginally possible.

In this section, I have shown that among the different views of the PT prosodic word, the sesquisyllabic view is the best hypothesis as it captures nicely the range of comparative data and at the same time offers a plausible scenario of the shift toward the monosyllabism in modern dialects. In contrast, the monosyllabic view is at first glance the most plausible hypothesis, if we only take the typology of modern Tai dialects into account. However, it fails to capture the whole range of sound correspondences among the dialects. On the other hand, the disyllabic can capture the comparative data quite nicely but it is very unlikely from the point of view of

chronological plausibility. Therefore, I argue that PT should be characterized as a sesquisyllabic language.

2.4 *The prosodic word in Southeast Asia*

As an integral part of the Southeast Asian linguistic area, Tai languages in all probability must have experienced some of the convergent currents that have swept through the area. It is therefore important in establishing the canonical shape of the PT prosodic word to examine prosodic types attested in languages of SEA. With respect to the prosodic word, Matisoff (2001) views syllable structure and word structure of Southeast Asian languages as constantly changing concomitantly with emergence and disappearance of tonal contrast. The Chamic branch of Austronesian is an excellent case that illustrates this constant restructuring of the prosodic word. Proto-Chamic evolved from a disyllabic language with stress mainly on the penult into the various types of Chamic languages found today (Lee 1966). Today, Chamic languages are either monosyllabic, sesquisyllabic, or disyllabic but always with ultimate stress (Lee 1966; Thurgood 1999). This difference among these closely-related languages highlights the fact that the canonical shape of the prosodic word is a property that is susceptible to change.

In this section, I review the three most common types of languages in SEA—monosyllabic, disyllabic, and sesquisyllabic. These types are defined by the canonical shape of their prosodic words. They will be discussed in relation to concepts relating to well-formedness in §2.2.2. It is crucial to remember that taking other languages in the area as models for reconstruction is not the same as forcing alien characteristics into the proto-language being reconstructed. Although the proto-language we are reconstructing may be structurally very different from other languages in the linguistic

area, areal typology helps contextualize the reconstruction and provide models against which the plausibility of the reconstruction can be checked.

2.4.1 *Monosyllabic languages*

Monosyllabic languages refer to linguistic varieties whose prosodic words are canonically monosyllabic words. Examples of monosyllabic languages in SEA include Dehong Tai (Luo 1997), Nhahuen (Ferlus 1998), Vietnamese (Đoàn 1992), Muong (Tài 1982), White Hmong (Ratliff 1992), Tsat (Thurgood 2005), among others. In these languages, both word maximality and word minimality must work together to ensure that each prosodic word is exactly one metrical foot. Logically a metrical foot can consist of one or two syllables, but obviously in monosyllabic languages a prosodic word canonically consists of only one syllable. According to the foot typology discussed in §2.2.2.3, only a heavy syllable can be a well-formed foot by itself. This means that the prosodic word in monosyllabic languages must be heavy. This indicates that there is an additional prosodic restriction in these languages that forces each syllable to be heavy.

Although monosyllabic languages share the word size requirements, they differ with respect to foot type, and permissible onsets and codas. With regards to foot typology, a monosyllabic language in theory can be either trochaic or iambic. However, evidence for analyzing a monosyllabic language as trochaic or iambic is usually not available. With respect to codas and onsets, they may differ in what consonant combinations are permissible as syllable onsets, and what kinds of codas are allowed. Monosyllabic languages in SEA may or may not allow complex onsets, but in the majority of languages only canonical clusters exist. In other words, the clusters in the overwhelming majority of monosyllabic languages in SEA respect SSP

and show the sonority distance of at least 2. To my knowledge, Khmer is the only language that potentially has marked clusters (Henderson 1952; Nacaskul 1978).

Some monosyllabic languages allow only simple onsets while others allow a wide range of consonant clusters. Dehong Tai (Luo 1999), spoken in Yunnan, is an example of a monosyllabic language that does not allow complex onsets. In addition to the constraints enforcing monosyllabicity, this Tai language also has a very strict constraint against complex onsets. As initial clusters do not exist in the language, the prosodic word in Dehong Tai is canonically CV(C) or CV(:)C. Table 2-6 provides the inventory of onsets in Dehong Tai.

Table 2-6 Permissible onsets in Dehong Tai (based on Luo 1999: xiii)

	labial	alveolar	palatal	velar	glottal
plain	p-	t-		k-	ʔ-
aspirated	p ^h -	t ^h -			
affricate		ts-			
fricative	f-	s-		x-	h-
nasal	m-	n-		ŋ-	
liquid		l-			
glide	w-		j-		

Nhahuen (Ferlus 1998) on the other hand allows two consonants to occur in the onset position, e.g. /cr-/ in /crɪəw/ ‘cliff’, /gj-/ in /gje:t/ ‘to kill’, /ɲw-/ in /ɲwɔ/ ‘bat’. Therefore, the Nhaheun word canon is C(C)V(C). Table 2-7 presents the inventory of complex onsets in Nhaheun.

Table 2-7: Complex onsets in Nhaheun (edited from Ferlus 1998: ii)

		labial	alveolar	palatal	velar
voiceless stops	-r-	pr-	tr-	cr-	kr-
	-l-	pl-			kl-
	-w-				kw-
voiced stops	-r-	br-	dr-	ʝr-	gr-
	-l-	bl-			gl-
	-w-		dw-	ʝw-	gw-
	-j-				gj-
aspirated stops	-j-	p ^h j-			k ^h j-
fricatives	-r-		sr-		
nasals	-r-	mr-		ɲr-	ŋr-
	-w-		nw-	ɲw-	ŋw-
geminate sonorants		mm-	nn- ll-	ɲɲ-	ŋŋ-
fricatives	-r-	wr-			

The inventory in Table 2-7 shows that this dialect of Nhaheun, except for geminates, obeys SSP outlined in §2.2.2.2. This means that in this language the first consonant in a complex onset is expected to be less sonorous than the consonants following it. For example, /gr-/ is possible because /g/ is less sonorous than /r/, cf. /grah/ ‘to bite’. On the same principle, the tautosyllabic cluster /rg/ does not occur in this language because /r/ is more sonorous than /g/. Similarly, /dw-/ is a permissible complex onset because /d/ is less sonorous than /w/. The geminate nasals are apparent exceptions to this generalization.

Unlike the onsets, the set of permissible codas in each language seems unrelated to the fact that the language is monosyllabic. While Vietnamese only allows

final stops, nasals, and glides (Đoàn 1992), prosodic words in Kammu may take either a stop, a fricative, a nasal, a liquid, or a glide as coda (Svantesson 1983). In the extreme case, Lahu (Matisoff 2001, 2006) disallows codas altogether. However, there are in fact some characteristics with regards to codas that are shared by monosyllabic languages of SEA. First, these languages lack phonation-type contrast in coda stops (Matisoff 2001; Pittayaporn 2005a; Rhee 2003). Second, it is perhaps safe to say that these languages quite consistently disallow complex codas (Rhee 2003). Nevertheless, these characteristics are independent of the monosyllabism because they are found in disyllabic and sesquisyllabic languages as well.

In sum, the canonical word shape in monosyllabic languages in SEA tends to be either CV(C) or CCV(C). That is, the prosodic word must be exactly one heavy syllable, but it is not clear what kind of a foot it is. When complex onsets are allowed, the second consonants almost always have relatively high sonority. Whether complex onsets are allowed or not depends on the individual languages. As for codas, there does not seem to be any correlation between monosyllabism and possible final consonants.

2.4.2 Disyllabic languages

Disyllabic languages refer to linguistic varieties in which the prosodic word predominantly consists of exactly two syllables. This type of languages is found only rarely in Mainland SEA but is widely attested in languages of Insular SEA. Examples of disyllabic languages in SEA include Acehnese (Durie 1985), Northern Roglai (Awoi-hathe et al. 1977), and Moken (Chantanakomes 1980; Pittayaporn 2005a). As in monosyllabic languages, both word maximality and word minimality must work together to keep each prosodic word exactly one metrical foot. For instance, Northern Roglai (Awoi-hathe et al. 1977), an Austronesian language of Central Vietnam, allows

only heavy monosyllables and disyllables consisting of a light syllable followed by a heavy one. Words that consist of only one light syllable, e.g. */bĩ/ or */mũ/, do not exist in this language.

In terms of foot typology, a bisyllabic foot can be either iambic or trochaic. Moken (Pittayaporn 2005a), Acehnese (Durie 1985), and Northern Roglai (Awoi-hathe et al. 1977; Lee 1966) are clear examples of iambic disyllabic languages. In these languages, the stressed syllable of a foot is always heavy, a fact which follows directly from Hayes's typology in §2.2.2.3¹⁴. Indonesian is an example of a disyllabic language with syllabic trochees (Cohn in progress). Although words can be longer than two syllables, an overwhelming majority of morphologically-simple words are bisyllabic. Interestingly, to my knowledge, no language with moraic trochees is attested in SEA. Also, note that these example languages all belong to the Austronesian language family.

Like their monosyllabic counterparts, disyllabic languages differ from one another by whether they allow complex onsets or not. Northern Roglai is an example of a disyllabic language with iambic feet. Morphologically simple words in Northern Roglai are predominantly disyllabic, although monosyllabic words are also common. Some disyllables are /rawak/ 'bridge', /publxj/ 'to sell' and /tjuh/ 'seven', contrasting with the following monosyllabic words /wak/ 'blade for weeding', /blxj/ 'to buy', and /joh/ 'to break'. According to my preliminary examination of the vocabulary, there is no restriction on what consonants can function as onsets of the two syllables. In other words, C₁ does not limit the choice of C₂ in disyllabic C₁V.C₂VC forms. Table 2-8 gives the inventory of northern Roglai complex onsets.

¹⁴ Hayes (1995: 82) acknowledges iambs consisting of two light syllables, i.e. even iambs, but states that they are rare and unstable.

Table 2-8 Complex onsets in Northern Roglai

		labial	alveolar	palatal	velar
C-	-r-	pr-	tr-	cr-	kr-
?C-					
Ç-		br-	dr-	ʝr-	gr-
C-	-l-	pl-	tl-		kl-
?C-		?bl-			
Ç-		bl-	dl-		gl-

Examples of Northern Roglai words with complex onset include /blop/ ‘hold in mouth’, /trom/ ‘trunk’, and /kacreh/ ‘sarong, skirt’. This means that onsets in Roglai also obey SSP and have the sonority distance of at least 2. In other words, the second element of a cluster must be more sonorous than the first consonant. In particular, only complex onsets that consist of a stop followed by a liquid are allowed.

In contrast to Roglai, Indonesian is a good example of a disyllabic language that by and large does not have complex onsets. The canonical shape of the Indonesian prosodic word is C₁VC₂.C₃VC₄. While only one consonant is allowed syllable-initially, there is no restriction on what consonants can function as onsets of the two syllables, i.e. C₁ and C₃. This lack of co-occurrence restriction of the two onsets is identical to the situation in Northern Roglai but is in sharp contrast with the situation in sesquisyllabic languages (see §2.4.3).

As in monosyllabic languages, the set of permissible codas in each language seem unrelated to the disyllabism. While Mishmi (Sastry 1984) only allows final stops and nasals, prosodic words in Indonesian may take a stop, a fricative, a nasal, a liquid, or a glide as coda (Rhee 2003). Moreover, disyllabic languages also share with monosyllabic languages the lack of phonation-type contrast in coda stops (Matisoff 2001; Pittayaporn 2005a; Rhee 2003) and the ban against complex codas (Rhee 2003).

These exact same characteristics are also found in monosyllabic and sesquisyllabic languages.

In short, the canonical shape of the prosodic word in disyllabic languages is C(C)V(C).C(C)V(C). When complex onsets are allowed, the second consonants have relatively high sonority. However, the onsets of the two syllables do not constraint each other. Many disyllabic languages do allow monosyllabic words although they often require their prosodic word to be no smaller than a heavy syllable.

2.4.3 *Sesquisyllabic languages*

Sesquisyllabic languages are languages whose prosodic words canonically consist of one syllable and a half. The term “sesquisyllable”, literally “one-and-a-half syllable”, was originally coined by Matisoff (1973) to refer to words that consist of an unstressed “minor” syllable followed by a fully stressed “major” syllable. The syllabicity of the minor syllable is carried by a neutral vowel or a syllabic consonant (Diffloth 1976: 232; Svantesson 1983: 27). Even though *a priori* the minor syllable can be located on either side of the major syllable, in the context of SEA the minor syllable always precedes the major syllable. Although Sloan (1988), Shaw (1994), Cho and King (2003), and I (Pittayaporn 2005c) have all argued that the minor syllable is a prosodically degenerate syllable of some kind, similar to the case of Polish clusters discussed in §2.2.2.2, it is likely that the so-called sesquisyllables differ structurally from one language to another. In this dissertation, I define a sesquisyllable as a prosodic word that consist of two syllables, the first of which is unstressed and lacks a phonological vowel. According to this definition, a sesquisyllable contrasts with a disyllable in that the first syllable of a disyllable does have a contrastive vowel.

An important typological observation is that sesquisyllabic languages usually allow for monosyllabic words while monosyllabic languages do not allow

sesquisyllabic forms. This difference is clearly illustrated by comparing the sesquisyllabic Kammu and its monosyllabic relative Vietnamese. A good example of a sesquisyllabic language is Kammu discussed above in relation to the importance of the prosodic word. Other examples of sesquisyllabic languages include Burmese (Green 2005), Palaung (Kasisopa 2003), Tariang (L-Thongkum 2001), Chru (Jrang et al. 1983), and Jingpho (Maran 1978), to cite a few.

With respect to the word size requirements, sesquisyllabic languages show effects of both word maximality and word minimality so that well-formed prosodic words are exactly one metrical foot. For example, a Kammu prosodic word is either C(C)VV, C(C)V(V)C, C(C).C(C)VV or C(C).C(C)V(V)C as discussed in §2.2.1. Assuming that the minor syllable is light, a sesquisyllable must be considered a prototypical iamb. This means that sesquisyllabic languages require their prosodic words to be exactly one iambic foot: either a single heavy syllable or a light syllable followed by a heavy syllable. Iambic feet consisting of two light syllables, also known as even iambs do not appear to be allowed in these languages (Pittayaporn 2005c).

Although the formal analysis of sesquisyllables may still leave room for discussion what sesquisyllables are formally, it is important to draw a distinction between sesquisyllables and disyllables (see §2.4.2). The crucial difference between the two types of prosodic word is that the onsets of the major and minor syllables interact very closely. While the onsets of the two syllables in disyllables are usually independent from each other, the onsets of the minor and major syllables in sesquisyllabic words may pose restrictions on each other. The Kammu language exemplifies such interaction very clearly. Prosodic words whose minor syllable and major syllable onsets are identical, i.e. */p.p-/ , */t.t-/ , */s.s-/ etc., are not attested in Kammu as described by Svantesson (1983). Second, sesquisyllabic words often violate SSP or show lower sonority distance than 2 (see §2.2.2.2). Therefore, the

minor syllable and the major syllable onsets together will be referred to in this study as “sesquisyllabic clusters”. This term is used in opposition to “tautosyllabic clusters”, which refers to true clusters that function as the onset of a single syllable.

Sesquisyllables also differ quite markedly from prototypical monosyllables. While complex onsets of monosyllables, i.e. monosyllables with unmarked initial clusters, obey SSP and show sonority distance of at least 2, complex onsets of sesquisyllables do not necessarily follow the generalizations. This distinction is shown in Table 2-9.

Table 2-9: Onsets and syllables in Kammu (from Pittayaporn 2005c)

Initials	monosyllabic	sesquisyllabic
obstruent+obstruent	✗	✓
obstruent+nasal	✗	✓
obstruent+liquid	✓	✓

We see that when a consonant sequence in Kammu violates SSP or has relatively low sonority distance among its members, it is parsed as sesquisyllabic. For instance, a sequence of two obstruents are always realized as sesquisyllabic, e.g. /p.téʔ/ ‘earth’, never */ptéʔ/. Similarly, a sequence of a liquid followed by an obstruent is also parsed as a sesquisyllabic cluster, e.g. /l.cò:t/ ‘to become snotty’, never */lcò:t/. On the other hand, if the consonant sequence obeys the SSP, it may be realized either as a sesquisyllabic or tautosyllabic cluster. Kammu in particular contrast /kl-/ as in /kló:k/ ‘bamboo bowl’ and /k.l-/ as in /k.ló:k/ ‘slit drum’. Independent evidence in Kammu for this contrast comes from tone patterning. Furthermore, in words whose major syllables have onset clusters, those onset clusters must respect the sonority hierarchy. For example, /s.kró:/ ‘to suddenly bend down head’ has an initial sesquisyllabic

cluster, which in turn consists of an initial /s-/ and a well-formed tautosyllabic cluster /kr-/. These generalizations are based only on Kammu, but to my knowledge they seem to hold for other sesquisyllabic languages as well. They play a crucial role in the reconstruction of PT complex onsets in §2.3.2 and Chapter 4.

In addition, many sesquisyllabic languages allow minor syllables that are more complex than a single consonant. For example, Jahai (Burenhult 2005), a Mon-Khmer language of Malaysia, has an abundance of sesquisyllabic words whose minor syllables consists of two consonants, e.g. /sl.tuh/ ‘to attack’, /cn.rɔs/ ‘nail’, /ʔn.teŋ/ ‘ear’. However, the second consonant has to be a sonorant and must agree in place of articulation with the onset of the major syllable if it is a nasal. Similarly, Kammu also allows for minor syllables that consist of two consonants but in addition to sonorants it also permits obstruents as the second consonant, cf. /ɿt.ʔət/ ‘plug’, /pɕ.rà:c/ ‘mesentry’. However, in Kammu these obstruents must be identical to the coda of the major syllable (Svantesson 1983: 28-29)¹⁵.

As in the other types of languages, the set of permissible codas in each language seem unrelated to the fact that they are sesquisyllabic. While Buyang only allows final stops and nasals (Li 1999), the coda of the major syllable of prosodic words in Nyah Kur may take a stop, a fricative, a nasal, a liquid, or a glide as coda (Diffloth 1984). Furthermore, sesquisyllabic languages also share with monosyllabic and disyllabic languages the lack of phonation-type contrast in coda stops (Matisoff 2001; Pittayaporn 2005a; Rhee 2003) and the ban against complex codas (Rhee 2003). These restrictions on codas are clearly a Southeast Asian phonological trait (Matisoff 2001; Rhee 2003).

¹⁵ This type of Kammu minor syllables is always morphologically complex, created either by prefixation or infixation (Svantesson 1983: 29).

In sum, the canonical shape of a prosodic word in sesquisyllabic languages is exactly one iambic foot, and may have one of the following shapes: C(C)VV, C(C)VC, C(C).C(C)VV, or C(C).C(C)V(V)C. While tautosyllabic clusters in sesquisyllabic languages obey SSP, and have the sonority distance of 2, sesquisyllabic clusters may or may not.

2.5 *PT as a sesquisyllabic language*

Having argued in §2.3.2 that the sesquisyllabic hypothesis is the best hypothesis for the reconstruction of PT, in this section I characterize PT as a sesquisyllabic language, very similar typologically to Kammu discussed in §2.4.3. The PT prosodic word can be either a monosyllable or a sesquisyllable, as argued for in §2.3. Moreover, comparative data discussed in detail in the following chapters permits us to view the PT word as having the following characteristics. First, each prosodic word contained exactly one iambic foot headed by a heavy syllable. Second, phonation-type and place contrasts in the onset were neutralized in the coda position. Last, each prosodic word is specified for one lexical tone. This section provides an explicit characterization of the PT prosodic word.

Diachronically speaking, this stage is one step away from the monosyllabism observed in modern languages along the evolutionary path taken by Chamic languages and others outlined above. In synchronic terms, the PT prosodic word was exactly one foot. In other words, word minimality and maximality both required that a prosodic word be a well-formed metrical foot, as in other sesquisyllabic languages. Furthermore, the metrical foot in PT was iambic as in sesquisyllabic languages (see §2.4.3). The prominent syllable of the PT prosodic word must have been the final syllable, i.e. the major syllable. Monosyllabic words were also permissible in PT

because a heavy syllable can by itself be an iamb. Possible shapes of PT prosodic words according to the current proposal are given in Table 2-10.

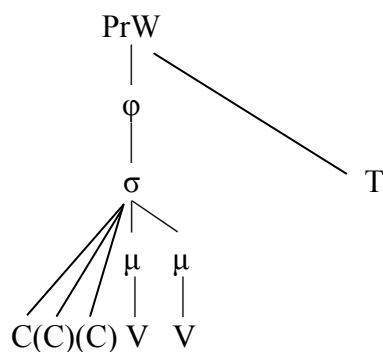
Table 2-10 Canonical shape of PT prosodic word

	Open syllable		Closed syllable
Monosyllable	*C(C)(C)V ^T	C(C)(C)VV ^T	C(C)(C)V(V)C ^T
Sesquisyllable	*C(C).C(C)(C)V ^T	C(C).C(C)(C)VV ^T	C(C).C(C)(C)V(V)C ^T

C = consonant, V = vowel, T = tone, * = not a possible word

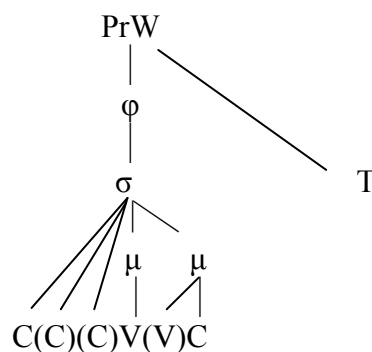
The PT prosodic word can be divided into three parts for the purposes of reconstruction: 1) onset, 2) rime, and 3) tone. The onset consists of any consonant that precedes the vocalic nucleus including the minor syllable. That is, it includes both the tautosyllabic and sesquisyllabic onsets as characterized above. The rime includes both the vocalic nucleus and the final consonant, if any. Lastly, the tone is the suprasegmental contrast linked to the major syllable; the minor syllable does not carry tone.

The PT prosodic word contained exactly one iambic foot. Comparative data discussed in §5.3 indicates that PT prosodic words ended in either open or closed syllables as schematized in Table 2-10. The only prohibition is against open syllables with one short vowel. In other words, only a heavy syllable can stand as a prosodic word by itself in the case of monosyllables, or function as the major syllable in the case of sesquisyllables. This is a clear manifestation of the minimality and maximality requirements. In terms of foot typology, this means that even iambs, i.e. iambic feet consisting of two light syllables, were not allowed. In the following subsections, I will outline the phonotactics of each of the three components according to the current proposal. An explicit characterization of the PT prosodic word is given in Figure 2-3.



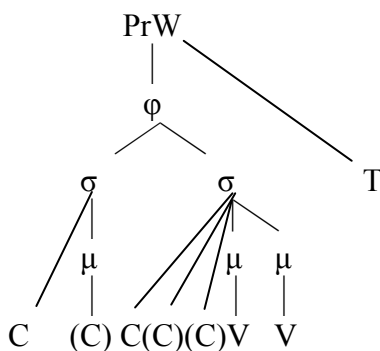
a. open monosyllable

Examples: *^hma:^A ‘dog’, *pi:^A ‘year’,
*^{da}:^B ‘to scold’



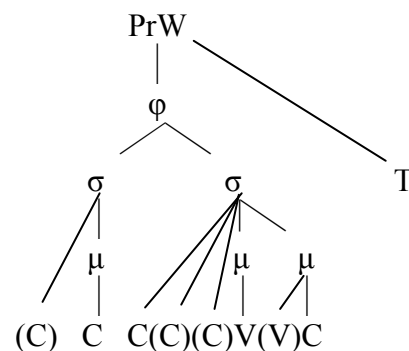
b. closed monosyllables

Examples: *^hmat^D ‘flea’, *^ɬlo:k^D
‘flower’, *^{ga}p^D ‘narrow’



c. open sesquyllable

Examples: *p.ta:^A ‘eye’, *k.ro:^A ‘to beg’,
*^Ç.ba:^B ‘shoulder’



d. closed sesquyllables

Examples: *p.ta:j^A ‘to die’, *k.tem^A
‘full’, *^{gm}.ra:^C ‘orphan’

Figure 2-3 Phonological representation of PT prosodic word

2.5.1.1 Onset

The PT prosodic word may have simple or complex onsets. A simple onset may be any of the consonants in the PT sound inventory¹⁶. PT had a three-way phonation type contrast among stops—voiceless, implosive, and voiced (see §3.5).

¹⁶ The proposed inventory of PT consonants is given in Table 3-1.

Moreover, it contrasted altogether six places of articulation—labial, alveolar, palatal, velar, uvular, and glottal (see §3.3). Complex onsets may be either tautosyllabic or sesquisyllabic. A tautosyllabic cluster is a sequence of consonants that functions as onset of a single syllable as illustrated in Figure 2-3a and Figure 2-3b. A tautosyllabic cluster in PT may consist of up to three consonants and must be canonical clusters. In other words, they respect SSP and show a minimum sonority distance of 2. Exceptions are tautosyllabic clusters containing a liquid followed by /-w-/, e.g. *lw-, *rw-, to cite a few. In these cases, the sonority distance between the liquid and the glide is only 1 on the sonority scale. Examples of permissible PT tautosyllabic clusters include *kr-, *kl-, *kw-, and *krw- (see §4.2).

In contrast, a sesquisyllabic cluster is a sequence of consonants that makes up the minor syllable plus the major syllable onset. A PT sesquisyllabic cluster in theory can have up to two consonants in the minor syllable, and up to three consonants in the major syllables, even though no sesquisyllabic clusters with five consonants is posited in this study. Examples of sesquisyllabic clusters are *k.t-, *p.k-, *k.r-, and *gm.r- (see §4.4). Crucially, in PT a contrast between *kt- and *k.t- is not possible but tautosyllabic *kr- did contrast with sesquisyllabic *k.r-.

2.5.1.2 Rime

The PT rime consists of two sub-parts—the vocalic nucleus and the coda. As represented in Figure 2-3, only the major syllable of a sesquisyllabic word has a vocalic nucleus. Therefore, the rime of a sesquisyllable refers specifically to the rime of the major syllable. PT vocalism shows a three-way vowel height contrast as well as a vowel length distinction (see §5.3). Any vowel in the PT phoneme inventory can occupy the nucleus of the syllable. However, the size requirement dictates that each prosodic word be exactly one iambic foot as represented in Figure 2-3. One

consequence is that open syllables with short vowels are banned. Open monosyllables with short vowels were banned because they are smaller than a foot. Similarly, sesquisyllabic forms with open major syllables ending in short vowels were not allowed because light syllables cannot head an iambic foot. Examples of words allowed by the PT phonotactics are *ka:^A, *kat^D, and *ka:t^D, but not *ka^A.

The second part of the rime is the coda. For each syllable, one coda consonant was allowed as represented in Figure 2-3. Not all PT consonants can occupy the final position of a prosodic word as phonation-type and place contrasts were neutralized in this position. Only the voiceless stops, the nasals, the lateral liquid, and the glides could function as codas. Similarly, not all places of articulation are found in the coda position. Of the six series, only labial, alveolar, palatal, and velar consonants can occupy the final position of a syllable (see §5.5).

2.5.1.3 Tone

Each PT prosodic word contained one lexical tone as represented in Figure 2-3. In sesquisyllabic words, only the major syllables carried distinctive lexical tones. A syllable ending in a vowel or a sonorant, “non-checked syllable”, could have either tone *A, *B, or *C. In contrast, a “checked syllable”, or stop-final syllable, is conventionally said to have had only one tone *D. From the point of view of phonological contrast, we can say that PT checked syllables did not contrast for tone. The distinction between *D on one hand, and *A, *B, and *C, on the other, was therefore one of syllable structure (§6.2). The three tones on non-checked syllables are tentatively posited as a modal tone with mid level contour, a creaky tone with low rising contour and relatively long vowel duration, and a finally glottalized tone with a high falling contour and short vowel duration (see §6.5).

2.6 *Summary*

In this chapter, I discussed the structure of the prosodic word in PT. I first discussed the importance of the prosodic word in the reconstruction of PT. I reviewed the monosyllabic, sesquisyllabic, and disyllabic views of PT and showed that the sesquisyllabic view can account for the range of comparative data as well as offer a plausible scenario of how PT developed into daughter languages. I have thus proposed that PT was a sesquisyllabic language whose prosodic word can be either a monosyllable or sesquisyllable. After providing a brief typological survey of the canonical shape of the prosodic word in SEA languages, I outlined the phonotactics of PT onsets, rimes, and tones. In contrast to earlier reconstruction, this dissertation addresses the issue of the PT prosodic word explicitly. The canonical shape of the PT prosodic word that has been established will be the framework for the segmental reconstruction. Detailed discussion of each part together with their reconstruction will be given in subsequent chapters.

CHAPTER 3

PROTO-TAI SIMPLE ONSETS

3.1 Introduction

Among the three components of the PT prosodic word, the onsets are probably the best understood. However, many issues await solutions. The best-known problem regarding the onsets of PT is the so-called “voicing alternation,” which involves discrepancies in the voicing of the onset initial between different groups of dialects as discussed in §1.3.2. In addition, one issue that has usually been overlooked is the number of places of articulation. In contrast to the conventional view that PT had labial, alveolar, palatal, velar, and glottal series, Haudricourt (1952), and Ostapirat (2007) posit a distinct series of uvular consonants, which seems to be preserved in PSWT (Pittayaporn 2008b, to appear-b). Moreover, I have elsewhere discussed gaps in the inventory of PSWT initial consonant clusters that need to be accounted for at the PT level (Pittayaporn 2008b, to appear-b).

In this chapter, I propose an inventory of simple PT onsets that effectively accounts for these outstanding problems. As a point of departure, I present the inventories of simple and complex onsets of PT and outline the phonotactics of the PT onsets. Then, I discuss three aspects of the present system: 1) the reconstruction of the uvular series, 2) the treatment of the so-called “voicing alternation,” and 3) the lack of contrastive aspiration. Lastly, I present the reconstructions of each of the PT simple onsets.

3.2 Inventory of PT simple onsets

Each PT syllable had an onset. I propose that there were 36 consonants in PT, all of which can occur alone as simple onsets. The proposed PT inventory is given in

Table 3-1. This section gives an overview of the onset system of PT, highlighting notable characteristics involving the manners and places of articulation as well as the phonation types. In the table, three consonants are put in parentheses because their reconstruction is tentative.

Table 3-1 PT consonant inventory

		labial	alveolar	palatal	velar	uvular	glottal
stops	voiceless	*p	*t	*c	*k	*q	
	voiced	*b	*d	*ɟ	*g	*ɢ	
glottalized		*ɓ	*ɗ	*ʝ			*ʔ
fricatives	voiceless		*s	(*ç)	*x	*χ	*h
	voiced		*z	(*ʒ)	*ɣ		
nasals	voiceless	* ^h m	* ^h n	* ^h ɲ	(* ^h ŋ)		
	voiced	*m	*n	*ɲ	*ŋ		
liquids and glides	voiceless	* ^h w	* ^h ɾ * ^h l				
	voiced	*w	*ɾ *l				

In this proposal, PT consonants can be divided into four groups according to their manner of articulation: 1) stops, 2) fricatives, 3) nasals, and 4) liquids and glides. Unlike Li who views the two palatal obstruents as affricates (Li 1977: 164), I consider them as stops phonologically. The choice is based mainly on the fact that these palatal consonants behave like stops in other places of articulation. First, the voiceless *c can occur as a coda, like *-p, *-t and *-k, as discussed in Chapter 5. Second, the voiced *ɟ always follow the same devoicing pattern as *b-, *d- and *g-. In languages where devoicing yields aspiration, *ɟ-, like *b-, *d-, and *g-, is reflected as an aspirated stop

/c^h-/ or its derivative /s-/, e.g. Siamese /cha:j^{A2}/ ‘male’ and Lao /sa:j^{A2}/ (from earlier *c^ha:j^{A2}). Similarly, these voiced obstruents are reflected as plain in languages that do not aspirate their reflexes of PT voiced *b-, *d-, *j- and *g-, e.g. Black Tai /ca:j^{A2}/ and Lungchow /ca:j^{A2}/. Lastly, treating *c and *j as stops makes the system economical and symmetrical. If they were considered affricates, a separate manner of articulation must be posited only for the two of them. Treating *c and *j as stops also makes the palatal series very similar to other places of articulation. However, this is not to argue that PT *c and *j were not affricated phonetically.

As for places of articulation, six different categories are posited for the proposed PT consonants: 1) labial, 2) alveolar, 3) palatal, 4) velar, 5) uvular, and 6) glottal. This reconstruction differs from most proposals (Haudricourt 1948; Li 1977; Nishida 1954; Sarawit 1973) by positing a distinct series of uvular consonants, but agrees with Haudricourt (1952)’s and Ostapirat (2007)’s proposals that uvular consonants be reconstructed for PT. The reconstruction of uvular consonants will be discussed in §3.3. Note that uvular and glottal places of articulation are defective in that they lack many consonant types found in the other places of articulation. Most notable is the lack of sonorants, which seems to be typologically common for places of articulation further back in the vocal tract.

As for phonation type, there was a three-way contrast in PT: 1) voiceless, 2) implosive/glottalized, and 3) voiced. When occurring as simple onsets, PT voiceless stops are reconstructed as unaspirated because they are reflected as unaspirated in all modern languages. In contrast, PT voiceless sonorants, except for *ʔj- are reconstructed as phonetically pre-aspirated because, as far as tonal development is concerned, they behave like aspirated stops in the modern varieties that have both plain and aspirated stops (Gedney 1972a; Li 1977). Moreover, the pre-aspiration is attested in loanwords from Tai to Austroasiatic languages, e.g. Red Palaung /^hla:n/

‘grandchild’ and /^hnoŋ/ ‘lake’ ultimately from PT *^hla:n[^] and *^hno:ŋ[^] respectively (Pittayaporn to appear-a). In addition, the implosives stops are viewed as half-voiced and half-voiceless. Specifically, they behave like voiceless consonants in terms of tonal development when occurred with certain tonal categories but like voiced consonant when occurred with others (Gedney 1972a).

The implosives *b- and *d- correspond to *^ʔb- and *^ʔd- in earlier constructions. Although the choice between implosive or glottalized stops does not have any implication to the rest of the system, reconstructing these sounds as implosives is preferable from a typological point of view. Specifically, cross-linguistically implosive /b/ and /d/ are more common than their counterparts further back (Croft 1990: 159; Maddieson 1984b: 111-114). However, these PT implosive stops pattern with *^ʔ- and *^j- in relation to tonal splits and mergers (see §6.2). For convenience, the series that consists of the implosives and the two glottalized consonants is referred to here as glottalized.

3.3 *A richer dorsal series*

Among the five contrastive series of consonants proposed in Li’s reconstruction, the velar series is the most problematic one. There are many inexplicable split reflexes for the many simple and complex velar onsets that he reconstructs. For example, each of Li’s *k^h- and *x- may be reflected as either /k^h-/ or /x-/ in White Tai, but the conditioning environment for the split cannot be identified. Another example is Li’s cluster *kr- which is reflected irregularly across dialects. Based on data from newly-described SWT languages, I have argued elsewhere that PSWT had two contrastive uvular consonants *q- and *χ- (Pittayaporn 2008b, to appear-b). At the PT level, Haudricourt (1952) proposes that there was one distinctive uvular stop *q- while Ostapirat (2007) believes that both the voiceless *q- and the

voiced *g- existed. Comparing the SWT data with data from NT dialects, I posit three uvular consonants for PT: *q-, *χ-, and *g-. Table 3-2 shows clearly that the velar series alone cannot account for the range of correspondences that involve dorsal onsets.

Table 3-2 Voiceless dorsal correspondence sets

	Li	Gloss	PT	PSWT	SWT		NT	
					White Tai	Kapong	Yay	Wuming
1	*x-, *k ^h -	‘rice’ ‘to open’ ‘guest’	-	*k ^h -	k ^h aw ^{C1} k ^h aj ^{A1} k ^h εʔ ^{DL1}	k ^h aw ^{C1} k ^h aj ^{A1} k ^h εʔ ^{DL1}	haw ^{C2} ha:j ^{A1} hek ^{DL1}	xaw ^{C2} ka:j ^{A1} hek ^{DS1}
2	*x-	‘white’ ‘galangal’ ‘green’	*x-	*x-	xa:w ^{A1} xa. ^{B1} xεw ^{A1}	k ^h a:w ^{A1} k ^h a. ^{B1} k ^h e:w ^{A1}	ha:w ^{A1} - hew ^{A1}	ha:w ^{A1} ha ^{A1} hew ^{A1}
3		‘to cross’ ‘to enter’ ‘ginger’	*χ-	*χ-	xa:m ^{C1} xaw ^{C1} xiŋ ^{A1}	ha:m ^{C1} haw ^{C1} hi:ŋ ^{A1}	ha:m ^{C1} haw ^{C1} hiŋ ^{A1}	ha:m ^{C1} haw ^{C1} -
4	*k ^h -	‘to kill’ ‘torn’ ‘arm’	*q-	*q-	xa: ^{C1} xa:t ^{D1} xεn ^{A1}	ha: ^{C1} ha:t ^{D1} hε:n ^{A1}	ka: ^{C1} ka:t ^{D1} cen ^{A1}	ka: ^{C1} ka:t ^{D1} ken ^{A1}

The etyma in set 1 must be reconstructed as *k^h- at the PSWT level. However, this PSWT aspirated velar stop cannot be projected back to PT because the etyma in this set either show voicing alternation (see §3.4), or are clearly relatively recent loans. The etymon for ‘rice’ along with many others have initials that go back to original voiceless onsets in SWT and CT but original voiced onsets in NT. The forms for ‘to open’ and ‘guest’ must be considered later borrowings from Chinese 開 *kāi* and 客 *kè* because they are irregular in NT. In addition to the aberrant initial /k-/ in Wuming, the vowels are unexpectedly long in ‘to open’ in both Yay and Wuming. The form for

‘guest’ in Wuming is also irregular as it shows DS1 tone rather than the expected DL1.

Although we can safely follow Li (1977: 207-214) in reconstructing /x-/ for etyma in set 2, Kapong forms in set 3 point to a different proto-phoneme. For both sets, White Tai, Yay, and Wuming show /x-/ , /h-/ , and /h-/ respectively. However, the Kapong dialect of Phu Thai, spoken in northeastern Thailand, shows two different reflexes for the two sets of etyma—/k^h-/ for set 2, and /h-/ for set 3. To account for Kapong /h-/ , which is produced further back in the vocal tract, I posit a uvular *χ- for etyma having the correspondence pattern of ‘to cross’, ‘to enter’, and ‘ginger’. PT *x- can be kept for ‘white’, ‘galangal’, and ‘green’. Although one may argue that the split reflex in Kapong is an innovation, neither a segmental nor a tonal conditioning environment can be identified. For example, the variation in vowels in these forms precludes positing a conditioned backing of *x- in Kapong.

The data in set 4 shows yet another pattern of correspondence. While the SWT dialects White Tai and Kapong have fricative reflexes /x-/ and /h-/ respectively, the NT dialects Yay and Wuming both show a stop /k-/. While one may be tempted to posit PT *k^h- to account for this discrepancy, data from the CT dialects of Long’an and Yongnan in central Guangxi rule out the possibility. In these two dialects, aspiration is contrastive, e.g. Yongnan /k^hat^{DS1}/ ‘section’ vs. /kat^{DS2}/ ‘tight’. Forms that are reconstructed with aspirated stop in Li’s system mostly show aspirated onsets as illustrated by Yongnan /p^hla:^{A1}/ ‘stone mountain’, /p^hlak^{DS1}/ ‘vegetable’, /t^hrən^{A1}/ ‘stone’, and /t^ha:n^{B1}/ ‘charcoal’ for Li’s *p^hra^A, *p^hrək^D, *t^hrin^A, and *t^han^B respectively. If the etyma in set 4 had had *k^h- as onsets in PT, we would expect to see aspirated reflexes in Long’an and Yongnan. However, the two dialects regularly show unaspirated reflexes in etyma belonging to set 4 as exemplified by Yongnan /ka:^{C1}/ ‘to

kill', /ka:t^{D1}/ 'to tear', and /ke:n^{A1}/ 'arm'. These two crucial dialects are therefore strong evidence for a lack of aspiration in the PT onsets of etyma belonging in set 4.

Therefore, I posit an unaspirated uvular *q- for this correspondence. Moreover, positing PT *q- for this set of etyma is also preferable from a typological perspective. That is, it is very unlikely for a language to have a fricative but lack the corresponding stop, except for the glottal series. The development of these voiceless initial dorsal consonants in White Thai, Kapong, and Yay is schematized in Figure 3-1.

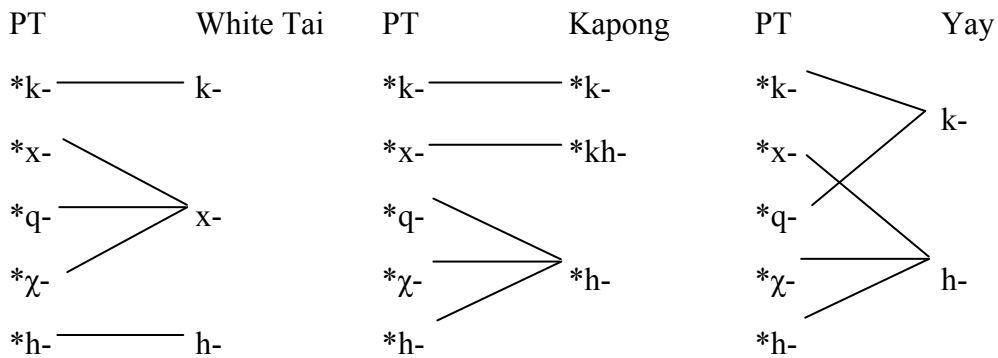


Figure 3-1 Development of dorsal onsets in White Thai, Kapong, and Yay

Not only is there evidence for the voiceless uvular set, data from modern dialects also point to voiced uvular onsets that contrasted with voiced velar onset in PT. Ostapirat (2007) shows that correspondence sets that Li assigns to *ɣ- in fact consist of two correspondence patterns. For one pattern, White Tai shows /x-/ regularly, pointing to PSWT *ɣ-. For the other pattern, the regular White Tai reflex is /k-/ suggesting PSWT *g-. He argues that there is no evidence for treating the two patterns as reflexes of two variants of the same sound and proposes that Li's *ɣ- must in fact be reconstructed as two different phonemes, *ɣ- and *g-. Data from a wider range of dialects given in Table 3-3 support this position.

Table 3-3 Voiced dorsal correspondence sets

	Li	Gloss	PT	PSWT	SWT	CT		NT
					White Tai	Leiping	Debao	Yay
1	*g-	‘pair’ ‘swollen’ ‘to chew’	*g-	*g-	ku: ^{B2} kauŋ ^{B2} kɛw ^{C2}	k ^h u: ^{B2} k ^h auŋ ^{B2} k ^h e:w ^{C2}	kow ^{B2} kɔj ^{B2} kew ^{C2}	ku: ^{B2} - cew ^{C2}
2	*ɣ-	‘carry pole’ ‘chin’ ‘human’	*G-		ka:n ^{A2} ka:ŋ ^{A2} kun ^{A2}	k ^h a:n ^{A2} k ^h a:ŋ ^{A2} k ^h ɯn ^{A2}	ka:n ^{A2} ka:ŋ ^{A2} kɔn ^{A2}	ha:n ^{A2} ha:ŋ ^{A2} hun ^{A2}
3			‘night’ ‘thatch grass’ ‘shin’	*ɣ-	*ɣ-	xɔ ^{A2} xa: ^{A2} xɛŋ ^{B2}	ho: ^{A2} ha: ^{B2} he:ŋ ^{B2}	vo: ^{A2} ja: ^{A2} -

The etyma in set 1 clearly go back to voiced velar stop *g-. This PT onset is reflected regularly as a stop in White Tai, Leiping, Debao, and Yay. Note that original voiced stops regularly yield aspirated stops in Leiping, cf. /p^{hi}:^{A2}/ ‘fat’ from *bwi:^A, and /t^ho:ŋ^{C2}/ ‘stomach’ from *dwu:ŋ^A. The etyma in set 2 show initial /k-/ and /k^h-/ in SWT and CT languages, indicating that the original onset was a stop. However, the regular Yay reflex for this set is /h-/, indicating that that stop cannot be *g-. Therefore, I follow Ostapirat (2007) in positing PT *G- for these items. Lastly, the etyma in set 3 are reflected as /x-/ in White Tai and /h-/ in Leiping and Yay, indicating that Li was correct to list them under *ɣ-. Note that the Debao reflex of PT *ɣ- is determined by the following segment. It is reflected as /v-/ before labial sounds and as /j-/ elsewhere. Interestingly, the corresponding Sui forms for ‘thatch grass’ /ja:/ and ‘dike between fields’ /jan/ cited as examples of PT *ɣ- in Ostapirat (2007) also have sonorant onsets. The development of the voiced dorsal onsets in White Tai, Debao, and Yay is schematized in Figure 3-2.

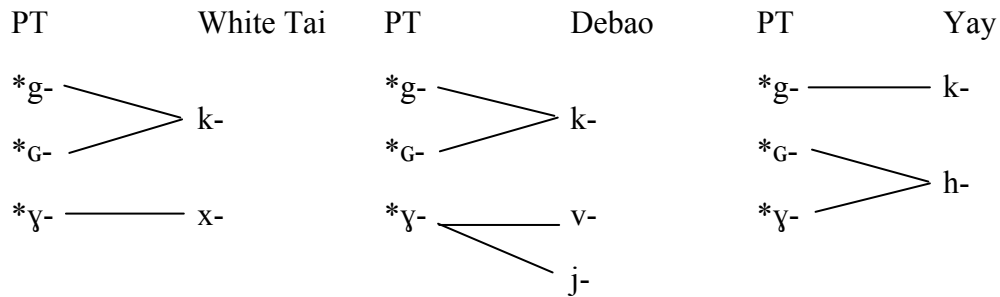


Figure 3-2 Development of dorsal onsets in White Thai, Debao, and Yay

Although none of the attested daughter languages have distinctive uvular consonants, I have provided solid comparative evidence for a distinct series of uvular consonants in PT. For this place of articulation, at least three phonemes must be posited: the voiceless stop *q-, the voiceless fricative *χ-, and the voiced stop *G-. Unlike its velar counterpart, the uvular place of articulation seems to be defective as it lacks sonorants altogether. Given that the acoustic structures of velar and uvular consonants are very similar (Ladefoged & Maddieson 1996: 36-37), we can view the lack of uvular sonorants as neutralization of dorsal consonants. Refer to §3.6.1.5, §3.6.3.5, and §3.6.4.4 for reflexes of individual etyma in different languages.

3.4 “Voicing alternation” revisited

One of the major problems in the field of Comparative Tai is the so-called “voicing alternation,” which refers to an apparent discrepancy regarding the voicing of onsets in different dialects. This discrepancy is generally described as a disagreement between SWT and CT on one hand, and NT on the other. For approximately 50 etyma found widely among Tai languages, tonal reflexes in SWT and CT points to earlier voiceless onsets while their NT counterparts indicate original voiced onsets (Gedney 1989a). For illustration, some cases of voicing alternation are given in Table 3-4.

Table 3-4 Examples of voicing alternation

Gloss	Li	Siamese	Lungchow	Po-ai
‘male’	*p ^h -	p ^h u: ^{C1}	p ^h u: ^{C1}	pu: ^{C2}
‘to carry’	*t ^h -	t ^h u: ^{A1}	t ^h u: ^{A1}	tu: ^{A2}
‘to ride’	*k ^h -	k ^h i: ^{B1}	k ^h wi: ^{B1}	kuj ^{B2}
‘ripe, cooked’	*s-	suk ^{DS1}	luk ^{DS1}	εuk ^{DS2}

The etymon ‘male’ is reflected in Siamese and Lungchow as /p^h-/ with tone C1 indicating earlier voiceless *p^h- but reflected in Po-ai as /p-/ with tone C2 pointing to original voiced *b-. Similarly, ‘to carry’ and ‘to ride’ are reflected in the two non-NT dialects as /t^h-/ and /k^h-/ with tones in the first series, indicating *t^h- and *k^h- at an earlier stage. For these same two etyma, Po-ai shows /t-/ and /k-/ with tones belonging to the second series, which are regular reflexes of voiced *d- and *g-. As for ‘ripe, cooked’, the /s-/ and /t-/ with tone DS1 in Siamese and Lungchow point to earlier *s-, but /ε-/ with tone DS2 in Po-ai indicates an original voiced initial¹⁷. These cases of etyma showing voicing alternation have remained a serious problem in the reconstruction of PT onsets.

Li (1977: 36-39) suggests that the voicing alternation was perhaps a derivational morphological process. He points out that there are doublets in Siamese that seem to show both the voiced and voiceless reflex. He cites Siamese pairs like /cam^{A1}/ ‘to confine, retain’ vs. /c^ham^{A2}/ ‘food susceptible of storage’, and /k^hiəw^{C1}/ ‘fang’ vs. /k^hiəw^{C2}/ ‘to chew’. Gedney (1989a), however, argues convincingly that the so-called voicing alternation cannot be morphological in nature. First, no specific

¹⁷ NT points to *j- rather than *z-.

morphological or syntactic function can be associated with these the alternating etyma. Second, it is rarely the case that a given modern language shows both the voiced and the voiceless variants. He also pointed out that it is rarely the case that both members of the pairs of Siamese words cited by Li are reconstructible at the PT level. In addition to Gedney's arguments, it should also be pointed out that even when both members of the alleged pairs go back to PT, often the similarity between the reconstructed forms is not as great as the similarity between their corresponding Siamese forms. In such cases, it is doubtful I speculate that voicing alternation that the two members of the alleged pairs actually go back to a single proto-form. For example, among Li's Siamese doublets /k^{hiəw}^{C1}/ 'fang' must be reconstructed with a uvular onset and a monophthong while /k^{hiəw}^{C2}/ 'to chew' goes back to a velar onset and a diphthong: PT *χe:w^C and *giəw^C respectively. In cases such as these, there is no reason to posit a voicing alternation because the words were completely distinct at the PT level. Therefore, facts like these cast doubt on the claim that the observed pattern of voicing alternation has a morphological origin.

In addition to Li (1977), a few proposals have been put forward. Ferlus (1990) posits sesquisyllabic forms to account for this voicing alternation, without providing any argument¹⁸. Gedney (1989a) in contrast, proposes that etyma showing the voicing alternation had a phonation type different from the four assumed in the conventional reconstruction. He speculates that this additional series might have been voiced and aspirated. Li (1989)'s later proposal closely resembles this position; the difference is largely notational. Lastly, Thurgood (2002a; 2007) proposes that forms showing the voicing alternation are post-PT borrowings, arguing that these etyma show internal

¹⁸ In the cited papers, he writes prenasalized stops as PT onsets of these voicing alternation items. However, he explains in a personal communication that the prenasalized stops are notational formulas for sesquisyllables of the type C.CVC or CN.CVC, where N represents any nasal.

irregularities and that many of them occur in other language families. Although a closer look at data from a wider range of Tai varieties seems to support Thurgood’s position, not all cases of voicing alternation can be explained away as loans. I conclude that voicing alternation items that cannot be treated as borrowings are native etyma that go back to PT sesquisyllables as Ferlus suggests. Examine the first set of data, given in Table 3-5.

Table 3-5 Voicing alternation items that show additional irregularities

Gloss	Li	Siamese	Lungchow	Po-ai
‘chopsticks’ ¹⁹	*t ^h -			tu: ^{B2}
‘log’	*k ^h -	k ^h ɔ:n ^{A1}		ku:n ^{B2}
‘to rake’	*f-	fɯə ^{A1}	p ^h u: ^{A1}	fɯ: ^{A1}
‘to cross the arms’	*k ^h -			ce:w ^{B2}
‘pole’	*s-	saw ^{A1}		la:w ^{A2}
‘boil, ulcer’	*f-	fi: ^{A1}		paj ^{A2}
‘to sharpen’	*f-	fon ^{A1}	p ^h ɯn ^{A1}	
‘eggplant’	*k ^h -	k ^h uə ^{A1}		kuə ^{A2}
‘cloth’	*p ^h -	p ^h a: ^{C1}		pu: ^{B2}
‘person’	*p ^h -	p ^h u: ^{C1}	pu: ^{C1}	pu ^{C2}
‘late in the morning’	*s-	sa:j ^{A1}	la:j ^{A1}	kwa:j ^{A2}

Among the etyma showing voicing alternation, many show other irregularities in addition to the initial consonants. Thurgood (2002a; 2007) shows convincingly that roughly a quarter of the etyma listed above should be eliminated. For example, ‘chopsticks’ points to *u: in SWT and CT but *u: in NT. This irregularity in the

¹⁹ This etymon is missing in Siamese and Lungchow. For SWT, see White Tai /t^hu:^{B1}/. For CT, see Lungming /t^how^{B1}/.

vowel suggests that these modern forms for ‘chopsticks’ are later borrowings. The etymon ‘log’ cannot be reconstructed for PT because the modern forms do not agree in tone and show various other minor problems. This suggests that they might not be related at all. The etymon ‘eggplant’ is reflected with an /e:/ in Saek instead of the expected /uə/. The SWT and NT forms for ‘cloth’ also disagree with respect to tones suggesting that they may not be related. Another example is ‘late in the morning’, which shows irregularities in vowel length, and tone, as well as onsets. The NT form, which suggests earlier *gw-, may not be related at all. Refer to Thurgood (2007: 246-248) for discussions of individual cases.

Thurgood also points out that many etyma were borrowings from outside of Kra-Dai. Although the etymologies he proposed for many cases are not convincing²⁰, it is clear that many of the voicing-alternation items are loans from other language families. The most important donor is of course Chinese. Table 3-6 presents etyma that have clear Chinese origins.

²⁰ For example, he erroneously links the Tai etymon for ‘ear’ (*k.rwu: in SWT and CT) to Chinese 耳 *ěr* (MC *ńzi*^B < OC *nə?*) and Proto-Tibeto-Burman *g-na. Another example is the Tai etymon for ‘compartment’ (*suəm^C in CT and SWT), which he connects to PAN *ɣumaq ‘house’. In this case, the Tai and Austronesian forms have very little in common. Their meanings also do not match well.

Table 3-6 Voicing alternation items that have clear Chinese origins²¹

Gloss	Li	Siamese	Lungchow	Po-ai	MC
‘bean’	*t ^h -	t ^h uə ^{B1}	t ^h u: ^{B1}	tu: ^{B2}	荳 dəu ^C
‘bowl’	*t ^h -	t ^h uəj ^{C1}	t ^h u:j ^{C1}	tu:j ^{C2}	斗 təw ^B
‘young male animal’	*t ^h -	t ^h uək ^{DS1}	tək ^{DS2}	tak ^{DS2}	特 dək
‘chopsticks’ ²²	*t ^h -			tu: ^{B2}	箸 djwo ^C
‘sugar’ ²³	*t ^h -		t ^h ɣ:ŋ ^{A1}		糖 daŋ
‘pond’	*t ^h -		t ^h um ^{A1}	tam ^{A2}	潭 dām
‘row’ ²⁴	*t ^h -	t ^h ɛw ^{A1}	te:w ^{A2}	te:w ^{A2}	條 diew
‘to ride’	*k ^h -	k ^h i: ^{B1}	k ^h wi: ^{B1}	ku:j ^{B2}	騎 gje
‘eggplant’	*k ^h -	k ^h uə ^{A1}		ku: ^{A2}	茄 gj
‘to dig’ ²⁵	*k ^h -	k ^h ut ^{D1}	ku ^{DS2}	hut ^{DS2}	掘 k ^h wət, or 掘 gjwət
‘cooked, ripe’	*s-	suk ^{DS1}	ɬuk ^{D1}	suk ^{D2}	熟 zjuk
‘ten’	*s-	sip ^{DS1}	ɬip ^{D1}	sip ^{D2}	十 zjəp
‘enemy, war’ ²⁶	*s-	suuk ^{DS1}			賊 dzək
‘right (hand)’	*k ^h w-	k ^h wa: ^{A1}	ɬa: ^{A1}	kwa ^{A2}	右 jəu ^B

The etyma in the table above all clearly have Chinese origins. Most items follow the regular tonal correspondences for shared Sino-Tai vocabularies (see §6.2), but many show unexpected tonal categories. For example, modern Tai forms for ‘to

²¹ The etymologies given in this column are from Thurgood, unless otherwise indicated.

²² For SWT, see White Tai, Black Tai, and Shan /t^hu^{B1}/. For CT, see Leiping /t^hu:^{B1}/, Lungming /t^how^{B1}/, Ningming /t^hau^{B1}/, among others.

²³ The MC form is from Pulleyblank (1991).

²⁴ Author’s etymology

²⁵ Author’s etymology

²⁶ The etymology is my own. For CT, see Western Nung /sɤk^{DS1}/, Ningming /sak^{DS2}/, and Qinzhou /ɛak^{DS2}/, etc. For NT, see Yay /sak^{DS2}/, Wuming /ɛak^{DS2}/, Yongbei /tsak^{DS2}/, etc. The forms in Ningming, Qinzhou, Wuming, Yongbei, and other NT dialects, all means ‘to steal’. Also see no. 685 in Appendix B.

ride’ all point to *B, which usually corresponds to MC Departing tone (^C). However, the MC form has the Level tone instead. Another case of tonal disagreement is ‘right’. Forms in Tai dialects all show *A, but the reconstructed MC form has the Departing tone. Note that the onset of this etymon is aberrant in CT varieties, i.e. /la:^{A1}/ instead of the expected /k^hwa:^{A1}/. The CT forms may have been contaminated by *za:j^C ‘left’, cf. Lungchow /la:j^{C2}/ (Gedney 1989a: 245).

The form ‘to dig’ deserves special attention. It is very likely that the modern Tai forms have different etymological sources. Siamese as well as other SWT varieties go back to original voiceless onsets, pointing to 掘 *kū* from MC *k^hwat*. In contrast, the Po-ai and other NT varieties, the etymological source for ‘to dig’ seems to be 掘 *jué* from MC *gjawat*. As for CT, some varieties side with SWT while others side with NT. For example, the Lungchow form /kut^{DS2}/ points to an earlier voiced onset but Yanshan Nung /k^hut^{DS1}/ indicates an earlier voiceless onset. One very important generalization regarding the etyma with clear Chinese etymologies is that their MC forms all have voiced onsets. In addition to the etyma found in Table 3-6, a number other forms may have also come from Chinese. Table 3-7 presents voicing alternation items with possible Chinese origins.

Table 3-7 Voicing alternation items with possible Chinese origins

Gloss	Li	Siamese	Lungchow	Po-ai	MC
‘person’	*p ^h -	p ^h u: ^{C1}	pu: ^{C1}	pu: ^{C2}	父 bju ^B or 夫 pju
‘male’	*p ^h -	p ^h u: ^{C1}	p ^h u: ^{C1}	pu: ^{C2}	夫 pju
‘to come into contact’	*t ^h -	t ^h u:k ^{DL1}	t ^h uk ^{DL1}	tuuk ^{DL2}	觸 t ^h juwok
‘to braid, harrow’	*f-	fua ^{A1}	p ^h u: ^{A1}	fu: ^{A2}	杷 ba ^(C)
‘to carry, to hold’ ²⁷	*t ^h -	t ^h u: ^{A1}	t ^h u: ^{A1}	tu ^{A2}	戴 tai ^C
‘hole, pit’	*k ^h	k ^h um ^{A1}	k ^h um ^{A1}	kum ^{A2}	坎 k ^h am ^B

The etymon ‘person’ may be related to Chinese 父 *fù* from MC *bju*^C (Thurgood 2007: 251). According to Schuessler (2007: 243), 父 was used to mark respected persons engaging in certain activities, e.g. 農父 *nóngfù* ‘minister of agriculture’. Phonologically, this Chinese etymon is a good match for Tai ‘person’ but the meaning does not match very well. Another possibility is suggested by Qin (1992), who argues that this etymon comes from 夫 *fū*, which functions as a suffix for men of various occupations, e.g. 農夫 *nóngfū* ‘farmer’. This etymology offers a better match in meaning but has the Level tone instead of the expected Rising tone. Similarly, the Tai etyma for ‘male’ most likely also goes back to 夫 *fū* but suffers the same sound correspondence problem as ‘person’. Tai ‘to hit, to touch’ may go back to Chinese 觸 *chù* but the onset is unexpected, though note that the reconstructed OC form has *t^h-. The etymon for ‘to braid’ may go back to Chinese 杷 *pá* but both the onsets and vowels are unexpected. Lastly, the etyma ‘to carry, to hold’, and ‘hole, pit’ may come from Chinese 戴 *dài*, and 坎 *kǎn* respectively but both the onset and tone are problematic.

²⁷ This is author’s etymology. Thurgood (2007) links this etymon to 帶 *dài* ‘belt, girdle’ (MC *tâi*^C).

The distribution of these etyma also reveals their post-PT origin. Both Gedney (1989a) and Li (1977: 36-39) claim that there is a correlation between the voicing alternation and the classification of Tai languages. In particular, reflexes in SWT and CT are thought to point to earlier voiceless onsets, while their NT counterparts indicate earlier voiced onsets. Luo (1996), however, shows that these generalizations do not always hold because there are many cases in which some non-NT dialects also show reflexes of earlier voiced onsets. Following Luo, I show in Table 3-8 that different etyma show different distribution patterns with regards to the original voicing of the onsets.

Table 3-8 Distribution patterns regarding original voicing of some voicing-alternation items

	SWT		CT		NT	
	Siamese	White Tai	Debao	Lungchow	Po-ai	Wuming
‘to carry’	t ^h u: ^{A1}	tu: ^{A2}		t ^h u: ^{A1}	tu: ^{A2}	tau ^{A2}
‘to ride’	k ^h i: ^{B1}	k ^h wi: ^{B1}	k ^h wɤj ^{B1}	k ^h wi: ^{B1}	ku: ^{jB2}	kuj ^{B1}
‘bean’	t ^h uə ^{B1}	t ^h o ^{B1}	t ^h u ^{B1}	t ^h u: ^{B1}	tu: ^{B2}	tu ^{B2}
‘thick, dense’	t ^h i: ^{B1}	t ^h i ^{B1}	tɤj ^{B2}	t ^h i: ^{B1}	ti: ^{B2}	toj ^{B2}

The reflexes of ‘to carry’ in White Tai as well as many other SWT dialects unexpectedly shows /t-^{A2}/ pointing to an earlier voiced onset, agreeing with NT rather than other SWT/CT dialects. For ‘to ride’, the Wuming form has /k-^{B1}/ indicating an earlier voiceless stop, disagreeing with other NT varieties. For ‘bean’, SWT/CT varieties have /t^h-^{B1}/ from earlier voiceless onsets while NT dialects show /t-^{B2}/ indicating an earlier voiced consonant. This is the expected pattern. Lastly, the modern

form for ‘thick, dense’ in Debao has /t^{-B2}/ indicating an earlier voiced stop, siding with NT rather than SWT/CT. This kind of overlapping distribution is a signature of diffusion, and demonstrates that many etyma showing voicing alternation should be considered post-PT borrowings that were propagated throughout the Tai-speaking area after PT broke up into different languages.

It has become clear that many voicing alternation items are post-PT borrowings as argued by Thurgood (2002a; 2007). It is now possible to consider how the Chinese etyma come to be reflected as voiced in some dialects but as aspirated in others. Taking the etymon ‘bean’ as a model case, I speculate that these Chinese etyma were borrowed into Tai when Chinese still preserved *NÇ- clusters. Sagart (1999: 185-187) proposes OC *N-tho[k]-s for 荳 *dou* < MC *dau^B. In this scenario, CT/SWT dialects simply dropped the *N- and borrowed this etymon as *t^huə^{B1} while NT dialects changed the prenasalized stop into a plain stop, and borrowed this word as *duə^{B1}. The borrowing may have happened after the OC period but definitely before the MC period because this *N-th- had become *d-* by the time of MC. If this scenario for ‘bean’ is correct, we can assume that other etyma in this set followed the same path.

However, there are many cases of voicing alternation that cannot be treated as post-PT loans. Firstly, they do not show other irregularities beside the voicing alternation. Secondly, they are also found in other branches of Kra-Dai. These etyma thus may have been part of the PT lexicon. Table 3-9 shows etyma that have Proto-Kam-Sui (Thurgood 1988), Proto-Hlai (Norquest 2007), or Proto-Kra (Ostapirat 2000) counterparts.

Table 3-9 Voicing alternation items that may have been part of PT their Kra-Dai cognates

	Siamese	Po-ai	Kam-Sui	Hlai	Kra
‘bitter’	k ^h om ^{A1}	ham ^{C2}	*kam ¹	*C.fɿ:m	*kəm ^A
‘excrement’	k ^h i: ^{A1}	hai ^{C2}	*ke ⁴	*fa:jʔ	*kaj ^C
‘to arrive’	t ^h uŋ ^{A1}	taŋ ^{A2}	*taŋ ¹		
‘ear’	hu: ^{A1}	ruə ^{A2}	*k ^h ra: ¹		*k.ra ^A

These four etyma seem to be good candidates for true Kra-Dai cognates. In Proto-Kam-Sui, they had voiceless stop onsets, except for *ke⁴ ‘excrement’. Similarly, the three Proto-Kra forms all had voiceless stops onsets. In contrast, the two Proto-Hlai forms cited both show voiced fricatives. This situation is in sharp contrast with those that can be shown to be Chinese loans, which had voiced onsets in MC. Therefore, it is unclear whether the voiceless reflexes in Siamese and other non-NT languages, or the voiced reflexes in Po-ai and the rest of NT preserved the original voicing. In fact, it is possible that the original onsets were complex and consisted of both a voiced and a voiceless consonant. Similar to Ferlus (1990), I tentatively posit sesquisyllabic clusters of the type *Ç.Ç- for voicing alternation items that cannot be explained as borrowings. In this scenario, ‘bitter’, ‘excrement’, and ‘to arrive’ should be reconstructed at the PT level as *Ç.qɿm^A, *Ç.qɿj^C and *Ç.tɿŋ^A respectively (see §4.4.3). One exception is ‘ear’, for which I reconstruct *k.rwu:^A and speculate that the *k.- was a prefix that was lost early on in NT but preserved in CT and SWT. Note that ‘to arrive’ might ultimately be related to PAN *datəŋ ‘to arrive’.

In this section, I have shown that the so-called voicing alternation is not a unified phenomenon. On the contrary, it consists of two distinct sets of data. The first set consists of those that are post-PT borrowings and must not be reconstructed for PT.

Not surprisingly, most of the etyma came from Chinese. The second set consists of etyma that cannot be explained away as loans and may have been part of the PT lexicon. I tentatively account for these cases by positing sesquisyllabic clusters as onsets of these etyma.

3.5 *Lack of aspiration in PT*

Aspiration has long been viewed as contrastive in PT (Haudricourt 1948; Li 1977; Nishida 1954, 1975a). In this conventional view, SWT and CT languages preserve the aspiration intact while their NT relatives lost it altogether. At first glance, this view of the PT consonant system seems to account for the dialectal diversity straightforwardly. However, Haudricourt (1963b: 157) and Ferlus (1990)²⁸ believe that aspiration secondarily developed in SWT and CT varieties out of original consonant clusters. A close examination of the data reveals that viewing contrastive aspiration as a characteristic of Proto-Tai is problematic.

The first problem is that loss of aspiration is not a common sound change. Recognizing that cases of de-aspiration are extremely rare, Ferlus (1990) argues that loss of aspiration is particularly implausible in Northern Tai languages, which have been heavily influenced by Chinese, where aspiration is very robust. The second problem is that the evidence for Li's simple aspirated stops *p^h-, *t^h-, *c^h-, and *k^h- is not as robust as it appears to be. Many of the etyma listed under these onsets show the so-called voicing alternation, and must be considered either later loans or reflexes of complex onsets (see §3.4). Moreover, etyma on Li's lists that invariably have aspirated onsets are attested by only a few examples, most of which are questionable or irregular (Gedney 1989a). Comparison with Kam-Sui also further casts doubt on the view that

²⁸ In a personal communication, Ferlus explains that aspiration in his reconstruction of PT is not contrastive but serves as a “junction” between the stop and the medial *-r-, i.e. *p^hr- is in fact *pr-.

PT aspirated consonants were contrastive. Li (1965: 55) noted that etyma with aspirated onsets in Kam-Sui do not often agree with the Tai aspirated onsets. More radically, Thurgood (1988) claims that aspiration in Kam-Sui and other branches is not an inherited trait from Proto-Kra-Dai. Similarly, Liang and Zhang (1993; 1996) claim that aspiration developed independently in different Tai varieties from various complex onsets, e.g. *pr-, *tl-, *tr-, *tr-, *pw-, *xp-, *xpl-, and *xt-.

In this section, I argue that PT lacked contrastive aspiration and that aspirated stops found widely among modern Tai languages are post-PT innovations. Specifically, I first show that etyma conventionally reconstructed with simple aspirated onsets that do not go back to original clusters were not part of the PT lexicon. I then propose that aspiration in CT and SWT developed mainly from aspiration of clusters with medial *-r-.

3.5.1 Li's simple aspirates as post-PT innovations

Although Li (1977) lists a number of items under simple aspirated stops *p^h-, *t^h-, *c^h-, and *k^h-, reconstructing contrastive aspiration for PT is in fact very problematic. As discussed earlier, the majority of these etyma show voicing alternation and therefore clearly do not go back to PT simple aspirates. Moreover, a number of etyma that Li reconstructed with *k^h- should in fact be reconstructed with uvular onsets (see §3.3). In this subsection, I show that most of the remaining items must be viewed as post-PT lexical innovations, either borrowings or forms derived after the establishment of the contrastive aspiration. These etyma can be divided into two groups—those that are found only in SWT and CT and those that are found in all the three groups. Table 3-10 lists etyma with aspirated onsets that are not found in NT.

Table 3-10 Forms with Li's simple aspirated onsets not found in NT

Gloss	Li	Siamese	Lungchow	Po-ai
'carpenter's square'	*c ^h -	c ^h a:k ^{DL1}		
'sing'	*k ^h -	k ^h ap ^{DS1}		
'cloudy, dark'	*k ^h -		k ^h am ^{A1}	
'turban'	*k ^h -		k ^h uun ^{A1}	
'to dry on fire'	*k ^h -	k ^h a:ŋ ^{A1}	k ^h a:ŋ ^{B1}	
'frog'	*k ^h -	k ^h iət ^{DL1}	k ^h i:t ^{DL1}	
'to coil'	*k ^h -	*k ^h ot ^{DS1}		
'mother's younger brother'	*k ^h -		k ^h u:C ¹	
'spicy hot'	*p ^h -	p ^h et ^{DS1}	p ^h it ^{DS1}	
'not exact, wrong'	*p ^h -	p ^h it ^{DS1}		
'to sweep' ²⁹	*p ^h -	p ^h ε:w ^{C1}		
'to knit' ³⁰	*t ^h -	t ^h ak ^{DS1}		

The etymon 'carpenter's square' is most likely a Siamese innovation. It is found only in Siamese and Lao. The Dioi form cited by Li (1977: 167) is probably not related. It may be related to Chinese 尺 *chǐ* (< MC *tʰjäk*) 'foot measure'. The forms 'to sing', 'turban', and 'to coil' are reflected with /x-/ in White Thai, pointing to either earlier *x-, *q-, or *χ-, rather than *k^h-. The /h-/ in the Saek reflex for 'to sing' also supports this interpretation. As for 'turban' and 'mother's younger brother', it is difficult to identify their original onsets because they are found exclusively in CT dialects, which have all lost the distinction. However, these etyma are ultimately from

²⁹ For Po-ai, Li (Li 1977: 64) gives /pew^{C1}/ 'to move residence'. In addition to the semantic problem, the short vowel is also unexpected. Therefore, this form is not related.

³⁰ For NT, Li (Li 1977: 102-103) cites Po-ai /tak^{DS1}/ 'knot', which may not be related.

Chinese 巾 *jīn* (< MC *kjen*) and 舅 *jiù* (< MC *gǝu^B*) respectively. Note that ‘turban’ is mostly likely borrowed via Vietnamese, cf. *xǎn*. The entry ‘to dry on fire’, found only in CT and SWT only, shows irregularity with respect to the tone. While the Siamese form goes back to *A, the Lungchow form indicates *B. The NT forms for ‘frog’ go back to PT **krwe*: but are not related to their CT/SWT counterparts. The form ‘to knit’ might not be found outside of Siamese. Manamaivibool (1975: 346) links this word to Chinese 織 *zhī* (< MC *tǝjək*) ‘to weave’. The forms for ‘spicy hot’ are found only in SWT and CT, assuming that the Saek forms for these etyma are loans from Lao or Thai. Similarly, ‘not exact’ is exclusively a SWT word. While ‘spicy hot’ might have been borrowed from an Austronesian source, cf. Malay /*pahit*/ ‘bitter’, the etymology of ‘not exact’ is unclear.

The crucial point is that the etyma in this set only occur in CT and SWT, never in NT. Many of them point to initials other than **k^h*-, i.e. **x*-, **q*-, or **χ*-. Although it is not possible to explain the etymological sources of every etymon, the words in this set cannot be regarded as cases of retention from PT on the part of CT and SWT. This is because they are the only instances of Li’s simple aspirates. If these particular forms were retentions, we also would expect to see in NT a sizeable number of other etyma shared with either CT or SWT that fit this correspondence pattern. Since no such item is found, these etyma must be considered post-PT vocabulary that were innovated after CT and SWT dialects had developed contrastive aspiration through independent processes (see §3.3 and §3.5). Unlike the set of data just discussed, the second set, given in Table 3-11 consists of etyma reconstructed with simple aspirated stops by Li that are shared by SWT, CT, and NT.

Table 3-11 Forms with Li's simple aspirated onsets that are found in all the three groups

Gloss	Li	Siamese	Lungchow	Po-ai	MC
'to split'	*p ^h -	p ^h a: ^{B1}	p ^h a: ^{B1}	pa ^{B1}	破 p ^h wâ ^C
'to roll, to twist'	*p ^h -	p ^h an ^{A1}	p ^h an ^{A1}	pan ^{A1}	翻 p ^h jwɛn
'clf. for boards, flat objects, etc.'	*p ^h -	p ^h ɛn ^{B1}	p ^h e:n ^{B1}	pe:n ^{B1}	片 p ^h ien ^C
'charcoal'	*t ^h -	t ^h a:n ^{B1}	t ^h a:n ^{B1}	ta:n ^{B1}	炭 t ^h ân ^C
'to pull out'	*t ^h -	t ^h ɔ:n ^{A1}		t ^h o:n ^{B1}	
'to retreat'	*t ^h -	t ^h ɔ:j ^{A1}	t ^h ɔ:j ^{B1}	t ^h ɔ:j ^{B1}	退 t ^h wâi ^C
'to remove'	*t ^h -	t ^h ɔ:t ^{DL1}	t ^h ut ^{DS1}	to:t ^{DL1}	脫 t ^h wât
'granary'	*c ^h -	c ^h a:ŋ ^{A1}	ɛa:ŋ ^{A1}	ɛa:ŋ ^{A1}	倉 ts ^h âŋ
'to tear' ³¹	*c ^h -	c ^h i:k ^{DL1}	ɛi:k ^{DL1}	li:k ^{DL1}	拆 t ^h ɛk
'to lie athwart, broad' ³²	*k ^h w-	k ^h wa:ŋ ^{A1}	k ^h wa:ŋ ^{B1}	kwa:ŋ ^{B1}	橫 ywaiŋ

Although the etyma in Table 3-11 are found in all three groups and are reflected with the expected onsets, they cannot be reconstructed at the PT level. First of all, a few of them show unexpected vowels or tones. For 'to pull out', the tones in Siamese and Po-ai do not agree. Similarly for 'to retreat' and 'to lie athwart, broad', Siamese points to *A while Lungchow and Po-ai indicate *B. As for 'to remove', the Lungchow form has short /u/ instead of the expected long /o:/. More crucially, it cannot be coincidental that all but one has clear Chinese etymologies. This strongly argues for treating these etyma as post-PT borrowings. In other words, they were not part of PT but were borrowed from Chinese separately after the break-up of PT into

³¹ This etymology is my own. Pulleyblank reconstruct *t^hre:jk*.

³² This etymon is not listed in Li (1977) but follows the same pattern.

different branches. In CT and SWT, the borrowing must have occurred after contrastive aspiration had emerged (see §3.5.2 below). This allows these Chinese loans to maintain their aspiration when incorporated into the recipient Tai dialects. On the other hand, NT dialects, which never acquired contrastive aspiration, lost the Chinese aspiration in the process of loanword adaptation.

3.5.2 *Aspiration of original PT clusters*

Having established in the previous section that Li's simple aspirated onsets in fact cannot be reconstructed for PT, we must then expect clusters with aspirated stops found in modern Tai languages to have developed secondarily. In the conventional reconstruction of PT onsets (Haudricourt 1948; Li 1977; Sarawit 1973), the aspirated stops can be combined with medial *-r-, *-l-, and *-w- to form complex onsets. In contrast to simple aspirated onsets discussed above, evidence for clusters consisting of an aspirated consonant followed by a medial *-r- are very robust. This is suspicious because clusters should consist of simple consonants that exist independently in the language. This indicates that Li's clusters with aspirated stops plus medial *-r- must in fact be reconstructed differently. Similarly, Ferlus (1990) does not reconstruct simple aspirates but posits clusters of type aspirated stops plus *-r-. However, he explains in a personal communication that he considers the aspiration in those clusters to be a "junction sound".

In this section, I argue that native Tai etyma that are reflected with aspirated onsets in CT and SWT languages must have had medial *-r- at an earlier stage. The sets of data in Table 3-12 illustrate this process. It is this medial *-r- that caused PT plain voiceless stops to become aspirated. From the point of view of this proposal, Ferlus's system represents an intermediate stage when the aspiration induced by the medial *-r- is still sub-phonemic.

Table 3-12 PT clusters with medial */-r-/

Gloss	Li	Siamese	Debao	Po-ai	Saek	PT
‘vegetable’	*p ^h r-	p ^h ak ^{DS1}	p ^h jak ^{DS1}	pjak ^{DS1}	p ^h rak ^{DS1}	*pr-
‘forehead’	*p ^h r-	p ^h a:k ^{DL1}	p ^h ja:k ^{DL1}	pja:k ^{DL1}	p ^h ra:k ^{DL1}	*pr-
‘head hair’	*p ^h r-	p ^h om ^A	p ^h jam ^A	pjɔm ^A	p ^h ram ^A	*pr-
‘headlouse’	*t ^h r-	haw ^A	t ^h aw ^A	law ^A	raw ^A	*tr-
‘tail’	*t ^h r-	ha:ŋ ^A	t ^h a:ŋ ^A	lu:ŋ ^A	ruəŋ ^A	*tr-
‘stone’	*t ^h r-	hin ^A	t ^h ɣn ^A	lam ^A		*tr-
‘plough’	*t ^h l-	t ^h aj ^A	t ^h aj ^A	saj ^A	t ^h aj ^A	*cr-
‘to wait’	*t ^h l-	t ^h a: ^C	t ^h a: ^C	sa: ^C	t ^h a: ^C	*cr-
‘to seek’	*xr-	ha: ^A	k ^h ja:	la: ^A	ra: ^A	*kr-
‘six’	*xr-	hok ^{DS1}	k ^h jok ^{DS1}	lɔk ^{DS1}	rɔk ^{DS1}	*kr-
‘to laugh’	*xr-	huə ^A	k ^h u: ^A	li:w ^A	ruəw ^A	*kr-

In the table above, Siamese represents SWT, Debao represents CT, and Po-ai and Saek represent NT. The first set of etyma must go back to PT *pr- rather than *p^hl/r- as reconstructed by Li. The medial *-r- had an effect on the preceding *p- such that it became aspirated in Siamese, Debao, and Saek. This process of aspiration apparently did not occur in Po-ai. For the second set, the PT cluster *tr-, which correspond to *t^hr- in Li’s system became aspirated in Debao before the loss of *-r-. In Siamese and Saek, the *tr- may have first also become aspirated to *t^hr- and then developed into a voiceless *^hr-. In contrast, it is clear that in Po-ai the *tr- did not aspirate but became simplified directly to *^hr-, which is now reflected as /l-/. In other words, the initial *t- was lost but the voicelessness is preserved on the medial *-r-. This is because, as evidenced in the case of PT *pr-, Po-ai did not participate in the

aspiration process. Skipping to the fourth set, for which Li reconstructs *xr-, we see a very similar situation. In all languages shown except for Debao the velar onset in the cluster *kr- was lost, resulting in a voiceless sonorant *hr-. Lastly, the third set presents cases of PT *cr- for Li's *thl-. The cluster simply lost the medial *-r- in Po-ai but became *c^hr- through the aspiration process in Siamese, Debao, and Saek. This palatal cluster later became *t^hr-, and then *t^h-.

Another source of aspiration in CT dialects is sesquisyllabic clusters that at some stage simplified to tautosyllabic cluster with medial *-r- (see Chapter 4). In Li's system, most of these clusters are reconstructed as plain stops plus medial liquids; however this reconstruction cannot account for the observed data. For example, Li's *tr- cannot account for the labiality of the Saek forms, which should in fact be reconstructed with a labial consonant. Following Haudricourt (1956), and Ferlus (1990), I reconstruct this PT complex onset as *p.t-, cf. Lakkja *plej^{A1} 'to die' and PAN *matáy 'to die'. In Debao, this *p.t- must have first become *tr- due to monosyllabicizing pressure and assimilation of the initial stop to *-r-, merging with original *tr-. In Saek, *p.t- became *pr- rather than *tr- but was kept apart from original PT *pr- which had already gone through the process of aspiration. Table 3-13 gives examples of sesquisyllabic clusters that are reflected as aspirated clusters in modern languages.

Table 3-13 Examples of PT sesquisyllabic clusters that are reflected with aspiration

Gloss	Li	Siamese	Debao	Po-ai	Saek	PT
‘to die’	*tr-	ta:j ^A	t ^h a:j ^A	ta:j ^A	pra:j ^A	*p.t-
‘eye’	*tr-	ta: ^A	t ^h a: ^A	ta: ^A	pra: ^A	*p.t-
‘grasshopper’	*tr-	tak ^{DS2} -t	t ^h ak ^{DS1}	tak ^{DS1}		*p.t-
‘centipede’	*k ^h r-	k ^h ep ^{DS1}		t ^h ip ^{DS1}	t ^h rip ^{DS1}	*q.s-
‘to imprison’	*k ^h l- ^A	k ^h aŋ ^{A1}	k ^h jaŋ ^{A1}	caŋ ^{A1}	t ^h raŋ ^{A1}	*k.r-
‘to drive away’	*k ^h l- ^A	k ^h ap ^{A1}		cap ^{A1}		*k.r-

Similar to the case of *p.t-, PT *q.s- in ‘centipede’ became *q.r- merging with original *q.r-, which is now reflected as /k^h-/ in Siamese and /k^hj-/ in Debao. In Po-ai, *q.s- simplified to *s-, which is now reflected as /t-/. In addition, some sesquisyllabic clusters that had *.r- as onset of the major syllable were reduced to tautosyllabic clusters with medial *-r- as illustrated by the case of ‘to imprison’ and ‘to drive away’ above. In these cases, the PT sesquisyllabic cluster *k.r- was reduced to a tautosyllabic cluster *kr- before becoming *k^hr-. This intermediate cluster is now reflected as /k^h-/ in Siamese, /k^hj-/ in Debao, /c-/ in Po-ai, and /t^hr-/ in Saek³³. For more discussion on reduction of sesquisyllabic clusters, refer to §4.5.

Further evidence for this proposal comes from Kam-Sui languages. Etyma that are reconstructed with *-r- clearly correspond to unaspirated clusters in Thurgood (1988)’s preliminary reconstruction of Proto-Kam-Sui. For example, PT *p^rɣm^A ‘head hair’, *pra: ‘stone mountain’, *cra:^C ‘to wait’, and *kruəw ‘to laugh’ correspond to *pram¹, *pra¹, *kra³, and *kru¹ in Proto-Kam-Sui respectively. These Kam-Sui forms

³³ Saek changed *k^hr- to /t^hr-/ regularly.

strongly support the view that the aspiration in clusters with medial *-r- developed secondarily in the CT and SWT of daughters of PT. The lack of aspiration in the NT reflexes of these etyma is therefore an inherited trait from Proto-Kra-Dai.

Aspiration induced by medial -r- is a not an unusual sound change. Many cases have been reported, especially in SEA. In Tai Yuan³⁴, the SWT dialect of Thailand, clusters consisting of an unaspirated stop followed by a medial *-r-³⁵ are now pronounced as aspirated stops, e.g. /p^hɛ:t^{DL1}/ from earlier *pre:t^{DL} (< Sanskrit *preta*-) ‘hungry ghost, /k^hiŋ^{B1}/ ‘to be suspicious’ from *kriŋ^B, /k^ha:p^{DL1}/ ‘to pay respect’ from *kra:p^{DL}, and /p^hɛ:A¹/ ‘to vary’ from *pre:A^A. These words are still pronounced as /pre:t^{DL1}/, /kriŋ^{B1}/, /kra:p^{DL1}/, and /pre:A¹/ in Siamese. Moreover, the first three etyma are spelled as *pre:t*, *kriŋ*, and *kra:p* in Tai Yuan palm-leaf manuscripts respectively. Another example of aspiration induced by medial *-r- is from Khmer. In the colloquial Phnom Penh dialect, initial stops /p-, t-, k-/ when followed by /-r-/ are pronounced as [p^h-, t^h-, k^h-], e.g. /pram/ ‘five’ pronounced as [p^heam], and /prap/ ‘to tell’ as [p^hɛap] (Guion & Wayland 2004; Wayland & Guion 2005). In other words, the /-r-/ aspirated the initial stops before dropping off, causing a change in the vowel quality.

Phonetically, aspiration of *-r- clusters is a very plausible sound change. Wayland and Guion (Guion & Wayland 2004; Wayland & Guion 2005) propose with respect to the Khmer case just discussed that aspiration in Cr- clusters helps facilitate the production of the medial /-r-/. From an acoustic point of view, a period of aspiration develops between the stop and the medial /-r-/ because the glottis remains open throughout the production of the cluster. Moreover, from an aerodynamic point of view, the aspiration increases the airflow needed in the production of the medial /-r-/. Specifically, spread glottis, which is associated with aspiration, allows for relatively

³⁴ The modern Tai Yuan forms are taken from Wichiankhiaw et al (1996).

³⁵ These clusters are found in loanwords only. Original PT clusters of this type had gone through several changes by the time of PSWT.

strong and high volume of airflow needed to initiate the trill. Therefore, the aspiration of original Cr- clusters found in CT and SWT dialects is a natural sound change.

The view that PT had no distinctive series of aspirated onsets has a few structural advantages over the conventional view that four contrastive phonation types existed at the PT level. Firstly, it explains curious gaps in the PSWT stop system I have discussed elsewhere that would otherwise be unexplained (Pittayaporn 2008b; to appear-b). Despite the rich PSWT consonant inventory, only labial and velar stops can be combined with medial *-l- and *-r- as illustrated in Table 3-14. The unaspirated stops were combined only with medial *-l- and never with *-r-. In contrast, aspirated stops only took medial *-r-, as the only aspirated cluster in PSWT is *k^{hr}-.

Table 3-14 Clusters consisting of a stop plus a liquid in PSWT

		labial	velar
*C ^h -	*-r-		*k ^{hr} -
*C-			
*C̣-		*br-	*gr-
*C ^h -	*-l-		
*C-		*pl-	*kl-
*C̣-		*bl-	*gl-

In the proposed reconstruction, these gaps are accounted for straightforwardly as consequences of the aspiration of PT *pr-, and *kr-. Specifically, these two clusters aspirated and became *p^{hr}-, and *k^{hr}-, each leaving a gap behind. These intermediate clusters then simplified to *p^h- and *k^h- in PSWT respectively. The gap in the velar series left behind by the simplification of *k^{hr}- was filled by original *qr-, *k.r-, and *q.s all of which became *k^{hr}- in PSWT (see Chapter 4). Similarly, the fact that

aspirated clusters with medial *-l- did not exist in PSWT is accounted for very nicely as in this proposal PT *-l- did not aspirate its preceding stops. Note that, on this view, the PSWT simple aspirated series came from different sources, including PT *cr-, PT *qr-, and later Chinese loans, etc. This scenario is schematized in Table 3-15.

Table 3-15 Schematization of the development of PSWT stop system from PT

		labial	velar	
*C-		p ^h -	*k ^h -	← loanwords, voicing alternation
*C ^h -		*p-	*k-	← loanwords
*C̣-		*b-	*g-	
*C-		*pl-	*kl-	
*C ^h -	-l-			
*C̣-		*bl-	*gl-	
*C-		(PT *pr-)	*k ^{hr} -	← PT *qr-, *q.r-, *ks-
*C ^h -	-r-		(PT *kr-)	← PSWT *hr-
*C̣-		*br-	*gr-	

The second advantage of the reconstruction proposed above is that the PT palatal series is no longer defective when compared to other places of articulation. Each of the five places of articulation in PT had a plain voiceless stop, and a voiced stop, except for the labial and alveolar series which also had an implosive. All these PT stops can be combined with either an *-l- or an *-r-. In this system, the fact that *c^h- is not reconstructible for PSWT is then explained as a result of the change from PT *cr- to *t^h- instead of *c^h-. This change can be understood as a ban against palatal clusters, which are phonologically marked.

Another appeal of the proposed PT reconstruction is that it offers an explanation for why some Chinese loanwords with aspirated initials unequivocally go back to earlier unaspirated stops in Tai. Among the Sino-Siamese vocabulary identified by Manomaivibool (1975) a few etyma have aspirated initials in Chinese but unaspirated initials in Siamese as shown in Table 3-16.

Table 3-16 Forms showing aspirated initials in Chinese but unaspirated in Tai.

Gloss	Siamese	Lungchow	Po-ai	MC	OC
‘seven’	cet ^{D1}	cit ^{D1}	ɕet ^{D1}	ts ^h jət	七 *tshit
‘lung’	pɔ:t ^{D1}	puut ^{D1}	puut ^{D1}	p ^h jwɛi ^C	肺 *phots
‘duck’	pet ^{D1}	puut ^{D1}	puut ^{D1}	p ^h jiet	鴨 *phit

Manomaivibool (1975: 124) notes that such alternations in the correspondences between Chinese and Thai forms occur frequently and hints that it may possibly indicate a morphological process yet to be identified. In contrast, the reconstruction proposed here explains this phenomenon quite nicely by treating these etyma as early loans into PT before aspiration was developed. Because PT did not have aspirated stops, Chinese words with aspirated stops must have gone through a process of loanword adaptation whereby the initials were converted into the most similar unaspirated stop. The implication is then that all Chinese borrowings in Tai languages that show aspiration were incorporated after the break-up of PT.

In sum, we can see that all of Li’s clusters with aspiration must be given an alternative set of reconstructions. For example, *p^hr- and *t^hr- must be replaced by *pr- and *tr- while *t^hl- must be rewritten as *cr-. Li’s clusters that consist of a plain stop followed by medial *-r- must be reconstructed as sesquisyllabic onsets. For instance, Li’s *tr- corresponds to *p.t- in the current system while some of the etyma

with *kr- must in fact be reconstructed with *q.s-. Some clusters with fricatives must also be replaced by complex onsets with plain stops followed by *-r-. For example, Li's *xr- is reconstructed as *kr- in the system proposed here.

3.6 Reconstructions of PT simple initials

Etyma from Siamese (SI), Lungchow (LC), and Yay (Y), are provided in each section to represent the three clusters of the modern varieties. Remember that this is not to say that they represent three genealogical subgroups of Tai (See §1.4.2.3).

3.6.1 Voiceless stops

There were altogether five voiceless stops that could function as onsets in PT: *p-, *t-, *c-, *k-, and *q-. Except for the palatal *c- and the uvular *q-, these PT stops are kept intact in all modern varieties. The glottal stop is *ʔ- is grouped with the glottalized/implosive consonants because it always pattern with them with respect to tonal splits and mergers. The etyma that go back to these PT voiceless stops all have tones in the first series.

3.6.1.1 *p-

PT voiceless labial stops are preserved as /p-/ without change in all attested modern Tai varieties. Etyma that go back to this PT initial show tones in the first series, which indicate that this PT phoneme was voiceless. PT *p- in the current reconstruction corresponds to Li's *p- (1977: 61-63). Examples of PT etyma with *p- are given in Table 3-17.

Table 3-17 Etyma with PT *p-

Gloss	PT	SI	LC	Y	Others
‘parent’s older sister’	*pa: ^C	pa: ^{C1}	pa: ^{C1}	pa: ^{C1}	
‘mouth’	*pa:k ^D	pa:k ^{DL1}	pa:k ^{DL1}	pa:k ^{DL1}	
‘year’	*pi: ^A	pi: ^{A1}	pi: ^{A1}	pi: ^{A1}	
‘eight’	*pe:t ^D	pe:t ^{DL1}	pe:t ^{DL1}	pet ^{DL1}	
‘to peel’	*po:k ^D	po:k ^{DL1}	po:k ^{DL1}	pok ^{DL1}	

3.6.1.2 *t-

PT voiceless alveolar stops are reflected as /t-/ in all attested modern Tai varieties. Etyma that go back to this PT initial show tones in the first series, which indicate that this PT phoneme was voiceless. It is also reconstructed as *t- in Li (1977: 97-102). Examples of PT etyma with *t- are given in Table 3-18.

Table 3-18 Etyma with PT *t-

Gloss	PT	SI	LC	Y	Others
‘to scoop’	*tak ^D	tak ^{DS1}	tak ^{DS1}	tak ^{DS1}	
‘low’	*tam ^B	tam ^{B1}	tam ^{B1}	tam ^{B1}	
‘door’	*tu: ^A	tu: ^{A1}	tu: ^{A1}	tu: ^{A1}	
‘to drop’	*tok ^D	tok ^{DS1}	tuk ^{DS1}	tok ^{DS1}	
‘to nibble’	*to:t ^D	to:t ^{DL1}	to:t ^{DL1}	tot ^{DL1}	

3.6.1.3 *c-

The reflexes of the PT voiceless palatal stop in modern varieties differ considerably. In most SWT and CT dialects and Saek, it is preserved as a palatal stop

/c-/, which is usually phonetically affricated [tɕ-]. Some examples of dialects that have this conservative reflex are Siamese, Lao, Black Tai, Aiton, Lungchow, Lungming, Western Nung, etc. In many dialects, however, it has become a true affricate /ts-/. Some modern dialects that show this reflex include Tai Ya, Debao, Shangsi, Chongzuo, Qinzhou, Fusui, Wenma, Liujiang, Rong'an, etc. Many Shan dialects as well as many NT varieties have gone one step further and turn the affricate into a fricative, either /s-/ or /ɕ-/. Southern Shan, Tai Mao, Po-ai, Yay, Longsheng, and Hechi are examples of varieties that show /s-/, while Wuming, Tiandong, Lingyue, and Du'an represent dialects that show /ɕ-/. Li (1977: 164-167) reconstructs an affricate *tɕ-³⁶ (see §3.2). Examples of PT etyma with *c- are given in Table 3-19.

Table 3-19 Etyma with PT *c-

Gloss	PT	SI	LC	Y	Others
'breath, heart'	*cau ^A	caj ^{A1}	cau ^{A1}	su ^{A1}	
'to kiss'	*cup ^D	cu:p ^{DL1}	cup ^{DS1}		Rong'an /tsup ^{DS1} /
'to sink'	*cɤm ^A	com ^{A1}		sam	Rong'an /tsam ^{A1} /
'insipid'	*cu:t ^D	cu:t ^{DL1}			Rong'an /tsit ^{DS1} /
'seven'	*cet ^D	cet ^{DS1}	cit ^{DS1}	sat ^{DS1}	Rong'an /tsat ^{DS1} /

3.6.1.4 *k-

PT voiceless velar stops are reflected as /k-/ in all attested modern Tai varieties. Etyma that go back to this PT initial show tones in the first series, which indicate that this PT phoneme was voiceless. It is also reconstructed as *k- in Li (1977: 186-192). Examples of PT etyma with *k- are given in Table 3-20.

³⁶ *č- in Li's notation

Table 3-20 Etyma with PT *k-

Gloss	PT	SI	LC	Y	Others
‘crow’	*ka: ^A	ka: ^{A1}	ka: ^{A1}		Wuming /ka: ^{A1} /
‘to take a bite’	*kat ^D	kat ^{DS1}	kat ^{DS1}	kat ^{DS1}	
‘old (for living being)’	*ke: ^B	kɛ: ^{B1}	ke: ^{B1}	ce ^{B1}	
‘frog’	*kɤp ^D	kop ^{DS1}	kup ^{DS1}	kap ^{DS1}	
‘to hug’	*ko:t ^D	kɔ:t ^{DL2}	ko:t ^{DL2}	kot ^{DL2}	

For ‘crow’, Yay and most other NT dialects has /ʔa^{A1}/. This form is probably not related to PT *ka:^A, but most likely a case of onomatopoeia. Yay regularly palatalized velar consonants preceding front vowels³⁷, thus showing /ce^{B1}/ for ‘old (for people)’.

3.6.1.5 *q-

The PT voiceless uvular stop is reflected either as /k^h-/, /x-/, /h-/, or /k-/ in different varieties. In CT, it is generally reflected as /k^h-/ as in Lungchow, Lungming, Debao, Chongzuo, Daxin, etc. In a few CT dialects including Shangsi, Yongnan, Fusui, Qinzhou, and Long’an, as well as all NT dialects, i.e. Yay, Wuming, Bouyei, Longsheng, etc., however, this PT onset merged with *k- to give modern /k-/. Note that *q- merged into *k- in Yay and became palatalized in front of front vowels. PT *q- remains *q- in PSWT but later became affricated to *ɣ- (Pittayaporn 2008b; to appear-b). This *ɣ- is now reflected as /k^h-/ in Siamese, Lao, and Black Tai, and /x-/ in some Shan dialects such as Dehong. Only in a handful of SWT dialects is *q- kept

³⁷ PT uvular consonants merged with the velar consonants in time to participate in this change (see §3.6.1.5).

distinct from *k^h-, which is the PSWT reflex of PT *Ç.k- and *Ç.q-. For example, in White Tai, PSWT *q- is reflected as /x-/ and PSWT *k^h- as /k^h-. In contrast, the Kapong dialect of Phu Thai shows /h-/ for *q- but /k^h-/ for PSWT *k^h- (< PT *Ç.k- and *Ç.q-). Refer to §3.3 and §3.4 for discussion. Table 3-21 presents examples of etyma reconstructed with PT *q-. Etyma with PT *q- are reconstructed with *k^h- or *x- in Li (1977: 192-197, 208-214).

Table 3-21 Etyma with PT *q-

Gloss	PT	SI	LC	Y	Others
‘to kill’	*qa: ^C	k ^h a: ^{C1}	k ^h a: ^{C1}	ka: ^{C1}	Kapong /ha:/
‘torn’	*qa:t ^D	k ^h a:t ^{DL1}	k ^h a:t ^{DL1}	ka:t ^{DL1}	Kapong /ha:t ^{DL1} /
‘horn’	*qaw ^A	k ^h aw ^{A1}		kaw ^{A1}	Kapong /haw ^{A1} /
‘needle’	*qem ^A	k ^h em ^{A1}	k ^h im ^{A1}	cim ^{A1}	
‘arm’	*qe:n ^A	k ^h ε:n ^{A1}	k ^h e:n ^{A1}	cen ^{A1}	Kapong /hε:n ^{A1} /

3.6.2 *Implosive/glottalized*

Two implosive stops are posited for PT—*ɓ-, and *ɗ-. These phonemes are interesting in terms of tonal development. They usually pattern with voiceless onsets in tonal splits but there are plenty of cases where they pattern with voiced onsets, especially for tone *A (see §6.2 and §6.4). In addition to these two implosives, two glottalized consonants *ʔ- and *ʔj- are reconstructed. These two consonants always pattern with the implosives with respect to tonal splits and mergers. Phonologically, the glottalized palatal glide *ʔj- can be considered a palatal counterpart of the implosives.

3.6.2.1 *ɓ-

The PT labial implosive *ɓ- remains an implosive or preglottalized stop in many dialects, such as Debao and Wuming. In most dialects, it is simplified to plain voiced /b-/. These dialects include Siamese, Lao, Black Tai, White Tai, Lungchow, Leiping, Saek, Yay, Qiubei, Huanjiang, Bouyei, etc. Some dialects have turned this sound into a sonorant. For example, Southern Shan, Lungming, Yongnan, Fusui, Yongbei, and Long'an all reflect PT *ɓ- as /m-/. Similarly, Longsheng's reflex of PT *ɓ- is /w-/. Some dialects such as Yuanyang and Menglian have gone one step further and turned this resulting /w-/ into /v-/. That the implosive *ɓ- became sonorants in some languages is consistent with Clements and Osu (2002)'s view that implosives are neither obstruents nor sonorants. Li (1977: 68-71) reconstructs *ʔb- for this onset. Table 3-22 gives examples of PT etyma with *ɓ-.

Table 3-22 Etyma with PT *ɓ-

Gloss	PT	SI	LC	Y	Others
'village'	*ɓa:n ^C	ba:n ^{C1}	ba:n ^{C1}	ba:n ^{C1}	Wuming /ʔba:n ^{C1} /
'unmarried man'	*ɓa:w ^B	ba:w ^{B1}	ba:w ^{B1}	ba:w ^{B1}	Wuming /ʔba:w ^{B1} /
'leaf'	*ɓau ^A	baj ^{A1}	bau ^{A1}	bau ^{A1}	Wuming /ʔbau ^{A1} /
'to fly'	*ɓin ^A	bin ^{A1}	bin ^{A1}	bin ^{A1}	Wuming /ʔbin ^{A1} /
'disgusted'	*ɓuə ^{B1}	buə ^{B1}	bu: ^{B1}	buə ^{B1}	Wuming /ʔbu ^{B1} /

Note a semantic change in Siamese from 'village' to 'house'. In addition, the Siamese form for 'unmarried man' now means 'servant'. It is also found in the compound /ca:w^{C1} ba:w^{B1}/ 'groom'.

3.6.2.2 *d-

The PT alveolar implosive *d- remains an implosive or preglottalized stop in many dialects, such as Debao and Wuming. In most dialects, it is simplified to plain voiced /d-/. These dialects include Siamese, Lao, Black Tai, White Tai, Lungchow, Leiping, Saek, Yay, Qiubei, Huanjiang, etc. Parallel to *b-, some dialects have turned this sound into a sonorant. For example, Southern Shan, Tai Nüa, Liujiang, Yishan, Rong'an, and Bouyei all reflect PT *d- as /l-/. In Qinzhou, Yongnan, and Long'an, the PT *d- is reflected as /n-/. In addition, Longsheng is quite unique in having /r/ as reflex of PT *d-. Like its labial counterpart, the fact that *d- is reflected as a sonorant in some dialects but an obstruent in others is consistent with the view that implosives are non-obstruent and non-sonorant (Clements & Osu 2002). Li (1977: 107-111) reconstructs *ʔd- for this onset. Table 3-23 gives examples of PT etyma with *d-.

Table 3-23 Etyma with PT *d-

Gloss	PT	SI	LC	Y	Others
‘to scold’	*da: ^B	da: ^{B1}	da: ^{B1}	da: ^{B1}	Wuming / ^ʔ da: ^{B1} /
‘nose’	*daŋ ^A	daŋ ^{C1} -t	daŋ ^{A1}	daŋ ^{A1}	Wuming / ^ʔ daŋ ^{A1} /
‘good’	*dɤj ^A	di: ^{A1} -v	daj ^{A1}	di: ^{A1}	Wuming / ^ʔ doj ^{A1} /
‘forest’	*doŋ ^A	doŋ ^{A1}	doŋ ^{A1}	doŋ ^{A1}	
‘late at night’	*dɯk ^D	duk ^{DS1}		dak ^{DS1}	Debao /dɤk ^{DS1} /

The meaning of ‘nose’ has been narrowed in Siamese to ‘nose bridge’, which also shows an unexpected tonal reflex, i.e. C1 rather than A1.

3.6.2.3 *ʔj-

This glottalized glide *ʔj- is the palatal counterpart of the implosives *ɖ- and *ɓ-. It is usually simplified to a simple /j-/, as in Siamese, Lao, Black Tai, White Tai, Lungming, Lungchow, Shangsi, Long'an, Jingxi, Wenma, etc. In Qiubei, the /j-/ has been hardened to /dz-/. Moreover, younger speakers of Sapa and Bao Yen, both spoken in Vietnam, now have a fricative pronunciation /z-/. In some dialects, it is still preserved as /ʔj-/. These conservative dialects include Wuming, Daxin, Pingguo Tiandong, Donglan, etc. Li (1977: 181-185) also reconstructs *ʔj- for this set of data. Examples of PT *ʔj- are given in Table 3-24.

Table 3-24 Etyma with PT *ʔj-

Gloss	PT	SI	LC	Y	Others
‘medicine’	*ʔjua: ^A	ja: ^{A1}	ja: ^{A1}	jiə ^{A1}	Wuming /ʔju ^{A1} /
‘hungry’	*ʔjuək ^D	ja:k ^{DL1}	ja: ^{DL1}	jiək ^{DL1}	Wuming /ʔjuək ^{DL1} /
‘to be at’	*ʔju: ^B	ju: ^{B1}	ju: ^{B1}	ju: ^{B1}	Wuming /ʔjaw ^{B1} /
‘to stretch’	*ʔjiət ^D	jiət ^{DL1}	ji:t ^{DL1}		Wuming /ʔjiət ^{DL1} /
‘to roast’	*ʔjua:ŋ ^C	ja:ŋ ^{C1}		jiəŋ ^{C1}	Wuming /ʔjuəŋ ^{C1} /

In addition to *ʔj-, Li (1977: 178-181) also reconstructs *j- for the following set of data. He includes more etyma than those listed in Table 3-25 but most of them show irregularities, only occur in SWT, or should be reconstructed with other onsets.

Table 3-25 Some etyma that Li reconstructs with *j-

Gloss	SI	LC	Y	Others
‘paternal grandmother’	ja: ^{B2}		ja: ^{B2}	
‘to mend’	ja: ^{A1}	ja: ^{A2}		
‘difficult’	ja:k ^{DL2}			Lingyue /ʔja:k ^{DL1} /
‘web’	jaj ^{A2}			
‘slack’	ja:n ^{A2}			Saek /ja:l ^{A2} /

The above data show clearly that the reconstruction of PT *j- is dubious as most etyma are found only in SWT. Etyma that are found outside of SWT either show irregularities in the correspondence, or can be identified as loans. For example, the etymon ‘difficult’ is also found outside SWT, but the NT forms point to *j- rather than *j-. As for ‘paternal grandmother’, it is found in both SWT and NT and clearly points to *j-. However, it most likely have a Mon-Khmer origin, cf. Proto-Mon-Khmer *yaʔ ‘grandmother’ (Shorto 2006). Therefore, Li’s *j- cannot be reconstructed for PT.

3.6.2.4 *ʔ-

The PT glottal stop is retained intact in all attested modern Tai varieties. Etyma that go back to this PT initial show tones in the first series, which indicates that this PT phoneme was voiceless. It is also reconstructed as *ʔ- in Li (1977: 243-250). Examples of PT etyma with *ʔ- are given in Table 3-26.

Table 3-26 Etyma with PT *ʔ-

Gloss	PT	SI	LC	Y	Others
‘to bathe’	*ʔa:p ^D	ʔa:p ^{DL1}	ʔa:p ^{DL1}	ʔa:p ^{DL1}	
‘to open the mouth’	*ʔa: ^C	ʔa: ^{C1}	ʔa: ^{C1}	ʔa: ^{C1}	
‘satiated’	*ʔi:m ^B	ʔim ^{B1}	ʔim ^{B1}	ʔim ^{B1}	
‘chest’	*ʔɤk ^D	ʔok ^{DS1}	ʔɤk ^{DS1}	ʔak ^{DS1}	
‘to exit’	*ʔo:k ^D	ʔo:k ^{DL1}	ʔo:k ^{DL1}	ʔok ^{DL1}	

3.6.3 Voiced stops

The PT voiced stops *b-, *d-, *ʃ-, *g-, and *ɣ- have been devoiced in most modern Tai varieties, except for a few dialects on the Sino-Vietnamese border. Dialects that have not gone through devoicing include Cao Bang, Wenma, and various Budai dialects (Haudricourt 1960; Kosaka 1997; L'Écôle Française d'Extême-Orient 1938; L-Thongkum 1997; Ross 1996; Zhang et al. 1999). For example, Cao Bang still retains voicing but has developed breathiness in all PT voiced stops. Therefore, PT *b- is realized as [b̥-] while PT *b- is now realized as [b-] (Pittayaporn 2007d). Other dialects reflect PT voiced stops either as plain /p-/ , /t-/ , /c-/ , and /k-/ , or as aspirated /p^h-/ , /t^h-/ , /c^h-/ , and /k^h-/ . Etyma that go back to PT voiced stops have tones in the second series.

3.6.3.1 *b-

This PT labial consonant has been devoiced in most Tai varieties. Most dialects including Shan, Black Tai, White Tai, Sapa, Lungming, Lungchow, Yay, Wuming, etc. have an unaspirated reflex /p-/ , but many others such as Siamese, Lao, Southern Thai, Phuan, Nyo, Phu Thai, Leiping, and Saek show an aspirated stop /p^h-/ .

Cao Bang and Wenma still keep this sound as voiced /ḃ-/ and /b-/, respectively. Examples of etyma with PT *b- are provided in Table 3-27. Li (1977: 65-67) also reconstructs *b- for this set of etyma.

Table 3-27 Etyma with PT *b-

Gloss	PT	SI	LC	Y	Others
‘expensive’	*be:ŋ ^A	p ^h ε:ŋ ^{A2}	pe:ŋ ^{A2}	peŋ ^{A2}	Cao Bang /ḃeŋ ^{A2} /
‘blistered’	*bo:ŋ ^A	p ^h o:ŋ ^{A2}		poŋ ^{A2}	
‘to fold’	*bap ^D	p ^h ap ^{DS2}		pap ^{DS2}	Cao Bang /ḃap ^{DS2} /
‘father, man’	*bo: ^C	p ^h o: ^{B2} -t	po: ^{C2}	po ^{C2}	Wenma /bɔw ^{C2} /
‘to lay across’	*ba:t ^D	p ^h a:t ^{DL2}		pa:t ^{DL2}	Cao Bang /ḃa:t ^{DL2} /

The etymon ‘father’ has *B in SWT but *C elsewhere. The original tone was probably *C but SWT changed to *B due to contamination by *me:^B ‘mother’

3.6.3.2 *d-

PT *d- has become devoiced in most Tai varieties. Most dialects including Southern Shan, Black Tai, White Tai, Sapa, Lungming, Lungchow, Yay, Wuming, etc. have an unaspirated reflex /t-/, but many others such as Siamese, Lao, Southern Thai, Phuan, Nyo, Phu Thai, Leiping, and Saek show an aspirated stop /t^h-/. Cao Bang, and Wenma still keep this sound as voiced /ḃ-/ and /d-/, respectively. Examples of etyma with PT *d- are provided in Table 3-28. Li (1977: 104-107) also reconstructs *d- for this set of data.

Table 3-28 Etyma with PT *d-

Gloss	PT	SI	LC	Y	Others
‘river’	*da. ^B	t ^h a. ^{B2}	ta. ^{B2}	ta. ^{B2}	Wenma /do ^{B2} /
‘ashes’ ³⁸	*daw ^B	t ^h aw ^{B2}		taw ^{B2}	Wenma /dɤw ^{B2} /
‘open field’	*doŋ ^B	t ^h uŋ ^{B2}		toŋ ^{B2}	Cao Bang /dɔŋ ^{B2} /
‘to pound’	*dup ^D	t ^h up ^{DS2}		tup ^{DS2}	Cao Bang /dɯp ^{DS2} /
‘stomach, belly’	*dwu:ŋ ^C	t ^h ɔ:ŋ ^{C2}	to:ŋ ^{C2}	tuŋ ^{C2}	Wenma /duŋ ^{C2} /

Note a semantic shift from ‘river’ to ‘pier’ in Siamese.

3.6.3.3 *j-

PT *j- has become devoiced in most Tai varieties. A very common reflex of this PT sound is /c-/, which merged with the original *c-, as in Tai Yuan, Lue, White Tai, Lungming, Western Nung, Lungchow, Guangnan, and Yishan. In a few dialects, after merging with the original PT *c-, the secondary *c- from PT *j- became affricated to /ts-/ as in Long’an, Debao, and Jingxi. In a number of dialects the affricate became a fricative /ç-/ as in Wuming, Tianlin, Lingyue, Donglan, Du’an, etc. Many CT and NT dialects including Shangsi, Chongzuo, and Daxin reflect this palatal stop *j- as /s-/ possibly through an intermediate stage *z-. Cao Bang, which still retains original voicing, has /z-/, attests this intermediate step. In contrast, Wenma has a voiced affricate /dz-/ as the reflex of this PT voiced stop. In dialects that reflect voiced stops as aspirated, the voiced *j- was also devoiced to *c^h-. In Siamese, Southern Thai, and Leiping, it is reflected simply as /c^h-. In contrast, Lao, Phu Thai,

³⁸ Li (1977: 95) cites Lunchow /pjaw^{B2}/ as the only example of *vl- but this form is in fact a different etymon— *braw^B ‘leaf ashes’. Debao has both /taw^{B2}/ from *daw^B and /pjaw^{B2}/ from *braw^B.

Phuan, Nyo etc. have changed the intermediate *c^h- to /s-/. This PT onset corresponds to Li's affricate *dz⁻³⁹ (Li 1977: 168-173). Table 3-29 provides examples of PT *j-.

Table 3-29 Etyma with PT *j-

Gloss	PT	SI	LC	Y	Others
'to weigh'	*jaŋ ^B	c ^h aŋ ^{B2}	caŋ ^{B2}	saŋ ^{B2}	Cao Bang /zaŋ ^{B2} /
'morning'	*ja:w ^C	c ^h a:w ^{C2}	caw ^{C2}	saw ^{C2}	Cao Bang /zaw ^{C2} /
'to taste'	*ji:m ^A	c ^h im ^{A2}	cim ^{A2}	sim ^{A2}	Cao Bang /zim ^{A2} /
'name'	*jɣ: ^B	c ^h u: ^{B2}		so ^{B2}	Lungming /cɣu ^{B2} /
'rope'	*jɣ:k ^D	c ^h uək ^{DL2}	cɣ:k ^{DL2}	sa:k ^{DL2}	Cao Bang /zuək ^{DL2} /

3.6.3.4 *g-

This PT consonant follows the same pathway as other voiced stops. It became devoiced in most dialects, except for Cao Bang and Budai dialects. In Cao Bang, it is reflected as a breathy /g̃-/. In most other varieties including Southern Shan, Dehong, Black Tai, White Tai, Lungchow, Debao, Yay, Qinzhou, Shangsi, Wuming, Longsheng, etc, its reflex is a voiceless unaspirated stop /k-/. In dialects like Siamese, Lao, Phuan, Phu Thai, and Leiping, the reflex of this voiced velar stop is aspirated /k^h-/. This PT onset corresponds to Li's *g- (1977: 198-203). Table 3-30 presents examples of etyma with PT *g-.

³⁹ *j̃- in Li's notation.

Table 3-30 Etyma with PT *g-

Gloss	PT	SI	LC	Y	Others
‘stuck’	*ga. ^A	k ^h a. ^{A2}	ka. ^{A2}	ka. ^{A2}	Wenma /go ^{A2} /
‘handle, rod’	*gal ^A	k ^h an ^{A2}		kan ^{A2}	
‘narrow, tight’	*gap ^D	k ^h ap ^{DS2}	kap ^{DS2}		Cao Bang /g̃ap ^{DS2} /
‘pair’	*gu. ^B	k ^h u. ^{B2}	ku. ^{B2}	ku ^{B2}	Wenma /gu ^{B2} /
‘pliers, thongs’	*gi:m ^A	k ^h i:m ^{A2}	kim ^{A2}	cim ^{A2}	

Note that the initial /c-/ in the Yay form for ‘tongs’ is a result of regular palatalization before front vowels.

3.6.3.5 *G-

The voiced uvular stop is reflected in a number of ways in modern languages. In SWT and most CT dialects, it merged with the velar *g-, which gives /k-/ or /k^h-/ depending on whether voiced stops are reflected as plain or aspirated. Black Tai, White Tai, Lungming, Lungchow, Debao, etc. all have /k-/. On the other hand, Siamese, Lao, Nyo, Phuan, Phu Thai, etc. show /k^h-/. In NT as well as a few CT dialects, PT *G- was spirantized and merged with PT *ɣ-. The modern reflex of the merger is /k^h-/ in Shangsi and Fusui; /x-/ in Wuming; /h-/ in Yay, Hechi, Tiandong, Longsheng, etc.; /ɣ-/ in Saek and Qiubei; and /g-/ in Huanjiang. Etyma with *G- are included under *ɣ- in Li (1977: 214-219). See §3.3 for discussion of the reconstruction of uvular consonant. Table 3-31 presents examples of etyma that go back to PT *G-.

Table 3-31 Etyma with PT *G-

Gloss	PT	SI	LC	Y	Others
‘carrying pole’	*ga.n ^A	k ^h a:n ^{A2}	ka:n ^{A2}	ha:n ^{A2}	Saek /ya:n ^{A2} /
‘chin, jaw’	*ga.ŋ ^A	k ^h a:ŋ ^{A2}	ka:ŋ ^{A2}	ha:ŋ ^{A2}	Saek /ya:ŋ ^{A2} /

It is interesting that only a few etyma with PT *G- as onset can be reconstructed.

3.6.4 Voiceless fricatives

There were altogether five voiceless fricatives that can function as onsets in PT: *s-, *ɕ-, *x-, *χ- and *h-. Except for *h-, the reflexes of these fricatives in modern dialects are usually different from their original values. The etyma that go back to these PT voiceless fricatives all have tones in the first series.

3.6.4.1 *s-

The PT voiceless alveolar fricative *s- shows a wide range of reflexes in modern varieties. It is preserved as /s-/ in most SWT dialects including Siamese, Lao, Lue, Black Tai, White Tai, etc.; in CT varieties such as Lungming, Western Nung, Jingxi, Yishan, etc.; and in NT dialects including Longsheng, Laibin, and most Bouyei dialects. It is fronted to /θ-/ in many NT varieties including Wuming, Hengxian, Liujiang, Lianshan, etc., and in the CT dialects of Debao and Jingxi. Another common reflex of *s- is the lateral /l-/ , which is found generally in CT dialects, including Qinzhou, Shangsi, Ningming, Lungchow, etc., as well as a few NT dialect such as Yongbei and Tiandong. In Bao Yen, the lateral /l-/ has become an aspirated stop /t^h-. A special development is found in Southern Shan, where PT *s- is now reflected as

aspirated /s^h-/ in contrast with /s-/ which comes from PT *c-. The distinction between PT *s- and *c- is thus still preserved. Another special development is the case of Hechi and Huanjiang, where PT *s- is regularly reflected as /r-/. Li (1977:152-161) also reconstructs this phoneme as *s-. Examples of PT *s- is given in Table 3-32.

Table 3-32 Etyma with PT *s-

Gloss	PT	SI	LC	Y	Others
‘unmarried woman’	*sa:w ^A	sa:w ^{A1}	la:w ^{A1}	θa:w ^{A1}	Bao Yen /t ^h a:w ^{A1} /
‘intestine’	*saj ^C	saj ^{C1}	la:j ^{C1}	θaj ^{C1}	Bao Yen /t ^h ɤj ^{C1} /
‘heel’	*sɤn ^C	son ^{C1}	lɤn ^{C1}	θan ^{B1}	Bao Yen /t ^h ɤn ^{C1} /
‘four’	*si: ^B	si: ^{B1}	li: ^{B1}	θi: ^{B1}	Bao Yen /t ^h i: ^{B1} /
‘to order’	*saŋ ^B	saŋ ^{B1}	laŋ ^{B1}	θaŋ ^{B1}	Bao Yen /t ^h aŋ ^{B1} /

3.6.4.2 *ɕ-

This voiceless palatal fricative is tentatively reconstructed to account for a discrepancy between SWT/CT on one hand and NT on the other. The only example is ‘owner’ which is reflected as /ca:w^{C1}/ in Siamese and /caw^{C1}/ in Lungchow, but /θu^{C1}/ in Yay and /su:^{C1}/ in Saek. However, there are parallel cases in the voiced series (see §3.6.5.2). In this case, SWT/CT dialects point to earlier *c- (see §3.6.1.3) while their NT counterparts point to *s- (see §3.6.4.1). In this account, *ɕ- strengthened to *c- in SWT/CT but merged with *s- in NT. Li (Li 1977: 165, 167) includes this etymon under *tɕ-⁴⁰.

⁴⁰ *ɕ- in Li’s notation

3.6.4.3 *x-

The PT voiceless velar fricative is rarely preserved unchanged in modern dialects. It merged with PT *h- in all NT dialects including Yay, Wuming, Po-ai, Longsheng, Rong'an, etc. as well as some CT varieties such as Fusui, Shangsi, and Long'an. It is now reflected a /h-/ in these dialects. In most SWT and CT varieties, in contrast, it merged with secondary *k^h- which came either from voicing alternation items (see §3.4), aspiration of *kr- (see §3.5.2), or borrowings (see §3.5.1). For instance, /k^h-/ in Siamese came from PT *x- (> PSWT *x-), PT *q- (> PSWT *q-), PT *C.k- (> PSWT *k^h-), *C.q- (> PSWT *k^h-), PT kr- (> PSWT *k^hr-), or PSWT borrowing (e.g. PSWT *k^haj^A ‘open’ from 開 *kāi*). Dehong and many other dialects in the area have /x-/ for *x-, but this /x-/ is an equivalent of Siamese /k^h-/, an end result of multiple mergers. Only in White Tai and some Lue dialects (Li 1964; 1977: 192, 207) is the velar fricative still retained as a distinct phoneme. Li (1977: 207-214) listed etyma with this PT fricative under *x-. Table 3-33 gives examples of etyma with PT *x-.

Table 3-33 Etyma with PT *x-

Gloss	PT	SI	LC	Y	Others
‘green’	*xiəw ^A	k ^h iəw ^{A1}	k ^h e:w ^{A1}	hew ^{A1}	White Tai /xɛw ^{B1} /
‘galangal’	*xa. ^B	k ^h a. ^{B1}		ha. ^{B1}	White Tai /xa. ^{B1} /
‘white’	*xa:w ^A	k ^h a:w ^{A1}	k ^h a:w ^{A1}	ha:w ^{A1}	White Tai /xa:w ^{A1} /
‘hook’	*xo. ^A	k ^h ɔ. ^{A1}			White Tai /xɔ ^{A1} /

Note the additional examples for ‘hook’: Debao /k^ho:^{A1}/, Lungming /k^ho:^{A1}/, Saek /hɔ:^{A1}/, and Lingyue /hɔ^{A1}/. They all point to PT *x-.

3.6.4.4 *χ-

The voiceless uvular fricative *χ- merged with *x- in most dialects. Therefore, reflexes of *χ- is generally identical to those of *x- (see §3.3 and §3.6.4.3). Exceptions are Phuan and Phu Thai dialects (2008b; Pittayaporn to appear-b), as well as some SWT dialects of central Vietnam (1999; Ferlus 2008)⁴¹. In these dialects, PT *χ- merged with *h- (see §3.3) rather than *x-. Li (1977: 207-214) includes this set of etyma under *x-. Examples of PT *χ- are provided in Table 3-34.

Table 3-34 Etyma with PT *χ-

Gloss	PT	SI	LC	Y	Others
‘to cross’	*χa:m ^C	k ^h a:m ^{C1}	k ^h a:m ^{C1}	ha:m ^{C1}	Kapong /ha:m ^{C1} /
‘to crow’	*χal ^A	k ^h an ^{A1}	k ^h an ^{A1}	han ^{A1}	Kapong /han ^{A1} /
‘fang’	*χe:w ^C	k ^h iəw ^{C1}	k ^h e:w ^{C1}	hew ^{C1}	Kapong /hε:w ^{C1} /
‘to go up’	*χu:n ^C	k ^h u:n ^{C1}	k ^h ɯn ^{C1}	hu:n ^{C1}	Kapong /hu:n ^{C1} /
‘ginger’	*χi:ŋ ^A	k ^h iŋ ^{A1}	k ^h iŋ ^{A1}	hiŋ ^{A1}	Kapong /hi:ŋ ^{A1} /

Data from the Kapong dialect of Phu Thai is given above. In this variety, *x- is reflected as aspirated stop /k^h-/ but *χ- is reflected as /h-/. Contrast /hε:w^{C1}/ ‘tooth’ from PT *χe:w^C ‘fang’ and /k^hiəw^{A1}/ ‘green’ from PT *xiəw^A.

3.6.4.5 *h-

The voiceless glottal fricative is retained in almost all modern dialects. There are a few varieties where it is reflected as /x-/, /ç-/, or /h-/. It is reflected as /x-/ in

⁴¹ Data from the SWT varieties in central Vietnam were not systematically examined in this study. Therefore, it does not appear on the map in Figure 1-3.

Wenma and Yishan, as /ɣ-/ in Qiubei and some Bouyei dialects (points 14, 18, 22, 23, 24, etc. in *Bùyīyǔ Diàochá Bàogào*), and as /fi-/ in some other Bouyei dialects (point 7, 15, 19, etc.). It is also reconstructed as *h- by Li (1977: 249-253). Etyma in Table 3-35 are examples of PT *h-.

Table 3-35 Etyma with PT *h-

Gloss	PT	SI	LC	Y	Others
‘five’	*ha: ^C	ha: ^{C1}	ha: ^{C1}	ha: ^{C1}	Bouyei(7) /fa ^{C1} /
‘goose’	*ha:n ^B	ha:n ^{B1}	ha:n ^{B1}	ha:n ^{B1}	Bouyei(7) /fan ^{C1} /
‘to give’	*haj ^C	haj ^{C1}	hu: ^{C1}	hau ^{C1}	Bouyei(7) /fau ^{C1} /
‘fragrant’	*ho:m ^A	hɔ:m ^{A1}	ho:m ^{A1}	hom ^{A1}	
‘to cover with cloth’	*hɤm ^B	hom ^{B1}	hum ^{B1}		Saek /ham ^{B1} /

3.6.5 Voiced fricatives

There were perhaps only three voiced fricatives that can function as onsets in PT: *z-, *ʒ-, and *ɣ-. There is no evidence for voiced fricatives at the uvular and glottal places of articulation, the two places farthest back in the vocal tract. When these voiced fricatives were devoiced, they generally merged with their voiceless counterparts *s-, *ç-, and *x-, respectively. The etyma that go back to these PT voiced fricatives all have tones in the second series.

3.6.5.1 *z-

The PT voiced alveolar fricative *z- shows reflexes identical with those of PT *s- in most dialects. This is because *z- became *s- after devoicing of original voiced

consonants. Cao Bang and Saek are two special cases. In Cao Bang, *z- became rhoticized to /r-/, while the original *j- became /z-/ to fill in the gap. Recall that Cao Bang retains the original voicing of PT obstruents. In Saek, *z- became /j-/ merging with PT *ʝ-, probably due to influence from North-Central Vietnamese, where original *z- has become /j-/ (Alves 2007)⁴². This PT fricative corresponds to Li's *z- (1977: 161-164). Table 3-36 gives examples of PT *z-.

Table 3-36 Etyma with PT *z-

Gloss	PT	SI	LC	Y	Others
‘to wash (clothes)’	*zak ^D	sak ^{DS2}	ɬak ^{DS2}	θak ^{DS2}	Saek /jak ^{DS2} /
‘straight’	*zɔː ^B	suuː ^{B2}	ɬuuː ^{B2}	θo ^{B2}	Saek /joː ^{B2} /
‘fish trap’	*zaj ^A	saj ^{A2}	ɬaj ^{A2}	θaj ^{A2}	Saek /jaj ^{A2} /
‘to repeat’	*zam ^C	sam ^{C2}			Saek /jam ^{C2} /

3.6.5.2 *z-

The PT voiced palatal fricative is tentatively reconstructed to account for a discrepancy between SWT/CT on one hand and NT on the other. For etyma in Table 3-37, SWT/CT dialects point to earlier *j- (see §3.6.1.3) while their NT counterparts point to *z- (see §3.6.5.1). In this account, *z- strengthened and merged with *j- in SWT/CT but merged with *z- in NT. Li (1977: 165, 167) includes these etyma under *dz-⁴³. Also, refer to the reconstruction of *ɕ- in §3.6.4.2.

⁴² According to Chamberlain (2000), Saek speakers in northeastern Thailand and Central Laos migrated from modern-day Thanh Hoa, Nghe An and Quang Binh Provinces of Vietnam. Also see Saek tales recorded by Gedney (Hudak 1993).

⁴³ *ʝ- in Li's notation

Table 3-37 Etyma with PT *z-

Gloss	PT	SI	LC	Y	Others
‘man, male’	*za:j ^A	c ^h a:j ^{A2}	ca:j ^{A2}	θa:j ^{A2}	
‘eaves’	*za:j ^A	c ^h a:j ^{A2}	ca:j ^{A2}	θa:j ^{A2}	Saek ^{C2} /ja:j ^{A2} /
‘to commission’	*zau _ɿ ^C	c ^h aj ^{C2}	cau _ɿ ^{C2}	θau _ɿ ^{C2}	

For ‘man’ and ‘to commission’, Saek has /sa:j^{A2}/ and /saj^{C2}/ respectively. The onset /s-/ is unexpected because we would expect Saek to show /j-/ (from *z- < PT *z). However, both cases can be viewed as influence from Lao or Siamese. The Lao forms for these two etyma are /sa:j^{A2}/ and /saj^{C2}/ respectively. The loan nature of the Saek form for the latter is especially obvious as the expected Saek vowel for this etymon is /ɿ:/, not /aj/.

3.6.5.3 *ɣ-

The PT voiced velar fricative generally merged with its voiceless counterpart after the devoicing of original voiced onsets (see §3.6.4.3). Only a few languages show special developments. Saek is conservative in that it retains PT *ɣ- intact. Black Tai and Lungchow reflect *x- as /k^h-/ but regularly hardened *ɣ- to *g- before devoicing the merged phoneme to /k-/. Wuming merged PT *x- with *h- but preserved PT *ɣ- until it became /x-/ after the devoicing. Many CT dialects such as Leiping, Lungming, and Western Nung now have /h-/ for PT *ɣ- but /k^h-/ for *x-. Debao is unique in having sonorant reflexes for PT *ɣ-. Now this PT fricative is reflected as /v-/ (from an intermediate *w-) before front vowels but /j-/ elsewhere. This PT fricative corresponds to Li’s *ɣ- (Li 1977: 214-219), but refer to the discussion of PT dorsal consonants in 3.3. Table 3-38 gives examples of PT *ɣ-.

Table 3-38 Etyma with PT *y-

Gloss	PT	SI	LC	Y	Others
‘night time’	*yam ^B	k ^h am ^{B2}	kam ^{B2}	ham ^{B2}	Saek /yam ^{B2} /
‘night’	*yɯ:n ^A	k ^h ɯ:n ^{A2}	kɯn ^{A2}	huɯn ^{A2}	Saek /yɯn ^{A2} /
‘neck’	*yo: ^A	k ^h ɔ: ^{A2}	ko: ^{A2}	ho ^{A2}	Saek /yɔ: ^{A2} /
‘dike between rice fields’	*yal ^A	k ^h an ^{A2}	kan ^{A2}	han ^{A2}	Saek /yal ^{A2} /
‘thatch grass’	*ya: ^A	k ^h a: ^{A2}	ka: ^{A2}	ha ^{A2}	Saek /ya: ^{A2} /

3.6.6 Voiced nasals

There were altogether four plain voiced nasals that could function as onsets in PT: *m-, *n-, *ɲ-, and *ŋ-. After their voiceless counterparts became voiced, the newly created voiced nasals merged with these original plain nasals. However, etyma that go back to PT *m-, *n-, *ɲ-, and *ŋ- are generally reflected with tones in the second series while the voiceless ones are reflected with tones in the first series.

3.6.6.1 *m-

Without exception, the plain nasal *m- is retained in all modern dialects. Now PT *m- is reflected as /m-/ with tones in the second series in all modern dialects. Li (1977: 71-73) also reconstructs *m- for this set of data. Table 3-39 presents example of PT *m-.

Table 3-39 Etyma with PT *m-

Gloss	PT	SI	LC	Y	Others
‘mother, woman’	*me: ^B	mɛ: ^{B2}	me: ^{B2}	me ^{B2}	
‘ant’	*mɤc ^D	mot ^{DS2}	mɤt ^{DS2}	mat ^{DS2}	
‘yam’	*man ^A	man ^{A2}	man ^{A2}	man ^{A2}	
‘mucus’	*mu:k ^D	mu:k ^{DL2}	muk ^{DS2}	muk ^{DS2}	

3.6.6.2 *n-

The plain alveolar nasal *n- is preserved in most modern dialects. Only in a few Shan varieties did it merge completely with PT *l-. For example, in the Mengshi and Chefang dialects of Tai Nüa, PT *n- is reflected as /l-/. Etyma with PT *n- now show tones in the second series. This alveolar nasal corresponds to Li’s *n- (Li 1977: 111-113). Table 3-40 presents example of PT *n-.

Table 3-40 Etyma with PT *n-

Gloss	PT	SI	LC	Y	Others
‘paddy field’	*na: ^A	na: ^{A2}	na: ^{A2}	na: ^{A2}	Mengshi /la: ^{A2} /
‘otter’	*na:k ^D	na:k ^{DL2}	na:k ^{DL2}	na:k ^{DL2}	
‘rotten’	*naw ^B	naw ^{B2}		naw ^{B2}	Mengshi /law ^{B2} /
‘slow’	*na:n ^A	na:n ^{A2}		na:n ^{A2}	
‘small’	*no:j ^C	no:j ^{C2}	no:j ^{C2}		

The etymon ‘small’ shows a tonal irregularity, probably due to confusion with *^hno:j^B ‘few, little’. Lungming has /no:j^{C2}/ for ‘minor’, and /no:j^{B1}/ for ‘few, little’.

Similarly, in Siamese /nɔːj^{C2}/ occurs with /lek^{DS2}/ ‘small’ in the co-ordinate compound /lek^{DS2} nɔːj^{C2}/ ‘small (size, degree)’ but /nɔːj^{B1}/ occurs with /nit^{DS2}/ ‘litte (amount)’ in the expression /nit^{DS2} nɔːj^{B1}/ ‘small (amount)’.

3.6.6.3 *ɲ-

In modern Tai languages, the plain palatal nasal *ɲ- is either preserved as /ɲ-/ or denasalized to a palatal glide /j-/. While Lao, Black Tai, White Tai, Western Nung, Debao, Yishan, Tianlin, Huanjiang, Guigang, Saek, etc. show the conservative reflex /ɲ-/, Siamese, Lue, Leiping, Lungchow, Qinzhou, Shangsi, Wuming, Hengxian, Tiandong etc. show the glide reflex /j-/. This palatal nasal corresponds to Li’s *ɲ- (Li 1977: 173-176). Table 3-40 presents example of PT *ɲ-.

Table 3-41 Etyma with PT *ɲ-

Gloss	PT	SI	LC	Y	Others
‘to sew’	*ɲɤp ^D	jɛp ^{DS2}	jap ^{DS2}	ɲip ^{DS2}	Lao /ɲip ^{DS2} /
‘to shoot’	*ɲuɲ ^{A2}	jiɲ ^{A2}	jaɲ ^{A2}	ɲiɲ ^{A2}	Lao /ɲiɲ ^{A2} /
‘tip, highest point’	*ɲoːt ^D	joːt ^{DL2}	joːt ^{DL2}		Lao /joːt ^{DL2} /
‘mosquito’	*ɲuɲ ^A	juɲ ^{A2}	juɲ ^{A2}	ɲuɲ ^{A2}	Lao /ɲuɲ ^{A2} /

3.6.6.4 *ŋ-

The plain velar nasal *ŋ- is preserved in almost all modern dialects. The only exceptions are Saek and certain Southern Thai dialects. While PT *ŋ- became /ɲ-/ in Saek, it became /h-/ with tones in the second series in some dialects of Southern Thai. This velar nasal corresponds to Li’s *ŋ- (Li 1977: 205-206). Table 3-42 presents example of PT *ŋ-.

Table 3-42 Etyma with PT *ŋ-

Gloss	PT	SI	LC	Y	Others
‘early meal’	*ŋa:j ^A	ŋa:j ^{A2}	ŋa:j ^{A2}	ŋa:j ^{A2}	
‘easy’	*ŋa:j ^B	ŋa:j ^{B2}		ŋa:j ^{B2}	
‘crocodile’	*ŋwək ^D	ŋwək ^{DL2}	ŋɣ:k ^{DL2}	ŋwək ^{DL2}	
‘elephant tusk’	*ŋa: ^A	ŋa: ^{A2}		ŋa: ^{A2}	
‘shadow’	*ŋaw ^A	ŋaw ^{A2}	ŋaw ^{A2}	ŋaw ^{A2}	

Note a semantic shift from ‘crocodile’ to ‘snake-like mythical water creature’ in a great number of SWT and CT dialects, including Siamese, Lao, Black Tai, Tai Nüa, Lungming, Lungchow, Bao Yen, Sapa etc. Many CT dialects in Guangxi and all NT varieties except Saek still preserve the original meaning. The Chinese source for this etymon 鱷 è (MC *ŋāk*) also means ‘crocodile’ (Pulleyblank 1991: 87; Schuessler 2007: 223).

3.6.7 *Voiceless nasals*

In addition to the plain nasals, PT had four voiceless nasals *^hm-, *^hn-, *^hŋ-, and *^hŋ-. These PT phonemes generally merged with their respective voiced counterparts so that etyma with these PT voiceless nasals have tones in the first series in modern dialects (see §6.2). Chronologically, the voicing of the voiceless nasals must have preceded the devoicing of voiced obstruents because of all the dialects that still preserve original voiced obstruents have voiced their sonorants. L-Thongkum (1997) argues convincingly that it is this voicing of voiceless sonorants that caused the sub-phonemic pitch difference to become a binary register split.

3.6.7.1 *^hm-

The PT voiceless labial nasal *^hm- fell together with its voiced counterpart *m-. Therefore, its reflex in each dialect is identical to that of PT *m- (see §3.6.6.1). This labial sonorant corresponds to Li's *^hm- (Li 1977: 74-76). Table 3-43 presents example of PT *^hm-.

Table 3-43 Etyma with PT *^hm-

Gloss	PT	SI	LC	Y	Others
‘dog’	* ^h ma: ^A	ma: ^{A1}	ma: ^{A1}	ma: ^{A1}	
‘pig’	* ^h mu: ^A	mu: ^{A1}	mu: ^{A1}	mu: ^{A1}	
‘pot’	* ^h mo: ^C	mɔ: ^{C1}	mo: ^{C1}	mo ^{C1}	
‘new’	* ^h mɣ:l ^B	maj ^{B1}	mauɣ ^{B1}	mo ^{B1}	
‘soot’	* ^h mi: ^C	mi: ^{C1}	mi: ^{C1}	mi ^{C1}	

3.6.7.2 *^hn-

The PT voiceless alveolar nasal *^hn- merged with its plain counterpart *n-. Therefore, its reflex in each dialect is identical to that of PT *n- (see §3.6.6.2). This alveolar nasal is also reconstructed by Li (Li 1977: 113-116) as *^hn-. Table 3-44 provides example of PT *^hn-.

Table 3-44 Etyma with PT *^hn-

Gloss	PT	SI	LC	Y	Others
‘heavy’	* ^h nak ^D	nak ^{DS1}	nak ^{DS1}	nak ^{DS1}	Mengshi /lak ^{DS1} /
‘to steam’	* ^h nuŋ ^C	nuŋ ^{C1}		naŋ ^{C1}	Mengshi /luŋ ^{C1} /
‘face’	* ^h na: ^C	na: ^{C1}	na: ^{C1}	na: ^{C1}	Mengshi /la ^{C2} /
‘skin’	* ^h naŋ ^A	naŋ ^{A1}	naŋ ^{A1}	naŋ ^{A1}	Mengshi /laŋ ^{A1} /
‘mouse’	* ^h nu: ^A	nu: ^{A1}	nu: ^{A1}	nu: ^{A1}	Mengshi /lu ^{A1} /

3.6.7.3 *^hŋ-

The PT voiceless palatal nasal *^hŋ- fell together with its voiced counterpart *ŋ-. Therefore, its reflex in each dialect is identical to that of PT *ŋ- (see §3.6.6.1). This palatal sonorant corresponds to Li’s *^hŋ- (Li 1977: 176-178). Table 3-45 presents example of PT *^hŋ-.

Table 3-45 Etyma with PT *^hŋ-

Gloss	PT	SI	LC	Y	Others
‘grass’	* ^h ŋuə ^C	ja: ^{C1}	ja: ^{C1}	ŋiə ^{C1}	Lao /ŋa: ^{C1} /
‘coarse, tough’	* ^h ŋa:p ^D	ja:p ^{DL1}	ja:p ^{DL1}	ŋa:p ^{DL1}	Lao /ŋa:p ^{DL1} /
‘to step on’	*ŋam ^B	jam ^{B2}	jam ^{B2}	ŋam ^{B2}	Lao /ŋam ^{B2} /
‘to stretch out’	* ^h ŋiət ^D	jiət ^{DL1}	ji:t ^{DL1}	ŋiət ^{DL1}	

3.6.7.4 *^hŋ-

The PT voiceless velar *^hŋ- has two main reflexes in modern Tai dialects: /ŋ-/ and /h-/. Many varieties simply voiced the nasal *^hŋ- to /ŋ-/ merging with the original

*ŋ-. These dialects include Siamese, Lao, Lue, Western Nung, Debao, Long’an, Guangnan Nung, Tiandong, Tianlin, Lingyue, etc. In contrast, many dialects merged *hŋ- with *h- before the voicing of voiceless sonorants took place. Dialects in this group include Black Tai, White Tai, Leiping, Lungchow, Shangsi, Ningming, Chongzuo, Saek, etc. This voiceless velar nasal corresponds to Li’s *hŋ- (Li 1977: 206-207). Table 3-46 presents example of PT *hŋ-.

Table 3-46 Etyma with PT *hŋ-

Gloss	PT	SI	LC	Y	Others
‘gum, gill’	*hŋwək	ŋwək ^{DL1}	hɣ:k ^{DL1}	ŋwək ^{DS1} -t	
‘gray (hair)’	*hŋo:k ^D	ŋo:k ^{DL1}			Saek /hu:k ^{DL1} /
‘to lie face up’	*hŋa:j ^A	ŋa:j ^{A1}		ha:j ^{A1}	
‘to tilt head up’	*hŋe:n ^A	ŋe:n ^{A1}			

Because examples of PT *hŋ- are very rare, its reconstruction is still tentative. It is possible that the set of etyma assigned to PT *hŋ- are secondary, as only a few examples exist, some of which are only found in SWT. The regular reflex in Yay is perhaps /h-/; the irregularity in tone in the etymon ‘gill, gum’ indicates that it might be secondary, i.e. dialect borrowing or late loan from a Chinese dialect, cf. 顎 è (< MC *ŋâk*, OC *ŋâk) ‘palate’. Perhaps the true PT *hŋ- merged with PT *h- early on as suggested by the onset *h- in many PT etyma that go back to *ŋ- (perhaps from earlier *hŋ-) in Chinese, e.g. PT *ha:^C from 五 *wǔ* (< MC *ŋuo^B*, OC *ŋâ?) ‘five’, and PT *ha:n^B from 雁 (< MC *ŋan^C*, OC *ŋrâns) ‘wild goose’

3.6.8 *Voiced liquids and glides*

There were two PT voiced liquids and one voiced glide in PT: *r-, *l-, and *w-. Each of them has a voiceless counterpart with which they usually merged. However, these plain voiced liquids and glides are generally reflected with tones in the second series while the voiceless ones are reflected with tones in the first series.

3.6.8.1 **r-*

The PT voiced alveolar trill *r- shows a wide range of reflexes in modern dialects, including /r-/, /l-/, /h-/, /^hl-/, /ɬ-/, /^hj-/, /j-/, /z-/ and /ð-/. Many languages from the three clusters have a conservative reflex /r-/, including Siamese, Southern Tai, Bao Yen, Cao Bang, Jingxi, Yishan, Yay, Saek, Huanjiang, Longsheng, and some dialects of Wuming. Most SWT dialects aspirated the trill into /h-/, including Lao, Phuan, Phu Thai, Tai Yuan, Lue, Aiton, White Tai, Sapa, etc. In many CT and NT dialects, *r- became a lateral /l-/, as in Lungming, Debao, Qinzhou, Chongzuo, Shangsi, Yongbei, Lingyue, and Donglan, to cite a few. A few varieties attest /^hl-/ from PT *r-, which is kept apart from original *l- which is reflected as plain /l-/ (see §3.6.8.2). Dialects in this group include Yongnan and Long'an. Some dialects in the NT group including Liujiang, Yishan, and Shanglin yodized PT *r- to /j-/. A few also have a yodized reflex but with pre-aspiration /^hj-/, distinct from /j-/. Many dialects, mostly in the NT group, e.g. Western Nung, Hengxian, Qiubei, and Guangnan Sha, show an interdental fricative /ð-/ as the reflex of the PT trill. Rong'an went one step further and devoiced its intermediate *ð- to /θ-/. Another group of dialects merged PT *r- with *z-, which is now reflected as /ɬ-/ (see §3.6.5.1). Dialects in this group include the CT dialects of Leiping, Daxin, Lungchow, etc. The intermediate step is attested in Wenma, which still retains /z-/. This PT phoneme corresponds to Li's *r- (Li 1977: 142-148). Table 3-47 presents example of PT *r-.

Table 3-47 Etyma with PT *r-

Gloss	PT	SI	LC	Y	Others
‘house’	*rɣ:n ^A	ruən ^{A2}	ɬɣ:n ^{A2}	ra:n ^{A2}	W. Nung /ðu:n ^{A2} /
‘hundred, to string’	*ro:j ^C	rɔ:j ^{C2}	ɬo:j ^{C2}	roj ^{C2}	W. Nung /ðoj ^{C2} /
‘to know’	*ru:w ^C	ru: ^{C2}	ɬu: ^{C2}	ro ^{C2}	W. Nung /ðu: ^{C2} /
‘long’	*ruj ^{A2}	ri: ^{A2}	ɬi: ^{A2}	raj ^{A2}	W. Nung /ði: ^{A2} /
‘strength’	*re:ŋ ^{A2}	rɛ:ŋ ^{A2}	ɬe:ŋ ^{A2}	reŋ ^{A2}	W. Nung /ðeŋ ^{A2} /

Note a semantic shift from ‘long’ to ‘oval’ in Siamese.

3.6.8.2 *l-

The PT voiced lateral *l- is retained as /l-/ in all attested Tai dialects. Li (1977: 133-136) also reconstructs *l- for this set of data. Table 3-48 presents example of PT *l-.

Table 3-48 Etyma with PT *l-

Gloss	PT	SI	LC	Y	Others
‘child (offspring)’	*lu:k ^D	lu:k ^{DL2}	luk ^{DL2}	luk ^{DS2}	
‘blood’	*luət ^D	luət ^{DL2}	ɬɜ:t ^{DL2}	luət ^{DL2}	
‘monkey’	*li:ŋ ^A	liŋ ^{A2}	liŋ ^{A2}	liŋ ^{A2}	
‘to fall’	*lɜm ^C	lom ^{C2}		lam ^{C2}	Debao /lam ^{C2} /
‘to forget’	*lu:m ^A	lu:m ^{A2}	lum ^{A2}	lum ^{A2}	

3.6.8.3 *w-

The PT labial glide has various reflexes in modern dialects but they can be considered end results of two main lines of development. The first line is an innovative one in which *w- became a fricative *v-. This intermediate step is attested in Cao Bang and Wenma as /v-/. Most dialects underwent further devoicing of this *v- to /f-/ when devoicing occurred. These dialects include Siamese, Black Tai, White Tai, Lao, Lungchow, Lungming, Shangsi, Qinzhou, Debao, Yay, Pinguo, Liujiang, Rong'an, etc. However, in many Shan varieties including Southern Shan, Phake, and Aiton, the /f-/ has become /p^h-/.

Another line of development is attested in Qiubei and Saek. These two varieties did not go through the early change from *w- to *v-. Qiubei is more conservative because the PT labial glide is kept as /w-/. Saek deserves special attention. Although PT *w- is now reflected as /v-/, this voiced fricative is not a retention of the intermediate *v- of the other line of development above. In fact, PT *w- was kept as intact in Saek at least until the devoicing, when it merged with PT *^hw- (see §3.6.9.3). The result of the merger between PT *w- and *^hw- was *w-. It is this merged *w- that became /v-/ in Saek. This change from *w- to /v-/ is a much later change than the change from PT *w- to *v- in discussed above. Table 3-49 presents examples of PT *w-.

Table 3-49 Etyma with PT *w-

Gloss	PT	SI	LC	Y	Others
‘fire’	*wi: ^A	faj: ^{A2}	faj: ^{A2}	fi: ^{A2}	Qiubei /wei: ^{A2} /
‘tooth’	*wan ^A	fan ^{A2}	fan ^{A2}	fan ^{A2}	
‘swollen’	*wok ^D	fok ^{DS2}		fok ^{DS2}	Qiubei /wok ^{DS2} /
‘firewood’	*wu:n ^A	fu:n ^{A2}	fɯn ^{A2}	fun ^{A2}	Qiubei /wən ^{A2} /
‘straw’	*wuəŋ ^A	fa:ŋ ^{A2}	fuəŋ ^{A2}	fuəŋ ^{A2}	Qiubei /yʉŋ ^{A2} /

These etyma also have *w- in Proto-Lakkja (L-Thongkum 1992), e.g. *wan^A ‘tooth’, and *wok^D ‘swollen’. This PT labial glide corresponds to Li’s *v- (Li 1977: 79-81). What he reconstructs as *w- (Li 1977: 81-82) is in fact not reconstructible for PT as there are very few etyma in this set. Crucially, those etyma do not have regular correspondence, or can be shown to be borrowings. Table 3-50 gives examples of etyma that Li assigns to *w-.

Table 3-50 Some etyma that Li reconstructs with *w-

Gloss	Siamese	Lungchow	Yay	Others
‘to say’	wa: ^{B2}	va: ^{B2}		
‘to place’	waj: ^{C2}	vaj: ^{C2}		
‘deep pool’	wəŋ ^{A2}	vəŋ ^{A2}		
‘work’		viək ^{DL2}	viək ^{DL2}	
‘free, idle’	wa:ŋ ^{B2} -t	va:ŋ ^{B1}		Longsheng /va:ŋ ^{B1} /

Among these, only two are found in NT. However, it is exactly those two that can be identified as loans from Chinese, i.e. ‘work’ from 役 *yì* (< MC *jiwäk*) ‘corvée duty’ perhaps through Vietnamese *việc*, and ‘free, idle’ possibly from 亡 *wáng* (< MC *mjwan*) ‘not present’ or 罔 *wǎng* (< MC *mjwan^B*) ‘not have’. The etymon ‘deep pool’ is also clearly a Chinese loan from, i.e. 汪 *wāng* (< MC *ʔwân*) ‘deep pool’. Further, the etymon ‘to say’ may also be a Chinese loan, cf. 話 *huà* (< MC *ɣwai^C*) ‘speech’.

3.6.9 *Voiceless liquids and glides*

There were two voiceless liquids and one voiceless glide in PT: **hr-*, **hl-*, and **hw-*. Each of them has a voiced counterpart with which they usually merged. However, these voiceless liquids and glides are generally reflected with tones in the first series while the voiced ones are reflected with tones in the second series.

3.6.9.1 **hr-*

In most dialects, the PT voiceless trill **hr-* fell together with its voiced counterpart **r-*. Therefore, its reflex is identical to that of **r-* (see §3.6.8.1). However, a number of dialects including Siamese, Lao, Lungchow, and Western Nung, underwent a change from **hr-* to **h-*, which occurred before PT voicedless sonorants became voiced. This PT trill corresponds to Li’s **hr-* (Li 1977: 148-151). Table 3-51 presents example of PT **hr-*.

Table 3-51 Etyma with PT *hr-

Gloss	PT	SI	LC	Y	Others
‘to yawn’	*hra:w ^A	ha:w ^{A1}	ha:w ^{A1}	ra:w ^{A1}	
‘to bark’	*hraw ^B	haw ^{B1}	haw ^{B1}	raw ^{B1}	
‘dried up’	*hre:ŋ ^C	hɛ:ŋ ^{C1}		reŋ ^{C1}	W. Nung /hɛŋ ^{C1} /
‘withered’	*hriəw ^B	hiəw ^{B1}	he:w ^{B1}	rew ^{B1}	
‘to shrink’	*hrot ^D	hot ^{DS1}		rut ^{DS1}	W. Nung /hot ^{DS1} /

3.6.9.2 *hl-

The PT voiceless lateral *hl- fell together with its voiced counterpart *l-. Therefore, its reflex in each dialect is identical to that of *l- (see §3.6.8.2). This PT lateral corresponds to Li’s *hl- (Li 1977: 136-141). Table 3-52 presents example of PT *hl-.

Table 3-52 Etyma with PT *hl-

Gloss	PT	SI	LC	Y	Others
‘many, much’	*hla:j ^A	la:j ^{A1}	la:j ^{A1}	la:j ^{A1}	
‘liquor’	*hlaw ^C	law ^{C1}	law ^{C1}	law ^{C1}	
‘sharp-pointed’	*hle:m ^A	le:m ^{A1}	le:m ^{A1}	lem ^{A1}	
‘to be left over’	*hlwə ^A	lwə ^{A1}	lw: ^{A1}	lwə ^{A1}	
‘to fall off’	*hlɣn ^B	lon ^{B1}		lan ^{B1}	Lungming /lɣn ^{B1} /

3.6.9.3 *hw-

Parallel to PT *w- (see 3.6.8.3), PT *hw- went through three lines of development. The first line of development is characterized by an early change from *hw- to *f-. This intermediate *f- has two main reflexes. The first reflex is /f-/ as found in Siamese, Lao,

Lue, Black Tai, Wuming, Pingguo, Tiandong, etc. The second reflex is /p^h-/. This further change from *f- to /p^h-/ is found in Aiton, Phake, etc. CT dialects, in general, show the second line of development. In Lungchow, Debao, Lungming, Shangsi, Daxin, Jingxi, etc., the reflex of *w- is /f-/ but the reflex of PT *^hw- is /p^h-/. This indicates that *^hw- did not become *w- when the original voiceless sonorants became voiced. Instead, PT *^hw- must have developed into /p^h-/ before the voicing took place.

Another line of development is represented by many NT dialects including Qiubei, Tianlin, Guangnan, Lingyue, Longsheng, Du'an, Shanglin, etc. In these dialects PT *w- is retained as /w-/. Liujiang, Yishan, Laibin and a few other dialects support an original value *^hw- as they preserve the pre-aspiration in the form of /h-/ in etyma with PT nucleus /wə/, e.g. Liujiang /hu^{C1}/ ‘cloud’ from *^hwuə^C < PT *fuə^C. In addition, Saek, Yay, and many Bouyei dialects have gone one step further and turned /w-/ into a fricative /v-/. This onset corresponds to Li’s *f- (Li 1977: 77-80). Examples of etyma with PT *^hw- are given in Table 3-53.

Table 3-53 Etyma with PT *^hw-

Gloss	PT	SI	LC	Y	Others
‘lid’	* ^h wa: ^A	fa: ^{A1}	p ^h a: ^{A1}	va: ^{A1}	Liujiang /wa: ^{A1} /
‘cloud’	* ^h wuə ^C	fa: ^{C1}	p ^h a: ^{C1}		Liujiang /hu ^{C1} /
‘dam’	* ^h wa:j ^C	fa:j ^{A1}	p ^h a:j ^{A1}		Liujiang /wa:j ^{A1} /
‘astringent’	* ^h wuət ^B	fa:t ^{DL1}		vuət	
‘soft-shelled turtle’ ⁴⁴	* ^h wuə ^A	fa: ^{A1}	p ^h a: ^{A1}		Liujiang /hwɿ ^{A1} /

⁴⁴ It is not clear why the rimes in Liujiang forms for ‘cloud’ and ‘soft-shelled turtle’ are different.

Zhang (1995) proposes that this phoneme should be reconstructed as *pw- at the PT level. This is probably not the case because Guigang and nearby dialects, which generally show /w-/ for this PT sound, have /f-/ in front of /uə/ as in ‘cloud’ and ‘soft-shelled turtle’. If the original PT sound was *pw-, we would expect to see /p-/ in that environment. Rather, the /f-/ in Guigang forms for ‘cloud’ and ‘soft-shelled’ should be considered a special development of *^hw- before /u/. The Proto-Lakkja (L-Thongkum 1992) correspondent of ‘cloud’ lends further support to this interpretation, cf. Proto-Lakkja *^hwa:^C ‘cloud’. It is however very likely that one source of PT *^hw- is *pw- at a level deeper than PT. Also, note an intriguing distributional fact about this PT voiceless glide. No instances of PT *^hw- attested so far precede a PT front vowel.

3.7 *Summary*

In this chapter, I have proposed a system of PT simple onsets. The proposed system resembles the conventional reconstruction (Li 1977) in that it is a rich system which symmetrically contrasts voice and voiceless consonants, both among obstruents and sonorants. However, it departs from the conventional reconstruction in major respects. First, the current system of PT onsets has six contrastive places of articulation. I have shown that a series of uvular onsets must be reconstructed for PT. Second, the proposed system of PT simple onsets views the so-called “voicing alternation” as two distinct phenomena—borrowings and reduction of sesquisyllables. While many etyma can be shown to be late borrowings from Chinese, others are viewed as results of reduction of C̣.C̣- sequences. Furthermore, this proposal holds that PT did not have contrastive aspiration. Aspirated consonants in modern Tai dialects have been shown to be either loans or to have had complex onsets in PT. PT clusters of the type *Cr- have been shown to be one of the sources that gave rise to modern aspirated onsets. The *Cr- and *C̣.C̣- sequences will be treated together with other PT complex onsets in the following chapter.

CHAPTER 4

PROTO-TAI COMPLEX ONSETS

4.1 *Introduction*

In addition to simple onsets, PT phonology also allowed a rich array of complex onsets. Previous works on PT reconstruction propose a sizeable inventory of clusters in PT. Though differing in details, Li (1977), Haudricourt (1948), Nishida (1954), and Sarawit (1973) hold that a PT cluster consisted of an obstruent or a nasal followed by a liquid or glide, e.g. *kr-, *gr-, *xr-, *ŋr-, *kw-, *ɣw-, *w-, *ŋw-, etc. In contrast, Luo (1997) believes that more complex clusters existed, e.g. *zr-, *zl-, *ɟl-⁴⁵, *ztr-, *ɕk^{hr}-⁴⁶, *sb-, etc. Similarly, Liang and Zhang (1996) posit clusters like *zd-, *xpl-, *xkw- etc. However, it is safe to say that these proposals all assume, at least implicitly, that these PT clusters functioned as onsets of monosyllabic words. In contrast to this monosyllabic view is the view that the PT prosodic word could be larger than a simple monosyllable.

As established in Chapter 2, the complex onset of a PT prosodic word, according to the current reconstruction, can be either a tautosyllabic or sesquisyllabic cluster. Each monosyllabic word consisted of an obligatory onset and a heavy rime. Similarly, the major syllable of a sesquisyllabic word was composed of an onset and a heavy rime. The minor syllable consisted of only one syllabic consonant or one onset plus an optional syllabic consonant. For convenience, the minor syllable and the onset of the major syllable in a sesquisyllabic word are treated as a single complex onset. A complex onset of a monosyllable is labeled here as a “tautosyllabic cluster” while a complex onset of a sesquisyllable is referred to here as a “sesquisyllabic cluster”.

⁴⁵ *ɟl- in Luo's notation

⁴⁶ *ɕk^{hr}- in Luo's notation

In this chapter, I propose a system of PT complex onsets according to the phonotactics of the PT prosodic word established in Chapter 2. I first present an inventory of tautosyllabic clusters in PT before proceeding to identify their sesquisyllabic counterparts. Then, I discuss the processes involved in the reduction of sesquisyllables to monosyllables.

4.2 *Tautosyllabic clusters*

PT tautosyllabic clusters could consist of up to three consonants. As discussed in §2.3, these consonants strictly obeyed the Sonority Sequencing Principle (Clements 1990). Most consonants could function as the first element in a tautosyllabic cluster but only the most sonorous segments can be in medial positions. Therefore, only liquids and glides were allowed to occupy the second position in a cluster. Table 4-1 presents the inventory of CC- clusters that can be reconstructed for PT.

Permissible tautosyllabic clusters in PT consisted of either two or three consonants. Most clusters consist of two consonants. In these two-consonant clusters, only stops can be combined with medial *-r-, and *-l-, e.g. *pr-, *br-, *bl-, *ʃl- etc. In contrast, consonants of any manner of articulation can all be combined with medial *-w- to form clusters, e.g. *pw-, *bw-, *^hmw-, *^hlw-, etc. In addition, a few three-consonant tautosyllabic clusters can be reconstructed, e.g. *krw-, and *klw-.

Table 4-1 CC- clusters in PT

		labial	alveolar	palatal	velar	uvular
voiceless stops	-r-	*pr-	*tr-	*cr-	*kr-	*qr-
	-l-	*pl-			*kl-	
	-w-	*pw-	*tw-		*kw-	*qw-
implosives	-l-	*ɓl-				
voiced stops	-r-	*br-			*gr-	
	-l-	*bl-			(*gl-)	
	-w-	*bw-				*gw-
fricatives	-w-		*sw		*xw-	
					*ɣw-	
nasals	-w-	* ^h mw-				
			*nw-	*ɲw-	*ŋw-	
liquid	-w-		* ^h rw-			
			*rw-			

Although the second consonant of a cluster can be either *-r-, *-l-, or *-w-, only the labial glide *-w- can occupy the third position in a tautosyllabic cluster. Some three-part tautosyllabic clusters in PT included *krw-, *klw-, *crw-, *trw-, to cite a few. This is predicted because only *-w- combines as a second member of biconsonantal clusters with initial sonorants. It is important that the medial liquids and glides in tautosyllabic clusters did not contrast for voicing. For example, there is no contrast between *k- plus *r-, and *k- plus *^hr-. Although the medial consonants are transcribed as voiced *-r-, *-l-, and *-w-, these sonorants were not specified for voicing phonologically. As far as the tonal register split is concerned, the first consonant of a cluster determined the series of the tone reflexes in modern dialects.

For example, PT *pr- is reflected with tones in the first series while *br- is reflected with tones in the second series. In the following section, Siamese (SI), Lungchow (LC), and Yay (Y) represent SWT, CT, and NT, respectively.

4.2.1 *Clusters with medial *-r-*

There were numerous clusters with medial *-r- in PT but only stops were allowed to occupy the first position of such a cluster. Moreover, voicing and places of articulation both seem to have been an important factor. While all voiceless stops occurred with a medial *-r-, among voiced stops, only the voiced labial and velar ones were allowed to combine with the liquid. PT tautosyllabic clusters with medial *-r- included *kr-, *tr-, *cr-, *kr-, *qr-, *br- and *gr-. Recall that the *C₀r- clusters went through aspiration due to the medial *-r- in CT and SWT (see §3.5.2).

4.2.1.1 **pr-*

This PT clusters developed along two paths in different dialects. In SWT, CT dialects, and Saek, it became aspirated to *p^hr- as discussed in §3.5.2. This aspirated cluster simplified to /p^h-/ in all SWT dialects including Siamese, Lao, Aiton, Black Tai, and White Tai, and in such CT dialects as Shangsi, Western Nung, Guangnan Nung, and Yanshan Nung. Moreover, the medial *-r- also yodicized to /-j-/ in many CT dialects including Lungchow, Leiping, Debao, Chongzuo, and Daxin. In Yongnan and Long'an, PT *pr- is reflected as /p^hl-/. Saek is unique in showing /p^hr-/, attesting the intermediate stage immediately after the aspiration of *C₀r- clusters. In NT, the cluster did not aspirate and is now reflected as /pl-/, /pj-/, /pɣ-/, or simple /p-/. This cluster corresponds to Li's *p^hl/r- (Li 1977: 87-89). Table 4-2 gives example of etyma with PT *pr-.

Table 4-2 Etyma with PT *pr-

Gloss	PT	SI	LC	Y	Others
‘vegetable’	*prak ^D	p ^h ak ^{DS1}	p ^h jak ^{DS1}	piək ^{DS1}	Saek /p ^h rak ^{DS1} /
‘head hair’	*prɤm ^A	p ^h om ^{A1}	p ^h jum ^{A1}	piəm ^{A1}	Saek /p ^h ram ^{A1} /
‘lean’	*pro:m ^A	p ^h ɔ:m ^{A1}	p ^h jo:m ^{A1}	pjom ^{A1}	Saek /p ^h rɔ:m ^{A1} /
‘bamboo’	*praj ^B	p ^h aj ^{B1}	p ^h aj ^{B1}		Saek /p ^h raj ^{B1} /
‘taro’	*pruək ^D	p ^h uək ^{DL1}	p ^h ɣ:k ^{DL1}	puək ^{DL1}	

Note that in Lungchow the medial *-j- from PT *-r- was lost if the PT rimes contained a high unrounded vowel *u or *i, or a final *-j, e.g. PT *pruək^{DL1} ‘taro’ > *p^hjuək^{DL1} > *p^huək^{DL1} > Lungchow /p^hɣ:k^{DL1}/. In addition, Yay -j- became absorbed into the rime if followed by /-a-/, e.g. *prak^D ‘vegetable’ > pjak^{DS1} > Yay /piək^{DS1}/, and *prɤm^A ‘head hair’ > pjam^{A1} > Yay /piəm^{A1}/.

4.2.1.2 *tr-

In parallel with *pr-, this PT cluster developed along two paths in different dialects. In SWT, NT, and some dialects of CT, *tr- simplified to *hr- merging with the original PT voiceless trill (see §3.6.9.1). Siamese, Black Tai, White Tai, Lue, Wuming, Bouyei, Saek, Hechi, Tiandong, Laibin, Shangsi, Leiping, Lungchow, Yay, etc. all show this development. The modern reflexes of the merger include /r-/, /l-/, /h-/, /j-/, /hj-/, /ð-/, /ɣ-/ and /h-/. In contrast, a number of dialects kept the initial *t- and later aspirated it under the influence of *-r-. Western Nung, Debao, Guangnan Nung, Bao Yen, Wenma, etc. developed along this path. The resulting cluster is now reflected as a plain aspirated /t^h-/. This cluster corresponds to Li’s *t^hr- (Li 1977: 121-125). Table 4-3 gives example of etyma with PT *tr-.

Table 4-3 Etyma with PT *tr-

Gloss	PT	SI	LC	Y	Others
‘carrying pole’	*tra:p ^D	ha:p ^{DL1}	ha:p ^{DL1}	ra:p ^{DL1}	Debao /t ^h a:p ^{DL1} /
‘to see’	*tran ^D	hen ^{A1} -v	han ^{A1}	ran ^{A1}	Debao /t ^h an ^{A1} /
‘stone’	*tri:l ^A	hin ^{A1}	hin ^{A1}	rin ^{A1}	Debao /t ^h ən ^{A1} /
‘head louse’	*traw ^A	haw ^{A1}	haw ^{A1}	raw ^{A1}	Debao /t ^h aw ^{A1} /
‘loom’	*truk ^D	hu:k ^{DL1}	huk ^{DL1}	rok ^{DL1}	Debao /t ^h o:k ^{DL1} /

4.2.1.3 *cr-

This PT cluster developed differently from the other *C_r- clusters. In SWT and most CT varieties, this cluster became aspirated to *c^{hr}- in parallel with *pr-, *tr-, and *kr-. Ferlus (1990)’s *c^{hr}- reflects this intermediate step. This cluster later became *t^{hr}- and is now reflected as /t^h-/. Siamese, Lao, Black Tai, Lue, Lungming, Lungchow, Jingxi, Ningming, Wenma, and Cao Bang are examples of varieties in this group. The only NT dialect that went through this development is Saek. In many NT dialects including Liujiang, Rong’an, Nandan, Lianshan, etc. *cr- dissimilated to *kr- after the original *kr- had become *h_r- (see §4.2.1.4). In these dialects, the cluster is now reflected /kj-/. Another development is found in some CT and NT varieties including Shangsi, Long’an, Wuming, Pinguo, Qiubei, Guigang, etc. In this group of dialects, the complex onset *cr- simplified and merged with PT *c- (see §3.6.1.3). It is now reflected as /c-/, /ts-/, /s-/, /ɕ-/. This cluster is attested only in a few examples in Table 4-4 and corresponds to Li’s *t^hl- (1977: 119-120).

Table 4-4 Etyma with PT *cr-

Gloss	PT	SI	LC	Y	Others
‘to wait’	*cra: ^C	t ^h a: ^{C1}	t ^h a: ^{C1}	sa: ^{C1}	Guigang /ca: ^{C1} /
‘bag’	*croŋ ^A	t ^h uŋ ^{A1}	t ^h uŋ ^{A1}	sɔŋ ^{C1}	
‘senior’	*craw ^C	t ^h aw ^{C1}	t ^h aw ^{C1}		

Note that ‘old’ is assigned here only tentatively because NT evidence is lacking.

4.2.1.4 *kr-

This PT cluster developed along two paths in different dialects. In SWT, NT, and many CT dialects, it was simplified to *^hr- merging with the original PT voiceless trill (see §3.6.9.1). Siamese, Black Tai, White Tai, Lue, Yay, Wuming, Bouyei, Saek, Hechi, Tiandong, Laibin, Shangsi, Leiping, Lungming, Lungchow, etc. all show this development. The modern reflexes of the merger include /r-/, /l-/, /^hl-/, /j-/, /^hj-/, /ð-/, /ɣ-/ and /h-/. Another path is taken by a number of CT dialects including Western Nung, Debao, Guangnan Nung, Bao Yen, and Wenma, to cite a few. In these dialects, *kr- did not simplify but later went through aspiration due to the medial *-r- (see §3.5.2), which gave *k^hr-. The resulting aspirated cluster is now reflected as /k^hj-/ or /c^h-. This cluster corresponds to *xr- in Li (Li 1977: 233-235)’s reconstruction.

Table 4-5 Etyma with PT *kr-

Gloss	PT	SI	LC	Y	Others
‘to seek’	*kra: ^A	ha: ^{A1}	ha: ^{A1}	ra: ^{A1}	Debao /k ^h ja: ^{A1} /
‘gust (of rain)’	*kra: ^B	ha: ^{B1}	ha: ^{B1}	ra: ^{B1}	Debao /k ^h ja: ^{B1} /
‘six’	*krok ^D	hok ^{DS1}	hok ^{DS1}	rok ^{DS1}	Debao /k ^h jok ^{DS1} /
‘to laugh’	*krwəw ^A	huə ^{A1}	hu: ^{A1}	riəw ^{A1}	Debao /k ^h u ^{A1} /
‘fish net’	*kre: ^A	hɛ: ^{A1}	he: ^{A1}	re ^{A1}	Debao /k ^h e ^{A1} /

4.2.1.5 *qr-

This cluster went through two main lines of development. In SWT as well as many CT and NT dialects, *qr- changed to *kr-, but did not merged with the original *kr- (see §4.2.1.4). The intermediate cluster is still retained as /kr-/ and /ky-/ in Guigang and Laibin respectively but became /kl-/ in Hengxian. In a number of other dialects, the medial *-r- yodicized to *-j-, which further palatalized the *k- to /c-/. Thus, while Pingguo, Long’an, Longsheng, Du’an, etc. have /kj-/ for this PT cluster, Yay, Tiandong, Tianlin, Lingyue, etc. show /c-/. In SWT and most CT dialects, the intermediate cluster *kr- became aspirated due to the influence of *-r-. Now it is reflected as /k^hj-/ in Leiping, Lungchow, Debao, and Ningming; as /c^h-/ in White Thai, Western Nung, Guangnan Nung, and Yanshan Nung; as /s-/ in Black Tai and Daxin; and as /k^h-/ in Siamese, Lao, Lue, etc. The NT language Saek also went through aspiration and now reflects this cluster as /t^hr-/ from an intermediate *k^hr-.

Another line of development is a full merger with *h^r- as illustrated by a number of CT dialects which simplified *qr- to *h^r- in parallel to PT *kr- (see §4.2.1.4). The reflex of PT *qr- in Lungming, Fusui, Qinzhou and Shangsi is a simple /l-/ while its reflex in Yongnan and Long’an is a /h^l-/. Etyma that have this PT cluster

are listed under *k^hl- in Li's reconstruction (Li 1977: 226-228). Only the two examples given in Table 4-6 have been attested for this PT cluster. Also see *k.r- in §4.4.4 for a similar development.

Table 4-6 Etyma with PT *qr-

Gloss	PT	SI	LC	Y	Others
‘egg’	*qraj ^A	k ^h aj ^{B1}	k ^h jaj ^{B1}	caj ^{B1}	Lungming /laj ^{B1} /
‘to sift’	*qrxŋ ^A	k ^h uŋ ^{A1}	k ^h jxŋ ^{B1}		Lungming /lxŋ ^{B1} /

4.2.1.6 *br-

The reflex of this cluster in modern dialects depends largely on whether PT voiced stops are reflected as plain or aspirated stops (see §3.6.3), and what the general reflex of medial *-r- is (see §4.2.1). In Lungchow, PT *b- was devoiced to voiceless *p-, and PT *-r- yodized to /-j-/. Therefore, the modern reflex of *br- in these varieties is /pj-/, i.e. *br- > *bj-, > /pj-/. Similarly, it is reflected in Lao as a simple /p^h-/ because *b- regularly became /p^h-/ and medial *-r- was regularly dropped, i.e. *br- > *p^hr- > /p^h-/. Reflexes of this cluster found in modern dialects include /bj-/, /p-/, /pl-/, /pr-/, /pɣ-/, /pj-/, /p^hr-/, /p^hj-/, and /p^h-/. This cluster is attested by only a few examples in Table 4-7 but they are attested in all groups of dialects and show regular correspondences. Li (1977: 89-91) also reconstructs *br- for these etyma.

Table 4-7 Etyma with PT *br-

Gloss	PT	SI	LC	Y	Others
‘to separate’	*bra:k ^D	p ^h ra:k ^{DL2}	pja:k ^{DL2}	pja:k ^{DL2}	Cao Bang /bja:k ^{DL2} /
‘to deceive’	*bra:ŋ ^A	p ^h ra:ŋ ^{A2}	pja:ŋ ^{A2}		Cao Bang /bja:ŋ ^{A2} /
‘deficient’	*bro:ŋ ^B	p ^h rɔ:ŋ ^{B2}	pjo:ŋ ^{B2}	pjo:ŋ ^{B2}	Cao Bang /bjɔŋ ^{B2} /

4.2.1.7 *gr-

Like *br-, the reflex of this cluster in modern dialects depends largely on what the reflexes of *g- (see §3.6.3.4) and *-r- (see §3.6.8.1) are. For example, PT is reflected as *kj- in Debao because *g- and *-r- are reflected as *k- and *-j- respectively. Furthermore, some dialects palatalized the velar stop due to the influence of /-j-/, e.g. Leiping *gr- > *k^hj- > /c^hj-/. Therefore, reflexes of this cluster found in modern dialects include /kj-/, /k-/, /kl-/, /kr-/, /kɣ-/, /kj-/, /k^hr-/, /k^hj-/, /k^h-/, /c-/ and /c^h-. A special development is found in Lungming, where *gr- is reflected as /l-/. This PT cluster is attested by only few examples in Table 4-8. Li (1977: 230-232) also reconstructs *gr- for these etyma.

Table 4-8 Etyma with PT *gr-

Gloss	PT	SI	LC	Y	Others
‘mortar’	*grok ^D	k ^h rok ^{DS2}	kjuk ^{DS2}	cok ^{DS2}	
‘to moan’	*gra:ŋ ^A	k ^h ra:ŋ ^{A2}			Guigang /kra:ŋ ^{A2} /
‘litter (of young animal)’	*gro:k ^D	k ^h rɔ:k ^{DL2}	kjo:k ^{DL2}	cok ^{DL2}	
‘half’	*gruŋ ^B	k ^h ruŋ ^{B2}			Wuming /klaŋ ^{B2} /
‘indigo’	*gra:m ^A	k ^h ra:m ^{A2}	kja:m ^{A2}	sa:m ^{A2} -i	

4.2.2 *Clusters with medial *-l-*

There were a number of clusters with medial *-l- in PT but only stops were allowed to occupy the first position of such a cluster. Like cluster with *-r-, places of articulation seems to have been an important factor. PT tautosyllabic clusters with medial *-l- included only *pl-, *kr-, *bl-, *gl-, and *ɸl-. Recall that the medial *-l- did not aspirate the preceding stop (see §3.5.2). Therefore, in many dialects the contrast between *C̣l- and *C̣r- is still preserved, e.g. Lungchow /pja:^{A1}/ from PT *pla:^A ‘fish’ vs. /p^hja:^{A1}/ from PT *pra:^A ‘stone mountain’.

4.2.2.1 **pl-*

This PT cluster is reflected as /p-/ , /pl-/ , /pj-/ , /pr-/ or /py-/ in modern dialects. While dialects like Siamese, Saek, Yongnan, and Hengxian still preserve the cluster as /pl-/ , Lao, White Tai, Black Tai, Western Nung, Qiubei, and Du’an simplified it to /p-/ . In a great number of varieties including Lungchow, Lungming, Bao Yen, Chongzuo, Jingxi, Tianlian, Rong’an, Nandan, etc., the medial *-l- yodized to /-j-/ , resulting in /pj-/ . In Laibin and Guigang, *pl- completely merged with *pr- and is now reflected as /py-/ and /pr-/ respectively. Li (Li 1977: 84-86) also reconstructs *pl- for this set of data. Table 4-9 presents examples of etyma with PT *pl-.

Table 4-9 Etyma with PT *pl-

Gloss	PT	SI	LC	Y	Others
‘fish’	*pla: ^A	pla: ^{A1}	pja: ^{A1}	pja: ^{A1}	
‘tip, end’	*pla:j ^A	pla:j ^{A1}	pja:j ^{A1}	pja:j ^{A1}	
‘to take down’	*ploŋ ^A	plon ^{A1}	pjuŋ ^{A1}	pjoŋ ^{A1}	
‘to awaken’	*plok ^D	pluk ^{DS1}	pjuk ^{DS1}	pjok ^{DS1}	
‘aquatic leech’	*pli:ŋ ^A	plin ^{A1}	piŋ ^{A1}	piŋ ^{A1}	

The Lungchow and Yay forms for ‘aquatic leech’ do not have medial /-j-/ because it was absorbed by the high front vowel /i/.

4.2.2.2 *kl-

This PT cluster is reflected as /k-/, /kl-/, /kj-/, /kr-/, /ky-/ or /c-/ in modern dialects. While dialects like Siamese, Yongnan, and Hengxian still preserve the cluster as /kl-/, Lao, White Tai, Black Tai, and Yongbei have simplified it to /k-/. Saek also still preserves the medial lateral but has changed the cluster to /tl-/. In a great number of varieties including Lungchow, Lungming, Bao Yen, Chongzuo, Jingxi, Rong’an, Nandan, etc., the medial *-l- vocalized to /-j-/, resulting in /kj-/. In Leiping, Guangnan Nung, Tiandong, and Qiubei the cluster further palatalized to /c-/. Yay, Laibin and Guigang, to cite a few, completely merged medial *-l- and *-r- so that *kl- is now reflected /ky-/ and /kr-/ respectively. Li (Li 1977: 220-225) also reconstructs *kl- for this set of data. Table 4-10 presents examples of etyma with PT *kl-.

Table 4-10 Etyma with PT *kl-

Gloss	PT	SI	LC	Y	Others
‘middle’	*kla:ŋ ^A	kla:ŋ ^{A1}	kja:ŋ ^{A1}	ca:ŋ ^{A1}	
‘banana’	*kluəj ^C	kluəj ^{C1}	ku:j ^{C1}	cuəj ^{C1}	
‘rice seedling’	*kla: ^C	kla: ^{C1}	kja: ^{C1}	ca: ^{C1}	
‘drum’	*klo:ŋ ^A	klɔ:ŋ ^{A1}	kjo:ŋ ^{A1}	coŋ ^{A1}	
‘last (year)’	*kla:j ^A	kla:j ^{A1}	kja:j ^{A1}	ca:j ^{A1}	

4.2.2.3 *ɬl-

There are three main lines of development for *ɬl-. The first one is simplification to *ɬ- in Black Tai, White Tai, and Shangsī. In these dialects, it is now reflected as plain /b-/. The second line of development is coalescence to *d-, merging with the original PT *d- (see §3.6.2.2). This development is found in Siamese, Lao, Western Nung, Debao, Yay, Rong’an, Hechi, etc. In these dialects, *ɬl- is now reflected as /d-/, /n-/, /l-/ or /d-/. The third line of development is retention of *ɬl-. Dialects in this group reflect the PT cluster as /bj-/, /mj-/, or /bl-/. In most dialects, the medial *-l- later yodized to /-j-/, creating /bj-/. Examples of these dialects include Lungming, Lungchow, and Chongzuo. Saek is unique in still retaining the medial -l-. Li (1977: 91-94) includes this set of data under *[?]bl/r-.

Table 4-11 Etyma with PT *ɸl-

Gloss	PT	SI	LC	Y	Others
‘moon’	*ɸluən ^A	duən ^{A1}	bɣ:n ^{A1}	duən ^{A1}	Saek /bliən ^{A1} /
‘flower’	*ɸlo:k ^D	dɔ:k ^{DL1}	bjo:k ^{DL1}	dok ^{DL1}	Saek /blɔ:k ^{DL1} /
‘to weed’	*ɸla:j ^A	da:j ^{A1}	bja:j ^{A1}	da:j ^{A1}	
‘gall bladder, bile’	*ɸli: ^A	di: ^{A1}	di: ^{A1} -i	di: ^{A1}	Saek /bli: ^{A1} /
‘to snap’	*ɸli:t ^D	di:t ^{DL1}			Saek /bli:t ^{DL1} /

Note an unexplained split in Shan dialects, in which PT *ɸl- is reflected as either /m-/ (< *ɸ-) or l- (< *d-) depending on the etyma, cf. Southern Shan /mɔk^{DL1}/ from *ɸlo:k^{DL1} ‘flower’ vs. /lɣn^{A1}/ from *ɸluən^{A1} ‘moon’. It is possible that the split represent a distinction in PT between *ɸr- as in ‘flower’ and *ɸl- as in ‘moon’. However, Saek, which normally preserves the distinction between medial *-l- and *-r-, only shows /-l-/. Therefore, I tentatively view the split reflex in Shan as an innovation⁴⁷.

4.2.2.4 *gl-

This cluster is tentative because only two etyma are found. The first one is ‘nearly alike’, which is reflected as /k^hlap^{DS2}/ in Siamese compound /k^hlap^{DS2} k^hla:j^{C2}/ ‘vaguely resembling’ (Li 1977: 230), and as /cap^{DS2}/ ‘just right’ in Po-ai, and /k^hjap^{DS2}/ ‘just right’ in Dioi. Another etyma is ‘to lasso’, whose modern forms are /k^hlo:ŋ^{C2}/,

⁴⁷ Li (1977: 91) says that Shan, Lao and Lue all show the split but disagree in their reflexes. In contrast, I believe that the regular reflex of *ɸl- in Lao and Lue is /d-/. The /b-/ reflected in some etyma is a special development. For example, Lao /bi:^{A1}/ for ‘gall bladder’ is probably contaminated by /bu:^{A1}/ ‘umbilical cord’.

/kjo:ŋ^{C2}/, and /kløŋ^{C2}/ in Siamese, Lungchow, and Wuming respectively. This cluster is also reconstructed as *gl- in Li's system (1977: 229-230).

4.2.3 *Clusters with medial *-w-*

Tautosyllabic clusters abound in PT as medial *-w- can be combined with many consonants including stops, fricatives, nasals, and liquids. However, it is important to note that some clusters are more robustly supported than others. While such clusters as *kw-, and *yw- are attested by many etyma, *hmw- is supported by only one example.

4.2.3.1 *Velar + *-w-*

The most common complex onsets with medial *-w- in PT had velar consonants as the first part of the cluster. These include *kw-, *xw-, *yw-, and *ŋw-. Interestingly, *gw- is not attested. It is possible that it merged with *yw- early on. While modern dialects usually preserve *kw- and *xw-, they have simplified one or more of the clusters *yw- and *ŋw-. For example, Siamese retains the medial *-w- in *kw-, *xw-, and *yw- but simplified *ŋw- to a simple /w-/. In contrast, Lungchow and Yay keep *kw- as /kw-/, but simplified the other three to a simple *w-, which is now reflected as /v-/. In Li's system, these four clusters are also reconstructed as *kw- (Li 1977: 236-237), *xw- (Li 1977: 240-241), *yw- (Li 1977: 241-242), and *ŋw- (Li 1977: 239-240). These clusters are exemplified in Table 4-12.

Table 4-12 Etyma with PT *kw-, *xw-, *ɣw- and *ŋw- respectively

Gloss	PT	SI	LC	Y	Others
‘to sweep’	*kwa:t ^D	kwa:t ^{DL1}		kwa:t ^{DL1}	Lungming /kwa:t ^{DL1} /
‘to cross over’	*kwa: ^B	kwa: ^{B1}	kwa: ^{B1}	kwa: ^{B1}	
‘deer’	*kwa:ŋ ^A	kwa:ŋ ^{A1}	kwa:ŋ ^{A1}	vʉəŋ ^{A2}	
‘axe’	*xwa:n ^A	k ^h wa:n ^{A1}	k ^h wa:n ^{A1}	va:n ^{A1}	White Tai /xwa:n ^{A1} /
‘putrid’	*xwi:w ^A	k ^h iəw ^{A1}	k ^h i:w ^{A1}		Bao Yen /xwi:w ^{A1} /
‘water buffalo’	*ɣwa:j ^A	k ^h wa:j ^{A2}	va:j ^{A2}	va:j ^{A2}	White Tai /xwa:j ^{A2} /
‘smoke’	*ɣwan ^A	k ^h wan ^{A2}	van ^{A2}		White Tai /xɔn ^{A2} /
‘penis’	*ɣwaj ^A	k ^h uəj ^{A1}			White Tai /xwaj ^{A2} /
‘daytime’	*ŋwan ^A	wan ^{A2}	van ^{A2}	van ^{A2}	Yongnan /ŋon ^{A2} /
‘dumb’	*ŋwam ^C		vam ^{C2}		Yongnan /ŋam ^{A2} /

The Yay form for ‘deer’ shows a special development. The cluster *kw- must have been simplified to *w- due to the following vowel *ʉə⁴⁸, i.e. PT *kwʉəŋ^A > *wʉəŋ^A > vʉəŋ^{A2}. Moreover, this form shows tone A2, indicating that *kw had simplified to *w- before the binary register split.

4.2.3.2 Uvular + *-w-

PT clusters with a uvular obstruent followed by *-w- are attested by a few examples in Table 4-13. *qw- completely merged with *xw- in most dialects but in Phu Thai and Phuan dialects they remain distinct. For example, the Kapong dialect of Phu Thai has /k^hw-/ for *xw-, but /w-/ for *qw- via an intermediate stage *^hw-. The voiced cluster *gw- is found only in *gwun ‘human’ The medial -w- simply dropped

⁴⁸ [k-], [w-] and [ʉ] are all produced at the velum.

out in CT and SWT dialects but left traces in the vowel in NT. In all SWT, including Siamese, White Tai, Lue, Lao, etc., the vowel *u lowered to *ɤ regularly before a non-velar nasal coda. This *ɤ is now reflected as /o/ in SWT (see §5.6.2.3). In many NT varieties including Yay, Saek, Tianlin, Lingyue, Donglan, etc., it had a rounding effect on the following vowel so that now the most common vocalic nucleus in this etymon is /u/. However, in a few dialects, including Pingguo and Nandan, the *gw- was simplified *w-, e.g. Pingguo Pingguo /wun^{A2}/. While the etyma reconstructed with PT cluster *qw- in the current system are included in Li (1977: 240-241) under *xw- (Li 1977: 240-241), *gw- is listed under and *ɣw- (Li 1977: 214-219, 272). Table 4-13 shows etyma with PT *qw-, and *gw-.

Table 4-13 Etyma with PT *qw-, and *gw- respectively

Gloss	PT	SI	LC	Y	Others
‘to hang’	*qwe:n ^A	k ^h wɛ:n ^{A1}		ve:n ^{A1}	Kapong /wɛ:n ^{A1} /
‘face down’	*qwam ^C	k ^h wam ^{C1}	k ^h um ^{C1}	ham ^{C1}	Kapong /wam ^{C1} /
‘spirit’	*qwan ^A	k ^h wan ^{A1}	k ^h wan ^{A1}	van ^{A1}	Kapong /wan ^{A1} /
‘human, person’	*gwun ^A	k ^h on ^{A2}	k ^h ɣn ^{A2}	hun ^{A2}	

Notice that the *qw- in ‘face down’ and the *gw- in ‘human’ became simplified early on in Yay, i.e. PT *qwam^C > *hwam^C > /ham^{C1}/ ‘face down’, and PT *gwun^A > *hwun^{A2} > /hun^{A2}/ ‘human’. In Yay, /h-/ is the regular reflex of *qw- and *gw-.

4.2.3.3 Labial + *-w-

Four labial clusters with *-w- have been identified for PT: *bw-, *pw-, *mw- and *^hmw-. The cluster *mw- is amply attested by reflexes in modern dialects.

Normally, PT *m- is reflected as /m-/ with tones in the second series in all dialects. However, there is a set of etyma which are reflected as /m-/ in SWT and CT but as /f-/ in most NT dialects. In other words, the SWT and CT reflex points to earlier *m- (see §3.6.6.1) but the NT reflex points to *w- (see §3.6.8.3). Following Li (1977: 73-74), I reconstruct *mw- for this set of etyma⁴⁹.

Table 4-14 Etyma with PT *mw-

Gloss	PT	SI	LC	Y	Others
‘clf. for tools’	*mwa:k ^D		ma:k ^{DL2}	fa:k ^{DL2}	
‘hand’ ⁵⁰	*mwu: ^A	mu: ^{A2}	mu: ^{A2}	fuŋ ^{A2}	
‘day’	*mwu: ^C	mu: ^{C2}	mu: ^{C2}	fu: ^{C2}	
‘drunk’	*mwi: ^A			fi ^{A2}	W. Nung /mi: ^{A2} /
‘spirit’	*mwa:ŋ ^A			fa:ŋ ^{A2}	Chongzuo /ma:ŋ ^{A2} /

While evidence for *mw- is the most robust, *bw-, *pw-, and *^hmw are supported by only a few examples. These etyma are shown in Table 4-15. In Li’s reconstruction, they are reconstructed as *b- (Li 1977: 66), *puŋ- (Li 1977, 278-279), and *^hmw- (Li 1977: 75-76, 263) respectively.

⁴⁹ It is also possible that ultimately this set of etyma had *r.m- in PT. Although Proto-Kra (Ostapirat 2000) *r-maŋ^A ‘spirit’ and *r-me^A ‘drunk’ suggest that these etyma also go back to PT *r.m-, Tai internal evidence alone does not indicate sesquisyllabic onsets.

⁵⁰ The final /-ŋ/ found in many NT dialects is a sporadic development, possibly PT *r.mu:^A > *r.mu:ŋ^A > Yay /fuŋ^{A2}/. Compare with PT ‘2s pronoun’ in Appendix B.

Table 4-15 Etyma with PT *bw-, *pw-, and *^hmw-

Gloss	PT	SI	LC	Y	Others
‘fat’	*bwi: ^A	p ^{hi} : ^{A2}	pi: ^{A2}	pi: ^{A2}	Huanjiang /pwi: ^{A2} /
‘lung’	*pwɤt ^D	pɔ:t ^{DL1}	pɤt ^{DS1}	put ^{DS1}	Nandan /puət ^{DS1} /
‘bear’	* ^h mwuj ^A	mi: ^{A1}	mi: ^{A1}	muəj ^{A1}	Huanjiang /mu:j ^{A1} /
‘pubic hair’	* ^h mwu:j ^A	mɔ:j ^{A1}			Po-ai /mi: ^{A1} /

For PT *bwi:^A ‘fat’, the medial *-w- is attested in Huanjiang as well as Rong’an and Longsheng. Further support comes from its Chinese source 肥 *fēi*, whose reconstructed MC form is *bjwei*. For PT *^Dpwɤt^D ‘lung’, the *-w- is most likely preserved as /u/ in Nandan. Moreover, the Late Han pronunciation of its Chinese source 肺 *fèi* is *p^huas*, which lends further support to the reconstruction of medial *-w-. As for *^hmwuj^A ‘bear’ and *^hmwu:j^A ‘pubic hair’, the medial *-w- is not attested as a separate segment but traces of its rounding are still found in the rimes.

4.2.3.4 Coronal + *-w-

Many coronal consonants in PT can be combined with medial *-w- but the clusters are usually simplified to simple onsets. At minimum, the following complex onsets must be reconstructed for PT: *tw-, *nw-, *rw-, *^hrw-, and *^hlw-. These clusters are reconstructed as simple onsets *t-, *n-, *r-, *^hr-, and *^hl- in Li. The most common coronal cluster with medial *-w- is perhaps *rw- shown in Table 4-17.

Table 4-16 Etyma with PT *rw-

Gloss	PT	SI	LC	Y	Others
‘to pour’	*rwa:t ^D	ra:t ^{DL2}			Kapong /hwa:t ^{DL2} /
‘to vomit’	*rwuək ^D		ła:k ^{DL2}	ruək ^{DL2}	Bouyei(4) /zvuo ^{DL2} /
‘gutter, trough’	*rwuəŋ ^A	ra:ŋ ^{A2}	ła:ŋ ^{A2}	ruəŋ ^{A2}	
‘chicken louse’	*rwɔj ^A	raj ^{A2}	łaj ^{A2}	ri ^{A2}	Yishan /hjwi ^{A2} /
‘dry field’	*rwɔj ^B	raj ^{B2}	łaj ^{B2}	ri ^{B2}	Yishan /hjwi ^{B2} /

For *rwɔj^A ‘chicken louse’ and *rwɔj^B ‘dry field’, the *-w- is attested in only a few varieties. In Yishan /hjwi^{A2}/ and /hjwi^{B2}/, PT *r- is reflected as /hj-/ and the *-w- is still preserved. For *rwuək^D ‘to vomit’, and *rwuəŋ^A ‘gutter, trough’, the medial glide is generally lost but its rounding is still preserved on the diphthong in Yay /ruək^{DL2}/, and /ruəŋ^{A2}/. Similarly, the medial /-v-/ in Bouyei(4) for ‘to vomit’. Moreover, /-w-/ in Kapong /rwa:t^{DL2}/ ‘to pour over’ and the rounded vowel in Po-ai /lu:t^{DL2}/ attest the PT medial *-w-. In addition to *rw-, a few other coronal clusters existed in PT. Examples of these clusters are given in Table 4-17.

Table 4-17 Etyma with PT *tw-, *nw-, *^hrw-, and *^hlw- respectively

Gloss	PT	SI	LC	Y	Others
‘to crawl over’	*twaj ^B	taj ^{B1}	taj ^{B1}		Bao Yen /twaj ^{B1} /
‘jackal’	*nwaj ^A	naj ^{A2}		naj ^{A1}	Bao Yen /nwaj ^{A1} /
‘way’	* ^h rwɤn ^A	hon ^{A1}		ran ^{A1}	Huanjiang /wan ^{A1} /
‘mushroom’	* ^h rwet ^D	het ^{DS1}	vit ^{DS1}	rat ^{DS1}	Yishan / ^h jot ^{DS1} /
‘comb’	* ^h rwu:j ^A	wi: ^{A1}	vi: ^{A1}	roj ^{A1}	Nandan /ruəj ^{A1} /
‘to flow’	* ^h lwaj ^A	laj ^{A1}	laj ^{A1}	laj ^{A1}	Bao Yen /lwaj ^{B1} /

For *twaj^B ‘to crawl’, *nwaj^A ‘jackal’, and *^hlwaj^A ‘to flow’, the medial *-w- is clearly attested in a number of dialects, cf. Bao Yen /twaj^{B1}/, /nwaj^{A1}/, /lwaj^{B1}/ respectively. For *^hrwɤn^A ‘way’, the medial *-w- is preserved as initial /w-/ as in Huanjiang /wan^{A1}/, or as rounding of the vowel as in Pingguo /ron^{A1}/). Similarly, the medial *-w- in *^hrwet^D ‘mushroom’ is preserved as /v-/ as in Lungchow /vit^{DS1}/ or as vowel rounding as in Yishan /^hjot^{DS1}/). Similarly, the medial *-w- in *^hrwu:j^A ‘comb’ is preserved as initial /w-/ in Siamese /wi:^{A1}/, /v-/ in Lungchow /vi:^{A1}/, and as rounding in most NT dialects as in Yay /roj^{A1}/ and Nandan /ruəj^{A1}/).

4.2.3.5 Three-consonant clusters *CCw-

PT allowed three consonants to form tautosyllabic clusters under the constraints discussed in §4.2. At a minimum, *krw-, *klw-, *qrw-, *trw- and *crw- must be reconstructed at the PT level. Examples of these clusters are given in Table 4-18. Li (Li 1977: 120, 227-228) reconstructs two-consonant clusters for these etyma.

Table 4-18 Etyma with PT *krw-, *klw-, *qrw-, *trw-, and *crw- respectively

Gloss	PT	SI	LC	Y	Others
‘spider’	*krwa:w ^A		k ^h ja:w ^{A1}	ca:w ^{A1}	Wuming /klwa:w ^{A1} /
‘small frog’	*krwe ^A			kwe ^{A1}	Wuming /klwe ^{A1} /
‘salt’	*klwuə ^A	klwə ^{A1}	ku. ^{A1}	kuə ^{A1}	Rong’an /kwɤ ^{A1} /
‘mountain stream’	*qrwɤj ^C	huəj ^{C1}	vu:j ^{C1}	vi ^{C1}	Longsheng /jwi ^{C1} /
‘to carry on the back’	*trwa:m ^A	ha:m ^{A1}	ha:m ^{A1}	ra:m ^{A1}	Longsheng /jwa:m ^{A1} /
‘plough, to plough’	*crwaj ^A	t ^h aj ^{A1}	t ^h aj ^{A1}	saj ^{A1}	Longsheng /kjwai ^{A1} /

For *krwa:w^A ‘spider’ and *krwe:^A ‘small frog’, the clusters usually simplified to either *kr- or *kw- but a few dialects including Wuming still retain all three consonants, cf. /klwa:w^{A1}/ and /klwe^{A1}/. For *klwuə^A ‘salt’, not only did the *-w- leave rounding on the rime in dialects like Yay, but the Rong’an form also retains the medial. For *qrwɤj^C ‘mountain stream’, the medial is reflected both in the /u/ in Siamese /huəj^{C1}/ and the /v-/ in Lungchow /vu:j^{C1}/ and Yay /vi^{C1}/. For *trwam and *crwaj, the medial *-w- is lost in most dialects, but is preserved in some dialects including Longsheng, cf. /jwa:m^{A1}/ and /kjwai^{A1}/. Note that both initial *r- and medial *-r- became /j/ in Longsheng.

4.3 *Speculative *-w-*

In addition to the complex onsets with medial *-w- above, I also speculate that medial *-w- is responsible for some vowel correspondences. This speculative *-w- is posited in both monosyllabic and sesquisyllabic PT etyma. There are altogether two paradigms of vowel correspondences for which I posit a *-w- despite a lack of unequivocal evidence.

The first paradigm involves etyma whose SWT and CT reflexes point to an earlier long mid vowel but whose NT reflexes point to earlier short high vowels. For example, Siamese /tʰɔːŋ^{C2}/ suggests earlier *doːŋ^C but Yay /tuŋ^{C2}/ suggests earlier *duŋ^C. In this case, I speculate that a medial *-w- had a lowering effect on following *uː, resulting in a long mid vowel *oː in Siamese. In Yay, the *-w- simply dropped out, leaving *uː as the reflex. This situation is in contrast with the PT *oː, which is reflected as /ɔː/ in Siamese, /oː/ in Lungchow, and /o/ in Yay. It also differs from PT uː, which is reflected as /uː/ in Siamese, /u/ in Lungchow, and /u/ in Po-ai. Table 4-19 presents some etyma for which I propose *wuː. Li reconstructs *uɔ (Li 1977: 278), which contrasts with his *wɔ (Li 1977: 277-278).

Table 4-19 Etyma with PT *wuː

Gloss	PT	SI	LC	Y	Others
‘stomach, belly’	*d <u>wuː</u> ŋ ^C	tʰɔːŋ ^{C2}	toːŋ ^{C2}	tuŋ ^{C2}	
‘young, soft’	*ʔ <u>wuː</u> n ^B	ʔɔːn ^{B2}	ʔoːn ^{B2}	ʔun ^{B2}	Guigang /ʔwun ^{B2} /
‘to dye’	*ɲ <u>wuː</u> m ^C	ɲɔːm ^{C2}	joːm ^{C2}	ɲum ^{C2}	
‘spool’	*l <u>wuː</u> t ^D	lɔːt ^{DL1}	loːt ^{DL1}		Po-ai /lut ^{DS1} /

Note that the NT reflex for ‘spool’ point to earlier short vowels rather than long vowels because of the regular shortening of PT long *uː in closed syllables (see §5.6.1.3). Although reconstructing *-w- for this set of data is very speculative, there seems to be some concrete evidence for two of the etyma. The first one is ‘soft’ whose reflex in the NT dialect of Guigang shows a medial /-w-/. The second one is ‘to dye’ whose Vietnamese counterpart *nhuộm* ‘to dye’ has a diphthong /uə/ as vocalic nucleus.

The speculative *-w- is also posited for the data involving front vowels in Table 4-20, for which Li (Li 1977: 274-275) reconstructs *iɛ and *je.

Table 4-20 Etyma with PT *-wi:-

Gloss	PT	SI	LC	Y	Others
‘sunshine’	*Ç.dwi:t ^D	dɛ:t ^{DL1}	dɛ:t ^{DL1}	dit ^{DS1}	
‘late meal’	* Ç.lwi:ŋ ^A	lɛ:ŋ ^{A2}	le:ŋ ^{A2}	riŋ ^{A2}	
‘red’	*Ç.dwi:ŋ ^A	dɛ:ŋ ^{A1}	dɛ:ŋ ^{A1}	diŋ ^{A1}	
‘wasp’	*b.twi:l ^A	tɛ:n ^{A1}	p ^h e:n ^{A1}	tin ^{A2}	

PT *e: is regularly reflected as /ɛ:/ in Siamese, /e:/ in Lungchow, and /e/ in Yay, while PT *i: is reflected as /i:/, /i/, and /i/ in the three languages respectively. The cases above, in contrast, show a different correspondence: Siamese /ɛ:/, Lungchow /e:/ but Yay /i/. I speculate that in these cases a medial *-w- lowered the following *i: to a long mid vowel *e: in CT and SWT before dropping out. In NT, the *-w- simply was lost, leaving *i: as the reflex. For example, Siamese /dɛ:ŋ^{A1}/ came from earlier *dɛ:ŋ^A, whose long *e: in turn resulted from lowering of PT *i: in *Ç.dwi:ŋ^A. In contrast, Yay /diŋ^{A1}/ came from earlier *diŋ^A, whose *i was a regular reflex of PT *i: after the speculated medial *-w-. Note that the Yay reflex for ‘sunshine’ has DS1 rather than DL1 because of the regular shortening of PT long *i: in closed syllables (see §5.6.1.1). So far, I have not been able to identify potential evidence for the medial *-w-. Therefore, the reconstruction of these cases remain speculative. Note that it is not possible, however, to posit *-j- instead of *-w- because this paradigm would be the only set of data that had the medial *-j-.

The second paradigm involves etyma whose SWT and CT reflexes point to an earlier long mid vowel but whose NT reflexes point to an earlier diphthong. For example, Siamese /nɔ:ŋ^{C2}/ suggests earlier *no:ŋ^C but Yay /nuəŋ^{C2}/ suggests earlier *nuəŋ^C. In this case, I speculate that the PT form of this etymon was *nwo:ŋ^C. In this scenario, the medial *-w- was lost in front of the long *o: yielding a simple long *o: in SWT and NT. This *o: is now reflected as /ɔ:/ in Siamese. In NT, the *-w- coalesced with the nucleus *o: to become *uə, which is still preserved in Yay.

In addition to this *o:~*uə correspondence, the paradigm also includes a front vowel *e:~*iə set and a back unrounded *a:~*uə set. Therefore, for this paradigm, I tentatively propose PT *we:, *wɤ: and *wo: for Li's *jɛ (1977: 274-275), *wa (1977: 276) and *wɔ (1977: 277-278) respectively. Table 4-21 presents etyma belonging to this paradigm.

Table 4-21 Etyma with PT *we:, *wɤ: and *wo:

Gloss	PT	SI	LC	Y	Others
‘cucumber’	*p.rwe:ŋ ^A	tɛ:ŋ ^{A1}	p ^h e:ŋ ^{A1}	tiəŋ ^{A1}	
‘tail’	* ^h rwx:ŋ ^A	ha:ŋ ^{A1}	ha:ŋ ^{A1}	ruəŋ ^{A1}	
‘downstairs’	*lwɤ:ŋ ^A	la:ŋ ^{B2}	la:ŋ ^{B2}		Po-ai /lu:ŋ ^{B2} /
‘crossbow’	* ^h nwx: ^C	na: ^{C1}		nuə ^{C1}	W. Nung /na: ^{C1} /
‘to disappear’	* ^h rwx:j ^A	ha:j ^{A1}		ruəj ^{A1}	Lungming /ha:j ^{C1} /
‘bad’	*rwx:j ^C	ra:j ^{C2}	la:j ^{C2}	ruəj ^{C2}	
‘younger sibling’	*nwo:ŋ ^C	nɔ:ŋ ^{C2}	no:ŋ ^{C2}	nuəŋ ^{C2}	
‘elbow’	* _ɕ .swo:k ^D	sɔ:k ^{DL1}		suək ^{DL1}	Lungming /so:k ^{DL1} /

The only potential support for the speculated *-w- in this paradigm comes from the MC form for 弩 *nǔ* (< MC *nuo^B*, LH *na^B*, OC **nâ?*) ‘crossbow’. Although the *u* in the MC forms may be viewed as concrete evidence for a medial *-w- in Proto-Tai **nwɔː^C*, the chronology is rather problematic. PT most likely predates Late Han Chinese at which stage the rime of 弩 was still simple *a*. Therefore, the reconstruction of PT *-w- for this set of data must be taken as tentative.

4.4 *Sesquisyllabic clusters*

In §2.3, I established that a PT prosodic word could be either monosyllabic or sesquisyllabic. I argued that this sesquisyllabic view of PT phonology best accounts for the wide range of sound correspondences observed among the daughter languages. According to the proposal, a sequence of consonants that makes up the minor syllable plus the major syllable onset, or a sesquisyllabic cluster, can have up to two consonants in the minor syllable, and up to three consonants in the major syllable. In this section, I identify sesquisyllabic clusters that can be reconstructed for PT before proceeding to discuss the process through which sesquisyllables consisting of these complex onsets reduced to monosyllables in modern languages. Note that some clusters posited here are more speculative than others. Many clusters are attested by only one or two etyma, and must therefore be taken as tentative.

4.4.1 *Voicless stop + voiceless stop*

Clusters that consist of two voiceless obstruents are perhaps the most common type of sesquisyllabic clusters in PT. In many of such **C̥.C̥-* clusters, both the minor syllable and the major syllable onsets are voiceless stops, e.g. **p.t-* and **k.t-*. When these complex onsets were reduced to fit the monosyllabic template, the resulting onsets stayed voiceless. In all modern languages, etyma with this type of PT cluster

have tones in the first series. The most robust set is PT *p.t- given Table 4-22. Li (1977: 118-119) posits a tautosyllabic *tr- for this set of data.

Table 4-22 Etyma with PT *p.t-

Gloss	PT	SI	LC	Y	Others
‘eye’	*p.ta: ^A	ta: ^{A1}	ha: ^{A1}	ta: ^{A1}	Saek /pra: ^{A1} /
‘to die’	*p.ta:j ^A	ta:j ^{A1}	ha:j ^{A1}	ta:j ^{A1}	Saek /pra:j ^{A1} /
‘grasshopper’	*p.tak ^D	tak ^{DS1}		tak ^{DS1}	Lungming /t ^h ak/

Following Ferlus (1990), PT *p.t- simplified to /t-/ in Siamese and Yay but went through weakening of the non-initial *t- in Saek and Lungchow. In Lungchow, original *p.t- became *p.r- after the non-initial *t- became weakened to *r-. The *p- was then lost leaving only a voiceless *hr-, which is now reflected as /h-/. In Saek, the original *p.t- became *p.r- after weakening but, rather than dropping the initial *p-, the sesquisyllabic cluster became a true tautosyllabic cluster /pr-/.

The second set of etyma that go back to a sesquisyllabic cluster consisting of two voiceless stops is presented in Table 4-23. Following Ferlus (1990), I reconstruct PT *k.t- for these etyma, for which Li (1977: 117-118) posits *tl-. In this scenario, PT *k.t- simplified to /t-/ in Siamese and Yay but became *t.r- in Lungchow through weakening of *t-. This intermediate cluster later simplified to *hr-, which is reflected now as /r-/ in Yay. In Long’an and a few other dialects, the modern reflex of *k.t- is /t^h-/ which attests the intermediate step *tr-, i.e. *k.t- > *t.r- > *tr- > /t^h-/ (see §4.4.4).

Table 4-23 Etyma with PT *k.t-

Gloss	PT	SI	LC	Y	Others
‘full’	*k.tem ^A	tem ^{A1}	tim ^{A1}	rim ^{A1}	Long’an /t ^h im ^A /
‘fart’	*k.tyt ^D	tot ^{DS1}	tyt ^{DS1}	rat ^{DS1}	Long’an /t ^h ut ^{DS1} /
‘big leaf’	*k.to:ŋ ^A	tɔ:ŋ ^{A1}	to:ŋ ^{A1}	roŋ ^{A1}	
‘to wake up’	*k.tu:n ^B	tu:n ^{B1}	tu:n ^{B1}	ru:n ^{B1}	Long’an /t ^h ɯn ^{B1} /

In addition to *p.t- and *k.t-, other sesquisyllabic clusters with two stops can also be reconstructed. However, their reconstruction is rather tentative because they are attested only by one or two examples. These clusters include *p.q-, and *q.p-. Table 4-24 presents data for these clusters. These cases are viewed as aberrancies in Li (Li 1977).

Table 4-24 Etyma with PT *p.q- and *q.p-

Gloss	PT	SI	LC	Y	Others
‘leg’	*p.qa: ^A	k ^h a: ^{A1}	k ^h a: ^{A1}	ka: ^{A1}	Saek /kwa: ^{A1} /
‘to sell’	*p.qa:j ^D	k ^h a:j ^{A1}	k ^h a:j ^{A1}	ka:j ^{A1}	Saek /kwa:j ^{A1} /
‘body hair, feather’	*q.pu:l ^A	k ^h on ^{A1}	k ^h on ^{A1}	pu:n ^{A1}	

As discussed in §2.3.2, the labiality in Saek /kwa:^{A1}/ and Saek /kwa:j^{A1}/ is a retention from PT, as these two etyma still preserve the initial *p- in Sui /pa:¹/ ‘leg’ and /pe¹/ ‘to see’. Therefore, I posit *p.qa:^A ‘leg’ and *p.qa:j^A at the PT level respectively. The former has been related to PAN *paqa ‘thigh’ (Ostapirat 2005). As for ‘body hair’, it is not clear if the CT/SWT from *qon are related to the NT form

from *puul. If they are, they should be reconstructed as *q.puul^A at the PT level. That is, SWT and CT simply dropped the non-initial *.p- but kept the labiality in the vowel. In contrast, NT lost the initial *q- when *q.p- simplified. Qinzhou, Yongnan, and Fusui show /p^hun^{A1}/, /p^hɿn^{A1}/, and /p^huən^{A1}/ respectively. The aspiration in /p^h-/ in these three dialects seems to be a trace of the initial *q.-

4.4.2 Voiceless obstruent + voiced stop

PT had a number of sesquisyllabic clusters with voiceless obstruents followed by voiced stops. These C̥.C̥- clusters became simple implosives in most dialects, except for Saek where they became weakened to simple sonorants. Note that the Saek reflex of *C̥.b- is /v-/ but this fricative must have relatively recently developed from *w-. The first set of data, presented in Table 4-25, can be reconstructed with *C̥.b-.

Table 4-25 Etyma with PT *C̥.b-

Gloss	PT	SI	LC	Y	Others
‘shoulder’	*C̥.ba: ^B	ba: ^{B1}	ba: ^{B1}	ba: ^{B1}	Saek /va: ^{B1} /
‘thin’	*C̥.ba:ŋ ^A	ba:ŋ ^{A1}	ba:ŋ ^{A1}	ba:ŋ ^{A1}	Saek /va:ŋ ^{A1} /
‘light’	*C̥.baw ^A	baw ^{A1}	baw ^{A1}	baw ^{A1}	Saek /vaw ^{A1} /
‘side’	*C̥.buəŋ ^C	buəŋ ^{C1}	bɿ:ŋ ^{C1}		Saek /viəŋ ^{C1} /
‘to poison’	*C̥.buə ^A	buə ^{A1}	bɿ: ^{A1}		Saek /viə ^{A1} /

In these etyma, most dialects point unanimously to an earlier *ɓ- but Saek /v-/ suggests *^hw-. This correspondence can be accounted for by positing PT *C̥.b-. In this scenario, the medial *-b- in most dialects became an implosive *ɓ- when the voiceless

*Ç- dropped out⁵¹, e.g. PT *Ç.ba:^B ‘shoulder’ > *ʃa:^B > Siamese /ba:^{B1}/ In contrast, the medial *.b- in Saek went through weakening and became *.w- before the voiceless *Ç- was lost, e.g. PT *Ç.ba:^B ‘shoulder’ > *Ç.wa:^B > *^hwa:^B > Saek /va:^{B1}/. Note that Ostapirat (2005) relates this etymon to PAN *qabaRa. In this account, the first consonant *q- in PAN *qabaRa corresponds to the initial *Ç- in PT *Ç.ba:^B. A second set of data is given in Table 4-26. These etyma go back to PT *Ç.d-.

Table 4-26 Etyma with PT *Ç.d-

Gloss	PT	SI	LC	Y	Others
‘black’	*Ç.dam ^A	dam ^{A1}	dam ^{A1}		Saek /ram ^{A1} /
‘bone’	*Ç.duk ^D	du:k ^{DL1}	du:k ^{DL1}	dok ^{DL1}	Saek /rɔ:k ^{DL1} /
‘raw’	*Ç.dip ^D	dip ^{DS1}	dip ^{DS1}	dip ^{DS1}	Saek /rip ^{DS1} /
‘catfish’	*Ç.dok ^D	duk ^{DS1}		dok ^{DS1}	Saek /rɔk ^{DS1} /
‘mountain’	*Ç.do:j ^A	dɔ:j ^{A1}		doj ^{A1}	Saek /rɔ:j ^{A1} /

In parallel to *Ç.b-, the PT sesquisyllabic cluster *Ç.d- became an implosive *d- in almost all dialects. Again, the only exception is Saek which has a weakened reflex *-r- in these etyma. For example, the non-initial *.d- in PT *Ç.dip^D became implosivized before the voiceless initial *Ç- was lost, and is now reflected as /d-/ in Siamese, Lungchow, and Yay, e.g. *Ç.dip^D > *d̥ip^D > Siamese /dip^{D1}/. In contrast, in Saek the non-initial *.d- went through weakening and became *.r- before the voiceless initial *Ç- was lost, e.g. *Ç.dip^D > *Ç.rip^D > *^hrip^D > Saek /rip^{D1}/. Among the etyma

⁵¹ It is not clear what phonetic reasons underly the change from non-initial voiced stop preceded by a voiceless onstruent became implosive. However, this change is consistent with the fact that implosives are tonogenetically half-voiced, half-voiceless (see §6.2).

shown in Table 4-26, two have been claimed to have Austronesian connections (Ostapirat 2005): PT *C̣.dip^D ‘raw’ and PAN *qudip ‘alive’; and *C̣.dam^A ‘black’ and PAN *tidem ‘black, dark’. These two Austronesian etyma lend support to the reconstruction of *C̣.d- clusters for these etyma.

In addition to *C̣.b- and *C̣.d-, a sesquisyllabic cluster with non-initial *ɟ- may have existed. The only potential case is ‘to stand’, which is reflected as /ju:n^{A1}/ in Siamese and Lungchow, and as /du:n^{A1}/ in Yay. While the former points to earlier *ʔju:n^A, the latter suggests *d̥u:n^A. These divergent developments probably reflect PT *C̣.ɟu:n^A. Specifically, the non-initial *ɟ- became implosivized when preceded by a voiceless obstruent as expected but the resulting implosive was not *ɟ̥-. Since *ɟ̥- has never been allowed in Tai languages, in the SWT and CT dialects, implosive/glottalized palatal is precisely *ʔj-. As for NT, the non-initial *ɟ- became implosive *d̥- rather than *ɟ̥- on account of this same ban against the palatal implosive.

The development of implosives from plain voiced stops in the non-initial position strongly suggests that there was no contrast between medial voiced stops and their implosive counterparts. In other words, PT only distinguished plain voiced stops *b, and *d from implosives *ɓ and *ɗ in the initial position.

4.4.3 *Voiced obstruent + voiceless stop*

As discussed in §3.4, a number of etyma with “voicing alternation” seem to go back to Proto-Kra-Dai. I have speculated that these items go back to sesquisyllabic clusters with voiced obstruents followed by voiceless stop in PT. These *C̣.C̣- clusters became aspirated in CT and SWT but became voiced in NT. However, it is not clear what the initial voiced consonants were. Most likely, the initial *C̣- is different in different etyma. Etyma in Table 4-27 go back to PT *C̣.t-.

Table 4-27 Etyma with PT *Ç.t-

Gloss	PT	SI	LC	Y	Others
‘to arrive’	*Ç.tɤŋ ^A	t ^h uŋ ^{A1}	t ^h ɤŋ ^{A1}	taŋ ^{A2}	
‘to come into contact’	*Ç.tu:k ^D	thu:k ^{DL1}	thuk ^{DS1}	tu:k ^{DS2}	
‘wild’	*Ç.tuəl ^B	t ^h uəl ^{B1}	t ^h ɤ:n ^{A1} -t	tuəl ^{B2}	
‘to flood’	*Ç.tuəm ^B	t ^h uəm ^{B1}	t ^h u:m ^{B1}	tum ^{B2}	

The PT sesquisyllabic cluster *Ç.t- was resolved differently in SWT and CT on one hand, and in NT on the other. In NT, the non-initial *.t- assimilated to *Ç- and became voiced, e.g. PT *Ç.tɤŋ^A > *dɤŋ^A > Yay /taŋ^{A2}/. In contrast, it became aspirated in SWT and CT dialects, i.e. *Ç.tɤŋ^A > *t^hɤŋ^A > Siamese /t^huŋ^{A1}/. This etymon is probably related to PAN *datenj ‘to arrive’ (Thurgood 2007: 255) which shows the expected sequence *Ç- + *.t-. Table 4-28 presents etyma that had *Ç.k-, and *Ç.q- at the PT level.

Table 4-28 Etyma with PT *Ç.k-, and *Ç.q- respectively

Gloss	PT	SI	LC	Y	Others
‘son-in-law’	*Ç.kuəj ^A	k ^h ɤ:j ^{A1}	k ^h ɤ:j ^{A1}	kuəj ^{A2}	
‘pit’	*Ç.kum ^A	k ^h um ^{A1}	k ^h um ^{A1}	kum ^{A2}	
‘rice’	*Ç.qaw ^C	k ^h a:w ^{C1}	k ^h aw ^{C1}	haw ^{C2}	
‘to bite’	*Ç.qɤp ^B	k ^h op ^{DS1}	k ^h op ^{DS1}	hap ^{DS1}	
‘excrement’	*Ç.quj ^C	k ^h i: ^{C1}	k ^h i: ^{C1}	haj ^{C1}	
‘bitter’	*Ç.cɤk ^D	k ^h om ^{A1}	k ^h um ^{A1}	ham ^{A2}	

In parallel with *Ç.t-, and *Ç.c-, the PT sesquisyllabic cluster *Ç.k- became a simple voiced *g- in NT but became voiceless aspirated *k^h- in CT and SWT. For example, *Ç.kuəj^A ‘son-in-law’ became *guəj^A > /kuəj^{A2}/ in Yay but became *k^huəj^A > /k^hɤ:j^{A1}/ in Siamese. Similarly, PT *Ç.q- became a simple voiced *g- in NT but became voiceless aspirated *k^h- in CT and SWT, merging with *Ç.k-. For example, *C.quj^C ‘excrement’ became *guj^C > /haj^{C2}/ in Yay but became *k^hi:^C > /k^hi:^{C1}/ in Siamese.

4.4.4 Voiceless stops + liquids/glides

As discussed in §2.4.3, it is possible in a sesquisyllabic language to have a contrast between tautosyllabic clusters and sesquisyllabic clusters consisting of a stop followed by a liquid, e.g. *pr- vs. *p.r-. In PT, a few sesquisyllabic clusters with medial liquids or glides can be reconstructed. Among these clusters, clusters of type *Ç.r- can be reconstructed with confidence. Table 4-29 presents example of PT *k.r-. This cluster corresponds to Li’s *k^hl- (Li 1977: 226-228).

Table 4-29 Etyma with PT *k.r-

Gloss	PT	SI	LC	Y	Others
‘to imprison’	*k.raŋ ^A	k ^h aŋ ^{A1}	haŋ ^{A1}	caŋ ^{A1}	
‘to drive away’	*k.rap ^D	k ^h ap ^{DS1}			Wuming /klap ^{DS1} /
‘top for spinning’	*k.ra:ŋ ^B	k ^h a:ŋ ^{B1}	ha:ŋ ^{B1}	ca:ŋ ^{B1}	
‘illness, fever’	*k.raj ^C	k ^h aj ^{C1}	haj ^{C1}	caj ^{C1}	
‘to beg’ ⁵²	*k.ro: ^A	k ^h ɔ: ^{A1}	ho: ^{A1}		Nandan /kjo ^{A1} /

⁵² Gedney (n.d.) and Hudak (2008) have /k^ho:^{A1}/ for Lungchow but Li (1940) and Zhang et al. (1999) has /ho:^{A1}/.

For the etyma given above, Lungchow points to a voiceless liquid **hr-* but Siamese and Yay suggest an earlier cluster **kr-*, which did not come from PT **kr-* (see §4.2.1.4). This otherwise unexplained correspondence strongly indicates PT sesquisyllabic **k.r-*, which either became a true cluster **kr-*, or simply lost the initial **k-* to become **hr-*, depending on the language. For example, PT **k.raŋ^A* became **kraŋ^A* before aspirating to **k^hraŋ^A*, which is now reflected as a simple /k^haŋ^{A1}/ in Siamese. In contrast, the same etymon PT **k.raŋ^A* became **hraŋ^A* in Lungchow and is now reflected as /haŋ^{A1}/. Another sesquisyllabic cluster with a medial liquid that can be reconstructed with confidence is **p.r-*, which contrasted with **pr-* (see §4.2.1.1). In addition, Wuming /k^{DS1}lap/ ‘to drive away’ and Nandan /kjo^{A1}/ ‘to beg’ shows that these two etyma are also found in NT. Li (1977: 86-87) reconstructs **pr-* for this set of data. Table 4-30 presents etyma that go back to PT **p.r-*.

Table 4-30 Etyma with PT **p.r-*

Gloss	PT	SI	LC	Y	Others
‘to be exposed’	<i>*p.ra:k^D</i>	ta:k ^{DL1}	p ^h ja:k ^{DL1}	ta:k ^{DL1}	Saek /pra:k ^{DL1} /
‘to burst’	<i>*p.re:k^D</i>	tɛ:k ^{DL1}	p ^h e:k ^{DL1}	tek ^{DL1}	Saek /pre:k ^{DL1} /
‘to hunt’	<i>*p.raw^B</i>		p ^h jaw ^{B1}	taw ^{B1}	
‘cucumber’	<i>*p.rwe:ŋ^A</i>	tɛ:ŋ ^{A1}	p ^h e:ŋ ^{A1}	tiəŋ ^{A1}	Saek /priəŋ ^{A1} /

In the above data set, the Lungchow reflex /p^hj-/ points to earlier **pr-*, but this **pr-* must be secondary. This is because Saek shows unaspirated /pr-/ rather than the expected aspirated reflex /p^hr-/ for PT **pr-* (see §4.2.1.1). Moreover, Siamese and Yay both show /t-/ indicating an earlier **t-*. To account for this correspondence, I posit **p.r-* which merged with PT **p.t-* to give /t-/ in Siamese and Yay, and /pr-/ in Saek

(see §4.4.1). In parallel to *k.r- above, this PT *p.r- became reduced to a monosyllabic cluster and merged with PT *pr- in Lungchow, which now reflect this complex onset as /p^hj-/. Note that the etyma ‘cucumber’ have the speculative *-w- in addition to *p.r- clusters, i.e. *p.rwe:ŋ^A (see §4.3).

In addition to *k.r- and *p.r-, we may tentatively posit *t.r- and *c.r- for the following etyma: ‘to break’ and ‘to ask’ respectively. The etymon ‘to break’ is reflected as /hak^{DS1}/, /t^hak^{DS1}/, /t^hak^{DS1}/, and /rak^{DS1}/ in Siamese, Ningming, Long’an, and Yay respectively. These reflexes point to earlier *tr-, but the reflexes in many CT dialects point to *t-, e.g. Leiping and Lungchow /tak^{DS1}/. It is likely that PT *t.r- merged with *tr- earlier on in most dialects (see §4.2.1.2) but became simplified to *t- in such CT dialects as Lungchow, etc. As for ‘to ask’, the reflexes in Siamese, Lungchow, and Yay are /t^ha:m^{A1}/, /t^ha:m^{A1}/, and /sa:m^{A1}/, all pointing to *cr-. However, their counterparts in Debao, Jingxi, Western Nung, and Guangnan Nung are /k^hja:m^{A1}/, /k^hja:m^{A1}/, /c^ha:m^{A1}/, and /c^ha:m^{A1}/ respectively. I propose that in parallel to *t.r-, PT *c.r- merged with *cr- early on in most dialects, i.e. *c.r- > *cr- > /t^h-/ or /s-/ (see §4.2.1.3). In Debao and like dialects, however, PT *c.r merged with *k.r- before both becoming *kr- > *k^hr- > /k^hj-/ (see §4.2.1.4).

There is also evidence for sesquisyllabic clusters of the shape *C̣.w-. This sesquisyllabic cluster corresponds to *^hw- in Li’s system (1977: 82). Table 4-31 lists some examples of PT *q.w-.

Table 4-31 Etyma with PT *Ç.w-

Gloss	PT	SI	LC	Y	Others
‘sweet, delicious’	*Ç.wa:l ^A	wa:n ^{A1}	va:n ^{A1}	va:n ^{A1}	Wuming /wa:n ^{B1} /
‘to sow’	*Ç.wa:l ^B	wa:n ^{B1}	va:n ^{B1}		
‘rattan’	*Ç.wa:j ^A	wa:j ^{A1}	va:j ^{A1}	va:j ^{A1}	

The etyma in this set of data all have tones in the first series and point to earlier *^hw-. However, it is not possible to posit the voiceless glide for these cases because *^hw- is regularly reflected as /f-/ in Siamese and Lungchow. Therefore, I posit a sesquisyllabic cluster *Ç.w- to account for this correspondence. It is not clear what voiceless consonant the *Ç.- might have been, but evidence from outside of Tai suggests *q.w- for ‘sweet’ and ‘rattan’. In particular, Proto-Lakkja (L-Thongkum 1992) has *k^hwa:n^A for ‘sweet’. The initial *k^h- seems to correspond to PT *q-, as Proto-Lakkja *k- corresponds regularly to Tai *k-. As for ‘rattan’, the reconstructed PAN form is *qúway, which also has *q- as initial. Therefore, it is possible that ‘sweet’ and ‘rattan’ had *q.w- as onset in PT.

4.4.5 Voiced consonant + liquid/glide

In parallel to sesquisyllabic clusters of the type *Ç.r- above, PT also had sesquisyllabic cluster with non-initial liquids or glides preceded by voiced consonants, including *Ç.r-, and *Ç.l-. Etyma that go back to these clusters all have tones in the second series, indicating that the initial consonants were voiced. Among these, *m.l- is the only cluster whose initial consonant can be recovered unequivocally.

Table 4-32 Etyma with PT *m.l-

Gloss	PT	SI	LC	Y	Others
‘slippery’	*m.lu:l ^B	lu:n ^{B2}	lɿ:n ^{B2}		Saek /mlu:l ^B /
‘grain’	*m.lec ^D	met ^{DS2}		nat ^{DS2}	Saek /mlɛk ^{DS2} /
‘insect’	*m.le:ŋ ^A	mɛ:ŋ ^{A2}	me:ŋ ^{A2}	neŋ ^{A2}	
‘body louse’	*m.lel ^A	len ^{A2}	min ^{A2}	nan ^{A2}	Saek /mlɛl ^{A2} /

The regular reflex of *m.l- in Siamese is /l-/ as in ‘slippery’ and ‘body louse’. The /m-/ in ‘seed’ and ‘insect’ is an instance of dialect borrowing as discussed in Pittayaporn (to appear-b) Except for *m.l-, there is no evidence for what their places of articulation or manners of *Ç.- might have been. The first set of etyma, given in Table 4-33, go back to PT *Ç.r-.

Table 4-33 Etyma with PT *Ç.r-

Gloss	PT	SI	LC	Y	Others
‘root’	*Ç.ra:k ^D	ra:k ^{DL2}	la:k ^{DL2}	ra:k ^{DL2}	
‘boat’	*Ç.rwuə ^A	ruə ^{A2}	lu: ^{A2}	ruə ^{A2}	
‘day after’	*Ç.ru: ^A	ru:n ^{A2}	lu: ^{A2}	ru ^{A2}	
‘tomorrow’ ⁵³					
‘bedbug’	*Ç.ruət ^D	ruət ^{DL2}	lɿ:t ^{DL2}		Longsheng /rɿ:t ^{DL2} /

⁵³ The Siamese form /-ru:n^{A2}/ is now part of a monomorphemic word /ma.ru:n^{A2}/ from an original compound /mw:^{C2} ru:n^{A2}/. The /-n/ was acquired under the influence of the deictic /ni:^{C2}/ ‘this’, because /ma.ru:n^{A2}/ often occurs in the expression /maru:n^{A2} ni:^{C2}/ ‘this coming day after tomorrow’. Also, see Li (1977: 128-129).

In the above data set, Siamese and Yay point to earlier *r- but Lungchow indicates *l-. To account for this discrepancy, Li (1977: 127-128) reconstructs *dr- for these etyma, but there is no evidence that the cluster had an initial *d-. In the current proposal, I propose that this set of etyma had sesquisyllabic clusters with a voiced initial followed by a non-initial liquid *r-. The non-initial liquid *r- in this environment merged with *l- in the same environment. The result of this neutralization in Lungchow and many CT dialects is /l-/ while its reflex in Yay and other NT dialects is /r-/. The /r-/ in Siamese is a retention of the contrast between *Ç.r- and *Ç.l-, as will be come clear in the discussion of *Ç.l- below.

Parallel to *Ç.r-, PT also had *Ç.l- clusters, exemplified by the etyma in Table 4-34. In these etyma, Siamese and Lungchow point to earlier *l- but Yay indicates *r-. This is the same case of neutralization that we have just seen with respect to *Ç.r-. In particular, *Ç.r- completely merged with *Ç.l- in CT and SWT but the reflexes are different in the two groups. Lungchow and many CT dialects show /l-/ as reflex of *Ç.l- while Yay and other NT dialects has /r-/. Again, PT *Ç.r- is also reflected as /l-/ in Lungchow and /r-/ in Yay. The contrast is preserved in Siamese and other SWT dialects where *Ç.l- is reflected as /l-/, contrasting with /r-/ from PT *Ç.r-.

Table 4-34 Etyma with PT *Ç.l-

Gloss	PT	SI	LC	Y	Others
‘to pull, drag’	*Ç.la:k ^D	la:k ^{DL2}	la:k ^{DL2}	ra:k ^{DL2}	
‘to steal’	*Ç.lak ^D	lak ^{DS2}	lak ^{DS2}	rak ^{DS2}	
‘fingernail, toenail’	*Ç.lep ^D	lep ^{DS2}	lip ^{DS2}	rit ^{DS2} -f	
‘wind’	*Ç.lum ^A	lom ^{A2}	lum ^{A2}	rum ^{A2}	
‘to lick’	*Ç.liə ^A	liə ^{A2}	li: ^{A2}	riə ^{A2}	

Li (1977: 124-125) reconstructs *dl- for these etyma, but cognates in Lakkja indicate that the initial *Ç.- was not the same in all etyma. The Lakkja forms for ‘to pull’, ‘to steal’, ‘fingernail’, and ‘wind’ are /kja:k^{D2}/, /la:k^{D2}/, /pli:p^{D1}/, and /jɣm^{A2}/ respectively. Although these Lakkja forms suggest different initial consonants for all of these etyma, the sound correspondence between Lakkja and Tai has not been worked out. Therefore, the sesquisyllabic clusters for this set of etyma are temporarily written as *Ç.l-.

4.4.6 *Clusters with non-initial nasals*

PT had a number of sesquisyllabic clusters with non-initial nasals. One of the most robustly supported is *t.n-. Almost all modern dialects show reflexes of *d- for this set of etyma, but Saek unexpectedly has /tr-/. Li (1977: 129-131) included etyma in this set under his *^ʔdr- together with etyma that are reflected as /hr-/ in Saek (see §4.4.2). Ferlus (1990) in contrast proposes that Saek /tr-/ came from original *^ʔl- but believes that this glottalized liquid had become *d- at the PT stage⁵⁴. Although I agree with Ferlus that Saek /tr-/ represents a retention of a stage before the formation of *d-, I claim that Saek /tr-/ goes back to a sesquisyllabic cluster *t.n- at the PT stage. Examples of etyma with PT *t.n- are given in Table 4-35.

⁵⁴ Ferlus (1990: 12) does not reconstruct distinctions because of Saek’s marginal position within Tai.

Table 4-35 Etyma with PT *t.n-

Gloss	PT	SI	LC	Y	Others
‘star’	*t.na:w ^A	da:w ^{A1}	da:w ^{A1}	da:w ^{A1}	Saek /tra:w ^{A1} /
‘earthworm’	*t.nuəɭ ^A	duən ^{A1}	dɤ:n ^{A1}	duən ^{A1}	Saek /truəɭ ^{A1} /
‘to transplant’	*t.nam ^A	dam ^{A1}	dam ^{A1}	dam ^{A1}	Saek /tram ^{A1} /
‘to be related by marriage’	*t.no:ŋ ^A	dɔ:ŋ ^{A1}		doŋ ^{A1}	Saek /trɔ:ŋ ^{A1} /
‘cockspur’	*t.nɤ:j ^A	duəj ^{A1}		da ^{A1}	Saek /tra: ^{A1} /

The etyma in this set show that PT *t.n- coalesced and became *d- in most dialects except for Saek, where the non-initial *n- became *r-. For example, PT *t.nam^A ‘to transplant’ became *dam^A > /dam^{A1}/ in Siamese but *tram^A > /tram^{A1}/ in Saek. Therefore, Saek still preserves the original distinction between PT *d-, *C̣.d-, and *t.n- (see §2.3.2). PT *t.nam^A ‘to transplant’ seems to be related to the PAN form *talem ‘to plant’⁵⁵, lending support for the reconstruction of PT *t.n- for this set of etyma.

In addition to *t.n, PT also had clusters that consisted of other voiced stops followed by a non-initial *n-. Examples of etyma with *C̣.n- are given in Table 4-36.

⁵⁵ This form corresponds to *taNem in Blust’s notation (Wurm & Wilson 1975).

Table 4-36 Etyma with PT *Ç.n-

Gloss	PT	SI	LC	Y	Others
‘water’	*Ç.nam ^C	na:m ^{C2}	nam ^{C2}	ram ^{C2}	
‘bird’	*Ç.nok ^D	nok ^{DS2}	nuk ^{DS2}	rok ^{DS2}	
‘bamboo shoot’	*Ç.na:ŋ ^A			ra:ŋ ^{A2}	Long’an /na:ŋ ^{A2} /
‘dew, mist’	*Ç.na:j ^A		na:j ^{A2}		Longsheng /rwa:j ^{A2} /

Although the tonal reflexes in modern forms unequivocally indicate that the the leftmost consonant of these *Ç.n- clusters was voiced, it not clear exactly what consonants they were. Li (1977: 131-132) posit *nl/r- for these etyma but Lakkja data reveals clearly that the *Ç.- in each etymon was different. Proto-Lakkja (L-Thongkum 1992) had *num^C, *ʔmlok^D, and *s-Na:ŋ^A for ‘water’, ‘bird’ and ‘bamboo shoot’, respectively. Furthermore, Ostapirat (2005) links the first two to PAN *daNum ‘water’⁵⁶ and *manuk ‘bird’, which also show different initial consonants. Therefore, I temporarily write *Ç.n- for these etyma. One exception is ‘bamboo shoot’, for which I posit *r.na:ŋ^A. The evidence for the initial *r- comes from some CT dialects such as Debao, Jingxi, Guangnan Nung, Western Nung. These varieties have either /r-/, /l-/, or /ð/, all of which point to *r- instead of the expected /n-/, cf. Debao /nam^{C2}/ ‘water’, /nok^{DS2}/ ‘bird’ but /la:ŋ^{A2}/ ‘bamboo shoot’.

4.4.7 Other clusters

In addition to the sesquisyllabic clusters discussed so far, some other combinations may have also existed in PT. Among these, some clusters consisted of two consonants, e.g. *r.t-, *t.h-, *q.s-, *m.p-, *s.c-, *z.j-, *g.r-, *m.n- etc. Others may

⁵⁶ This form corresponds to *daNum in Blust’s notation (Wurm & Wilson 1975).

have been more complex, e.g. *gṃ.r-, *jṃ.r-, *c.pl-, *g.lw- etc. However, they are sometimes attested by only one or two etyma. For example, the only case of *q.s- so far is *q.sep^D ‘centipede’ which is reflected as Siamese /k^hep^{DS1}/ (< PSWT *k^hrep^{DS1}), Lungchow /k^hip^{DS1}/, Jingxi /k^hjap^{DS1}/. Qinzhou /lyp^{DS1}/, Wuming /θip^{DS1}/, Du’an /θip^{DS1}/. No other etymon shows this correspondence. Although one may argue that the etyma claimed to have these complex onsets in fact are not reconstructible, the fact that they are found throughout Tai and that they are usually regular regarding the rimes and the tones suggests that they might have been part of PT. Therefore, these clusters are tentatively reconstructed until more evidence becomes available. Refer to Appendix B for the reconstructions of individual etyma.

Moreover, the inventory of reconstructed simple and complex onsets suggests that many more sesquisyllabic clusters may have existed. First of all, diverse types of sesquisyllabic onsets have been reconstructed for PT, but only a limited number of combinations have been reconstructed. For example, only one clusters consisting of a voiceless stop followed by a nasal has been reconstructed—*t.n-. Similarly, *k.r-, *p.r-, *t.r, and *c.r- have been reconstructed for sesquisyllabic clusters containing non-initial *-r-, but *q.r- is still missing. Second, an abundance of voiceless nasals have been reconstructed for PT despite the presence of sesquisyllabic clusters. In SEA, voiceless sonorants are usually results of loss of a voiceless consonant at the left edge of tautosyllabic or sesquisyllabic clusters. For example, voiceless nasals in Burmese historically came from clusters with non-initial nasals, e.g. *hna* ‘face’ from **sna* (Nishida 1975b). Therefore, it is very plausible that many etyma reconstructed with voiceless sonorants as onsets might in fact have had sesquisyllabic clusters.

Taken together, these two facts about the inventory of reconstructed simple and complex onsets in PT seem to suggest that many more sesquisyllabic clusters existed but their traces have been lost in modern dialects. The Kam-Sui language

Lakkja seems to support this point of view. One crucial case is that of PT voiceless sonorants that correspond to Lakkja forms with initial /k^h-/. These forms include pairs like PT *^hma:^A ~ Lakkja /k^hũə^A/ ‘dog’, PT *^hmu:^A ~ Lakkja /k^hũ:^A/ ‘pig’, PT *^hmat^{DS1} ~ Lakkja /k^hũət^{D1}/ ‘flea’, etc. This suggests that these etyma might have in fact been sesquisyllabic, i.e. *q.ma:^A, *q.mu:^A, and *q.mat^A respectively at the PT level. This would be perfectly in parallel with the reconstruction of *q.w- discussed in §4.4.4. However, in the current reconstruction, I opt to posit *^hm- and *^hr- for these etyma rather than *q.m- and *q.r- because Tai-internal evidence is lacking.

4.5 *Reduction of sesquisyllabic clusters*

Even though I have argued that PT was a sesquisyllabic language, none of its daughter languages attested so far shows this typological profile. The biggest question is then what the processes through which PT sesquisyllables reduced to monosyllables were. The process of a sesquisyllabic language becoming monosyllabic can be referred to as “monosyllabicization.” In many languages of SEA, the monosyllabicization can be captured quite straightforwardly by a unified process in which sesquisyllables, disyllables, or polysyllables become reduced to monosyllables. Such “monosyllabification” is governed by a set of rules that determines the outcome of the reduction. For example, Tsat, an Austronesian language of Hainan Island, went through a monosyllabification process in which the first syllable was completely lost unless the medial was a liquid.

Table 4-37 Monosyllabification in Tsat (Sagart 1993)

	Old Cham	Tsat
‘eye’	mata	ta ²
‘bee’	hani	ni ²
‘fish’	ʔikan	ka:n ²
‘wet’	pasah	sa ¹
‘white’	putih	ti ¹
‘plant’	pala	pia ² (< pla)
‘village’	palei	piei ² (< plei)
‘moon’	bulan	p ^h ian ³ (< blan)
‘new’	birou	p ^h iə (< brou)
‘shoulder’	bara	p ^h ia ³ (< bra)

The first set of data above shows that Old Cham disyllables were normally reduced to monosyllables by simply dropping the first syllables. In contrast, the second set of data shows that the resulting monosyllables still retained the initial consonants if they could combine with the medial consonants to form a cluster. A similar process of reduction must have operated in Tai languages but the details may have been different. Although one is also tempted to look for a set or sets of principles that make up a unified process of monosyllabification for Tai, comparative data in modern languages indicate that PT sesquisyllables had gone through various processes of cluster modification before the actual process of reduction took place. In this section, I outline a preliminary account of the reduction of PT sesquisyllables to monosyllables and illustrate how the reduction happened with case studies from four dialects: Siamese, Lungchow, Yay, and Saek.

4.5.1 *Processes of cluster adjustment*

Unlike the case of Tsaic, monosyllabification in Tai did not involve only a process of monosyllabification. Rather, it is best characterized as a series of changes that ended with a process of monosyllabification. In other words, the sesquisyllabic PT had gone through sound changes that modified the original sesquisyllabic clusters before the monosyllabification of sesquisyllables finally took place. The sound changes that resulted in adjustments of sesquisyllabic clusters prior to the monosyllabification include weakening, implosivization, metathesis, assimilation, and simplification.

The process of weakening can be characterized as a change in which a segment becomes less consonant-like (Trask 2000: 190). Many of the PT sesquisyllabic clusters underwent weakening. For example, the non-initial **t-* in PT **p.t-* was weakened to **r-* in many dialects including Lungchow and Saek (see §4.4.1). The result of this weakening of **t-* is that PT **p.t-* became **p.r-*. In Lungchow, the resulting **p.r-* clusters lost the initial **p-* and became **hr-*. This **hr-* merged with the original **hr-* and is now reflected as /h-/ with tones from the first series. In Saek, the **p.r-* cluster that resulted from the weakening merged with the original PT **p.r-* so that both PT **p.r-* and **p.t-* are now reflected as a monosyllabic cluster /pr-/. Another example of weakening is that of PT non-initial voiced stops in **C̣.C̣-* clusters (see §4.4.2). In Saek, a non-initial **d-* that occurred after an initial voiceless stop was weakened to **r-*. The resulting cluster **C̣.r-* later went through monosyllabification and became **hr-*, which is now reflected as /r-/. These cases of weakening are schematized in Figure 4-1.

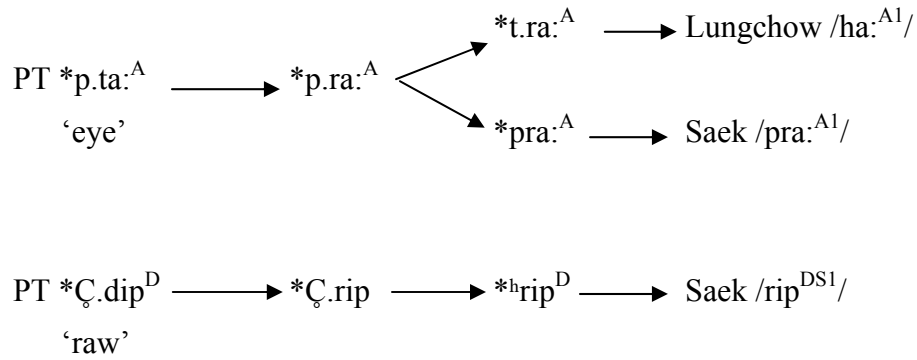


Figure 4-1 Weakening of *p.t- and *C.d-

The second process of cluster adjustment is implosivization, a change in which a segment becomes implosive. In Tai varieties, implosivization occurred mainly in two types of clusters: *t.n- and *C.d-. In the former, the two consonants in PT *t.n- were collapsed together yielding an implosive *ɗ- in almost all dialects. The only language that did not undergo the implosivization is Saek, in which the non-initial *.n- weakened to /-r-/ (see §2.3.2). In *C.d-, PT non-initial *.d- became implosivized to PT *ɗ-. Later the resulting cluster *C.ɗ- was simplified to simple *ɗ-. Again, all Tai dialects except for Saek went through this change. The non-initial *.d- was weakened to *.r- in Saek (see §4.4.2). The implosivization of *t.n- and *C.d- is schematized in Figure 4-2.

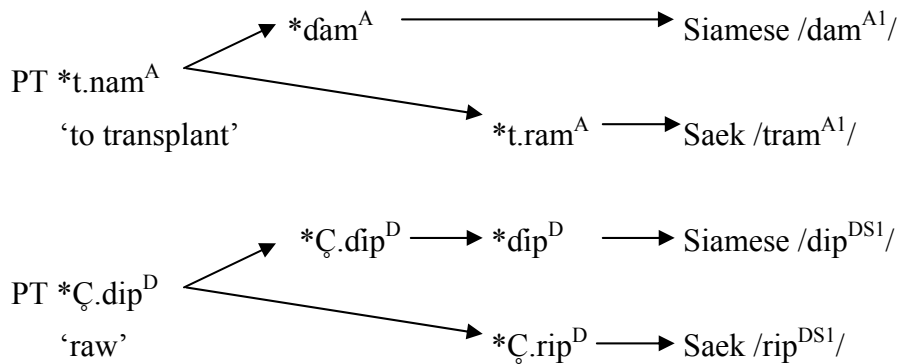


Figure 4-2 Implosivization of *t.n- and *C.d-

The third process of cluster adjustment is metathesis, which refers to a change in which the order of segments in a word is altered (Trask 2000: 211). Two examples of PT sesquisyllabic clusters that were metathesized are *r.t- and *p.q-. The PT cluster *r.t- can be found in the etymon *r.tuə^B ‘sweat’. It was metathesized to *t.r- in SWT and most CT dialects. The resulting cluster merged with PT *t.r- which later became monosyllabified to *tr-. Now the cluster is reflected as /h-/ in both Siamese and Lungchow (see §4.2.1.2 and §4.4.4). Shangsi and a few other CT dialects did not metathesize but voiced the non-initial *t- under the influence of *r-, yielding *r.d-. This intermediate cluster was later reduced and became *d-, e.g. Shangsi /ty^{B2}/ from *r.tuə^B < PT *r.tuə^B. The second example is the development of PT *p.q- in Saek. In this NT language, PT *p.q- is reflected /kw-/, suggesting that *p.q- metathesized to *q.p- before the non-initial *p- was weakened to *w-. In other dialects, the initial *p- simply dropped out due to the monosyllabification. The metathesis of *r.t- and *p.q- is schematized in Figure 4-3.

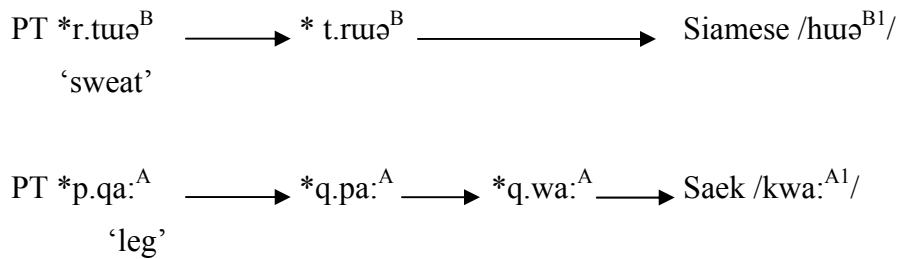


Figure 4-3 Metathesis of *r.t-, and *p.q-

The fourth process to be discussed is assimilation, which refers to a change in which a segment becomes more similar to another segment in the same phonological sequence (Trask 2000: 30). An example of assimilation in PT sesquisyllabic clusters is the case of PT *p.t-. In Debao and a number of other CT dialects, this PT cluster went through weakening so that its non-initial *t- became *r-, resulting in a cluster *p.r-. This new *p.r- clusters subsequently went through assimilation such that its *p.- assimilated to *r- resulting in *t.r-. This *t.r- merged with the original PT *tr- and is now reflected as /t^h-/ in Debao (see §4.2.1.2 and §4.4.4).

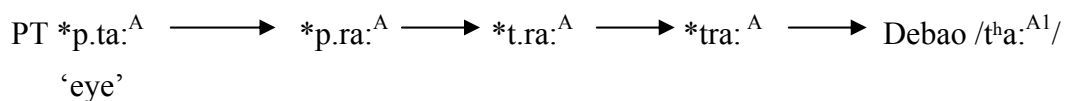


Figure 4-4 Assimilation of *p.t-

The last process that targeted PT sesquisyllabic clusters is simplification, which refers to the reduction of a consonant cluster by dropping one or more of its constituents (Trask 2000: 310). An example of simplification in Tai occurred with PT

*t.h- in *t.haj^C ‘to weep’. In all NT and a few CT dialects, the cluster *t.h- lost its non-initial *h- and became a simple *t-, merging with the original PT *t-. In Yay, *t.haj^C is now reflected as /taj^{C1}/. Note that the process of monosyllabification is essentially a simplification process although for purposes of the discussion it is treated as a distinct process in this account. The simplification of *t.h- is schematized in Figure 4-5.

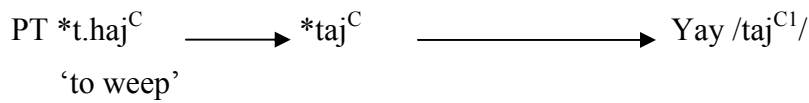


Figure 4-5 Simplification of *t.h-

In this subsection, I have shown that at least five processes of cluster adjustment are commonly found among Tai languages. It is not the case that all dialects underwent all of these processes. Nor is it true that all dialects went through the same subsets of processes. In the following subsection, I discuss the process of monosyllabification which reduced the sesquisyllabic clusters that had undergone the cluster modification to monosyllabic clusters.

4.5.2 *Monosyllabification of sesquisyllables*

After undergoing the processes of cluster adjustments described above, PT sesquisyllables subsequently went through a process of monosyllabification. The general result of this process was that sesquisyllabic clusters were reduced to true tautosyllabic clusters. Like the case of Tsat briefly discussed earlier, the monosyllabification of Tai languages can be captured with a set of developments that determines the outcome of the cluster simplification. These developments conspire so

that the results of the reduction obeyed the Sonority Sequencing Principle (see §2.2.2.2).

The first development involves sesquisyllabic clusters with medial liquids. A PT sesquisyllabic *C̣.l- or *C̣.r- cluster whose initial *C̣- was a voiceless stop kept both consonants but became a tautosyllabic cluster. For example, PT *k.r- became *kr- in SWT and NT when the monosyllabification took place. PT *k.r- became *kr- and then aspirated to PSWT *k^hr-. It is now reflected as /k^h-/ in Siamese. In some CT dialects, PT *k.r- had undergone simplification and became *^hr- before the monosyllabification took place. In Lungchow, it is now reflected as /h-/ (see §4.4.4). Another example is PT *p.r- which became *pr- in Lungchow when the monosyllabification took place, merging with the original *pr-. Subsequently, aspiration of clusters with medial *-r- took place and transformed it to *p^hr-, which is now reflected as /p^hj-/ (see §4.4.4). The monosyllabification of *k.r- and *p.r- is schematized in Figure 4-6.

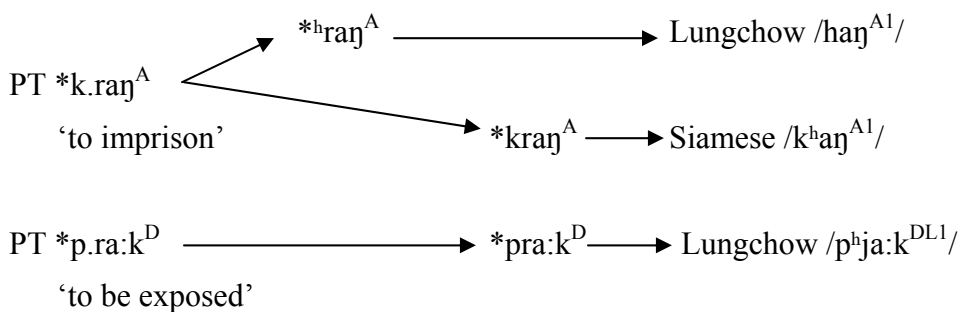


Figure 4-6 Monosyllabification of *k.r- and *p.r-

The second development determines more generally which consonant within a cluster was dropped. As a rule, the first syllable of a *C.C- cluster is dropped while the

second is retained. For example, PT *k.t- went through two lines of development. NT dialects went through the first one which involved weakening of non-initial *t-, i.e. *k.t- > *k.r-. CT and SWT, on the other hand, did not go through weakening and remained *k.t-. Therefore, at the time of monosyllabification, PT *k.t- had two distinct reflexes: *k.r- in NT and *k.t- in CT and SWT. According to this first principle, the second consonant of each of their reflexes was kept while the first was dropped. Therefore, NT *k.r- was monosyllabified to *^hr-, but CT and SWT *k.t- went to *t-. The *^hr- is now reflected as /r-/ in Yay (see §3.6.9.1) while the *t- is reflected as /t-/ in Siamese and Lungchow (see §3.6.1.2). Another illustration of this development is the case of *t.h-. This PT sesquisyllabic cluster went through simplification and became *t- in Long'an but remained *t.h- in CT and SWT. At the time of monosyllabification, the *t.h- preserved in SWT and CT was simplified to *h- as this fricative is the second consonant of the cluster. In contrast, the *t- in Long'an did not go through monosyllabification because it was already monosyllabic. The development of these two clusters can be schematized in Figure 4-7.

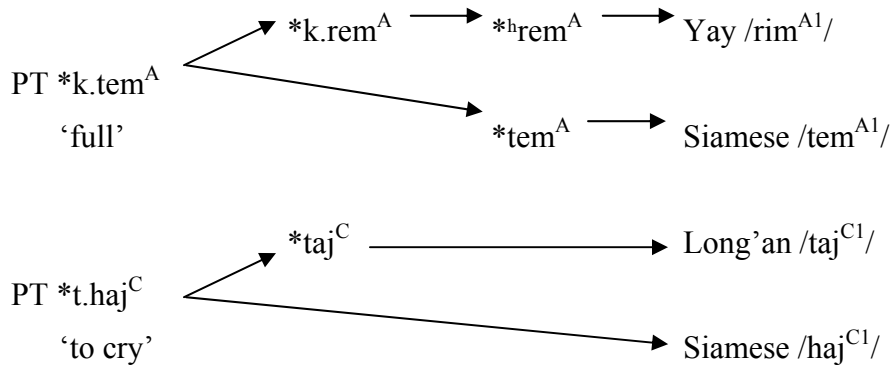


Figure 4-7 Monosyllabification of *k.t- and *t.h-

The last development determines the phonation type of the resulting clusters. Like many other languages of SEA, it is the phonation type of the first consonant that is preserved on the resulting clusters, cf. Chamic (Brunelle 2008; Thurgood 1999). For example, PT *p.r- became *p.t- in SWT and some NT dialects. The resulting *p.t- cluster was eventually monosyllabified to a voiceless *t- because both *p- and *t- were voiceless. In contrast, PT *ᶑ.d- became *ᶑ.r- in Saek through weakening (see §4.4.2). At the time of monosyllabification, the initial *ᶑ- was lost but its phonation type was preserved on the resulting *hr. Therefore, modern Saek has /r-/ with the first series of tones for PT *ᶑ.d-, reflecting the voicelessness of the initial *ᶑ-. Another example is the case of the voicing-alternation items (see §3.4 and §4.4.3). Voicing alternation items that had *d- in an earlier stage of NT came from PT *ᶑ.t-. At the time of monosyllabification, the voiced *ᶑ- was dropped but the non-initial *.t- took on its voicing and became *d-. The reflex of PT *ᶑ.t- in Yay is /t-/ with tones in the second series, reflecting the phonation type of the first consonant in PT *ᶑ.t-. The developments of *p.t-, *ᶑ.d-, and *ᶑ.t- are schematized in Figure 4-8.

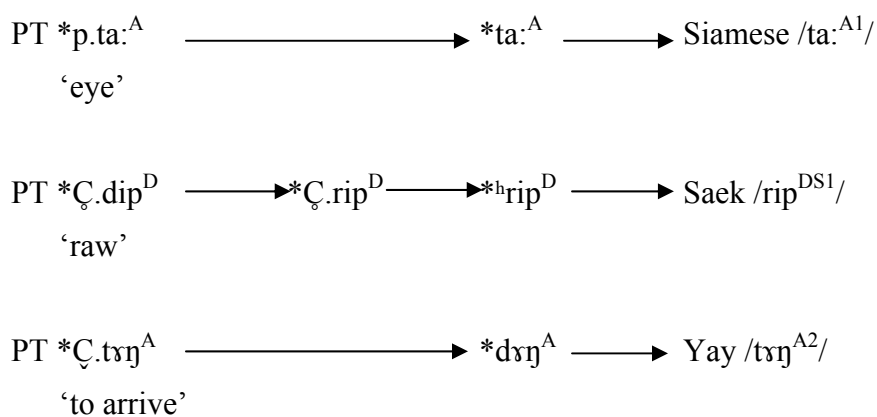


Figure 4-8 **Monosyllabification of *p.t-, *ᶑ.d-, and *ᶑ.t-**

In this subsection, I have discussed the monosyllabification of PT sesquisyllables. I have shown that it is not the original PT sesquisyllabic clusters that underwent the monosyllabification because the process occurred after a number of sound changes, e.g. weakening, assimilation, simplification etc., had taken place (see §4.5.1). Therefore, the sesquisyllabic clusters that were materials for the monosyllabification were the result of those changes. The monosyllabification itself consisted of three developments that work together to transform sesquisyllables to monosyllables whose onset obeys the Sonority Sequencing Principle.

4.6 Summary

In this chapter, I have discussed the two types of PT complex onsets: tautosyllabic clusters and sesquisyllabic clusters. I have shown that tautosyllabic clusters in PT can consist of up to three consonants and must obey the Sonority Sequencing Principle. Sesquisyllabic clusters, on the other hand, usually violated the Sonority Sequencing Principle. In modern Tai dialects, PT sesquisyllables have been reduced to monosyllables through a long process of reduction. I have shown that this process cannot be viewed as a unified process of monosyllabification because PT sesquisyllabic clusters had undergone various phonological changes before they were finally reduced to tautosyllabic clusters. In sum, this chapter is the first systematic attempt at positing sesquisyllables for PT and their developments into monosyllabic words in modern Tai languages.

CHAPTER 5

PROTO-TAI RIMES

5.1 *Introduction*

Perhaps the least understood aspect of PT reconstruction is its vocalic system. The central problem seems to be that the number of correspondence sets that have been identified is so great that one is led to account for it by positing a complex vocalism (Li 1977: 297-299). While Li (1977) holds that PT vowel system lacked vowel length but consisted of a large number of diphthongs and triphthongs, Sarawit (1973) posits a vowel-length contrast and thus a smaller set of complex vowels. This problem has also lead Haudricourt (1975) and Strecker (1983) to envision PT as a disyllabic language with a relatively simple vowel system.

In contrast to the vowels, the system of final consonants is by consensus judged to be straightforward. This view is reflected in the work of Li (1977: 58) who states that the PT final system is preserved with very little change in most dialects. The inventory of codas that Li proposes includes only stops, nasals, and glides, with a possible liquid *-l as attested in Saek. However, approaching these two components of PT phonology separately is probably ill-conceived. Given that vowels and codas frequently function together as one unit cross-linguistically, it is highly likely that their mutual influence is at work in the observed correspondence patterns among modern languages. In this study, the PT vowels and codas are treated together as rimes.

In Chapter 2, I established that PT was a sesquisyllabic language in which prosodic words can be either monosyllabic or sesquisyllabic. Due the maximality and minimality constraints, each PT prosodic word must be exactly one iambic foot. In this chapter, I propose an inventory of PT rimes made up of a comparatively small number of vowels and a relatively rich set of final consonants. Three aspects of the rime

system are discussed: 1) the reconstruction of the vowel-length contrast, 2) the solution to the so-called “Gedney’s Puzzle,” and 3) the existence of final*-l, *-c, and possibly *-ŋ. Lastly, I present the reconstructions of each PT rime.

5.2 *Inventory of PT rimes*

In §2.3, I argued that PT is best characterized as a sesquisyllabic language, which allowed both simple monosyllabic and sesquisyllabic words. Obviously, each monosyllable can only have one rime. This is also the case for sesquisyllables because the first “half-syllable” allows neither a contrastive vowel, nor a real coda consonant, as discussed in §2.4.3 and §2.5. Therefore, the rime of a sesquisyllable refers more specifically to the rime of the major syllable.

As outlined in §2.5.1.2, size requirements on the prosodic word dictate that each PT rime must contain exactly two timing units, allowing for long open rimes (-V:), short closed rimes (-VC), and long closed rimes (-V:C). Examples of words allowed by the PT phonotactics are *ka:^A ‘crow’, *kat^D ‘to take a bite’, and *ka:t^D ‘mustard green’, but not *ka^A. Because PT rimes cannot take up more than two timing units, the coda in CV:C syllables likely share the second mora with the vowel as is the case in Siamese as analyzed by Morén and Zsiga (2006).

I propose that the PT vowel system consists of 7 simple vowels, each occurring as long or short, and five diphthongs. The PT inventory of simple vowels is given in Table 5-1. This triangular system is typologically well-attested (Maddieson 1984b).

Table 5-1 Inventory of PT simple vowels

	front	back	
	unrounded	unrounded	rounded
high	*i, *i:	*ɯ, *ɯ:	*u, *u:
mid	*e, *e:	*ɤ, ɤ:	*o, *o:
low		*a, *a:	

Overall, this inventory is identical to that proposed by Sarawit (1973) but differs substantially from the conventional reconstruction. Two aspects that should be highlighted are the contrastive vowel length and the three-way height contrast. For each vowel quality, both short and long vowels were contrastive. As far as vowel height is concerned, a vowel can be either high, mid, or low. The low series consists only of a long and short *a: and *a. Turning to complex vowels, the inventory of PT diphthongs is given in Table 5-2.

Table 5-2 Inventory of PT diphthongs

	front	back	
	unrounded	unrounded	rounded
rising	*iə	*ɯə	*uə
falling		*ɤɯ *aɯ	

An outstanding characteristic of the proposed diphthong inventory illustrated above is its relative small size, especially when compared to the system of Li (1977) and Sarawit (1973). This system consists of three rising diphthongs and two falling diphthongs. These diphthongs do not contrast for length. The lack of length contrast in

diphthongs is expected because long diphthongs would exceed the two timing units allowed. As short diphthongs have two timing units by definition, long diphthongs would have to be linked to three timing units. Except for *ɾuɨ and *auɨ posited here, surface rising diphthongs are treated as vowels plus glides phonologically because they do show length contrast. For example, *[ai] is phonologically a vowel *a followed by *-j contrasting with *[a:i], which is phonologically *a:j.

Of the PT consonants that can occur in onset position, fewer than half can function as codas. In each syllable, only one final consonant is allowed. The inventory of PT codas is given in Table 5-3.

Table 5-3 Inventory of PT codas

	labial	alveolar	palatal	velar	uvular	glottal
stop	*p	*t	*c	*k		
fricative						
nasal	*m	*n	(*ɲ)	*ŋ		
liquid		*l				
glide	*w		*j			

The set of PT final consonants can be characterized as the onset inventory reduced by synchronic positional neutralizations. In terms of manners of articulation, PT did not distinguish between fricatives and stops so that the only final obstruents allowed were voiceless stops. As for sonorants, only nasals and the lateral *-l could occupy the coda slot. This manner neutralization can be understood as a neutralization of the feature [+/-continuant] found widely across languages. In terms of places of articulation, the dorsal places of articulation were extremely impoverished when

compared to onset position. Although velar, uvular, and glottal consonants could all function as onsets, only the velar stop and nasal can be in coda position. This can be considered another case of positional neutralization. As for phonation type, most noticeable is the absolute lack of contrast. While PT onsets can be either voiceless, implosive/glottalized, or voiced, only one phonation-type is possible for each manner of articulation in coda position: all final obstruents are voiceless but all final sonorants are voiced. This can be understood as laryngeal neutralization. This generalization agrees with Li's and Sarawit's reconstructions. This restricted inventory of coda consonants is typologically very common and is one of the most widespread SEA areal traits (Matisoff 2001; Rhee 2003). Cross-linguistically, neutralizations of manner, place, and laryngeal features are robustly attested.

5.3 *PT vowels: quality and quantity*

The question whether there was a quantity contrast among PT vowels is one of the most crucial factors in understanding PT phonology. This question also ties very closely to the issue of vowel qualities, especially vowel heights. Opinions among researchers differ with regard to whether both quality and quantity were contrastive in PT vocalism. While Li (1977) and Luo (1997) consider vowel length to have developed secondarily in daughter languages, Haudricourt (1948; 1966-1974), Wang (1966), Nishida (1954; 1955), and Sarawit (1973) hold that it was contrastive in PT. In this section, I argue in favor of the latter position. As a point of departure, the inventories of PT simple vowels according to Li and Sarawit are illustrated in Table 5-4.

Table 5-4 PT vowel system according to Li (1977) and Sarawit (1973)

*i	*ɯ	*u	*i, *i:	*ɯ, *ɯ:	*u, *u:
*e	*ɤ	*o	*e, *e:	*ɤ, ɤ:	*o, *o:
*ɛ	*a	*ɔ		*a, *a:	
Li (1977)			Sarawit (1973)		

Li (1977) proposes a PT vowel system that distinguished three degrees of vowel height but lacked vowel length contrast. He argues that in many dialects long and short vowels do not contrast, except for *a and *a:. He interprets the contrast in the low vowel as one of quality rather than one of quantity (Li 1977: 259). Taking Siamese as a case study, he further argues that the quantity contrast among non-high vowels was introduced by borrowing or secondary lengthening. He only recognizes Siamese short mid vowels /e/, and /o/, together with long low vowels /ɛ:/ and /ɔ:/ as having PT origins. Although he admits that the length contrast among high vowels in Siamese does point to a distinction in PT, he argues that Siamese long high vowels /i:/, /ɯ:/, /u:/ resulted from monophthongization of PT *ie, *ue, *ue.

Luo (1997: 56-63) assumes Li's position that PT did not have a vowel-length contrast but also argues that the low vowels *ɛ and *ɔ did not exist. Although he correctly identifies some cases of low vowels in modern dialects as secondary developments, he fails to recognize that many of these cases may have been conditioned by vowel length. For example, he claims that PT *e and *o became *je and *wo in closed syllables in the SWT dialect of Tai Ya, spoken in the upper reaches of the Red River, e.g. /kjem^{C1}/, and /swoŋ^{A1}/ from what he would reconstruct as *kem^C 'cheek'⁵⁷ and *soŋ^A 'two'⁵⁸ Although Luo is correct in pointing out that the

⁵⁷ Reconstructed as *kem^C in Li's system and as *ke:m^C in the current reconstruction.

breaking seen in Tai Ya is secondary, he ignores the fact that there are equally many cases of closed-syllable items that show simple /i/ and /o/ (presumably from earlier *e and *o respectively), e.g. /mit^{DS2}/ and /hok^{DS1}/ from what he would also have to set up *mlet^D ‘grain’⁵⁹ and *xrok^D ‘six’⁶⁰. Moreover, he does not discuss cases of minimal pairs such as /pit^{D1}/ ‘duck’ from PT *pet^D vs. /pjɛt^{D1}/ ‘eight’ from *pe:t^D, and /nok^{D1}/ ‘bird’ from PT *Ḷ.nok^D vs. /nwok^{D1}/ ‘outside’ from PT *ɭ.no:k^D. It is impossible to explain what conditioned the alleged split if the two sets of words come from the same PT sources⁶¹.

Prominent proponents of contrastive vowel length in PT are Haudricourt (1948; 1966-1974), Sarawit (1973), Nishida (1954; 1955) and Wang (1966), among others. Although Haudricourt does reconstruct contrastive short and long vowels, he did not provide a picture of the entire system nor discuss it explicitly. The next major study to posit a primary length contrast in the PT vowel system is Sarawit (1973) who meticulously reconstructs PT vocalism bottom-up by first reconstructing the intermediate proto-languages of SWT, CT, and NT. In contrast to Li, Sarawit proposes that the low vowels /ɛ:/ and /ɔ:/ in Siamese and other dialects resulted from lowering of long mid vowels */e:/ and */o:/ respectively⁶². Furthermore, she views the quantity contrast among high vowels as a retention from PT rather than the result of simplification of original diphthongs as Li proposes.

Although Li’s and Sarawit’s views are both appealing, only the latter can account for the distributional gaps in PSWT. Elsewhere (2008b; to appear-b) I have shown that both length and height were contrastive in PSWT only for high vowels: the

⁵⁸ Reconstructed as *klɔŋ^A in Li’s system and as *klo:ŋ^A in the current reconstruction.

⁵⁹ Reconstructed as *mlet^D in Li’s system and as *mlec^D in the current reconstruction.

⁶⁰ Reconstructed as *xrok^D in Li’s system and as *krok^D in the current reconstruction.

⁶¹ I believe that the change from PT to Tai Ya is as follows: PT */e:/, */o:/ > PSWT */ɛ:/, */ɔ:/ > /je/, /wo/.

⁶² In contrast to Luo, she correctly posits this development only for long vowels.

distribution of height in mid and low vowels is predictable. Specifically, the PSWT mid vowels *e, *ɤ, and *o are always short and cannot occur in open syllables, but the low vowels *ɛ: and *ɔ: are always long and can occur in open syllables. This distribution is illustrated in Table 5-5.

Table 5-5 Distribution of PSWT non-high vowels

	open syllables		closed syllables	
	long	Short	long	short
mid				*ok
low	*ɛ:		*ɔ:k	

The distributional pattern in Table 5-5 argues against Li’s account and lends strong support to Sarawit’s proposal for PT. That is, Li cannot account for why the PSWT mid vowels *e, and *o did not occur in open syllables. In his scenario (Li 1977), PT *ɛ and *ɔ were long because they were subject to regular open-syllable lengthening. In the same way, PT *e and *o were not lengthened because they did not occur in open syllables. However, it is not clear why PT mid vowels *e and *o were banned from open syllables in the first place.

Sarawit (1973)’s system, in contrast, offers a very powerful explanation for these distributional gaps. Although she also believes that PT vowels in open syllables are always long, she holds that there was a quantity contrast in PT closed syllables. It is precisely this quantity contrast that explains why short *e and *o were not allowed in open syllables. That is, only long vowels were allowed in open syllables. This ban against short vowels in open syllables is exactly an effect of word minimality discussed in §2.2.2.1. Word minimality dictates that a prosodic word must be at least

one iambic foot. An open syllable with a short vowel is thus not allowed because it is light and cannot head an iambic foot.

In this scenario, PT long mid *e: and *o: lowered and became PSWT low vowels *ɛ: and *ɔ:, retaining their original length. Because PT vowels in open syllables were all long, all mid vowels in open syllable became low vowels in PSWT. In contrast, their short counterparts *e, and *o did not lower and were kept as short mid vowels. Because PT short vowel only occurred in closed syllables, there was no PT source for PSWT mid vowels in open syllables. The lack of PT *e and *o thus results in the observed gaps in PSWT. Therefore, these otherwise peculiar gaps are strong evidence for quantity contrast in PT. Dhananjayananda (1997)'s study illustrates how these gaps were filled in Siamese.

Not only does this system of contrast nicely account for the comparative data, but it is also consistent with the typology of the prosodic word in SEA. As shown in §2.4.3, sesquisyllabic languages impose a bimoraic minimum on the prosodic word so that all monosyllables and the major syllables of sesquisyllabic words are heavy. This is exactly the system reconstructed for PT. Furthermore, the proposed vowel inventory is a very likely one from a typological point of view. More specifically, the vowel system posited here is a triangular system, which has been shown to be typologically more common than square systems like that proposed by Li (Maddieson 1984b).

5.4 Solving Gedney's Puzzle

First discussed in detail by Gedney (1972b), the so-called "Gedney's Puzzle" refers to particular sets of vowel correspondences involving high vowels. A number of Tai etyma unequivocally go back to PT simple vowels *i, *ɯ, and *u in Li's system, which correspond to long *i:, *ɯ:, and *u: in the system proposed here. Another set of etyma clearly go back to PT rimes *aj, *aw, and *au. In addition to these two sets,

there are a number of items that seem to point to earlier *i:, *u:, and *u: in some languages but *aj, *aw, and *auj in others. The reverse patterns also exist. To illustrate the issue, first examine the two sets of data provided in Table 5-6 and Table 5-7.

Table 5-6 Items that unequivocally go back to PT */i:/, */u:/, and */u:/, i.e. the ‘year’ set

Gloss	Li	Siamese	Lungchow	Yay
‘year’ ‘four’	*i	pi: ^{A1} si: ^{B1}	pi: ^{A1} fi: ^{B1}	pi: ^{A1} θi: ^{B1}
‘day’ ‘book’	*u	mu: ^{C2} su: ^{A1}	mu: ^{C2} fu: ^{A1}	fu: ^{C2} θu: ^{A1}
‘pig’ ‘door’	*u	mu: ^{A1} tu: ^{A1}	mu: ^{A1} tu: ^{A1}	mu: ^{A1} tu: ^{A1}

In Table 5-6 above, Siamese, Lungchow, and Yay represent SWT, CT, and NT respectively. These three varieties all have /i:/, /u:/, and /u:/ for these correspondences. However, some varieties, e.g. Debao and Lungming, have diphthongized their original long high vowels, e.g. Lungming /pej^{A1}/ ‘year’, /sɤu^{A1}/ ‘book’, and /mow^{A1}/ ‘pig’. Having established that vowel quantity was contrastive in PT, it is clear that the data provided above points to original simple long high vowels. Li reconstructs *i (Li 1977: 262-263), *u (Li 1977: 264-265), and *u (Li 1977: 266-267) for this set of etyma. For convenience, this set of data may be referred to as the ‘year’ set.

Table 5-7 Items that unequivocally go back to PT */aj/, */au/, and */aw/, i.e. the ‘egg’ set

Gloss	Li	Siamese	Lungchow	Yay
‘egg’ ‘to cough’	*ɣj	k ^h aj ^{B1} ʔaj ^{A1}	k ^h aj ^{B1} ʔaj ^{A1}	raj ^{B1} ʔaj ^{A1}
‘leaf’ ‘clear’	*ɣu	baj ^{A1} saj ^{A1}	bau ^{A1} lau ^{A1}	bau ^{A1} θau ^{A1}
‘old’ ‘head louse’	*ɣw	kaw ^{B1} haw ^{A1}	kaw ^{B1} haw ^{A1}	kaw ^{B1} raw ^{A1}

In Table 5-7, all the three varieties unequivocally show reflexes of earlier *aj, *au, and *aw for the etyma above. In dialects where PT high vowels diphthongized, the contrast between the two sets of PT rimes is still preserved. For example, Debao has /ɣj/ for PT *i: but /aj/ for PT *aj, cf. /pɣj^{A1}/ from PT *pi:^A ‘year’ and /paj^{A1}/ from *paj^A ‘to go’. Note that in Li’s system the rimes of ‘to cough’ and ‘egg’ are reconstructed as *ɣj (Li 1977: 285-286), the rime of ‘leaf’ and ‘clear’ as *ɣu (1977: 288-289), and the rime of ‘old’ and ‘head louse’ as *ɣw (Li 1977: 290). This set of data may be referred to as the ‘egg’ set. In contrast to the two sets shown so far, for other items, modern dialects seem to give conflicting evidence on how the PT rimes for these words should be reconstructed.

Table 5-8 The ‘fire’ set of Gedney’s Puzzle correspondences

Gloss	Li	Siamese	Lungchow	Yay
‘fire’ ‘dry field’	*ɛj	faj ^{A2} raj ^{B2}	faj ^{A2} ɬaj ^{B2}	fi ^{A2} ri ^{B2}
‘breath’ ‘dry’	*eɰ/ɛɰ	caj ^{A1}	caɰ ^{A1} k ^h auɰ ^{B1}	su ^{A1}
‘empty’ ‘nine’	*jɤw	pla:w ^{B1} ka:w ^{C1}	pjaw ^{B1} kaw ^{C1}	pju ^{B1} ku ^{C1}

We may refer to the data in Table 5-8 as the ‘fire’ set. Siamese lacks ‘dry’ but White Tai has /xauɰ^{B1}/. Similarly, Yay does not have this form but Po-ai has /hu:^{B1}/. It is generally held that the data in this set pattern with the ‘egg’ set in SWT and CT, but pattern with the ‘year’ set in NT. That is, the CT and SWT reflexes for the etyma in this set seem to suggest complex rimes *aj, *auɰ, and *aw in PT but their NT counterparts seem to go back to simple rimes *i:, *u: and *u:. Li reconstructs *ɛj (Li 1977: 286) for ‘fire’ and ‘dry field’, *eɰ or *ɛɰ for ‘breath’ and ‘dry’ (Li 1977: 289), and *jɤw for ‘empty’ and ‘nine’ (Li 1977: 291). To complicate the situation, the reverse pattern also exists.

Table 5-9 The ‘long’ set of Gedney’s Puzzle correspondences

Gloss	Li	Siamese	Lungchow	Yay
‘long’ ‘excrement’ ⁶³	*ei	ri: ^{A264} k ^h i: ^{C1}	ɦi: ^{A2} k ^h ɣj ^{C1}	raj ^{A2} haj ^{C2}
‘crab’ ‘paternal grandfather’	*ew	pu: ^{A1} pu: ^{B1}	pu: ^{A1}	paw ^{A1}

The set of etyma in Table 5-9 may be referred to as the ‘long’ set. CT lacks the etymon ‘paternal grandfather’ altogether while some NT languages still retain it, cf. Po-ai /paw^{B1}/. The received generalization is that it patterns with the ‘year’ set in SWT and CT, but with the ‘egg’ set in NT. In other words, the CT and SWT reflexes of etyma belonging to this set point to earlier *i:, *u:, and *u: but their NT reflexes suggest *aj, *auj, and *aw. Li posits *ej for ‘long’, and ‘excrement’ (Li 1977: 286), and *ew for ‘crab’ and ‘paternal grandfather’ (Li 1977: 291). Notice that the pattern /u/~~/auj/, though expected, is not attested.

As Gedney argues (1972b), it is clear that this puzzle does not indicate doublets in PT, since no language is known to have inherited both the monophthongal and diphthongal forms for any given etymon. However, it remains unclear what these data indicate about the rime system of PT. Confronted with such puzzling sets of data, various researchers have proposed different accounts. Li (1977), followed by Liang and Zhang (1996) accounts for these “Gedney’s Puzzle” items by positing very fine contrasts among complex vowels as we have seen in our discussion of the data sets. Sarawit (1973) similarly posits complex vowels for these etyma but differs in the actual forms posited. In contrast, Haudricourt (1975), followed by Strecker (1983),

⁶³ This is a ‘voicing alternation’ item. See §3.4.

⁶⁴ This form means ‘oval’ in Thai

proposes that this is a kind of “umlaut” phenomenon. Specifically, the vowel of the first syllable in a disyllabic etymon affected the vowel of the second syllable differently in the process of monosyllabicization. Li’s, Sarawit’s, and Haudricourt’s solutions are compared in Table 5-10. Note that it is not clear what Haudricourt would posit for the ‘year’, ‘egg’, and ‘long’ sets.

Table 5-10 Comparison of solutions to “Gedney’s Puzzle”

	Li	Sarawit	Haudricourt
‘year’	*i	*i:	
‘egg’	*ɣj	*aj	
‘fire’	*ɛj	*ia:j	*CaCi
‘long’	*ej	*ɣj	

Although it is extremely likely that some kind of “umlaut”-like sound changes were operating at some point in the history of these etyma, it cannot be the case that they occurred at the PT stage or later. First of all, this issue is closely tied to the canonical shape of the PT prosodic word. As discussed in §2.3.2, it is unlikely that the PT prosodic word was disyllabic. Haudricourt’s view is thus ruled out by the sesquisyllabic theory that we have earlier argued for. Moreover, for the umlaut view to work, we must assume two separate one-step reductions of disyllables to monosyllables in SWT and CT on one hand, and in NT on the other. That is, *CaCi must have been reduced directly to *Cai in SWT/CT and to *Ci: in NT without any intermediate step. This is because such an intermediate step would have to involve the same sort of complex vowels that Li and Sarawit posit.

Another piece of evidence against Haudricourt’s view is that it falsely predicts only two types of languages. Languages pertaining to the first type are those that

merged the ‘long’ set with the ‘year’ set, and merged the ‘fire’ set with the ‘egg’ set. SWT and CT are generally said to belong to this type. The second type consists of those varieties that merged the ‘long’ set with the ‘egg’ set, and merged the ‘fire’ set with the ‘year’ set. NT is viewed as belonging to this type. In contrast, Li’s and Sarawit’s accounts allow at least eight types of languages as schematized in Table 5-11.

Table 5-11 Types of languages allowed by Li’s and Sarawit’s account

	Type 1	Type 2	Type 3	Type 4	Type 5	Type 6	Type 7	Type 8
‘year’								
‘fire’								
‘long’								
‘egg’								

Tai varieties of type 1 do not show any mergers and have four different reflexes for the four sets. Varieties belonging to type 2 merged the ‘year’ set with the ‘fire’ set, and the ‘egg’ set with the ‘long’ set. Yay represents this type of language. Dialects belonging to type 3 show two reflexes for the four sets. They merged the ‘year’ set with the ‘long’ set, and the ‘egg’ set with the ‘fire’ set. Siamese and Lungchow both belong to this type. Types 4 and 5 merged three of the four sets together. Type 4 has the same reflex for the ‘fire’, the ‘long’, and the ‘egg’ sets while type 5 merged the ‘year’, ‘fire’, and ‘long’. Similar to type 3, types 6 and 7 have three reflexes for the four sets. Type 6 merged the ‘year’, and the ‘fire’ sets, but kept the ‘egg’ and the ‘long’ sets apart. In contrast, type 7 has one reflex for the ‘long’ and the ‘egg’ sets, but keeps the ‘year’ and the ‘fire’ sets distinct. Lastly, languages belonging to type 8 show only one reflex for the four sets. This last type is not informative as it can be arrived at from different paths. In short, Li’s and Sarawit’s accounts allow for

patterns of mergers that differ from the two found in Siamese, Lungchow on one hand, and Yay on the other. Even though it was formerly believed that only languages belonging to type 2 and type 3 exist, languages belonging to other types are in fact attested in southern Guangxi as shown in Table 5-12.

Table 5-12 Tai varieties that do not belong to type 2 or type 3

	Shangsi	Chongzuo	Ningming	Wuming
‘year’	po:y ^{A1}	pɿj ^{A1}	pɿj ^{A1}	pi: ^{A1}
‘gall’	bo:y ^{A1}	bɿj ^{A1}	wɿj ^{A1}	bi: ^{A1}
‘fat’	po:y ^{A2}	pɿj ^{A2}	pɿj ^{A2}	pi: ^{A2}
‘fire’	fo:y ^{A2}	fɿj ^{A2}	fɿj ^{A2}	foj ^{A2}
‘dry field’	lo:y ^{C2}	lɿj ^{C2}	lɿj ^{C2}	roj ^{C2}
‘chicken louse’	lo:y ^{A2}	lɿj ^{A2}	laɿj ^{A2} -v	roj ^{A2}
‘long’	lo:y ^{A2}	lɿj ^{A2}	lɿj ^{A2}	raɿj ^{A2}
‘excrement’	k ^h o:y ^{C1}	k ^h ɿj ^{C1}	k ^h ɿj ^{C1}	xaj ^{C2}
‘egg’	laɿj ^{B1}	laɿj ^{B1}	laɿj ^{B1}	raɿj ^{B1}
‘tree’	faj ^{C2}	faj ^{C2}	faj ^{C2}	faj ^{C2}
‘intestine’	ɬaj ^{C1}	ɬaj ^{C1}	ɬaj ^{C1}	θaj ^{C1}
	type 5			type 7

Shangsi, Chongzuo, and Ningming are all type-5 languages because they merged the ‘year’, ‘fire’, and ‘long’ sets but kept ‘egg’ set apart. Wuming may belong to type 7. It seems to have merged the ‘long’ and the ‘egg’ sets but left the ‘year’ and ‘fire’ sets distinct. However, some etyma in the ‘year’ set show /oj/ rather than the expected /i/, e.g. ‘older sibling’ and ‘bladder’. Note that the Ningming form for ‘chicken louse’ unexpectedly shows /aj/. Based on the actual attestation of these types,

I adopt Li's and Sarawit's general position and propose that the four sets go back to four different rime categories as schematized in Figure 5-1.

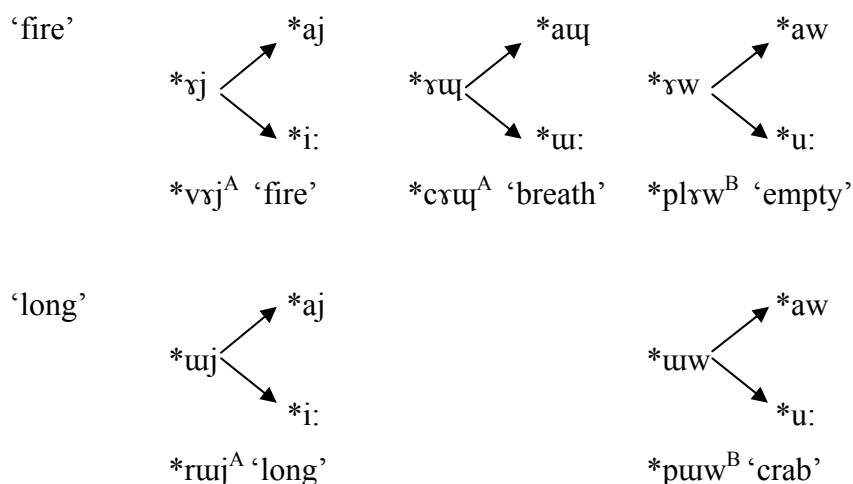


Figure 5-1 Proposed solution for “Gedney’s Puzzle”

In this proposal, each of the four sets of etyma goes back to a distinct rime category. The etyma in the ‘year’ set had either *i:, *u:, or *u: while members of the ‘egg’ set go back to PT *aj, *auɥ, *aw. Etyma that belong to the ‘fire’ set are reconstructed with either *ɤj, *ɤuɥ, and *ɤw. These three rimes merged with *i:, *u:, and *u: in NT languages as well as Shangsi, Ningming, and Chongzuo but merged with *aj, *auɥ, and *aw in SWT, most CT languages, and Wuming. In contrast, etyma belonging to the ‘long’ set are reconstructed with *uj or *uw. These two rimes merged with *i: and *u: respectively in SWT and CT, including the three type-5 dialects but merged with *aj, and *aw in NT varieties including Wuming.

In addition, this proposal neatly accounts for the fact that etyma showing *u: in SWT/CT but *auɥ in NT are not attested. If they existed, these items would belong to the ‘long’ set. According to the proposal, they would have to be reconstructed with

PT *uuɿ, which would not be distinguishable from PT *uɿ:. We see, then, not only that this proposal accounts for “Gedney’s Puzzle” by positing relatively simple rime categories, it also correctly predicts a gap in the data.

Often lumped together with the “Gedney’s Puzzle” items are some etyma that show irregular reflexes. Treating those items as belonging to the “Gedney’s Puzzle” paradigm gives an impression that modern dialects freely choose between the monophthongal or diphthongal forms of the so-called “Gedney’s Puzzle” items (Gedney 1989b: 60-61). In fact, these aberrant cases should be treated separately as exceptions, as they are cases of function words or words that occur with high frequency. It is largely unpredictable which vocalic variant a given dialect has, as illustrated in Table 5-13.

Table 5-13 Etyma with irregular reflexes usually discussed under “Gedney’s Puzzle”

	PT	Siamese	Tai Nüa	Lungchow	Shangsi	Wuming
‘1s pron.’	*ku: ^A /*kaw ^A	ku: ^{A1}	kaw ^{A1}	kaw ^{A1}	kaw ^{A1}	kaw ^{A1}
‘2s pron.’	*muɿ ^A /*mauɿ ^A	muɿ ^{A2}	mauɿ ^{A2}	muɿ ^{A2}	maɿ ^{A2}	muɿ ^{A2}
‘good’	*dɿj ^A	di: ^{A1} -v	di: ^{A1} -v	da ^j ^{A1}	do:y ^{A1}	da ^j ^{A1}
‘to give’	*hauɿ ^C	hauɿ ^{C1}	huɿ: ^{C1} -v	huɿ: ^{C1} -v	do:y ^{A1}	hauɿ ^{C1}
‘to go’	*paj ^A	pa ^j ^{A1}	pa ^j ^{A1}	pa ^j ^{A1}	po:y ^{A1} -v	pa ^j ^{A1}

The first-person singular pronoun seems to have been two separate forms in PT, *ku:^A and *kaw^A. Similarly, its second-person counterpart seem to have also been two separate forms *muɿ^A, and *mauɿ^{A65} (Refer to Strecker (1984) for a detailed

⁶⁵ The distinction between *ku:^A and *muɿ^A, on one hand, and *kaw^A and *mauɿ^A on the other, may have been prosodically conditioned allomorphy. In particular, *ku:^A

study of PT personal pronouns). The etymon ‘good’ is generally rendered in modern Tai dialects with the reflex of PT *ɽj, except in SWT dialects, represented here by Siamese and Tai Nüa, where it shows the reflex of *uj. Similarly, the verb ‘to give’ seems to go back to PT *au, but shows reflexes of *u: in the CT dialects of Ningming, Longzhou, and Daxin, as well as some SWT dialects, e.g. Tai Yuan. Lastly, the verb ‘to go’ clearly had *aj in PT but unexpectedly has the reflex *i: in Shangsi, and Chongzuo. These unexpected cases must be treated as irregular and must be kept separate from the more regular “Gedney’s Puzzle” sets. Also note that the etyma in this set are either function word, or lexical items with a high frequency of occurrence.

5.5 *Enriching the codas*

The conventional reconstruction of PT holds that only eight consonants can function as codas: *-p, *-t, *-k, *-m, *-n, *-ŋ, *-j, and *-w (Li 1977: 58-59). This characterization is consistent with the syllable typology of modern Tai languages. However, the discovery of the displaced NT language Saek (Gedney 1989c; Haudricourt 1958; 1960; 1963a) has opened up the possibility that the coda system of PT may have been richer than we originally assumed. In this section, I argue that, in addition to final *-l preserved in Saek, final *-c and possibly *-ɲ must also have existed in PT. Moreover, I show that these three consonants play a vital role in the development of the vowels.

One of the intriguing characteristics of Saek is its final /-l/, unattested elsewhere in Tai. Li (1977: 58-59) does not reconstruct final *-l for PT but notes that some etyma that he reconstructs with *-n may in fact got back to *-l. The reason for

and *muŋ^A were possibly the unstressed forms while *kaw^A and *mau^A were stressed.

his reservation is lack of Saek data. The situation changed dramatically after Gedney's extensive study of this language (Hudak 1993). By examining the distribution of /-l/ in Saek, he argues convincingly that Saek /-l/ is a case of retention (Gedney 1993). First of all, he points out that most Saek words with final /-l/ are native Tai words. It is impossible to find a possible conditioning environment that changes final *-n to /-l/ in a great number of words but left it unaffected in many others. More importantly, he shows that final /-l/ in Saek is amply attested in rimes directly inherited from PT but is absent from rimes that were introduced later. Therefore, it is safe to conclude that PT had rimes ending in *-l.

Comparison between Saek and Laha, a Kra language, also supports Gedney's view that PT had final */-l/ (Ostapirat 1995). Many forms found in both Saek and Laha agree in having /-l/ as shown in Table 5-14. In addition to the cognates cited by Ostapirat (1995), I also cite 'body louse' as another Kra-Dai etymon with final /-l/. The Laha and Proto-Kra forms are taken from Ostapirat (2000). The etymon 'grandchild' is problematic because the Saek form has /-n/ instead of the expected /-l/. The final *-l may have become /-n/ in Saek due to dissimilation from the onset /l-/.

Table 5-14 Shared etyma with final /-l/ in Saek and Kra

	Saek	Laha	Proto-Kra
'husked rice'	sa:l ^{A1}	sa:l ^{A1}	*sal ^A
'shake'	sɛl ^{B1}	sal ^{B1}	*səl ^B
'body louse'	mlel ^{A2}	mdal ^{A2}	*m-drəl ^A
'grandchild'	la:n ^{A1}	kla:l ^{A1}	*klal ^A
'to snore'	trel ^{A1}	kal ^{A1}	

Although Saek is the only Tai language that has final /-l/, traces of PT*-l can also be found in the vowel correspondences among its sister languages. Most revealing are the data sets given in Table 5-15. The PT mid front vowel /e/ stays front in SWT and CT e.g. *cet^D > Siamese /cet^{D1}/ and Lungchow /cit^{D1}/. In contrast, it normally became *a in NT varieties, e.g. PT *cet^D > Yay /sat^{D1}/ and Saek /cɛt^{D1}/. Except for a few cases, Saek has changed earlier short *a to /ɛ/. The unexpected /-n/ in ‘grandchild’ is perhaps result of dissimilation or dialect mixing.

Table 5-15 Etyma with PT */en/ and */-el/

	Siamese	Lungchow	Qinzhou	Yay	Saek	PT
‘seven’	cet ^{D1}	cit ^{D1}	cit ^{D1}	sat ^{D1}	cet ^{D1}	*cet ^D
‘mushroom’	het ^{D1}	vit ^{D2} -i	lyt ^{D1}	rat ^{D1}	(het ^{D1})	* ^h rwet ^D
‘great-grandchild’ ⁶⁶	le:n ^{A1} -t		lin ^{C1}		(le:n ^{A1})	* ^h len ^C
‘body louse’	len ^{A2}	min ^{A2}	man ^{A2}	nan ^{A2}	mlel ^A	*mlel ^A
‘to become’ ⁶⁷	pen ^{A1} -i	pin ^{A1} -i	pan ^{A2}	pan ^{A2}	phal ^{A2} -v	*bel ^A
‘civet cat’	hen ^{A1}	hin ^{A1}		ɲan ^{A1}	ɲel ^{A1}	* ^h ɲel ^A

In Qinzhou, however, there is a split between ‘seven’, ‘mushroom’, and ‘great-grandchild’ on the one hand, and ‘body louse’, ‘to become’, and ‘civet cat’ on the other. The vowels in the first set of etyma have been raised but still retain their front quality. The /y/ in ‘mushroom’ results from a coalescence of *-w- and *i. In contrast, the vowels in the second set have not been raised but were retracted and lowered to /a/. As all these items have final /-l/ in Saek, the backing of PT *e in these words should

⁶⁶ The original tone for this etymon is most likely *C but it became *A in SWT due to contamination from *^hla:n^A ‘grandchild’. The vowel in the Siamese form has been secondarily lengthened.

⁶⁷ This form is found in /pan^{A2} jeŋ^{B2}/ ‘how’.

be considered to have been conditioned by PT *-l. Note the contrast between ‘great-grandchild’ and ‘body louse’. Although they are both reflected with /-n/ in Qinzhou, the vowels are different. Therefore, these two sets of data show that final /-l/ caused PT *e to split into /i/ and /a/ in Qinzhou. Also note that the Saek forms for ‘mushroom’ and ‘great-grandchild’ are borrowings from Siamese or Lao.

In addition to final *-l, evidence from Saek as well as from non-Tai languages points to the existence of final *-c in PT. Even though final palatals are common in SEA, especially among Mon-Khmer languages, they are not attested in any modern Tai language. However, Gedney (1989c: 961) notes that some of the Saek etyma that show a change from *at to /ɛt/ sometimes show a further change to /ɛk/. These etyma are shown in Table 5-16. I propose that etyma belonging to this set in fact go back to final *-c.

Table 5-16 Etyma with final */-c/

	Siamese	Lungchow	Yay	Saek	PT
‘fish scale’	klet ^{D1}	kit ^{D1}	cip ^{D1} -f ⁶⁸	tɛk ^{D1}	*klec ^{D1}
‘ant’	mot ^{D2}	mɔt ^{D2}	mat ^{D2}	mɛk ^{D2}	*mɔc ^{D2}
‘grain’	met ^{D2}		nat ^{D2}	mɛk ^{D2}	*mlec ^{D2}
‘to cut’	tat ^{D1}	tat ^{D1}	tat ^{D1}	tɛk ^{D1}	*tac ^{D1}
‘to dust off’	pat ^{D1}	pat ^{D1}	pat ^{D1}	pɛk ^{D1}	*pac ^{D1}

One may argue that the final /-k/ in Saek is an innovation as it is preceded in all these examples by a front vowel /ɛ/. However, this cannot be the case for two

⁶⁸ There exist many cases of sporadic /-p/ instead of the expected /-t/, e.g. Southern Shan /hep^{DS1}/ and Ningming /lap^{DS1}/ from *hrwit^D ‘mushroom’, and Lungming /kip/, Western Nung /cip/, and Wuming /klip/ from *klec^D ‘fish scales’

reasons. First, the front quality of the vowel in Saek cannot be said to have been the condition for a change from *-t to /-k/ because many words with the rime /ɛt/ do exist, e.g. cɛt^{D1} ‘seven’ and /rɛt^{D1}/ ‘fart’. Remember that Saek /ɛ/ before coronal codas corresponds regularly to /a/ in other NT varieties. Second, the Siamese and Lungchow forms show clearly that this mysterious coda consonant was preceded by many different vowels including *ɤ, *e, and *a. These vowels do not form a natural class that could potentially condition a change from *-t to /-k/. Therefore, the “aberrant” final /-k/ in Saek should be considered evidence for PT *-c.

Evidence from outside of Tai also supports the hypothesis that PT had final *-c. First of all, Be, one of Tai’s closest kin, shows unique reflexes in items given in Table 5-17. PT *-t corresponds regularly to Be /-t/ as in ‘fart’, and PT *-k regularly to Be /-k/, as ‘bird’. However, the etyma that show the /-t/~/-k/ correspondence in Tai have final /-ʔ/ in Be. The fact that Be shows this three-way contrast supports the hypothesis that PT had a distinct final *-c. This correspondence set has been noted by Ostapirat (2005: 119), who believes that PT *-c was inherited from Proto-Kra-Dai.

Since it has been established that PT allows palatal consonants in the coda, one would also expect to find the palatal nasal occurring in coda position. Although the reconstruction of PT *-c is unequivocal, there is rather little evidence for final *-ɲ. The only potential case I have identified so far is ‘to eat’, which is reflected as /kin^{A1}/ in all SWT varieties but as /kɯn^{A1}/ in NT dialects like Wuming and Yay. We can speculate that the PT form for ‘to eat’ was *kɯɲ^A but the vowel was fronted so that the PSWT form for this etyma was *kin^A. Therefore, I tentatively hypothesize that PT had both *-c and *-ɲ.

Table 5-17 Final *-c in Tai and Be

	Siamese	Saek	Be
‘fart’	tot ^{D1}	ret ^{D1}	dfut ^{D1}
‘flea’	mat ^{D1}	mat ^{D1}	mat ^{D1}
‘bedbug’	ruət ^{D2}	ruət ^{D2}	liət ^{D2}
‘bird’	nok ^{D2}	nək ^{D2}	nok ^{D2}
‘deep’	luuk ^{D2}	lak ^{D2}	lak ^{D2}
‘mucus’	mu:k ^{D2}	mu:k ^{D2}	muk ^{D2}
‘fish scale’	klet ^{D1}	tɬək ^{D1}	liʔ ^{D1}
‘ant’	mot ^{D2}	mɛk ^{D2}	muʔ ^{D2}
‘grain’	met ^{D2}	mɬək ^{D2}	moʔ ^{D2}
‘to cut’	tat ^{D1}	tɛk ^{D1}	
‘to dust off’	pat ^{D1}	pɛk ^{D1}	ɬaʔ ^{D1}

5.6 *Reconstructions of PT rimes*

Etyma from Siamese (SI), Lungchow (LC), and Yay (Y), are provided in each section to represent the three clusters of the modern varieties. Remember that this is not to say that they represent three genealogical subgroups of Tai (See §1.2 and §7.3 for discussion).

5.6.1 *Rimes with high vowels*

The PT high vowels *i, *i:, *u, *u:, *u, and *u: are relatively stable, especially when compared to the mid vowels. Except in a few environments, these high vowels normally remain high in all languages. There seems to have been no restriction on the codas that particular vowels could co-occur with. However, short vowels were not allowed in open syllables due to the minimality requirement discussed in §2.2.2.1.

5.6.1.1 Rimes with *i and *i:

PT high vowels *i and *i: are usually kept as high vowels in modern dialects. In open rimes, the long vowel *i: retains its original quality in most dialects, including Siamese, Lao, Black Tai, Yay, and Wuming. However, in some CT and NT dialects such as Lungming, Debao, Yongnan, Shangsi, Hengxian, and Longsheng, PT *i: became diphthongized in open syllables. An interesting case of a diphthongized reflex is /o:y/ in Shangsi, which is quite different from /ɤj/, /ej/, /uj/, or /uj/ in other dialects. PT *i:# corresponds to *i# in Li's reconstruction (1977: 262-263). Examples of PT *i:# are given in Table 3-50.

Table 5-18 Etyma with PT *i:#

Gloss	PT	SI	LC	Y	Others
‘soot’	* ^h mi: ^C	mi: ^{C1}	mi: ^{C1}	mi: ^{C1}	Lungming /mej ^{C1} /
‘four’	*si: ^B	si: ^{B1}	ɬi: ^{B1}	θi: ^{C1}	Lungming /sej ^{C1} /
‘year’	*pi: ^A	pi: ^{A1}	pi: ^{A1}	pi: ^{A1}	Lungming /pej ^{A1} /
‘to have’	*mi: ^A	mi: ^{A2}	mi: ^{A2}	mi: ^{A2}	Lungming /mej ^{A2} /

In closed syllables, both *i and *i: can occur. In some dialects such as Siamese, Lao, and Saek, the original quantity contrast is retained. In some other dialects, the contrast has been lost but the tonal reflexes on checked syllables provide clues to the original length of the etyma. For example, Southern Shan only has a length contrast for /a/ and /a:/ but shows different tonal reflexes for checked syllables with PT *i: and *i, e.g. /lip⁵⁵/ from *dip^D ‘raw’ vs. /nip²¹/ *hni:p^D ‘to pinch’ (see §6.5.1 for discussion). Li reconstructs *iC (Li 1977: 261-262) and *iəC (Li 1977: 263-264) for

*iC and *i:C in the current reconstruction, respectively. Table 5-19 provides examples of PT *iC and *i:C.

Table 5-19 Etyma with PT *iC and *i:C

Gloss	PT	SI	LC	Y	Others
‘to taste’	*jim ^A	ch ^h im ^{A2}	cim ^{A2}	sim ^{A2}	
‘short’	*tin ^B		tin ^{B1}	tin ^{B1}	Black Tai /tin ^{B1} /
‘pangolin’	*lil ^B		lin ^{B1}	lin ^{B1}	Black Tai /lin ^{B1} /
‘raw’	*C̣.dip ^D	dip ^{DS1}	dip ^{DS1}	dip ^{DS1}	
‘foot’	*ti:n ^A	ti:n ^{A1}		tin ^{A1}	Lungming /tin ^{A1} /
‘to climb’	*pi:n ^A	pi:n ^{A1}		pin ^{A1}	Lungming /pin ^{A1} /
‘knife’	*mi:t ^D	mi:t ^{DL2}	mit ^{DL2}	mit ^{DS2}	
‘to pinch’	*ni:p ^D	ni:p ^{DL1}	nip ^{DL1}	nip ^{DS1}	

Siamese normally preserves the original quantity contrast. In other words, PT *i is reflected as /i/ while PT *i: is reflected as /i:/ in Siamese. However, it has undergone a secondary shortening. In particular, PT *i: became shortened in Siamese words with the rising tone⁶⁹, e.g. /hin^{A1}/ ‘stone’ rather than */hi:n^{A1}/, and /jin^{A1}/ rather than */ji:n^{A1}/.

This is in sharp contrast with Lungchow and Yay, which both lost the distinction between the two high vowels without trace. Note that *i: in checked syllables regularly shortened in Yay as well as most other NT dialects, e.g. *mi:t^D > /mit^{DS2}/ instead of /mit^{DL2}/.

⁶⁹ This tone goes back to PT *A on syllables with PT fricatives, and voiceless sonorants in word-initial position.

5.6.1.2 Rimes with *u and *u:

The high back unrounded vowels *u and *u: are the least stable among the PT high vowels. Like their front counterparts, only long *u: was allowed in open rimes. They are preserved as back unrounded vowels in most dialects including Siamese, Black Tai, Lue, Lungchow, Western Nung, Daxin, Yay, Tianlin, Du'an, Saek, etc. Similarly to *i:#, this PT back vowel shows a diphthongal reflex /ɤu/ in such dialects as Lungming, Debao, Hengxian, and Longsheng. However, many dialects show special developments. For example, Qinzhou, Yongnan, Fusui, and Shangsi merged PT *u:# with *i:# and now have /uj/, /ej/, /uj/ and /o:y/ respectively as reflex for both. PT *u:# corresponds to *u# in Li's reconstruction (1977: 262-263). Examples of PT *u:# are given in Table 5-20.

Table 5-20 Etyma with PT *u:#

Gloss	PT	SI	LC	Y	Others
‘hand’ ⁷⁰	*mwu: ^A	mu: ^{A2}	mu: ^{A2}	fuŋ ^{A2}	
‘day’	*mwu: ^C	mu: ^{C2}	mu: ^{C2}	fu ^{C2}	
‘day after tomorrow’ ⁷¹	*ɕ.ru:n ^A	ru:n ^{A2}	lu: ^{A2}	ru:n ^{A2}	
‘to buy’	*z.ju: ^C	su: ^{C2}	tu: ^{C2}	su ^{C2}	
‘writing, book’	*su: ^A	su: ^{A1}	tu: ^{A1}	θu ^{A1}	

In closed rimes, both the short *u and the long *u: were allowed. Similarly to their front counterparts, the original quantity contrast is retained only in some varieties. However, unlike the other high vowels, the quality of PT *u and *u: have been

⁷⁰ The final /-ŋ/ in Yay and other NT dialects for this etymon is aberrant.

⁷¹ Refer to footnote 53 for discussion on the aberrant /-n/ in the Siamese form.

affected by the codas. For example, they are lowered to /ɣ/ in many dialects including Lungchow, Leiping, and Lungchow. This development is seen most clearly when the rimes had an alveolar coda. PT *u:n and *u:l are reflected as /ɣn/ in Lungchow. For these two sets of etyma, Li (1977: 265) reconstructs *uən. In Siamese and other SWT languages, short *un lowered to *ɣn (> PSWT *on) but long *u:n did not. Li reconstructs (Li 1977: 274) *wun for this PT short rime⁷².

Table 5-21 Etyma with PT *u:n, *u:l respectively

Gloss	PT	SI	LC	Y	Others
‘sky’	*ɖun ^A	bon ^{A1}		bun ^{A1}	Po-ai /mun ^{A1} /
‘rain’	*q.wun ^A	fon ^{A1}	p ^h ɣn ^{A1}	hun ^{A1}	Po-ai /hun ^{A1} /
‘human, person’	*gwun ^A	k ^h on ^{A2}	kɣn ^{A2}	hun ^{A2}	Po-ai /hun ^{A2} /
‘night’	*ɣu:n ^A	k ^h u:n ^{A2}	kɣn ^{A2}	hu:n ^{A2}	
‘other’	*ʔu:n ^B	ʔu:n ^{B1}	ʔɣn ^{B1}		Wuming /ʔun ^{B1} /
‘firewood’	*wu:l ^A	fu:n ^{A2}	fɣn ^{A2}	fun ^{A2}	
‘to slide’	*mlu:l ^B	lu:n ^{B2}			Saek /mlu:l ^{B2} /

This development of rimes with alveolar codas above is in contrast to rimes with final labial consonants. PT *um and *u:m became *um and *u:m in many dialects including Longzhou, Qinzhou, Guangnan Nung, Yay and Wuming. The rounding of *um and *u:m must have been due to the influence of *-m. Li

⁷² Coincidentally, both ‘human’ and ‘rain’ had medial *-w- in PT. However, this medial *-w- is not the same as the *w in Li’s *wun. The etymon ‘sky’ clearly illustrates this fact as it did not have *-w- in PT, i.e. PT *ɖun^A > ɖɣn^A > PSWT ʔbon^A > Siamese /bon^{A1}/. The rounding in the Siamese forms was not caused by the *-w-. Rather, it is a result of a regular change from *ɣ to PSWT *o, cf. PT *sɣn^C ‘heal’ > Siamese /son^{C1}/.

reconstructs *uum (1977: 265) and *wom (1977: 273-274) for these two PT rimes. The development of *uŋ and *u:ŋ is in turn quite different from that of *um and *u:m. In contrast, PT *u and *u: before velar codas are usually lowered to /ɤ/ or /a/ in SWT, CT, and NT dialects. Both *uŋ and *u:ŋ are reflected as /uŋ/ in Siamese, but as /aŋ/ and /u:ŋ/ in Ningming. Since the only etymon with PT *u:ŋ is not found in NT, it is not clear what its reflex is. Li (1977: 264) reconstructs *uŋ for the short rime but did not give a reconstruction for its long counterpart. Etyma in Table 5-22 show the development of PT *um, *u:m, *uŋ, and *u:ŋ. These rimes are illustrated by ‘wind’, ‘to forget’, ‘to steam’, and ‘one’ respectively.

Table 5-22 Etyma with PT *um, *u:m, *uŋ, and *u:ŋ respectively

Gloss	PT	SI	LC	Y	Others
‘wind’	*Ç.lum ^A	lom ^{A2}	lum ^{A2}	rum ^{A2}	
‘to forget’	*lu:m ^A	lu:m ^{A2}	lum ^{A2}	lum ^{A2}	
‘to steam’	*hnuŋ ^C	nuŋ ^{C1}		naŋ ^{C1}	Ningming /naŋ ^{C1} /
‘one’	*hnu:ŋ ^B	nuŋ ^{B1}	nɤŋ ^{B2} -t		Ningming /nu:ŋ ^{B1} /

In addition to the effect of places of articulation discussed above, manners of articulation of the coda also came into play in the development of the high back unrounded vowel. For example, *ut became *ut in Po-ai, in contrast to *un above which remains /un/. Another example is the development of PT *u: before the velar stop *-k. This rime became /u:k/ in Siamese, and /uk/ in Lungchow, but remains unrounded in Po-ai and Yay /uk/. Notice the regular shortening of high vowel *u: preceding stop codas in NT, parallel to *i: and *u: (see §5.6.1.1 and §5.6.1.3). The correspondents of PT *u:t, and *u:k in Li’s system are *uut (1977: 265), and *uuk

(1977: 268) respectively. Notice the contrast between *uu and *u in Li's system. In Table 5-23, 'distance from thumb to fingertip' represents *u:p, 'insipid' represents *u:t, and 'child' and 'to come into contact' represent *u:k. Also, note that *u:p rounded to /up/ in Yay because of the labial coda, in parallel to *um and *u:m above.

Table 5-23 Etyma with PT *u:p, *u:t, and *u:k respectively

Gloss	PT	SI	LC	Y	Others
'distance from thumb to fingertip'	*y <u>u</u> :p ^D	k ^h <u>u</u> :p ^{DL2}		hup ^{DS2}	
'insipid'	*c <u>u</u> :t ^D	c <u>u</u> :t ^{DL1}			Po-ai /ɛut ^{DS1} /
'child'	*l <u>u</u> :k ^D	l <u>u</u> :k ^{DL2}	luk ^{DL2}	luuk ^{DS2}	
'to come in contact'	*ç. <u>tu</u> :k ^D	t ^h <u>u</u> :k ^{DL1}	t ^h uk ^{DS1}	tuuk ^{DS2}	

The coda is not the only component of the syllable that can affect the development of the vowel; the medial glide could also play a crucial role. One very important change of PT *u: that is conditioned by the medial *-w-. In particular, PT *u:# fused with the medial *-w- that preceded it, resulting in special developments as shown in Table 5-24. The sequence *-wu: became simplified to *u: in CT and SWT but became a diphthong *uə in most NT dialects. The NT dialect of Lingyue seems to be the only dialect that still retains the medial *-w-, cf. /lwa^{A2}/ 'ear' and /ŋwa^{A2}/ 'snake'. Li (1977: 268) posits *u# for this set of data.

Table 5-24 Etyma with PT *wu:#

Gloss	PT	SI	LC	Yay	Others
‘ear’	*k.rwu: ^A	hu: ^{A1}	hu: ^{A1}	rwə ^{A2} -i	Lingyue /lwa ^{A2} /
‘snake’	*ŋwu: ^A	ŋu: ^{A2}	ŋu: ^{A2}	ŋwə ^{A2}	Lingyue /ŋwa ^{A2} /

Also refer to §5.4 for the reconstruction of PT *uw and *uj.

5.6.1.3 Rimes with *u and *u:

PT high vowels *u and *u: are usually kept as high vowels in modern dialects. In open rimes, the long vowel *u: retains its original quality in most dialects, including Siamese, Lao, Black Tai, Yay, and Wuming. However, some CT and NT dialects such as Lungming, Debao, Yongnan, Shangsi, Hengxian, and Longsheng, PT *u: became diphthongized to /ɤw/, /ow/, or /aw/ in open syllables. PT *u:# corresponds to *u# in Li’s reconstruction (Li 1977: 266-267). Examples of PT *u:# are given in Table 5-25.

Table 5-25 Etyma with PT *u:#

Gloss	PT	SI	LC	Y	Others
‘pig’	* ^h mu: ^A	mu: ^{A1}	mu: ^{A1}	mu: ^{A1}	Lungming /mow ^{A1} /
‘rat’	* ^h nu: ^A	nu: ^{A1}	nu: ^{A1}	nu: ^{A1}	Lungming /now ^{A1} /
‘to be at’	* [?] ju: ^B	ju: ^{B1}	ju: ^{B1}	ju: ^{B1}	Lungming /jow ^{B1} /
‘door’	*tu: ^A	tu: ^{A1}	tu: ^{A1}	tu: ^{A1}	Lungming /tow ^{A1} /
‘pair’	*gu: ^B	k ^h u: ^{B2}	ku: ^{B2}	ku: ^{B2}	Lungming /kow ^{B2} /

In closed syllables, both *u and *u: can occur. In some dialects such as Siamese, Lao, and Saek, the original quantity contrast is retained. In some other dialects, the contrast has been lost but the tonal reflexes on checked syllables provide a clue to the original length of the etyma. For example, Southern Shan only has a length contrast for /a/ and /a:/ but shows different tonal reflexes for checked syllables with PT *u and *u:, e.g. /tup⁵⁵/ from *dup^D ‘to beat’ vs. /lup²¹/ *Ç.lu:p^D ‘to stroke’ (see §5.3 for discussion). Li reconstructs *uC (Li 1977: 266) and *uəC (Li 1977: 267) for *uC and *u:C in the current reconstruction respectively. Table 5-26 provides examples of PT *uC and *u:C.

Table 5-26 Etyma with PT *uC and *u:C respectively

Gloss	PT	SI	LC	Y	Others
‘to cook’	* ^h ruŋ ^A	huŋ ^{A1}	huŋ ^{A1}	ruŋ ^{A1}	
‘to be entangled’	* ^h ɲuŋ ^B	juŋ ^{B2}	juŋ ^{B2}	ɲuŋ ^{B2}	
‘puppet’	*hun ^B	hun ^{B1}		hun ^{B1}	W. Nung /hun ^{B1} /
‘to beat’	*dup ^D	t ^h up ^{DS2}		tup ^{DS2}	W. Nung /tup ^{DS2} /
‘to carry in the arm’	*ʔu:m ^C	ʔum ^{C1}	ʔum ^{C1}	ʔum ^{C1}	
‘to lead’	*cu:ŋ ^A	cu:ŋ ^{A1}	cu:ŋ ^{A1}		
‘mucus’	*mu:k ^D	mu:k ^{DL2}	mu:k ^{DL2}	muk ^{DS2}	
‘to stroke’	*Ç.lu:p ^D	lu:p ^{DL2}		rup ^{DS2}	W. Nung /lup ^{DS2} /

Siamese normally preserves the original quantity contrast. PT *u is reflected as /u/ while PT *u: is reflected as /u:/. However, the Siamese form for ‘to carry in the arm’ underwent a secondary shortening, possibly conditioned by tone. Notice the regular shortening of *u: to /u/ before stop codas in NT, illustrated by Yay /muk^{DS2}/

‘mucus’ and /rup^{DS2}/ ‘to stroke’. This shortening also occurred with long *i:C (see §5.6.1.1) and *u:C (see §5.6.1.2). Note that *uk went through special developments in SWT and NT as illustrated by the data in Table 5-27.

Table 5-27 Etyma with PT *uk

Gloss	PT	SI	LC	Y	Others
‘bone’	*Ç.duk ^D	du:k ^{DL1}	duk ^{DL1}	dok ^{DL1}	Lungming /nok ^{DS1} /
‘loom’	*truk ^D	hu:k ^{DL1}	huk ^{DL1}	rok ^{DL1}	Lungming /lok ^{DS1} /
‘tomorrow’ ⁷³	*ʃm.ruk ^D	p ^h ru:ŋ ^{DL2}	pjuk ^{DL2}	sok ^{DL2}	Lungming /cok ^{DL2} /

Corresponding to *uok in Li’s system (1977: 267-268), PT *uk was lengthened to *u:k in all SWT and many CT dialects including Lungchow, Leiping, Ningming, etc. In all NT varieties including Yay, *uk became lengthened and lowered to *o:k. The modern forms in Siamese, Lungchow, and Yay all show DL tones, indicating an earlier long vowel. There are some CT dialects that retained the original short vowel. For example, Lungming and Western Nung have DS tones in these etyma, pointing to an earlier short vowel.

5.6.2 Rimes with mid vowels

PT had altogether three pairs of short and long mid vowels: *e, *e:, *ɤ, *ɤ:, *o, and *o:. Like their high vowel counterparts, short vowels were not allowed in open syllables due to the minimality requirement discussed in §2.2.2.1. However, unlike PT

⁷³ The Siamese form shows /-ŋ/ instead of the expected /-k/ due to assimilation to /ni:^{C2}/, cf. expression /p^hru:ŋ^{DL2} ni:^{C2}/ ‘tomorrow’ in present-day Siamese.

high vowels, the long *e: and *o: regularly lowered to *ɛ: and *ɔ: in a great number of SWT and NT dialects.

5.6.2.1 Rimes with *e and *e:

PT had two mid front vowels *e and *e:, which contrasted with each other in quantity. PT *e:# was lowered to *ɛ:# in SWT and some CT dialects. Examples of languages that underwent the lowering include Siamese, Lao, Southern Shan, Black Tai, Lue, Sapa, Bao Yen, Cao Bang, and the NT language Saek. The PT *e: is generally preserved as a mid vowel in the remaining CT and NT dialects. Like their vowel counterparts, only long *e: can occur in open rimes due to the minimality constraint on the prosodic word as discussed in §2.2.2.1. PT *e:# corresponds to *ɛ# in Li's reconstruction (1977: 273-274). Examples of PT *e:# are given in Table 5-28.

Table 5-28 Etyma with PT *e:#

Gloss	PT	SI	LC	Y	Others
‘old (of living beings)’	*ke: ^B	kɛ: ^{B1}	ke: ^{B1}	ce ^{B1}	Wuming /ke: ^{C1} /
‘to untie’	*ke: ^C	kɛ: ^{C1}	ke: ^{C1}		
‘mother’	*me: ^B	mɛ: ^{B2}	me: ^{B2}	me ^{B2}	
‘raft’	*be: ^A	p ^h ɛ: ^{A2}	pe: ^{A2}	pe ^{A2}	
‘fish net’	*kre: ^A	hɛ: ^{A1}	he: ^{A1}	re ^{A1}	

In closed syllables, both *e and *e: can occur. In SWT dialects including Siamese, Lao, Black Tai, etc., PT *e in closed syllables is generally retained as a mid vowel. However, it is raised to /i/ before nasal codas in some dialects such as Lue and White Tai. Among CT varieties, it is raised to /i/ in most dialects including

Lungchow, Lungming, Leiping, Shangsi, and Qinzhou. It is, however, preserved as /e/ in a few varieties including Daxin and Guangnan Nung. In all NT dialects including Yay, Saek, and Wuning, PT *eC is regularly raised to /iC/. Li (1977: 269-270) reconstructs *je for this set of data. Table 5-29 illustrates the development of PT *eC.

Table 5-29 Etyma with PT *eC

Gloss	PT	SI	LC	Y	Others
‘full’	*k.tem ^A	tem ^{A1}	tim ^{A1}	rim ^{A1}	
‘needle’	*qem ^A	k ^h em ^{A1}	k ^h im ^{A1}	cim ^{A1}	
‘child (young person)’	*døk ^D	dek ^{DS1}	dik ^{DS1}		
‘to pick up’	*kep ^D	kep ^{DS1}	kip ^{DS1}	cip ^{DS1}	
‘fingernail’	*Ç.lep ^D	lep ^{DS2}	lip ^{DS2}	rit ^{DS2} -f	

However, there is a special development shared by most NT languages including Saek, Yay, Po-ai, Yongbei, Hechi, Rong’an, Guigang, etc. In these varieties, PT *e when followed by an alveolar coda became /a/ so that PT *et, *en, and *el are now reflected as /at/, /an/ and /an/ respectively. Table 5-30 provides example of *et, *en, and *el. Notice that the Lungchow forms evidence the general raising of *eC to *iC discussed above.

Table 5-30 Etyma with PT *et, *en, and *el respectively

Gloss	PT	SI	LC	Y	Others
‘seven’	*cet ^D	cet ^{DS1}	cit ^{DS1}	cat ^{DS1}	
‘mushroom’	* ^h rwet ^D	het ^{DS1}	vit ^{DS1}	rat ^{DS1}	
‘porcupine’	* ^h men ^C	men ^{C1}	min ^{C1}		Saek /man ^{C1} /
‘great-grandchild’	*len ^C	le:n ^{A1} -t			Po-ai /lan ^{C1} /
‘body louse’	*mlel ^A	len ^{A2}	min ^{A2}	nan ^{A2}	Saek /mlel ^{A2} /
‘civet cat’	* ^h nel ^A	hen ^{A1}	hin ^{A1}	ɲan ^{A1}	Saek /ɲel ^{A1} /

In contrast to *e, PT *e: in closed rimes follow the same course of development as *e: in open rimes. That is, it is reflected as /ɛ:/ in SWT and some CT varieties but preserved as /e:/ in most dialects. Examples of PT *e:C are given in Table 5-31. Li (1977: 273-274) reconstructs *ɛC for this set of data.

Table 5-31 Etyma with PT *e:C

Gloss	PT	SI	LC	Y	Others
‘cheek’	*ke:m ^C	kɛ:m ^{C1}	ke:m ^{C1}	cem ^{C1}	
‘drought’	* ^ç .le:ŋ ^C	lɛ:ŋ ^{C2}	le:ŋ ^{C2}	reŋ ^{C2}	
‘strong, strength’	*re:ŋ ^A	rɛ:ŋ ^{A2}	ʔe:ŋ ^{A2}	reŋ ^{A2}	
‘eight’	*pe:t ^D	pɛ:t ^{DL1}	pe:t ^{DL1}	pet ^{DL1}	
‘to burst’	*p.re:k ^D	tɛ:k ^{DL1}	p ^h e:k ^{DL1}	pek ^{DL1}	

5.6.2.2 Rimes with *ɤ and *ɤ:

PT mid back unrounded vowels *ɤ and *ɤ: show a wide range of reflexes in modern dialects. Like the other mid vowels, *ɤ and *ɤ: only contrasted in closed syllables. In a number of CT dialects including Western Nung and Jingxi, as well as most NT dialects including Yay, Saek, Qiubei, Shanglin, Guigang etc., *ɤC lowered to /aC/. In all SWT varieties including Siamese, Lao, Lue, Black Tai, and White Tai, PT *ɤC is generally reflected as /oC/. A number of CT and NT varieties such as Shangsi, Yongnan, Laibin, and Yishan also underwent the rounding of *ɤ. In contrast, in a number of CT dialects including Lungchow, Leiping, Lungming, Debao, and Chongzuo, to cite a few, the short *ɤC remains /ɤC/, except before labial codas, where they became rounded and raised to /up/ or /um/. Table 5-32 gives examples of etyma that go back to PT *ɤC. The current interpretation of the data largely agrees with Sarawit (1973) and Wang (1966) but differs from Li (1977: 271-272), who posits *oC for this set of etyma.

Table 5-32 Etyma with PT *ɤC

Gloss	PT	SI	LC	Y	Others
‘ant’	*mɤc ^D	mot ^{DS2}	mɤt ^{DS2}	mat ^{DS2}	
‘heel’	*sɤn ^C	son ^{C1}	ɭɤn ^{C1}	θan ^{C1}	
‘fart’	*k.tɤt ^D	tot ^{DS1}	tɤt ^{DS1}	rat ^{DS1}	
‘chest’	*ʔɤk ^D	ʔok ^{DS1}	ʔɤk ^{DS1}	ʔak ^{DS1}	
‘frog’	*kɤp ^D	kop ^{DS1}	kup ^{DS1}	kap ^{DS1}	
‘sour’	*sɤm ^C	som ^{C1}	ɭum ^{C1}	θam ^{C1}	

The reflex of PT long *ɤ: in closed rime is in sharp contrast with *ɤ. Following Sarawit (1973) and Wang (1966), *ɤ:C have two main reflexes. In NT and a few CT dialects, *ɤ:C simply lowered to /a:C/ in parallel with *ɤC. Languages that show this development include Yay, Saek, Guigang, Longsheng, Qiubei, Qinzhou, Yongnan etc. The second group of dialects includes SWT and CT varieties such as Siamese, Dehong, Lao, Lue, Longzhou, Longming, Dabao, Daxin, etc. In these dialects, PT *ɤ:C diphthongized to /uəC/ in contrast to the other long mid vowels *e: and *o: which simply lowered to /ɛ/ and /ɔ/ (see §5.6.2.1 and §5.6.2.3). Table 5-33 provides examples of etyma with PT *ɤ:C. Li (1977: 282) posits *ua for this set of data. Note that Lungchow /ɤ:C/ comes from earlier *uəC (< PT *uaC and *ɤ:C).

Table 5-33 Etyma with PT *ɤ:C

Gloss	PT	SI	LC	Y	Others
‘house’	rɤ:n ^A	ruən ^{A2}	ɬɤ:n ^{A2}	ra:k ^{A2}	
‘rope’	ɟɤ:k ^D	c ^h uək ^{DL2}	cɤ:k ^{DL2}	sa:k ^{DL2}	
‘boil (intr.)’	dɤ:t ^D	duət ^{DL1}	dɤ:t ^{DL1}	da:t ^{DL1}	

In open syllables, the minimality requirement allowed long *ɤ: but bars short *ɤ from occurring. PT *ɤ:# is raised, merging to *u: in SWT and CT dialects including Siamese, Lao, Lungchow, Lungming, Debao, Jingxi, etc. In all NT dialects and a few CT varieties including Yongnan, Long’an, Fusui, Shangsi, and Chongzuo, it became rounded and merged with PT *o:#. Qinzhou is special in that it is the only dialect that simply lowered *ɤ:# to /a:/. A handful examples of *ɤ:# are available, as given in Table 5-34. Note that ‘meat’ shows a peculiar development in SWT and CT

due to its onset, i.e. *n.mɿ:C > PSWT *nuə^C > Siamese /nuə^{C2}/ vs. *n.mɿ:C > *mlɿ:C > Yay /no^{C2}/.

Table 5-34 Etyma with PT *ɿ:# and *ɿ:l

Gloss	PT	SI	LC	Y	Others
‘straight’	*zɿ:B	su: ^{B2}	ɬu: ^{B2}	θo ^{B2}	Qinzhou /la: ^{B2} /
‘name’	*ɟɿ:B	c ^h u: ^{B2}	cu: ^{B2}	so ^{B2}	
‘meat’	*n.mɿ:C	nuə ^{C2}	nu: ^{C2}	no ^{C2}	Qinzhou /na: ^{C2} /
‘new’	*mɿ:l ^B	maj ^{B1}	mau ^{B1}	mo ^{B1}	Qinzhou /mo ^{B1} /
‘to put’	*s.cɿ:l ^B	saj ^{B1}		so ^{B1}	

Also included in Table 5-34 are etyma with *ɿ:l, which shows a different development from *ɿ: with other codas. In SWT and most CT, PT *ɿ:l became *ɿu earlier on, merging with original *ɿu. It is now reflected as /aj/ in Siamese but /au/ in Lungchow (see §5.4). In NT and a few CT dialects, PT *ɿ:l merged with *ɿ:# so that it is now reflected as /o/ in Yay. Also refer to §5.4 for the reconstruction of PT *ɿj, *ɿu, and *ɿw.

5.6.2.3 Rimes with *o and *o:

The PT mid back rounded vowels *o and *o: went through different paths of development. In parallel with *e:#, PT long *o:# was lowered to /ɔ:/ in SWT and some CT dialects. Examples of languages that underwent the lowering include Siamese, Lao, Southern Shan, Black Tai, Lue, Sapa, Bao Yen, Cao Bang, and the NT language Saek. PT *o:# is still generally preserved as a mid vowel in the remaining CT and NT dialects. Like its high vowel counterparts, only the long *o: can occur in open rimes

due to the minimality requirement on PT prosodic word as discussed in §2.2.2.1. PT *o:# corresponds to *ɔ# in Li's reconstruction (1977: 277-278). Examples of PT *o:# are given in Table 5-28.

Table 5-35 Etyma with PT *o:#

Gloss	PT	SI	LC	Y	Others
‘stump’	*ko: ^A	kɔ: ^{A1}	ko: ^{A1}	ko ^{A1}	
‘well, spring’	*bo: ^B	bɔ: ^{B1}	bo: ^{B1}	bo ^{B1}	
‘shaman’	*hmo: ^A	mɔ: ^{A1}	mo: ^{A1}	mo ^{A1}	
‘bee’	*to: ^B	tɔ: ^{B1}	to: ^{B1}	to ^{B1}	
‘neck’	*ɣo: ^A	kʰɔ: ^{A2}	ko: ^{A2}	ho ^{A2}	

In closed syllables, both *o and *o: can occur. The *o however did not develop exactly like its front counterpart *e. In SWT and CT dialects including Siamese, Lao, Black Tai, Leiping, Lungming, etc., PT *o in closed syllables is generally retained as a mid vowel. However, it is raised to /u/ before nasal codas in some dialects such as Lue and White Tai. In contrast, a few CT varieties raised *oC to /uC/ across the board. Some examples of these dialects are Lungchow, Western Nung, and Qinzhou. In parallel with *eC, in most NT dialects including Yay, Tianlin, Donglan, and Wuming, PT *oC is regularly raised to /uC/. Li (1977: 272-273) reconstructs *wo for this set of data. Table 5-36 illustrates the development of PT *oC.

Table 5-36 Etyma with PT *oC

Gloss	PT	SI	LC	Y	Others
‘to boil (tr.)’	*tom ^C	tom ^{C1}	tum ^{C1}	tum ^{C1}	S. Shan /k ^h om ^{A2} /
‘buttock’	*kon ^C	kon ^{C1}	kun ^{C1}		
‘to itch’	*yom ^A			hum ^{A2}	
‘to shrink’	* ^h rot ^D	hot ^{DS1}		rut ^{DS1}	
‘crooked’	*got ^D	k ^h ot ^{DS2}	kut ^{DS2}		

However, there is a special development shared by most NT languages including Yay, Po-ai, Yongbei, Hechi, Rong’an, Guigang, etc. In these varieties, PT *o when followed by a velar coda did not raise to *u so that PT *ok, and *oŋ remain /ok/ and /oŋ/. Table 5-37 provides examples of *o followed by velar codas. This set of data is reconstructed with *ok, and *oŋ in Li’s system (1977: 271-272).

Table 5-37 Etyma with PT *ok, and *oŋ respectively

Gloss	PT	SI	LC	Y	Others
‘bird’	* ζ .nok ^{DS}	nok ^{DS2}	nuk ^{DS2}	rok ^{DS2}	
‘to drop’	*tok ^D	tok ^{DS1}	tuk ^{DS1}	tok ^{DS1}	
‘dry land’	* b ok ^D	bok ^{DS1}	buk ^{DS1}	bok ^{DS1}	
‘to take down’	*ploŋ ^A	plŋ ^{A1}	pjuŋ ^{A1}	pjoŋ ^{A1}	
‘forest’	*doŋ ^A	doŋ ^{A1}	duŋ ^{A1}	doŋ ^{A1}	
‘lost’	* ^h loŋ ^A	loŋ ^{A1}	luŋ ^{A1}	loŋ ^{A1}	
‘knee’	* χ ow ^B	k ^h aw ^{B1}	k ^h aw ^{B1}	ho ^{B1}	
‘to blow’	*m.pow ^B	paw ^{B1}	paw ^{B1}	po ^{B2}	

Note a special development of *ow. In SWT and CT dialects, PT *ow became *aw merging with the original PT *aw (see §5.4). The change in vowel quality is probably a dissimilation in contact with final *-w. In NT, the rime *-ow became *o:# merging with original PT *o:# discussed above. Li (1977: 291) also posits *ow for this set of data.

In contrast to *o, PT *o: in closed rimes followed the same course of development as *o: in open rimes. That is, it is reflected as /ɔ:/ in SWT and some CT varieties but preserved as /o:/ in most dialects. Examples of PT *o:C are given in Table 5-38. This set of data is reconstructed with *ɔC in Li's system (1977: 277-278, 287-288).

Table 5-38 Etyma with PT *o:C

Gloss	PT	SI	LC	Y	Others
‘to exit’	*ʔo:k ^D	ʔɔ:k ^{DL1}	ʔo:k ^{DL1}	ʔok ^{DL1}	
‘blind’	*ʔo:t ^D	bɔ:t ^{DL1}	bo:t ^{DL1}	bot ^{DL1}	
‘fragrant’	*ho:m ^A	hɔ:m ^{A1}	ho:m ^{A1}	hom ^{A1}	
‘drum’	*klo:ŋ ^A	klɔ:ŋ ^{A1}	kjo:ŋ ^{A1}	coŋ ^{A1}	
‘sugarcane’	*ʔo:j ^C	ʔɔ:j ^{C1}	ʔo:j ^{C1}	ʔoj ^{C1}	

5.6.3 Rimes with low vowel *a and *a:

The low vowels *a and *a: are the most stable of all the PT vowels. Like other vowel pairs, only the long *a: was permitted in open syllables due to the minimality

requirement. PT *a:# is generally retained as long /a:/⁷⁴. Li (1977: 275-276) reconstructs *a for this PT long vowel. Examples of PT *a:# are given in Table 5-39.

Table 5-39 Etyma with PT *a:#

Gloss	PT	SI	LC	Y	Others
‘to kill’	*qa: ^C	k ^h a: ^{C1}	k ^h a: ^{C1}	ka: ^{C1}	
‘dog’	* ^h ma: ^A	ma: ^{A1}	ma: ^{A1}	ma: ^{A1}	
‘fish’	*pla: ^A	pla: ^{A1}	pja: ^{A1}	pja: ^{A1}	
‘to scold’	*dā: ^B	da: ^{B1}	da: ^{B1}	da: ^{B1}	
‘to open mouth’	*ʔa: ^C	ʔa: ^{C1}	ʔa: ^{C1}	ʔa: ^{C1}	

The length contrast between PT *a and *a: in closed rimes is usually retained even in varieties that have lost it elsewhere. For example, high and mid vowels in Wuming do not contrast for length but low vowels /a/ and /a:/ do, cf. /plak^{DS1}/ from PT *prak^D ‘vegetable’ vs. /pla:k^{DL1}/ from *pra:k^D ‘forehead’. Moreover, varieties that have lost the length contrast completely still have different reflexes for *a and *a:. For example, Bouyei dialects substituted the original length in low vowels by a quality contrast so that it has /a/ for *a and /a/ for *a:, cf. /pjak^{DS1}/ from PT *prak^D ‘vegetable’ vs. /pjak^{DL1}/ from *pra:k^D ‘forehead’. Table 5-39 provides examples of PT *aC and *a:C. These two rimes correspond to *əC (Li 1977: 270-271) and *aC (Li 1977: 275-276) in Li’s system.

⁷⁴ Since /a/ and /a:/ never contrast in open syllables, data from Chinese sources usually analyze [a:#] as /a/. I systematically re-transcribe it to /a:/.

Table 5-40 Etyma with PT *aC and *a:C respectively

Gloss	PT	SI	LC	Y	Others
‘to take’	*ʔaw ^A	ʔaw ^{A1}	ʔaw ^{A1}	ʔaw ^{A1}	
‘black’	*dam ^A	dam ^{A1}	dam ^{A1}	dam ^{A1}	
‘to spin’	*pan ^B	pan ^{B1}	pan ^{B1}	pan ^{B1}	
‘to steal’	*h ^l lak ^D	lak ^{DL2}	lak ^{DL2}	rak ^{DL2}	
‘to take a bite’	*kat ^D	kat ^{DS1}	kat ^{DS1}	kat ^{DS1}	
‘father’s younger brother’	*ʔa:w ^A	ʔa:w ^{A1}	ʔa:w ^{A1}	ʔa:w ^{A1}	
‘village’	*ba:n ^C	ba:n ^{C1}	ba:n ^{C1}	ba:n ^{C1}	
‘handle’	*da:m ^A	da:m ^{A1}	da:m ^{A1}	da:m ^{A1}	
‘to drag’	*ç.la:k ^D	la:k ^{DL2}	la:k ^{DL2}	ra:k ^{DL2}	
‘mustard green’	*ka:t ^D	ka:t ^{DL1}	ka:t ^{DL1}	ka:t ^{DL1}	

The above sets of data provide a clear illustration of the contrast between *aC and *a:C. Not only are the reflexes still distinguished by vowel length, their tonal reflexes also attest the original vowel length contrast. For example, Siamese /lak^{DS2}/ from PT *ç.lak^D ‘to steal’ has DS2 tone but /la:k^{DL2}/ from PT *ç.la:k^D ‘to drag’ shows DL1. Also, note a special development of *aC in etyma with medial *-w-. In many CT and NT dialects including Debao, Jingxi, Yay, Wuming, Yongbei and Saek, the medial *-w- and the vowel *a usually fused together to become /ɔ/ or /o/, e.g. Saek /yɔn^{A2}/ from *ywan^A ‘smoke’. This special rounding apparently did not occur with *a:, cf. Saek /va:n^{A2}/ from *xwa:n^A ‘axe’.

5.6.4 Rimes with diphthongs

Among the five diphthongs *iə, *uə, *uə, *ɤu, and *au in PT, the three rising diphthongs generally pattern together. Generally, they are still preserved in some SWT dialects including Siamese, Lao, and Black Tai but have been monophthongized in all other dialects. The results of the monophthongization are mid vowels /e/, /ɤ/, and /o/ in Phu Thai, White Tai, Southern Shan, Dehong, and Lue, but high vowels /i/, /u/, and /u/ in Lungming, Leiping, Debao, Yongnan, Chongzuo, Wuming, Yishan, Qiubei etc. Some varieties such as Yay, Lingyue, and Tiandong however only monophthongized PT *uə but kept *iə, and *uə in tact. In Lungchow, *uə gives /ɤ:/ (< *u:, see §5.6.1.2) in closed syllables but /u:/ in open syllables. Table 5-41 gives examples of etyma with PT *iə, *uə, *uə. These PT diphthong corresponds to *ie, (Li 1977: 280-281) *ue (Li 1977: 281-282), and *ue (Li 1977: 283-284) in Li's system.

Table 5-41 Etyma with PT *iə, *uə, and *uə respectively

Gloss	PT	SI	LC	Y	Others
‘to lick’	*ɕ.liə ^A	liə ^{A2}	li: ^{A2}	riə ^{A2}	
‘sticky’	* ^b niəw ^A	niəw ^{A1}		niəw ^{A1}	
‘disgusted’	*ɖuə ^B	bua ^{B1}	bua: ^{B1}		
‘moon’	*ɖluən ^A	duən ^{A1}	bɤ:n ^{A1}	duən ^{A1}	
‘blood’	*luət ^D	luət ^{DL2}	lɤ:t ^{DL2}	luət ^{DL2}	
‘deaf’	*nuək ^D	nuək ^{DL1}	nu:k ^{DL1}	nu:k ^{DL1}	
‘to scald’	*luək ^D	luək ^{DL2}	lu:k ^{DL2}	lu:k ^{DL2}	
‘coconut grub’	*duəŋ ^C	duəŋ ^{C1}		duŋ ^{C1}	

Note a special development widely found in all the three groups. PT *iəw simplified to *ɛ:w in many dialects including Lungchow, Leiping, Yay, Wuming etc. For example, Lungchow has /de:w^{A1}/ for PT *d̪iəw^A ‘single’ rather than */di:w^{A1}/ indicating that the change *iəw > *ɛ:w occurred before the change from *iə to /i:/. Siamese still retains the original rime, cf. Siamese /diəw^{A1}/. Table 5-42 gives examples of PT *iəw.

Table 5-42 Etyma with PT *iəw

Gloss	PT	SI	LC	Y	Others
‘single, only one’	*d̪iəw ^A	diəw ^{A1}	de:w ^{A1}	dew ^{A1}	
‘withered’	*h̪riəw ^B	hiəw ^{B1}	he:w ^{B1}	rew ^{B1}	
‘green’	*xiəw ^A	kh̪iəw ^{A1}	khe:w ^{A1}	hew ^{A1}	

The falling diphthong *au developed very differently from the rising ones. In a number of languages across the three dialect clusters, *au merged with *aj so that it is now reflected as /aj/ in Siamese, Lao, Tai Yuan, Yongnan, Qinzhou, Tiandong, Hechi, etc.; /ɔj/ in Debao and Jingxi, or /o:y/ in Shangsi and Chongzuo. In contrast, the PT *au simplified to /ɤ/ in some NT dialects including Saek, Liujiang, Yishan, and Rong’an. Of course, there are varieties that still retain the diphthong as /au/, including Black Tai, White Tai, Lungchow, Lungming, Daxin, Yanshan Nung, Yay, Wuming, Shanglin, Laibin, etc. Table 5-43 presents some etyma that go back to PT *au. This PT diphthong corresponds to Li’s *ɤu (1977: 288-289).

Table 5-43 Etyma with PT *auɥ

Gloss	PT	SI	LC	Y	Others
‘leaf’	*ɓauɥ ^A	baj ^{A1}	bauɥ ^{A1}	bauɥ ^{A1}	
‘clear, clean’	*sauɥ ^A	saj ^{A1}	sauɥ ^{A1}	sauɥ ^{A1}	
‘gizzard’	*p.tauɥ ^A	taj ^{A1}	tauɥ ^{A1}	tauɥ ^{A1}	
‘near’	*k.rauɥ ^C	klaj ^{C1}	k ^h rauɥ ^{C1}	cauɥ ^{C1}	
‘daughter-in-law’	*bauɥ ^C	p ^h aj ^{C2}		paɥ ^{C2}	

The diphthong *ɣuɥ will not be discussed here because it has already been discussed in relation to the so-called “Gedney’s Puzzle” (see §5.4). Note that PT *ɣuɥ and *auɥ are written as if they were sequences of vowel plus glide to reflect the fact that they pattern with *ɣj, *ɣw, *aj, and *aw with regards to “Gedney’s Puzzle”. However, phonologically they are clearly complex vowels. First of all, they do not show length contrast. If they had been sequences of vowel plus glide, we would expect *ɣ:uɥ and *a:uɥ to have existed. Moreover, *uɥ never occurs with other vowels, i.e. sequences like *euɥ, *ouɥ, *iuɥ and *uuɥ did not exist. If *uɥ had functioned as a final consonant, it should have been found in combination with other vowels as well.

5.7 Summary

In this chapter I have proposed a system of PT rimes that differ markedly from the conventional reconstruction (Li 1977). First, I have shown that PT had a full-fledged vowel length contrast which interacted closely with syllable structure. Second, I have accounted for the so-called “Gedney’s puzzle” by positing four sets of rimes that do not add any complexity to the system. Lastly, I proposed that final *-l, *-c, and perhaps *-ŋ also existed in PT in addition to those assumed in the conventional

reconstruction. The evidence for these codas comes from Saek as well as less aberrant varieties of Thai. Unlike previous proposals, the current system takes seriously the effect of medial and coda consonants on the development of the vowels. These improvements eliminate the need to posit improbably fine contrasts or unrealistically complex vowels in the PT phonology. Not only does this proposal account better for the whole range of data, it is also simpler and more symmetrical.

CHAPTER 6

PROTO-TAI TONES

6.1 Introduction

A reconstruction of a proto-language is an attempt to present a concrete picture of that linguistic system. With this in mind, the reconstruction of PT cannot be complete without addressing the poorly understood issue of PT tone shapes. This statement is at odds with the general impression that the tonal system of PT has already been firmly established. It is true that a great amount of research has clearly shown that PT had four tonal categories, conventionally labeled *A, *B, *C, and *D (Gedney 1972a; Li 1977; Nishida 1954; Wulff 1934). However, the reconstructed system is still abstract, as the concrete characteristics of each tone are still unknown.

One major problem is the lack of an adequate methodology for tonal reconstruction. Recently, some attempts to reconstruct the phonetic characteristics of proto-tone systems have been made, e.g. Proto-Southwestern Tai (Brown 1965), Old Black Tai (L-Thongkum 2002), Proto-Shan (Morey 2004), and Proto-Northern Min (Handel 2003), but most of them are confined to a group of closely-related dialects whose tonal systems are almost identical, both in terms of contrasts and realizations. This methodological inadequacy in turn stems from poor theoretical and empirical understanding of tone change in general. Faced with this challenge, this chapter approaches the question not only from a comparative perspective but also from the standpoint of systematic similarities between tonal categories across varieties.

In this chapter, I first review the literature on PT tones highlighting those findings that will be crucial in the reconstruction of the PT tone shapes. Then, I present a brief theoretical discussion of the mechanisms that underlie tonal changes. Then, I propose a concrete reconstruction of PT tonal categories with respect to pitch

height, pitch contour, vowel duration, and voice quality. The evidence for such a system comes mainly from modern reflexes of the tonal categories. I claim that PT belongs to the “pitch-plus-voice quality” type of tone language. In other words, both pitch and voice quality were crucial to the identity of PT tones. Lastly, I present an inventory of PT etyma classified according to their reconstructed tones.

6.2 3+1: The PT tonal system

One of the most important advancements in the field of Comparative Tai is the successful reconstruction of tonal categories in PT. Because the methodology and theory behind the reconstruction has been discussed in great detail in great many fora (Edmondson & Solnit 1997; Gedney 1972a; Li 1977), in this section I only briefly highlight those aspects of the PT tonal system that are crucial to the understanding of PT phonology as well as review the earlier reconstruction of PT tones.

The PT tonal system can be characterized as a “three-plus-one” tonal system. Most, if not all, researchers working on the history of Tai languages agree that the proto-language had three contrastive tones, conventionally labeled *A, *B, and *C. However, the contrast among these tonal categories only existed in “non-checked” syllables, or syllables that ended in a vowel or a sonorant consonant. In addition to these three categories, a fourth category *D can be said to have existed in checked syllables, or syllables that have a stop coda. Since the *D tone is the only tone that occurs in checked syllables, one may say that checked syllables neutralize the contrast observed in non-checked syllables. Phonologically, what distinguished *D from *A, *B, and *C was its syllable structure. Despite this lack of tonal contrast, it was probably not the case that *D did not have characteristic tonal properties associated with it. Rather, it is likely that its phonetic characteristics resembled those of one of

the three tones on non-checked syllables. However, for convenience I follow standard practice in referring to as if it was a separate tonal category.

Clearly, it is not an accident that the *ABCD system reconstructed for PT resembles those systems found in Chinese (Baxter 1992; Karlgren 1964; Li 1980; Pulleyblank 1991) and other language groups in SEA (Haudricourt 1961), e.g. Hmong-Mien (Chang 1975) and Vietnamese (Haudricourt 1954b), etc. It is not trivial that all these groups have historically been in intense contact with Chinese. The connections between the PT and Chinese tonal systems are especially striking. Not only are the two systems structurally identical, their tones also show robust correspondences. Specifically, vocabulary shared between Tai and Chinese permits us to establish connections between Chinese tones and Tai tones (Li 1976; Manomaivibool 1975; Wulff 1934) as schematized in Table 6-1. The Middle Chinese forms are from Pulleyblank (1991). The correspondences between Tai and Chinese tones will provide useful information for the reconstruction of PT tones (see §6.5).

Table 6-1 Correspondences between Tai and Chinese tones

	PT	Middle Chinese	
*A	*so:ŋ ^A ‘two’ *bwi: ^A ‘fat’	雙 sǎŋ 肥 bjwei	Even 平
*B	*ha:n ^B ‘goose’ *ye:ŋ ^B ‘shin’	雁 ŋan ^C 脛 yieŋ ^C	Departing 去
*C	*ha: ^C ‘five’ *ma: ^C ‘horse’	五 ŋuo ^B 馬 ma ^B	Rising 上
*D	*pe:t ‘eight’ *ŋuək ‘crocodile’	八 păt 鱷 ŋāk	Entering 入

Table 6-1 illustrates the tonal correspondence between PT and MC. Etyma that had *A in PT regularly showed Even tone in MC, as exemplified by ‘two’ and ‘fat’. Similarly, etyma that show *B in PT had Departing tone in MC. Lastly, etyma that had *C in PT had Entering tone in MC. This systematic correspondence between Tai and Chinese is useful in the reconstruction of PT tones as well as in the dating of Chinese borrowings into Tai. However, this is not to say that all Chinese loans in PT came from a single stage of Chinese. As shown convincingly by Manomaivibool (1975; 1976), many Chinese etyma in Tai must have been borrowed prior to the MC period. Moreover, it is also true that there are many cases where the generalization does not hold. For example, *kaj^B ‘chicken’ had *B in PT but its Chinese source 雞 *jī* (< MC *kiei*) had Even tone. These cases have been analyzed by Manomaivibool (1975). Nevertheless, why Old Chinese and Middle Chinese loans behave similarly in this respect requires an in-depth study.

Even though the tonal systems among modern Tai varieties differ greatly, they can always be shown to have developed from this *ABCD tonal system (Gedney 1972a). However, none of the modern dialects keeps the three-plus-one system intact. This is because daughter languages have all undergone tonal splits conditioned by the laryngeal features of the onsets. In all Tai languages, the four original tones have split into at least two series, those reflecting original voiceless and those reflecting original voiced onsets. This kind of voicing-based tonal split across *A, *B, *C, and *D tones may be referred to as “register split”.

In addition to voicing, other laryngeal features that condition tonal splits in Tai languages include aspiration, glottalization, and frication. In checked syllables, vowel length may introduce a tonal split in the *D category. It is crucial to note that it is the laryngeal features of the onset at the time of splits rather than the original PT onset features that were the determining factors in tonal splits. Presumably, an onset feature

at any period could have triggered a tonal split, but many of the features that did this were not present in the proto-language. For example, many SWT and CT dialects including Siamese, Southern Thai, Lungming, etc. have gone through tonal splits based on aspiration but this does not mean that PT had contrastive aspiration. Rather, the aspiration-based tonal splits indicate that by the time the splits occur those individual dialects had already developed a series of contrastive aspirated stops. If the tonal splits had to refer necessarily to the original PT laryngeal features, aspiration-based tonal split could not have arisen in any Tai dialects, as PT did not contrast for aspiration (see §3.5).

To complicate the picture, tones that resulted from such splits may later merge, further obscuring the tonal correspondence among different dialects. The best tool for capturing the history of tonal splits and mergers in individual Tai varieties is generally known as the “tone-box” method (Gedney 1972a), briefly discussed in §1.4.3. Table 6-2 shows phonological factors that have been commonly found to have conditioned tonal splits and mergers in Tai dialects including voicing, aspiration, and glottalization. Because tonal reflexes provide clues to the phonation types of the onsets, tones in modern languages are referred to according to their etymological sources (see §1.4.3). For instance, using “1” and “2” to represent original voiced and voiceless onsets respectively, tone A2 refers to tone in a given daughter language that developed from PT *A tone conditioned by voiced onsets. Similarly, tone DS1 refers to tone in a given dialect whose etymological source is tone *D with short vowel conditioned by voiceless onsets.

Table 6-2 Phonological factors commonly found to condition tonal splits and mergers (modified from Gedney 1972a)

		*A	*B	*C	*DS	*DL
1	aspirated *f-, hm-, ph-,...					
	unaspirated *p-,...					
	glottalized/implosive *ɓ-,...					
2	voiced *v-, m-, b-,...					

Most modern varieties show a simple register split, i.e. bipartition across the four tonal categories based on voicing. Tones in modern dialects that were conditioned by earlier voiceless and glottalized onsets are referred to as A1, B1, C1, and D1 tones. Similarly, tones that were conditioned by earlier voiceless onsets are label A2, B2, C2, and D2 tones. The tonal systems of White Tai and Western Nung, given in Table 6-3 and Table 6-4, respectively, represent dialects with this kind of basic split. The difference between the two is that the latter further split *D according to vowel length, but the former did not. The *D tone with short vowels is labeled *DS while the *D tone with long vowels is usually referred to as *DL. Because tones in checked syllables are analyzed as belonging to the same categories as those in non-checked syllables, these types of dialects are expected to show six tones altogether. However, the Southern Shan dialect described by Cushing (Hudak 2008) given in Table 6-5 later merged B2 and C1 to yield five tones instead of six.

Table 6-3 Pattern of tonal splits and mergers in White Tai

	*A	*B	*C	*DS	*DL
aspirated	1	2	3	2	2
unaspirated					
glottalized					
voiced	4	5	6	5	5

Table 6-4 Pattern of tonal splits and mergers in Western Nung

	*A	*B	*C	*DS	*DL
aspirated	1	2	3	6	2
unaspirated					
glottalized					
voiced	4	5	6	4	5

Table 6-5 Pattern of tonal splits and mergers in Southern Shan

	*A	*B	*C	*DS	*DL
aspirated	1	2	3	4	2
unaspirated					
glottalized					
voiced	4	3	5	5	3

One fact that is crucial to the reconstruction of the shapes of the PT tones is that *DL very often patterns with *B for the purposes of tonal mergers and splits (Chamberlain 1975; Gedney 1989a). In other words, the two categories often are sensitive to the same conditioning features in the onsets, and are often reflected as the same tones in modern dialects. An example is the case of Western Nung in Table 6-4, among others. In this variety, there was a simple register split in both *B and *DL. Moreover, the two categories agree in having tones 2 and 5 as reflexes of the first and second series, respectively. White Tai and Southern Shan also share this pattern of tone mergers with Western Nung. This strong connection between *B and *DL suggests that *D shared many phonetic properties with *B.

Some varieties show tonal splits based on aspiration or glottalization in addition to the basic register split. The Southern Thai dialect of Surat Thani (L-Thongkum 1978) in Table 6-6 is a clear case of a variety that went through an aspiration-based tonal split. This variety underwent a three-way split across *A, *B and *C but later merged tones that developed from *A1 with tones that developed from *B1. In contrast, the NT dialect of Po-ai in Table 6-7 is an example of a variety that underwent a tonal split based on glottalization. In this case, the *A tone was split

into three tones—one from original voiceless onsets, one from original glottalized/implosive onsets, and one from voiced onsets.

Table 6-6 Pattern of tonal splits and mergers in Southern Thai

	*A	*B	*C	*DS	*DL
aspirated	1		2	2	2
unaspirated	3		7	7	7
glottalized					
voiced	4	5	6	6	5

Table 6-7 Pattern of tonal splits and mergers in Po-ai

	*A	*B	*C	*DS	*DL
voiceless	1	5	3	2	5
glottalized	6			3	
voiced	2	6	4		6

Of particular interest are dialects that still retain the original PT voiced onsets despite having undergone register splits. These varieties include many CT dialects of northeastern Vietnam and southwestern Guangxi (Haudricourt 1960; Kosaka 1997; L-Thongkum 1997; Ross 1996) as illustrated by the dialect of Cao Bang in Table 6-8. Varieties of this type indicate that the binary register split and the devoicing of the voiced onsets, though closely connected, are two separate sound changes. The pattern of tone splits and mergers in Cao Bang is given Table 6-8, where each tone is also

annotated with Chao's tone numbers (Chao 1930). Examples of the retained voicing contrast are provided. In this variety, the vowels are breathy after voiced onsets (Pittayaporn 2007d).

Table 6-8 Pattern of tonal splits and mergers in Cao Bang

	*A	*B	*C	*DS	*DL
aspirated	1 = [53]	3 = [43]	2 = [34]	5	3
unaspirated					
glottalized	5 = [21]				
voiced		4 = [33]	6 = [25]	5	6

A1 = pi:¹ 'year' < *pi:^A

A2 = bi:⁵ 'fat' < *bwi:^A

B1 = tɿw³ 'turtle' < *taw^B

B2 = dɿw⁴ 'ashes' < *daw^B

C1 = pa:² 'turtle' < *pa:^C

C2 = bja:⁶ 'knife' < *jm.ra:^C

Moreover, Cao Bang and other dialects that retain voicing reveal how the basic binary register split and neutralization of phonation type in onsets occurred. In particular, they show that voicing of voiceless sonorants caused the original tones to split into two series (L-Thongkum 1997). Originally, the PT tonal categories were influenced by the phonation type of the onsets so that they had different phonetic realizations after voiced onsets and voiceless onsets. In the second step, these phonetic differences became phonologized, creating two categorically distinct pitch registers. These registers were redundant because they were still predictable from the phonation type of the onset. Subsequently, the voiceless sonorants became voiced, merging with the original voiced sonorants. It is this merger that caused the redundant registers to

become phonemic, resulting in a complete binary register split. The number of tones in each dialect consequently doubled. Cao Bang and other conservative dialects have only completed this step of the evolutionary path. The fourth step is the devoicing of the original voiced obstruents. Most Tai dialects have completed this step but the reflexes of the devoiced onsets could either be plain voiceless consonants or aspirated consonants. While Lue, Black Tai, Lungchow, Lungming, Yay, and most dialects show unaspirated reflexes, Siamese, Lao, Leiping, and Saek have aspirated reflexes. The interrelated processes of binary register split and neutralization of phonation type in onsets discussed is schematized in Figure 6-1.

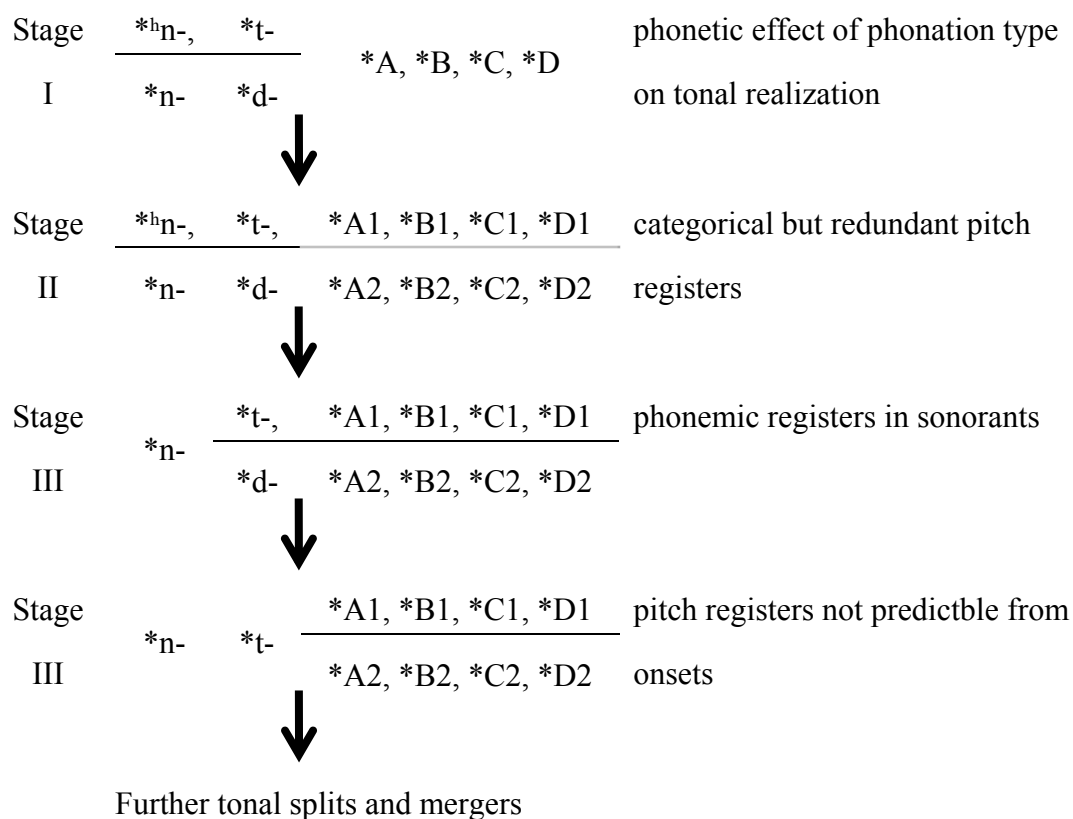


Figure 6-1 Binary register split and neutralization of phonation types

Despite the fact that register split and devoicing are found through out Tai, it is not the case that these changes had already occurred before the diversification of PT. On the contrary, the binary register split and the devoicing took place relatively recently in the history of individual Tai dialects. Chamberlain (1975; 1991) dates the register split back to the PT period but such a position is untenable. The original distinction between voiced and voiceless onsets clearly survived until late in the history of modern dialects. For example, Ayutthaya inscriptions dated to the 15th century CE (Fine Arts Department 1986) still reflect the distinction between voiceless and voiced consonants (Gedney 1991; Pittayaporn 2007a). Moreover, feeding and bleeding relationships among changes indicate that the devoicing occurred only after Black Tai had changed *ɣ- to *g- (Pittayaporn 2007a). These kinds of evidence demonstrate beyond doubt that the interrelated process of register split and devoicing are relatively late events in the history of individual Tai languages.

In talking about the realizations of tonal categories, one needs to seriously consider the role of voice quality in distinguishing among the tones. It is extremely common typologically for tones to be distinguished by voice quality in addition to pitch. A well-known case is Northern Vietnamese, in which three out of six tones exhibit salient glottal properties (Gedney 1989a; Maspéro 1912; Nguyễn & Edmondson 1998; Thompson 1965). The “curve” tone (*hỏi*) is falling-rising and usually creaky. The “broken” tone (*ngã*) starts with an abrupt fall followed by a sharp rise. The abrupt fall is accompanied by strong glottalization of vowel nucleus. Similarly, the “drop” tone (*nặng*) is a falling tone which ends in a glottal constriction. In fact, some researchers have claimed that PT had contrastive phonation types (Sagart 1988), or lacked prosodic contrast (Gedney 1989d). In the remainder of this section, I review evidence for voice quality in *B and *C.

There is evidence that the *ABCD tonal system of PT may have made use of voice quality. Gedney (1989d) points out that in a great number of CT and SWT varieties tones that have developed from PT *C are characterized by glottalization. According to Gedney's impressionistic observations, this glottalization is usually realized as extreme tension of the glottis throughout the articulation of the vowel, ending in a complete glottal constriction at the end of the syllable. Unfortunately, to date there have been no systematic instrumental studies of glottalization across dialects. Moreover, most dialect descriptions use the term glottalization very loosely without distinguishing between creakiness and final glottal constriction (see §6.5.5 for the distinction between creakiness and glottal constriction). However, from Gedney's discussion it is safe to assume that the glottalization he proposes for *C was a final glottal constriction. Western Nung in Table 6-9 is a clear example of a dialect that has been described as having glottalized tones.

Table 6-9 Western Nung tonal inventory

A1	low rising [14]	A2	mid-high level [44]
B1	low falling [21]	B2	mid falling [31]
C1	mid-low level, glottalized [22ʔ]	C2	high level, glottalized [55ʔ]

In Western Nung, the two tones that developed from PT *C are realized as mid-low level and high level tones, both with clear glottalization. In contrast, the other tones, which come from PT *A and *B, do not show such voice quality. This strong tendency for tones that have developed from *C to be glottalized suggests that at least one of the PT tones may have been characterized by voice quality in addition to pitch.

Gedney (1989d) specifically proposes that PT *C tone had a final glottalization or final glottal stop at an earlier stage before the tonal splits occurred.

Gedney's proposal that PT *C was a glottalized tone is in contrast with Sagart's hypothesis (1988) that *C was a creaky tone. Although Sagart recognizes that modern dialects usually show final glottalization for tones that developed from *C, he refers to it as "creakiness" and does not reconstruct *C with final glottalization. He accounts for glottalization in modern reflexes of *C by positing creaky voice quality. This choice of reconstruction is motivated by his view that it was *B that ended with a glottal stop. I will show shortly that this could not have been the case.

As for PT *B, Gedney (1989d) speculates that it had a final *-h. First of all, he draws on the connection between *B and *DL to argue that *B may have had a final consonant like its *DL counterparts. Moreover, he shows that Indic words that end in /-h/, when borrowed into Siamese, lost the final fricative, and were assigned to tone *B. This leads him to speculate that *B tone in PT developed from final *-h. If this theory is correct, PT must be considered a completely non-tonal system. Specifically, the categories *A, *B, and *C would have to be characterized as plain syllables, syllables ending in *-h, and syllables ending in *-ʔ, respectively. Although this view that *B syllables may have had a final *-h at the PT level is not impossible, the tone assignment in Indic borrowings cannot be taken as evidence for this account. First of all, Gedney (1989d) himself points out that his account for *B is based only on data from a relatively recent layer of Siamese vocabulary. Second, he admits the possibility that the Indic words ending in /-h/ were assigned to the *B category because they were phonetically somehow similar to syllables ending in stops.

In contrast to Gedney's view of *B, Sagart (1988) hypothesizes that syllables with PT *B tone had a modal voice rime that ended with a glottal stop. He reasons that the fact that *B patterned with *D with respect to tonal mergers and splits indicates

that they ended with the same type of consonants, namely stops, cf. Vietnamese tonogenesis (Haudricourt 1954b). He also argues that positing final *-ʔ for *B accounts for Gedney's Indic loanword data discussed above as well as the fact that Siamese words ending in /-ʔ/ belong to the DS category. Basically, he hypothesizes that Siamese incorporated Indic and Mon-Khmer borrowings with /-h/ by converting the final glottal fricative to a final glottal stop /-ʔ/. Consequently, those loanwords were naturally considered as belonging to the *D category. According to this explanation, the glottal stop was lost after a long vowel but is kept after a short vowel in Siamese. It is crucial to note that Sagart does not distinguish between creakiness on vowels and final glottalization.

However, this explanation is unlikely for two reasons. Firstly, if PT *B was in fact a glottal stop /-ʔ/, we would expect to see a contrast between PT etyma with long vowel followed by a glottal stop, i.e. *CV:ʔ, and etyma with short vowel followed by a glottal stop, i.e. *CVʔ. However, there is no evidence for such contrast in PT. Sagart himself cites a personal communication with Ferlus, who correctly points out that many of the Siamese /CVʔ^B/ words are loanwords, especially from Khmer. Secondly, the final /-ʔ/ in Siamese is clearly epenthetic in nature and usually does not have an etymological source. Bennett (1995) shows that the final glottal stop in Siamese is phonologically inserted to satisfy word minimality which dictates that a stressed syllable must be heavy (see §2.2.2.1).

In contrast to Gedney's and Sagart's interpretations, I view this set of Siamese borrowing data as a result of loanword adaptation. In particular, Indic etyma with final /-h/ are assigned to the *B category because they follow the same pattern as Indic words ending in other obstruents. Based on data in Gedney (1965), Indic words ending

in obstruents⁷⁵, including the voiceless stops *-p*, *-t*, *-c*, and *-k*, and the fricatives *-s*, *-ś*, and *-ṣ*, were regularly assigned to the *D category when borrowed into Siamese, e.g. รถ /rot^{DS2}/ ‘car’ from Sanskrit *rath(a-)*, ทิศ /tʰit^{DS2}/ ‘direction’ from *diś*, บาท /ba:t^{DL1}/ from *pād(a-)*, and รอก /ro:k^{DL2}/ from *rog(a-)*. Given that */-h/* is an obstruent, we would expect it to follow the same pattern. For instance, we expect **/pʰa:h^{DL2}/* ‘bearer’ from *vāh(a-)*, **/ʔut^{DS1}sa:h^{DS1}/* ‘to make an effort’ from *utsāh(a-)*. However, the final */-h/* was in fact dropped in the process of loanword adaptation because final fricatives are not permissible in Siamese, leaving these etyma to end with an open syllable, cf. ฟาห์ /pʰa:^{B2}/ and อุทศำหึ /ʔut^{DS1}sa:^{B1}/. Because Siamese B1 and B2 are the open-syllable counterparts of DL1 and DL2 respectively (refer to Table 6-12), these forms are considered to have tone B1 and B2 rather than DL1 and DL2.

This explanation also accounts for Sagart’s observation that Siamese words ending in */-ʔ/* belong to the DS category. According to this scenario, the final */-h/* in loanwords with short vowels was dropped when incorporated into Siamese, which does not allow final fricatives. By dropping the */-h/*, the loanwords are left with a short vowel without a coda. This kind of light syllable was banned by word minimality, triggering the */-ʔ/* epenthesis so as to preserve the original short vowel. For example, Khmer *brah* /preah/ ‘honorific prefix’ would have been incorporated into Siamese as **bra*, but the phonology of PT automatically inserted */-ʔ/* so that the etymon was actually borrowed as **braʔ*, which is now pronounced พร๑ /pʰraʔ/ in Siamese. This interpretation is supported by the fact that loanwords that do not have a final consonant in the source Indic language also have final */-ʔ/*, e.g. Siamese ๑๑๑ /ʔa:^{A1}yuʔ^{B2}/ ‘age’ from *āyu-*, ๑๑๑ /kʰanaʔ^{B2}/ ‘group’ from *gana-* etc.⁷⁶ Therefore, the tonal

⁷⁵ Indic etyma were usually borrowed into Siamese with the final vowel truncated (Gedney 1965: 77). For example, Sanskrit लेख *lekha-* gives /le:k^{DL2}/ without the final *-a*.

⁷⁶ The fact that Khmer, through which Siamese borrowed most of its Indic vocabulary, has /ʔa:yuʔ/ for ‘age’, and /kʰanaʔ/ for ‘group’ indicates that the Indic sources did not

assignment of Indic loanwords in Siamese cannot be used as evidence for positing a laryngeal coda for PT *B as proposed by Gedney and Sagart.

In sum, PT consisted of three contrastive tonal categories *A, *B and *C in non-checked syllables. Checked syllables, which had no tonal contrast, and are usually labeled *D for convenience, most likely had the same tone as *B. While it is unclear what the phonetic characteristics of those tonal categories were, modern reflexes of *C suggest that voice quality played an important role in the contrast between PT tones. In the following section, a discussion of mechanisms of tonal change is presented to provide the basis for the reconstruction of PT tones.

6.3 *Directionality of tone change*

The typology of sound changes has long played a role in historical phonology, although this role is sometimes controversial (see (Hock 1991: 617-629)). The essential usefulness of an understanding of the typology of sound change is to clarify likely asymmetries in the directionality of change, e.g. [ki] > [ci] is more likely than [ci] > [ki]. Unfortunately, it is still a mystery how tonal contours change after tonal categories have been established, despite advances in research on tonogenesis and register split (e.g. Abramson & Erickson 1978; Haudricourt 1954b; Hombert, Ohala, & Ewan 1979; Matisoff 1973; Thurgood 2002b). In order to discuss the development of PT tones, it is necessary to have a framework that addresses the way tonal contours change over time. As a solution to this lack of research, I adopt the framework that I myself developed in Pittayaporn (2007b). Even though the framework is tentative, it is useful as a synthesis of earlier work on tone.

end in a final consonant. Final consonants in Sanskrit are normally preserved in Khmer, e.g. /tut/ ‘direction’ from *diś*, and /kruəh/ ‘(mis)fortune’ from *grha-*.

Drawing on the advances in the phonetic study of lexical tone (e.g., Abramson 1978; Gandour 1994; Xu 1997; 2004), this framework attempts to establish a theory of tone change by positing mechanisms that predict the directions of tonal contour change. Following Ohala (2003), I assume that each phonological unit is realized with a great deal of orderly phonetic variation, as shown by fine-grained instrumental analyses. Studying this phonetic variation and speakers' accommodation of ambiguity that it entails can provide a principled way of determining the likelihood of sound changes.

In this framework, I claim that a change in tonal contour (i.e. change in the overall f₀ trajectory) results from a reorganization of the tonal variants that occur in different environments (phonetic, phonological, stylistic, sociolinguistic, etc.). Drawing on synchronic studies of tonal variation, I propose three major mechanisms of tone change that make predictions about the directionality of tone change: 1) segment-tone interaction, 2) contextual variation, and 3) perceptual maximization.

6.3.1 Segment-tone interaction

It is well-established that segments may cause variation in the f₀ trajectory of tones (e.g. Hombert, Ohala, & Ewan 1979; Maddieson 1984a). I argue that the interaction between tones and segments may cause the tonal onset to shift. The most common effect of onset on f₀ is the f₀ lowering following voiced consonants, but other effects exist as well. Solnit and Kingston (1989) and Kingston (2004) discuss how higher f₀ can develop after voiced onsets. In addition, Xu (2003) shows that the onsets of Mandarin tones are phonetically lowered after aspirated onsets. It seems that whether a given type of segment would raise or depress f₀ does not follow automatically from the phonetics but depends on the phonology of the language (Kingston 2004; Kingston & Diehl 1994).

Under this view, any change in overall contour shape is an incidental consequence of the onset shift. Furthermore, the mechanism of segment-tone interaction is predicted not to affect the tone offsets. For example, for a case in which aspiration in the onset induces pitch lowering, a mid level tone would more likely yield a low rising tone than a low level one (see Figure 6-2a). Similarly, a mid level tone is more likely to become a high falling tone in cases where aspiration causes pitch raising.

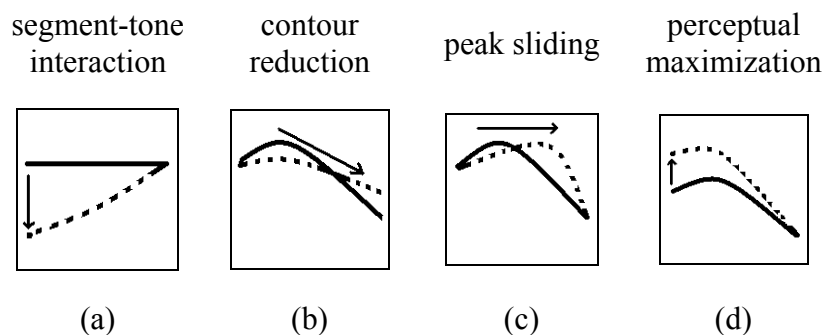


Figure 6-2 Schematization of the mechanisms of tone change

This claim is supported by the fact that the greatest effect is observed at the onset of the vowel. Note that the effects on the f_0 values of vowels after initials of different laryngeal configurations may be significantly different long after the onset of the vowel (Hombert, Ohala, & Ewan 1979).

6.3.2 Contextual variation

The phonetic realizations of lexical tones have been shown to vary across phonetic and phonological contexts (e.g. Gandour 1994; Xu 1997). I claim that the

fact that the underlying targets of lexical tones are often not fully reached due to such effects as contour reduction, peak delay, etc. (Xu 2004) causes “distorted” variants to be generalized and taken as the best approximation of the underlying representation. The intuition is that the frequency of contextually-affected contours is greater than that of “ideal” contours.

6.3.2.1 Contour reduction

Dynamic tones are realized with reduced contours when the relevant syllables are in non-final positions. I claim that the mechanism of contour reduction only targets the contour shape and leaves the tonal onset intact. A change due to contour reduction can therefore be viewed as a change from a dynamic tone to a static tone without modification of the tonal onset. For example, a change from high falling tone to high level tone is more likely than a change in the opposite direction (see Figure 6-2b).

This is supported by the fact that reduced variants show a flatter contour characterized by a smaller amount of f_0 excursion, and a less extreme f_0 value for the tonal offset. However, in these contours the tonal onset stays in the range expected for the onset of the full contour. Under this view, the less extreme f_0 in the tonal offset is a result of the flattening of the contour shape.

6.3.2.2 Peak sliding

I argue that it is more likely for the peak of a tonal contour to slide rightward than leftward. In such cases of peak sliding, the tonal onset should stay intact but the contour shape should change as a result of the rightward sliding of the peak. For example, a change from a falling tone to a convex tone is more likely than one in the opposite direction (see Figure 6-2c).

Xu (2001) finds that the f0 peak of a Mandarin syllable carrying a high or a rising tone may in fact occur in the following syllable. He argues that because sharp rises take a relatively long time to terminate, the f0 turning point may occur somewhat after the syllable boundary in certain environments and in fast speech. As f0 peaks tend to be delayed⁷⁷ rather than early, tone peaks have a greater tendency to shift rightward than leftward.

6.3.3 *Perceptual maximization*

As avoidance of perceptual confusions between phonologically distinct categories is usually hypothesized to be an important ingredient for successful communication (Flemming 2004), I claim that pitch height and pitch excursion may enhance each others to achieve maximal perceptual contrasts. Specifically, the greater the f0 excursion, the less likely a dynamic tone is to be perceived as static. For example, a change from a mid falling tone to a high falling tone is more likely than one in the opposite direction (see Figure 6-2d).

This mechanism is based on Flemming's (2004) argument that perceptual difficulties are not derived from properties of particular sounds but from constraints on categorization of speech sounds. In addition, Stevens et al. (1986) observe that contrastive sounds are often enhanced by redundant features that help listeners to perceive the distinctions. With respect to tones, Yip (2001) argues that the existence of dynamic tones allows a great number of contrasts without placing static tones tightly together in the perceptual space (also see Hombert 1977). As argued by t'Hart et al. (1990), the size of a pitch change must exceed a certain threshold to play a part in communicative functions.

⁷⁷ While "peak sliding" refers to the diachronic displacement of the f0 peak, "peak delay" refers the synchronic phenomenon of late occurrence of the f0 peak.

While contour reduction tends to convert contour tones into static tones, perceptual maximization works within dynamic tones to heighten peak-trough contrasts. They need not have contradictory effects. If a dynamic tone has variants that differ in terms of the pitch value of the tone onset, this mechanism predicts that the variant that shows the greatest f_0 excursion is likely to be generalized.

The three mechanisms of tonal change discussed in Pittayaporn (2007b) as summarized above will be the basis for the discussion of the development of PT tones in the following section. In particular, they will be referred to in discussing how modern tonal reflexes can be projected back to their respective three-plus-one systems. Having established the theoretical background for tonal changes, we now turn to the modern reflexes of the PT tonal categories.

6.4 *Tonal reflexes of PT in daughter varieties*

Comparing modern reflexes of PT tones is a challenging task because of the magnitude of the variation among modern tone systems. Various changes that bring about splits and mergers of tonal categories have produced a very complex set of correspondences which makes strict application of the comparative method a daunting task. However, the picture becomes clearer when the data are organized from the point of view of the *ABCD system. The introduction of internal reconstruction helps reduce the number of correspondences, facilitating the reconstruction of the proto-tonal system. First, we arrive at the tonal system of the patterns of tonal splits and mergers of the proto-language in question. Then, we apply internal reconstruction to individual varieties to discover properties that they inherited from each of the proto-tones. These steps in effect reduce modern tone systems to *ABCD systems, whose tonal categories can be more conveniently compared.

In the following discussion, I will focus on the structural characteristics of each system that are helpful in arriving at an internal reconstruction. I will discuss the phonetic details of the change only when necessary. For our purposes, only seven tone systems are compared. The particular varieties are chosen because 1) their first-series tones and second-series tones can be straightforwardly collapsed into the original three-plus-one system, and 2) they represent different geographical regions of the Tai-speaking area. These varieties include Siamese, Black Tai, Leiping, Cao Bang, Wuming, Liujiang, and Lianshan. Because the CT dialect of Cao Bang in northern Vietnam, spoken the center at the Tai-speaking territory, still preserves the original voiced onsets, and its first-series tones and their corresponding second-series tones still have similar tone shapes, its tonal system (given again in Table 6-10) is the first to be examined.

Table 6-10 Pattern of tonal splits and mergers in Cao Bang⁷⁸

	*A	*B	*C	*DS	*DL
aspirated	1 = [53]	3 = [43]	2 = [34]	5	3
unaspirated					
glottalized	5 = [21]				
voiced		4 = [33]	6 = [25]	5	6

The main mechanism that brought about the observed tone system in Cao Bang is segment-tone interaction (see §6.3.1). The two tones that developed from *A both

⁷⁸ This analysis is based on my own fieldwork. It differs from Hoang (1997)'s analysis in which the glottalized onsets patterned with the voiceless onsets in the *A column.

have falling contours, clearly pointing back to a high falling tone. Assuming that in Cao Bang's register split voiced consonants depressed the onset of the tones, tone A2 (= tone 5) can be derived by lowering the onset of tone *A1 (= tone 1). Interestingly, the glottalized/implosive onsets pattern with the voiced onsets in this column, as in many other dialects. However, the phonetic motivation for this patterning is irrelevant. Similarly, the two tones that come from *B indicate an earlier mid falling tone. Lowering the onset of B1 (= tone 3) straightforwardly gave B2 (= tone 4). Notice that for *B the implosives patterned with voiceless initials. In contrast to *A and *B, Cao Bang *C must have a rising tone. Lowering of C1 (= tone 2) also straightforwardly gives C2 (= tone 6). Therefore, the original Cao Bang tonal system consisted of a high falling, a mid falling, and a mid rising tone. Note the connection between *B1 and *DL1. The next variety to be examined is the CT dialect of Leiping, spoken in southwestern Guangxi. Its tone system is given in Table 6-11

Table 6-11 Pattern of tonal splits and mergers in Leiping

	*A	*B	*C	*DS	*DL
aspirated	1 = [454]	2 = [44]	3 = [13ʔ]	2	2
unaspirated	2 = [44]	5 = [11]			
glottalized					
voiced	4 = [21]			4	5

The easiest tone to reconstruct for Leiping is *C because it does not show any split. Therefore we can posit an earlier glottalized rising tone for this Leiping tone. For tones *A and *B, it is apparent that A1 (= tones 1 and 2) and B1 (= tone 2) are

distinguished from A2 (= tone 4) and B2 (tone 5) respectively in terms of register. While tones A1 and B1 are high tones, tones A2 and B2 are low tones. Therefore, we can posit a falling tone and a level tone for Leiping *A and *B respectively. In addition, *DS and *DL are both connected to *B. We can consider tone 4 [21] in DS2 to have secondarily developed from the phonetically similar tone 5 [11] due to syllable structure. We can thus see that *DS in Leiping is also linked to *B. In sum, Leiping had a falling tone, a level tone, and a rising tone. The next variety to be discussed is the SWT dialect Siamese in Table 6-12.

Table 6-12 Pattern of tonal splits and mergers in Siamese

	*A	*B	*C	*DS	*DL
aspirated	5 = [24 [?]]	2 = [21]	3 = [42]	2	2
unaspirated	1 = [33]				
glottalized					
voiced		3 = [42]	4 = [453 [?]]	4	3

The mechanism of tone-segment interaction can also derive the modern Siamese five-tone system from the *ABCD system. In this case, the first-series tones consistently have lower onsets than their second-series counterparts. As tones B1 (= tone 2), B2C1 (tone 3), and C2 (tone 4) are all falling tones, we can posit falling tones for both *B and *C. Siamese *B thus must have been a mid falling tone, while *C must have been a high falling tone. In addition, *C must have been glottalized as indicated by the presence of glottalization in C2. Similarly, tone A1 (= tone 5) most likely resulted from lowering of the onset of tone A2 (= tone 1) due to the preceding aspiration, therefore indicating that Siamese *A was mid level tone. As for *D, we

again see a link between *DL and *B. *DS, however, does not seem to have had connection to *DL. However, if we consider tone 2 [21] in DS1 to have resulted from a reduction of tone 3 [42] due to syllable structure, we arrive at a new connection between *C and *DS. In sum, Siamese had mid level, mid falling, and high falling tones before the register split. The next variety to be discussed is Black Tai, a SWT dialect of northwestern Vietnam, given in Table 6-13.

Table 6-13 Pattern of tonal splits and mergers in Black Tai

	*A	*B	*C	*DS	*DL
aspirated	1 = [22]	2 = [45]	3 = [21 [?]]	2	2
unaspirated					
glottalized					
voiced	4 = [55]	5 = [44]	6 = [31 [?]]	5	5

The Modern Black Tai tone system is a straightforward case of a binary register split. The second-series tones A2 (=tone 4), B2 (= tone 5), and C2 (=tone 6) can be viewed as a higher version of the first-series tones A1 (= tone 1), B1 (= tone 2), and C1 (= tone 3) respectively⁷⁹. As tones A1 and A2 are all level tones, we can posit a level tone for the original Black Tai system. Similarly, both C1 and C2 are glottalized falling tones, indicating that Black Tai *C must have been a glottalized falling tone. In the case of *B, tone B2 can be considered tone B1 with a raised onset

⁷⁹ Although voiced onsets usually induce lower pitch, cases of high tones from earlier voiced onsets are amply attested as discussed in Kingston and Solnit (1989) and Kingston (2004).

resulting from the register split⁸⁰. Therefore, we can reconstruct Black Tai *B as a rising tone. Therefore, Black Tai *A, *B, and *C were a mid, a rising, and a falling tone, respectively. Notice that both *DL and *DS are tied to *B. The next variety to be discussed is the NT dialect of Wuming in central Guangxi given in Table 6-14.

Table 6-14 Pattern of tonal splits and mergers in Wuming

	*A	*B	*C	*DS	*DL
aspirated	1 = [33]	5 = [24]	3 = [55]	5	5
unaspirated					
glottalized					
voiced	2 = [31]	6 = [13]	4 = [51]	6	6

The pattern of tonal splits and mergers in Wuming is identical to that of Black Tai but the contour shapes are very different. Columns A and C are parallel. Tones A1 (= tone 1) and C1 (= tone 3) are both level but C1 is higher. Similarly, tones A2 (= tone 2) and C2 (= tone 4) are both falling but C2 is higher. Therefore, we can reconstruct two falling tones for Wuming *A and *C respectively. The difference between the two must have been pitch height. *A must have been a mid tone while *C must have been a high tone. As for *B, both the first-register and second-register reflexes, i.e. B1 (= tone 5) and B2 (= tone 6) respectively, are rising tones in the lower pitch region. Therefore, we can reconstruct Wuming *B as a low rising tone. From a

⁸⁰ In Gedney's impressionistic transcription, tone B1 is described as high rising [45]. To consider tone B2 as tone B1 with a raised onset, we need to speculate that this high rising tone in Black Tai may in fact be something between [34] and [45].

systemic point of view, the original contrastive property of Wuming tones seem to have been pitch height—mid for *A, low for *B, and high for *C. In addition, the identity between *B and *D in modern Wuming suggests that the two tones were the same tone, or had properties in common. Now consider the NT dialect of Liujiang in the northeast of Guangxi, given in Table 6-15.

Table 6-15 Pattern of tonal splits and mergers in Liujiang

	*A	*B	*C	*DS	*DL
aspirated	1 = [42]	5 = [33]	3 = [53]	3 ⁸¹	5
unaspirated					
glottalized					
voiced	2 = [231]	6 = [22]	4 = [24]	4	6

Liujiang’s tonal system is quite similar to that of Wuming both in terms of splits and mergers and in terms of the actual tone shapes. Like Wuming, A1 (= ton3 1), and A2 (= tone 2) are both falling tones, even though A2 is in fact rising-falling. We can thus posit a falling tone for Wuming *A. However, Liujiang differs from Wuming in that C1 is a falling tone but C2 is rising. Therefore, it is not clear what tone to posit for *C. As for *B, Liujiang has level tones both for B1 and B2, indicating an earlier level tone. Again, *B is linked to *DL as in most modern Tai dialects. Moreover, there is a very clear connection between *C and *DS, as in Siamese. The

⁸¹ Zhang et al. (1999) describe Liujiang DS1 as [55] and follow the Chinese tradition in considering this checked-syllable tone as a separate tone. However, I view [55] as a checked-syllable allophone of tone 3 [53].

last tone system to be examined is that of Lianshan, a NT variety spoken in Guangdong, provided in Table 6-16.

Table 6-16 Pattern of tonal splits and mergers in Lianshan

	*A	*B	*C	*DS	*DL
aspirated	1 = [42]	5 = [35]	3 = [55]	3	5
unaspirated					
glottalized					
voiced	2 = [132]	6 = [13]	4 = [214]	6	6

The Lianshan tone system shows both the basic binary register split and the quantity-based split in *D syllables. Both tone A1 (= tone 1) and A2 (= tone 2) are falling tones but they differ in terms of onset. While the former starts in the upper part of the pitch range, the latter starts very low and rises before falling slightly. This early low pitch possibly resulted from lowering through the mechanism of segment-tone interaction. Therefore, we can posit a falling tone for *A. The reconstruction of Lianshan *B is straightforward since both B1 (= tone 5) and B2 (= tone 6) are rising tones. Therefore, the *B tone must have been a rising tone in an earlier stage of Lianshan. As for tone *C, it is not clear what tone shape it may have had because C1 (= tone 3) and C2 (= tone 4) agree neither in pitch height or contour shapes. Like many other Tai dialects, there is a connection between *B and *DL. In addition, the link between *C and *DS identified for Siamese and Liujiang is also found in Lianshan. Although we expect tone 4 for DS2 rather than tone 6, but it is very likely that tone 6 [13] in short checked syllables resulted from a contour reduction of tone 4 [214].

In this section, I have applied internal reconstruction to individual dialects to arrive at the tone system before the binary register splits occurred. We will next use the Comparative Method to arrive at the formal characteristics of the PT tonal categories. In the following section, the seven *ABCD systems will be compared in order to infer the characteristics of the common ancestral *ABCD system.

6.5 *Reconstructing the contrasts*

Having internally reconstructed the shapes of the *A, *B, *C, and *D tones in individual varieties, the next step is to compare these reflexes to arrive at a concrete reconstruction of the PT tonal categories. In order to put forward a proposal about the nature of the PT tonal contrast, I first reconstruct the following aspects of the PT tonal categories: pitch height, pitch contour, vowel duration, and voice quality. Then, I proceed to discuss the four phonetic parameters and the claim that PT tones should in fact be considered register complexes.

6.5.1 *Comparing *ABCD in daughter languages*

In this sub-section, I propose a concrete reconstruction of PT tones by comparing reflexes of PT tones in individual varieties before the binary split as established by internal reconstruction in §6.4. In comparing the reflexes, if all the varieties agree, the choice of reconstruction is obvious. In contrast, if the languages present two possible reconstructions, other considerations come into play. For example, by comparing modern reflexes alone, *A can be either falling or level but I posit a level contour for *A because *C is clearly falling (see below). Lastly, if the languages present more than two possibilities, no conclusion can be made. The first PT tone to be considered is *A, whose reflexes in individual varieties discussed earlier are given in Table 6-17.

Table 6-17 PT *A in individual varieties

	Siamese	Black Tai	Cao Bang	Leiping	Wuming	Liujiang	Lianshan
height	mid		high		mid		
contour	level	level	falling	falling		falling	falling
voice quality							

The reflexes of PT *A in individual Tai varieties suggest two characteristics of this proto-tone. First, PT *A must have been either a mid or a high tone. In other words, it must have been in the upper region of the pitch range. This is because varieties whose pre-register split reflexes can be internally reconstructed either had a mid or a high tone for *A, never a low tone. Second, the contour of PT *A must have been either level or falling as reflected in individual varieties except for Wuming. None of the varieties discussed show a rising tone. The situation for tone *B is very different as summarized in Table 6-18.

Table 6-18 PT *B in individual varieties

	Siamese	Black Tai	Cao Bang	Leiping	Wuming	Liujiang	Lianshan
height	mid		mid		low		
contour	falling	rising	falling	level		level	rising
voice quality							

The reflexes of *B in individual varieties only point to possible pitch heights but not contours of PT *B. Individual dialects had either mid or low tones as reflexes for PT *B. In other words, they indicate that PT *B occupied the lower range of the pitch range. As for tonal contour, there is no clear tendency among individual dialects. Level, falling, and rising contours are attested among individual varieties. Therefore, the only characteristic of PT *B that we can posit at this point on the basis of reflexes in the daughter dialects compared is the pitch height. The case of *C, summarized in Table 6-19, is clearer.

Table 6-19 PT *C in individual varieties

	Siamese	Black Tai	Cao Bang	Leiping	Wuming	Liujiang	Lianshan
height	high		mid		high		
contour	falling	falling	rising	rising			
voice quality	?	?		?			

Like PT *A, tone *C was reflected either as a high or a mid tone in the pre-register split stage of individual languages. This suggests that tone *C must have occupied the upper range of the pitch range. As for the contour shape, individual varieties show falling or rising reflexes, suggesting that the PT tone *C may have a dynamic tone, not a level tone. In addition, Gedney (1989d)'s observation that reflexes of *C tones are often glottalized is true for the pre-register split stage of the dialects


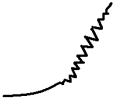

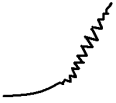
discussed. This fact supports his proposal that PT *C had glottalization⁸² as an important characteristic.

Although we cannot make completely conclusive characterizations of the shapes of PT tones at this point, non-trivial relationships among the categories have become apparent. First of all, tones *A and *C must have been higher in pitch than tone *B because they most likely occupy the upper part of the pitch range. The *B tone, in contrast, must have been relatively low in pitch. Second, we also see a difference in terms of pitch contour among the PT tones. *A is unlikely to go back to a rising tone while *C could not have been a level tone. Moreover, we also see that, at least for one tone, voice quality was an important property. Specifically, tone *C mostly was characterized by glottalization.

Based on the result of the comparison above, I now put forward a reconstruction of the phonetic characteristics of the PT tonal categories. The three contrastive tones in the PT *ABCD system differed from each other in four dimensions: 1) pitch level, 2) tonal contour, 3) vowel duration, and 4) voice quality. A proposed reconstruction of PT *A, *B, *C, and *D is presented in Table 6-20 .

⁸² Gedney (1989d) seems to view this glottalization as a segment *-ʔ while I view it as part of the tone.

Table 6-20 Phonetic characteristics of PT tones

	*A	*B	*C	*D
pitch height	mid	low	high	low
contour	level	rising	falling	rising
vowel duration		long	short	
voice quality	modal	creaky	glottal constriction	
				

I propose that PT *A was a modal tone with mid level contour. In contrast, *B was a creaky tone with a low rising contour, and relatively long vowel duration. Lastly, *C had a high falling contour ending in a glottal constriction, and a relatively short vowel duration. Again, *D is viewed as belonging to the same tone as *B. In the following sub-sections, I discuss each of the phonetic properties that characterized the PT tonal categories.

6.5.2 *Vowel duration*

In tonal system, different tones co-vary with duration of the vowels (Abramson 1962; Gandour 1977; Kong 1987; Yu to appear). For PT tonal categories, the tonogenetic behavior of *D reveals more relationships of just such a systemic nature. I propose that PT *B had a relatively long vowel duration but PT *C was relatively short. Table 6-21 summarizes the connection of *D to *B and *C in the individual varieties.

Table 6-21 Connection of *D to *B and *C

	*DL	*DS
Siamese	*B	*C
Black Tai	*B	
Cao Bang	*B	
Leiping	*B	
Wuming	*B	
Liujiang	*B	*C
Lianshan	*B	*C

As seen above, PT *D has connections with both *B and *C. The strongest link is that between *B and *D. In general, *D followed the same pattern of splits and mergers as *B. The clearest case are Black Tai, Leiping, and Wuming, where *D belongs to the same tone as *B synchronically. In cases where *D split into *DL and *DS, *DL patterns with *B as is the case for Siamese, Cao Bang, Liujiang, and Lianshan. As Gedney discusses (1989d), this suggests that tone *D was phonetically very similar to the *B tone at the PT level (see §6.2). As for *DS, it patterns with *C in many cases suggesting that when *D split into two tonal categories, *DS came to be identified with *C instead of *B.

These inter-relationships among *B, *C, and *D reveal that duration may have played an important role in the contrast between *B and *C. If *B and *DL were phonetically similar, and *C and *DS were similar, then the difference between *B and *C on one hand and the difference between *DL and *DS on the other must be of the same nature. Since the difference between *DL and *DS is clearly that of vowel length, the difference between *B and *C therefore may have been one of duration. In other words, PT *B must have been in some way longer than *C.

Because PT is proposed to have had a vowel-length contrast in closed syllables, vowel duration, which functioned as one of the phonetic correlates of tones, must have interacted directly with contrastive segmental vowel length. In non-checked syllables *A, *B, and *C, the duration of the phonologically short vowels was probably roughly half the duration of the phonologically long vowels. When comparing across tones, the short and the long vowels with tone *B were longer than their respective counterparts with tone *C. Most crucially, long vowels that were shortened due to *C were still significantly longer than short vowels in any of the tones. Similarly, lengthened short vowel with tone *B were significantly shorter than long vowel with any of the tones. This type of interaction between tone and vowel length is exactly what we find in Siamese (Abramson 1962) illustrated in Table 6-22.

Table 6-22 Siamese tones correlated with average vowel durations in msec (Abramson 1962: 107)⁸³

	short vowel	long vowel
low tone	163	524
high tone	97	436

The Siamese data in Table 6-22 presents a comparison between Siamese long and short vowels in syllables with the low tone (= B1) and the high tone (= C2). It shows clearly that the phonetic vowel duration that functions as a property of tones does not interfere with the phonological vowel length contrast. Short vowels and long vowels are kept distinct from each other on syllables with different tones. For any given tone, the duration of long vowels in Siamese is more than twice as great as the duration of their respective short counterparts.

⁸³ Only the data for speaker W.N. are shown here.

The situation may have been a little different for checked syllables, or tone *D. First of all, bear in mind that PT *D did not contrast for tones. Second, it is not uncommon for vowels in checked syllables to be phonetically shorter than vowels in non-checked syllables (Gordon 2001; Zhang 2002). Therefore, it is not surprising that a phonologically short vowel in a *D syllable would have roughly the same duration as the short vowel in a *C syllable. It is also not impossible for the duration of a long vowel on *D syllables to be so short as to resemble that of a short vowel on *B syllables. This kind of situation would give rise to a length-based tonal split in which *DS patterns with *C and *DL with *B, as is observed among Tai dialects. Such split is exactly what we happened in so many Tai varieties including Siamese, White Tai, Western Nung etc. (see §6.2).

6.5.3 *Pitch height*

Pitch height is one of the most important phonetic properties of tones. The pitch height characteristics of the PT tonal categories have been partially uncovered by comparing their reflexes in daughter dialects (see §6.5.1). The comparison reveals that *C had a higher pitch than *B. In addition, the pattern of tonal splits and mergers also provides crucial information. In particular, the behavior of *D with respect to tonal mergers and splits points to a shorter duration for *C than for *B (§6.5.2). Based on this intriguing pitch-duration patterning, I propose that the pitch heights of PT *A, *B, *C were in the mid, low, and high respectively. As for *D, I claim that pitch height was not among its distinctive phonetic properties.

This patterning of higher pitch with short vowels on one hand, and lower pitch with long vowels on the other, is a very common phenomenon cross-linguistically. Diachronically, this pitch-duration patterning has been shown to be the source of tonal contrasts in two Mon-Khmer languages: U (Diffloth 1991) and Hu (Diffloth 1991;

Svantesson 1991). In these cases, low tones have been shown to have developed from long vowels while high tones developed from short vowels. Synchronically, various studies on Thai (Abramson 1962), English (Lehiste 1970), Kui (L-Thongkum 1989), and Chong (L-Thongkum 1991) among others have shown that vowel duration inversely relates to the pitch.

Therefore, positing a low and high tone respectively for PT *B and *C nicely accounts for both the fact that *B had a lower tone than *C, and the fact that the *B was somehow longer than *C. As for PT *A, I posit a mid tone for two reasons. The first reason is that dialect data suggest that *A was higher than *B. Therefore, it must have been either a mid or a high tone. However, treating *A as a mid tone is more likely because *C is already a high tone. As for the vowel duration, *A does not usually pattern with either *DS or *DL with respect to tonal splits and mergers. Therefore, it is unclear whether the height of *A had an effect on the duration of the vowel. As for *D, the fact that it is a checked-syllable variant of *B suggests that it also occupied the lower part of the pitch range.

6.5.4 *Pitch contour*

In addition to pitch height and vowel duration, the comparative data discussed in §6.5.1 also suggest that pitch contours were important phonetic properties of PT tones. Here I propose that PT *A, *B, *C, and *D had level, rising, falling, and rising contours respectively.

If we base our analysis on the contours in individual daughter dialects alone, PT *C could be reconstructed with either a rising or a falling contour. Similarly, there is no agreement among dialects whether tone *A had a level, or a falling contour. However, the glottalization, which is clearly indicated by the comparative data, favors a falling contour for *C. Phonetically, a glottal closure may have an effect of either

raising or lowering of the pitch of the preceding vowel (e.g. Hombert, Ohala, & Ewan 1979; Mazaudon 1977; Thurgood 2002b). This means that a level contour is not a likely candidate for a glottalized tone. Therefore, I posit a falling contour for *C, reserving level contour for *A.

Since *A and *C are reconstructed with a level contour and a falling contour respectively, we can thus posit a rising contour for PT *B. This hypothesis for *B is supported by two facts. First, languages that point to a rising contour in *B are spoken very far apart. Black Tai is a SWT dialect spoken in the northwest of Vietnam while Lianshan is a NT dialect spoken in the northwest corner of Guangdong. This makes the rising contour seem like a case of retention. Second, both experimental and typological studies (Gordon 2001; Kong 1987; Yu to appear; Zhang 2002) have shown that vowels under rising tones tend to be longer than vowels on level and falling tones. The longer duration of tone *B discussed earlier is thus very suggestive of a rising contour. Therefore, I posit a rising contour for PT *B.

Finally, I also reconstruct a rising contours for *D because of its identity to *B. One may argue that, typologically, checked syllables with rising contour are typologically marked (Gordon 2001; Zhang 2002), casting doubt on the view that *D had a rising contour. However, many modern Tai dialects attest rising tones on checked syllable, cf. Wuming, Lianshan, and Cao Bang in §6.4. Wuming, crucially, has only rising tones on checked syllables. Moreover, the reflexes of *DS in numerous modern dialects, e.g. Tai Yuan, the Lao dialect of Nong Khai, Yay, Liujiang, etc., are rising tones. Given the typological restriction against rising tones on checked syllables, it is very implausible that these diverse daughter languages would independently developed rising contours on checked syllables with short vowels. Therefore, the rising contours in reflexes of tone *D must be considered a retention from PT.

6.5.5 *Voice quality*

The last phonetic property of the PT tones to be discussed is voice quality. In addition to the glottalization of tone *C that is robustly supported by the reflexes in modern dialects, we can hypothesize about the voice quality of PT *A and *B. I propose that PT *A, *B, *C, and *D had modal, creaky, finally glottalized, and creaky voice qualities respectively.

Before proceeding to discuss the voice qualities of PT tones, a distinction between “creakiness” and “final glottal constriction” must be made. “Creakiness”, or “laryngealization”, refers to the irregular, low-frequency vibration of the vocal folds extended over a relatively long period of the duration of the vowel (Gordon & Ladefoged 2001; Michaud 2004). It is produced with the vocal folds tightly adducted but open enough to allow voicing. In contrast, “final glottal constriction” refers to a complete or almost complete closure of vocal fold at the end of the rime (Michaud 2004). The vocal fold vibration of the vowel is not irregular. The differences between the two is demonstrated by Michaud (Michaud 2004: 137-139), who clearly compares the “broken” and the “drop” tones in Vietnamese. While the former is creaky, the latter is characterized by a glottal constriction.

The reconstruction of voice quality for *A is straightforward because there is no evidence to date that would indicate a special voice quality. Therefore, it is likely that at the PT level *A had a modal voice quality. For *C, reflexes in modern dialects robustly indicate that it must be reconstructed with a characteristic glottalization (see §6.5.1). However, the nature of the glottalization needs clarification. In most dialect descriptions, the term glottalization is never explicitly defined. However, it is clear from Gedney (1989d: 210)’s discussion that the glottalization found in reflexes of *C is not creakiness but a final glottal constriction preceded by a transitional period of

intermittent voicing following modal voicing. Therefore, I propose that PT *C characteristically had a final glottal constriction. The regular correspondence of *C to the Rising tone in Chinese also lends additional support to the reconstruction of final glottal constriction as part of PT *C. This is because the Rising tone is believed to have developed from OC final glottal stop (Baxter 1992; Haudricourt 1954b) and to have modal voice quality ending in a glottal stop in MC (Sagart 1986).

Additional support for creakiness in *B comes from the durational difference between PT *B and *C. That PT *B had a relatively long duration while *C had a relative short duration suggests that they had creaky voice and final glottal constriction respectively. As discussed by Michaud (2004), final glottal constriction tends to end the syllable abruptly, resulting in a shorter vowel duration. For example, Vietnamese rimes carrying the ‘drop’ tone are impressionistically described as “immediately cut off by a strong glottal stop (Thompson 1965: 41). Acoustically, the “drop” tone is shorter than the other tones and ends in a glottal stop (Pham 2003). Similarly, Burmese tone 4, which developed from earlier /-ʔ/, is extra-short phonetically (Bradley 1982). Therefore, the association of *C with final glottal constriction fits together very nicely with the shorter duration of the proto-tone.

As for *B, the reconstruction of voice quality is not unequivocal. That *C is reconstructed with final glottal constriction and *A with modal voice suggests that *B might have also had a characteristic voice quality at the PT level. Therefore, I hypothesize that PT *B had a creaky voice quality. Support for this hypothesis comes from a few dialects not included in §6.4 and §6.5.1. Dialects that show non-modal voice quality for reflexes of PT *B are found in three geographical areas: the Shan-speaking area, Lao-speaking area, and the CT area of northeastern Vietnam. Various dialects from these areas have been described as having glottalized reflexes of PT *B. In the Shan-speaking area, we have the dialects of Hsi Paw (Strecker 1979), Kengtung

(Hudak 1994), and Hsenwi (Strecker 1979) in Myanmar; the dialect of Mae Sot (Sarawit 1979) in Thailand; and the dialect of Chefang in Yunnan (Hudak 1994). In the Lao-speaking area, we have the Nyo dialect of Tha Uthen in northeast Thailand (Hudak 1997). Lastly, in northeast Vietnam, we have Nung Fan Slihng (Freiberger 1964; Freiberger & Bé 1976; Saul & Freiberger Wilson 1980) and two other CT dialects from Po Muc (Hudak 1995), and Bac Va (Hudak 1995)⁸⁴.

Because the dialect descriptions usually do not distinguish between the two types of glottalization, i.e. final glottal constriction and creaky voice, we can only conclude that these dialects show glottalized voice quality in tones that developed from PT *B. Another complication is that the creak-like quality in B tones in some dialects such as tone B1 in Chefang may in fact have developed secondarily as an enhancement property of low tones, cf. Mandarin tone 3 (Belotel-Grenié & Grenié 1994; Davison 1991). Among the varieties mentioned, Nung Fan Slihng, shown in Table 6-23, is the strongest piece of evidence for the creakiness of tone *B due to the clarity of its description (Freiberger 1964; Freiberger & Bé 1976; Saul & Freiberger Wilson 1980).

⁸⁴ Sagart (1988) was the first to use the glottalized tones in Hsi Paw Shan, Mae Sot Shan, and Nung Fan Slihng as evidence for reconstruction of PT tones. Also see §6.2 for arguments against his proposal that *B had a final glottal stop.

Table 6-23 Pattern of tonal splits and mergers in Nung Fan Slihng

	*A	*B	*C	*DS	*DL
aspirated	1 low-rising	6	5 low falling	4	6
unaspirated, vel. sonorants	4 high-rising	4		1	4
glottalized		6 low glottal			
voiced	3 mid level		2 high glottal	3	

As shown above, the reflexes of PT *B2 and *C2 in Nung Fan Slihng are both glottal tones. As expected, the C2 tone ends with glottal constriction. It is described as having a “high rising glide and a final glottal stop (Freiberger 1964).” Tone B2 ends with a glottal constriction similar to C2, but additionally shows creakiness in the vowel. According to the description, it “ends with a final glottal stop, but laryngealization may occur over all the voiced phonemes (Freiberger 1964).” I claim that this optional laryngealization in tone B2 is a remnant of the creaky voice characteristic of PT *B. It is also possible that the laryngealization is a synchronic enhancement property of low tones which does not have anything to do with PT voice quality. However, the fact that the low falling tone, i.e. C1, is not also described as showing laryngealization suggests that the creakiness is not an enhancement feature of low tones.

Further support for the creakiness of PT *B comes from the regular correspondence to Chinese Departing tone (see §6.2). The most obvious explanation for the correspondence between PT *B and the Chinese Departing tone is that were phonetically similar. There is a consensus among Chinese historical linguists that the Departing tone in MC developed from OC final *-h from an earlier *-s (Baxter 1992;

Haudricourt 1954b; Pulleyblank 1991; Sagart 1999). However, opinion varies with respect to the intermediate stage between non-tonal OC and the modern Chinese dialects which are fully tonal. Haudricourt (1954a) and Pulleyblank (1984) believe that the intermediate stage of the Departing tone was a final aspiration. In contrast, Sagart (1986; 1988) argues that the Departing tone in MC was a short tone with creaky phonation. Reconciling these hypotheses about the development of the Departing tone, we arrive at the following evolutionary path: *-h (from earlier *-s) > creaky or aspirated > pitch-based tone. Together with the glottalized reflexes of *B in modern dialects, the correspondence between PT *B and the Departing tone can therefore be taken as additional support for creaky voice in PT *B.

Additional support for creakiness in *B comes from the durational difference between PT *B and *C. That PT *B had a relatively long duration while *C had a relative short duration suggests that they had creaky voice and final glottal constriction respectively. According to Gordon and Ladefoged (Gordon & Ladefoged 2001), non-modal voice qualities tend to increase the duration of vowels cross-linguistically. For example, Jalapa Mazatec (Silverman et al. 1995) creaky vowels are longer than their modal counterparts. Similarly, the rime of Vietnamese “broken” tone syllable is creaky and has greater duration. Therefore, the association of *B with greater vowel duration supports the hypothesis that *B was a creaky tone.

Further, I propose that PT *D did not have a voice quality characteristic of the tone itself. Instead, I claim that voice quality was neutralized in checked syllables. However, I believe that the vowel in *D syllables were phonetically glottalized due to the following final stops. This glottalization explains why *B and *D patterned together. In particular, glottalization in *D resembled the creakiness characteristics of *B. As discussed by Hombert et al. (1979), Mazaudon (1977), and Thurgood (2002b),

glottalization of vowels preceding final stops has been reported for numerous languages of SEA, e.g. Khaling, Kiranti etc.

6.5.6 *PT as a “pitch-plus voice quality” tone language*

I have proposed above a concrete reconstruction of PT tones *A, *B, *C, and *D. In the proposal, PT tone *A had mid level pitch contour with modal voice. In contrast, tone *B had a low rising contour with by a relatively long vowel duration, and creaky voice. *C had a high falling contour with short vowel duration, and final glottal constriction. Lastly, *D, which contrasted with the other tones in terms of syllable structure, had a low rising pitch contour similar to *B. Not coincidentally, this proposed PT tone system resembles closely that proposed for Proto-Kra (Ostapirat 2000). This bundling of phonetic features closely resembles the situation in Northern Vietnamese. In this dialect of Vietnamese, tones are not distinguished solely by pitch but are complex bundles of pitch contour and voice quality (Brunelle 2009; Maspéro 1912; Michaud 2004; Pham 2003). For example, the “broken” tone shows salient laryngealization, while the “drop” tones are characterized by a final glottal constriction. Moreover, the “broken” tone has a greater rime length in contrast with a relatively short rime length in the “drop” tone. Therefore, I claim that PT, like Vietnamese, was a “pitch-plus-voice quality” tone language⁸⁵.

6.6 *From PT tones to modern reflexes*

Having proposed a concrete reconstruction of PT tones, I now illustrate how modern reflexes can be derived from the reconstructed tones. Based on the framework outlined in §6.3, the discussion will be limited to the seven varieties analyzed in §6.4: Cao Bang, Leiping, Black Tai, Siamese, Wuming, Liujiang, and Lianshan.

⁸⁵ I adopt this term from Michaud (2004).

As discussed earlier, the *A tone was a modal tone with a mid level contour in PT but is reflected as a level or falling tone in daughter languages. In Siamese and Black Tai, PT *A was kept as a level tone before the register split. In contrast, Cao Bang changed tone *A to a high falling tone before the register split. As for Liujiang, Lianshan, and Leiping, *A tone was also a falling tone but the pitch height is unknown. Perhaps, this is because the level contour of PT *A was in fact slightly falling as is the case for the modern Siamese mid tone (Morén & Zsiga 2006). This phonetic fall may have been reanalyzed in these four varieties so that *A became a mid falling tone in these four dialects. The high falling contour in Cao Bang can then be accounted for by the mechanism of perceptual maximization (§6.3.3), which raised the onset from mid to high so that the fall is perceptually more distinct.

PT *B is reconstructed here as a creaky tone with a low rising contour. However, the creakiness has been completely lost in all the varieties discussed here. Moreover, its reflexes in individual varieties vary tremendously. Black Tai and Lianshan both have a rising tone, which can be regarded as a case of retention. Leiping and Liujiang had a level tone, which may have been derived from the PT rising contour by contour reduction (§6.3.2.1). That is, the rising contour was flattened, yielding a level contour. For Siamese and Cao Bang, the *B tone had a falling contour before the register split, the opposite of the original PT rising contour. We can account for the radical change by two simple steps, keeping in mind that a tone with a low onset usually starts in the mid region before falling to reach the low target (Xu 2004). The first step was the loss of pitch rise through peak sliding (§6.3.2.2). This mechanism left *B with a simple fall that started in the mid region. In the next step, the fall may have become reanalyzed as a distinctive characteristic of *B.

As for *C, I reconstruct a high falling contour with a final glottal constriction. Wuming kept this PT tone as a high tone but it is not clear what the contour was like.

In modern Wuming, the glottal constriction has been lost. The CT dialects of Leiping and Cao Bang both had rising tones for *C. The change can be explained in a few steps. First, the initial rise, which was originally, a transition from the mid region, became more prominent as the peak was realized later in the rime through the mechanism of peak delay. After this prominent rise became phonologized, it became a target of perceptual maximization such that the tone onset was lowered to create a steeper pitch rise, yielding a true rising tone. This series of changes is attested in the case of modern Siamese tone 4 (Pittayaporn 2007b; Theranon 2002a, 2002b). Note that Cao Bang has lost its original glottal constriction while Leiping maintains it. As for Siamese and Black Tai, both the falling contour and glottal constriction were still present at time of the register split.

PT *D was phonetically similar to *B even though the truly contrastive characteristic was the syllable structure. This means that *D had a low rising contour as discussed above. However, in most dialects *D later split into *DS and *DL based on vowel length. *DS, which had short vowels at the time of this split, remained unified with *DL and *B in some dialects including Black Tai, Leiping, and Wuming. In contrast, it was re-associated with *C in many dialects including Siamese, Liujiang, and Lianshan. In these varieties, *DS came to be the same tone as *C because its short vowel resembled the relatively short vowel duration characteristic of *C. In addition, there are also some cases where *DS possibly became associated with *A, e.g. Cao Bang, Leiping, and Southern Shan⁸⁶.

In contrast to *DS, *DL, which developed from checked syllables with long vowels, in most cases remained the same tone as *B. The seven varieties used as representatives in §6.4, including Siamese, Black Tai, Cao Bang, Leiping, Wuming,

⁸⁶ In Yay and most NT varieties, the shortening of high vowels before stops (see §5.6.1) occurred before the length-based split occurred, as indicated by the fact that etyma that had these rimes in PT all show DS tones rather than DL tones in these languages.

Liujiang, and Lianshan, all retained the connection until the binary register split. In a handful of languages, however, there is no clear trace of such a link. These varieties include Lao dialects in Laos and northeastern Thailand, as well as Southern Tai dialects in Southern Thailand and northern Malaysia.

In this section, I have shown that the proposed concrete reconstruction of PT tones can straightforwardly account for the wide range of modern reflexes using a framework of tone change that was independently motivated. This is a good indication that the proposed reconstruction is a plausible one.

6.7 Summary

Although it has long been established that PT had four tonal categories *A, *B, *C, and *D, this chapter constitutes a major advancement in uncovering the specific phonetic content of the four PT tonal categories. In particular, I proposed that PT was a tone language that made use of both pitch and voice quality for tonal contrast. I posited a modal tone with a mid level contour for *A; a creaky tone with a low rising contour and a relatively long vowel duration for *B; and a high falling contour ending with a glottal constriction with a relative short vowel duration for *C. Lastly, *D patterned with *B because they were phonetically similar. Specifically, I posited a low rising contour for *D. Not only is the proposed tonal system very plausible from a phonetic and areal point of view, it can also account for the range of variation among reflexes in different modern dialects in a simple way.

CHAPTER 7

DISCUSSION

7.1 *Introduction*

The reconstruction of PT that I have motivated in this dissertation holds that PT was a sesquisyllabic language in which phonological words were either monosyllables or sesquisyllables. Diachronically speaking, this stage is one step away from the monosyllabism observed in modern Tai languages. This view of PT phonology differs markedly from Li's reconstruction (1997) and offers a very different picture of the historical development of Tai languages. This chapter discusses the implications of this reconstruction. I first compare the proposed reconstruction of PT phonology with Li's reconstruction. Subsequently, I propose a preliminary subgrouping schema of Tai languages. I then relate my reconstruction of PT phonology to the most recent findings about Kra-Dai historical phonology, and also discuss two factors that contribute to the convergence among modern Tai dialects.

7.2 *Comparison of PT reconstructions*

As discussed in Chapter 1, Li's reconstruction of PT (Li 1977) holds that PT was a monosyllabic language with a rich set of onsets, and a massive vowel system, but a relatively impoverished set of codas. As for tone, previous reconstructions only provide abstract labels for the reconstructed tonal categories. The current reconstruction differs markedly from that conventional view of PT phonology as it claims that PT was a sesquisyllabic language that had both monosyllabic and sesquisyllabic prosodic words. This view leads to a very different reconstruction of PT onsets and rimes, as we have seen. Moreover, the current system proposes a concrete reconstruction of the tonal categories established by previous research. In this section,

I provide a comparison between the current reconstruction and the heretofore prevalent view of PT phonology.

According to previous reconstructions, PT had a rich consonant system consisting of a four-way phonation type contrast among stops and five contrastive places of articulation. The current reconstruction, on the other hand, posits only three contrastive phonation types, but posits six contrastive places of articulation. With respect to phonation types, the current system reconstructs plain voiceless, implosive, and voiced stops, but lacks the aspirated stop series that is central to previous reconstructions. For example, Li (1977) reconstructs altogether four labial stops *p, *p^h, *ʔb, and *b, but the view of PT phonology adopted here only posits three: *p, *ɓ, and *b. As for places of articulation, my reconstruction posits a distinctive uvular series, in addition to the labial, alveolar, palatal, velar, and glottal series proposed in earlier reconstructions. For example, while Li reconstructs four different plain stops *p, *t, *k, and *ʔ, plus an affricate *tɕ, the current reconstruction of PT phonology posits altogether six voiceless stops *p, *t, *c, *k, *q, and *ʔ.

As for manner of articulation, the current and previous reconstructions agree in having stops, fricatives, nasals, liquids, and glides, but differ in three respects. First, the current system treats the palatal affricates of previous reconstructions as stops: *c and *j. Second, it lacks the labial fricatives *v, and *f but has a uvular fricative *χ and possibly palatal fricatives *ɕ and *z. The palatal glide *j does not exist in the current system. The only palatal glide in the system is *ʔj but it belonged to the glottalized series rather than the sonorants. Note that *w- and *^hw- in the current proposal do not correspond to *w- and *^hw- in previous reconstructions. Table 7-1 gives the PT consonants posited in the current reconstruction and their equivalents in Li's reconstruction (1977) as well as alternative proposals (Ferlus 1990; Sarawit 1973).

Table 7-1

PT consonants and their correspondents in previous reconstructions

	PT	Li	Sarawit	Ferlus
1	*p-	*p-	*p-	*p-
2	*t-	*t-	*t-	*t-
3	*c-	*tɕ-	*c-	*c-
4	*k-	*k-	*k-, *kj-	*k-
5	*q-	*k ^h -	*k ^h -	*ŋk ² -
6	*ʔ-	*ʔ-	*ʔ-	*ʔ-
7	*ʃ-	*ʔb-	*ʔb-	*ʃ-
8	*dʃ-	*ʔd-	*ʔd-	*dʃ-
9	*ʃj-	*ʔj-	*ʔj-	*ʃj-
10	*b-	*b-	*b-	*b-
11	*d-	*d-	*d-	*d-
12	*J-	*J-	*J-, *JW-	*J-
13	*g-	*g-	*g-	*g-
14	*G-	*ɣ-	*ɣ-	*g-/ *ɣ-
15	*s-	*s-	*s-	*s-
16	(*ɕ-)	*tɕ-	*c-	*c-
17	*x-	*x-	*x-	*x-
18	*χ-	*x-	*x-	*ŋx ² -
19	*h-	*h-	*h-	*h-
20	*z-	*z-	*z-, *zW-	*z-
21	(*z-)	*dz-	*J-	
22	*ɣ-	*ɣ-	*ɣ-	*ɣ-
23	* ^h m-	* ^h m-	* ^h m-	* ^h m-, *C.m-
24	* ^h n-	* ^h n-	* ^h n-	* ^h n-
25	* ^h ŋ-	* ^h ŋ-	* ^h ŋ-	* ^h ŋ-

Table 7-1 (Continued)

	PT	Li	Sarawit	Ferlus
26	(* ^h ŋ-)	* ^h ŋ-	* ^h ŋ-	* ^h ŋ-
27	*m-	*m-	*m-	*m-
28	*n-	*n-	*n-	*n-
29	*ɲ-	*ɲ-	*ɲ-, *ŋj-	*ɲ-
30	*ŋ-	*ŋ-	*ŋ-	
31	* ^h w-	*f-	*f-	*f-
32	* ^h r-	* ^h r-	* ^h r-	* ^h r-
33	* ^h l-	* ^h l-	* ^h l-	* ^h l-
34	*w-	*v-	*v-	*v-
35	*r-	*r-	*r-	*r-
36	*l-	*l-	*l-, *lw-	*l-

As in Li's reconstruction, two or more consonants can combine to form complex onsets. The biggest difference with regard to complex onsets is that the current system, like Ferlus's, posits sesquisyllabic clusters while Li's reconstruction allows only tautosyllabic clusters. This means that many sesquisyllabic clusters in this proposal correspond to tautosyllabic clusters in Li's reconstruction. For example, *p.t- in the current system corresponds to *tr- in Li's reconstruction. The PT tautosyllabic clusters that I have posited are given in Table 7-2 together with their equivalents in Li's (1977) and other reconstructions (Ferlus 1990; Sarawit 1973).

Table 7-2

PT consonants and their correspondents in previous reconstructions

	PT	Li	Sarawit	Ferlus
1	*pr-	*p ^h l/r-	*p ^h l/r-	*p ^h r-
2	*tr-	*t ^h r-	*t ^h r-	*t ^h r-
3	*cr-	*t ^h l-	*t ^h l-	*c ^h r-
4	*kr-	*xr-	*x ^h r-	*xr-
5	*qr-	*k ^h l-	*k ^h l-	*k ^h r ² -
6	*br-	*br-	*br-	*br-
7	*gr-	*gr-	*gr-	*gr-
8	*pl-	*pl-	*pl-	*pl-
9	*kl-	*kl-	*kl-	*kl-
10	*ɸl-	* [?] bl/r-	* [?] bl-	*ɸl-
11	*gl-	*gl-	*gl-	*ɣl-
12	*pw-	*pu-	*p-	*p-
13	*bw-	*b-	*b-	*b-
14	*mw-	*mw-	*mw-	
15	* ^h mw-	* ^h mw-	* ^h mw-	
16	*kw-	*kw-	*kw-	*kw-
17	*xw-	*xw-	*xw-	*xw-
18	*ɣw-	*ɣw-	*ɣw-	*ɣw-
19	*ŋw-	*ŋw-	*ŋw-	*ŋw-
20	*qw-	*xw-	*xw-	*xw-
21	*ɣw-	*ɣw-	*ɣ ^h w-	*ɣw-
22	*sw-	*s-	*swj-	
23	*rw-	*r-, rw-	*rw-, *rwj-	
24	*tw-	*t-		
25	*nw-		*nw-	

Table 7-2 (Continued)

	PT	Li	Sarawit	Ferlus
26	* ^h rw-	* ^h r-	* ^h rw-, * ^h ŋrw-	
27	* ^h lw-	* ^h l-	* ^h lw-	
28	*krw-	* ^h krw-	* ^h klw-	* ^h krw-
29	*klw-	*kl-	*klw-	
30	*qrw-	*xr-	* ^h xrw-	
31	*trw-	* ^h tr-	* ^h trw-	
32	*crw-	* ^h tl-	* ^h tlw-	

PT sesquisyllabic onsets posited in the current reconstruction and their equivalents in Li's reconstruction (1977) as well as alternative proposals (Ferlus 1990; Sarawit 1973) are provided in Table 7-3. Note that it is not possible, at least at this stage, to give an exhaustive list of sesquisyllabic clusters in PT.

Table 7-3 Examples of PT sesquisyllabic clusters and their correspondents in previous reconstructions

	PT	Li	Sarawit	Ferlus
1	*p.t-	*tr-	*tr-	*p.t ² -
2	*k.t-	*tl-	*tl-	*k.t-
3	*p.q-	*x-	*xw-	
4	*q.p-	*x-	*x-	
5	*q.t-	*tr-	*tr-	
6	* _ç .b-	*b-	* [?] bw-	*b- (<* [?] v-)
7	* _ç .d-	* [?] dl/r-	* [?] dl-	*d- (<* [?] r-)
8	* _ç .j-	* [?] j-		

Table 7-3 (Continued)

	PT	Li	Sarawit	Ferlus
9	*Ç.t-	*t ^h -	*t ^h -	*n ³ -
10	*Ç.c-	*s-	*s-	
11	*Ç.k-	*k ^h -	*k ^h -	* ^h x ³ -
12	*Ç.q-	*k ^h -	*k ^h -	* ^h k ³ -
13	*k.r-	*k ^h l-	*k ^h l-, *x ^h rw-	*k ^h r ¹ -
14	*p.r-	*pr-	*pr-	*p.t ¹ -
15	*t.r-	*t ^h r-		
16	*c.r-	*t ^h l-	*t ^h l-	*c ^h r-
17	*q.w-	* ^h w-	* ^h w-	* ^h w-
18	*m.l-	*ml/r-	*ml-	*ml-
19	*Ç.r-	*dr-	*dr-	*r/l
20	*Ç.l-	*dl-	*dl-	*l/r
21	*t.n-	* [?] dl/r-	* [?] dr-	*d- (<* [?] l-)
22	*Ç.n-	*nl/r-	*nr-	*C.n-
23	*g.r-	*yr-	*yr-	*yr-
24	*q.s-	*kr-	*kr-	
25	*m.p-	*p-	*p-	* ^m p ¹ -
26	*s.c-	*s-		
27	*z.j-	*z-		
28	*r.t-	*t ^h r-		
29	*t.h-	*t ^h r-		
30	*n.m-	*n-		
31	*gm.r-	*br-		*br-
32	*jm.r-	*vr-	*vr-	*mr-
33	*c.pl-	*pl-		

The system of PT rimes proposed in this reconstruction also differs markedly from previous reconstructions both with respect to the vowels and the codas. Although the current system agrees with Li's proposal that PT contrasted three vowel heights, the two systems disagree with the respect to vowel quantity. Unlike Li's reconstruction, the current system, like Sarawit's, posits a vowel length contrast for each vowel quality. Consequently, the quality distinction in the conventional system between mid and low vowels is replaced by a length distinction. As for the codas, the current system proposes a richer set of coda consonants that includes *-l, *-c, and *-ŋ in addition to those posited in Li's reconstruction. Most importantly, this reconstruction crucially differs from Li's in not reconstructing the typologically suspect contrast between pairs like *wo and *uo, or pairs like *ouŋ and *ɤuŋ. Table 7-4 shows clearly that the rime system of the current reconstruction is simpler and more symmetrical than those of Li's and Sarawit's systems.

Table 7-4 PT rimes and their correspondents in previous reconstructions

	PT	Li	Sarawit
1	*-i:#	-i:#	*-i:#
2	*-iC	*-iC	*-iC
3	*-i:C	*-iəC	*-i:C
4	*-u:#	*-u#, *-uŋu#	*-u:#
5	*-uC	*-wum, *-uum, *-uŋ, *-ej, *-ew	*-uC, *-oim, *-ɤŋ, *-ɤw, *-ɤj
6	*-u:C	*-uəŋ, *-uun, *-wom, *-uut, *-uŋuk	*-u:C, *-waC
7	*-u:#	*-u#	*-u:#
8	*-uC	*uC, *-ouk	*-uC
9	*-u:C	*-uəC	*-u:C, *-eu:k, *-eu:#
10	*-e:#	*-ɛ#	*-e:#

Table 7-4 (Continued)

	PT	Li	Sarawit
11	*eC	*-eC, *-jeC	*eC, *ei:, *-ɣn
12	*-e:C	*εC	*-e:C
13	*-ɣ:#	*-uɰ#,	*-wo:
14	*-ɣC	*oC, *-εj, *-eɰ/-εɰ, *-jɣw	*-ɣC, *-oC, *-ia:j, *-ia:w, *ɣ:w
15	*-ɣ:C	*-uaC, *-ouɰ#,	*-ɣ:C, *-ua:C, *-ua:ɰ, *ua:#
16	*-o:#	*-ɔ#	*-o:#
17	*-oC	*-oC, *-woC, *-ow	*-ok, *-oŋ, *-oit, *-oi:C, *-euC
18	*-a:#	*-a#	*-a:#
19	*-aC	*-ɣC	*-aC
20	*-a:C	*-aC	*-a:C
21	*-iə	*-ie, *-iaw, * -iaj	*-ia
22	*-uə	*-ue, *-uaj, *-uu	*-ua, *-ua:C, *-ua:C
23	*-uə	*-ue, *-uaj	*-ua
24	*-auɰ	*-ɣɰ	*-auɰ

Last but not least, the tonal system of PT in the current system is identical to earlier reconstructions in terms of number of categories. In PT, there were four tonal categories—*A, *B, *C, and *D. However, the current reconstruction goes beyond Li and other researchers in proposing a concrete reconstruction of these categories. Pitch, voice quality, and vowel duration are all claimed to be part of the identities of the tones. In particular, PT *A was a modal tone with a mid level contour, *B was a creaky tone with a low rising contour and a relatively long vowel duration, *C was a high falling contour ending in a glottal constriction with a relatively short vowel duration, and *D was a low rising contour restricted to checked syllables.

My reconstruction nicely accounts for the three outstanding problems in the historical development of Tai languages discussed in §1.3.2. The first problem is the large number of correspondences not satisfactorily accounted for by earlier reconstructions. This problem is overcome by distinguishing loanwords from native PT words, and by positing sesquisyllabic roots alongside the conventionally reconstructed monosyllabic roots. This is clearly illustrated by the case of the “voicing alternation” discussed in §3.4. The current reconstruction shows that a number of etyma standardly cited as examples of the voicing alternation are in fact not reconstructible at the PT level, e.g. ‘bean’, ‘ten’ etc. For etyma that cannot be explained as borrowings, the current reconstruction proposes that they were derived from sesquisyllabic clusters, e.g. *C.tɿŋ^A ‘to arrive’. Other Kra-Dai languages show plain voiceless stops for these sesquisyllabic etyma, e.g. Proto-Kam-Sui *taŋ^A ‘to arrive’, supporting the claim that they go back to clusters containing voiceless stops.

The second problem faced by earlier reconstructions, Li’s in particular, is the complexity of the vowel system. My reconstruction approaches this problem by taking seriously the conditioning environments for vowel changes. The “Gedney’s Puzzle” paradigm illustrates clearly how the proposed vocalism handles the problem. Rather than positing complex rimes for etyma in this paradigm as in Li’s reconstruction, the current approach argues that the observed pattern of reflexes resulted from divergent changes of the vowels conditioned by final glides. The vowel *u in *puw^A ‘crab’ became rounded in Siamese due to the influence of the final glide *-w. The resulting *-uw then merged with PT *-u: to give /pu:^{A1}/ in modern Siamese. The situation is different from Yay, where *-u lowered to *-a probably through dissimilation from *-w. The resulting *-aw then merged with the inherited PT *-aw to give modern Yay /paw^{A1}/.

The third problem is the aberrancy of Saek, which shows numerous idiosyncrasies irreconcilable within all earlier reconstructions. The current reconstruction views most of these properties as clues to the PT phonology not retained elsewhere. For example, etyma that show reflexes of earlier *w- in Saek but *ɸ- in all other Tai languages are reconstructed with sesquisyllabic clusters in PT, e.g. *C.ba:ŋ^A ‘thin’ for Saek /va:ŋ^{A1}/ (< *wa:ŋ^A/) but Siamese /ba:ŋ^{A1}/ (< *ɸa:ŋ^A/). Another example is Saek /-l/, which is viewed in this reconstruction as a retention of PT final *-l. It should be noted that not all idiosyncrasies in Saek are retentions from PT. The /j-/ reflex of PT *z- is clearly an innovation on the part of Saek, most likely due to the influence of Vietnamese.

Further, Li’s and other previous reconstructions suffer from the abstractness of the reconstructed tones. They do not offer a concrete characterization of the tonal categories *A, *B, *C, and *D. The current reconstruction adopts an explicit framework of tone change and combines internal reconstruction and the Comparative Method to arrive at a concrete reconstruction of PT tones. The proposed tonal reconstruction describes PT tones in terms of pitch level, pitch contour, voice quality, and vowel duration. Typologically, it characterizes PT as a pitch-plus-voice quality tonal system.

In sum, the current reconstruction differs markedly from Li’s reconstruction with respect to all aspect of the phonology: structure of prosodic words, onsets, rimes, and tones. It accounts for outstanding problems that cannot be resolved within Li’s system including the large number of onset correspondences, the complexity of the vowel system, the aberrancy of Saek, and the abstractness of the tonal categories.

7.3 *Tai subgrouping*

The conventional subgroup structure of the Tai family is that proposed by Li (1960; 1977), as discussed in §1.2. According to this view, the Tai language family comprises three primary branches, Southwestern Tai (SWT), Central Tai (CT), and Northern Tai (NT), as illustrated in Figure 7-1. Alternative views (Chamberlain 1975; Ferlus 1990; Gedney 1989b; Haudricourt 1956; Strecker 1985) hold that CT and SWT form one primary branch which is a sister to NT. Note that Haudricourt (1956) further notes the existence of some understudied dialects that are intermediate between his two primary branches.

DIAGRAMMATIC CLASSIFICATION OF TAI DIALECTS

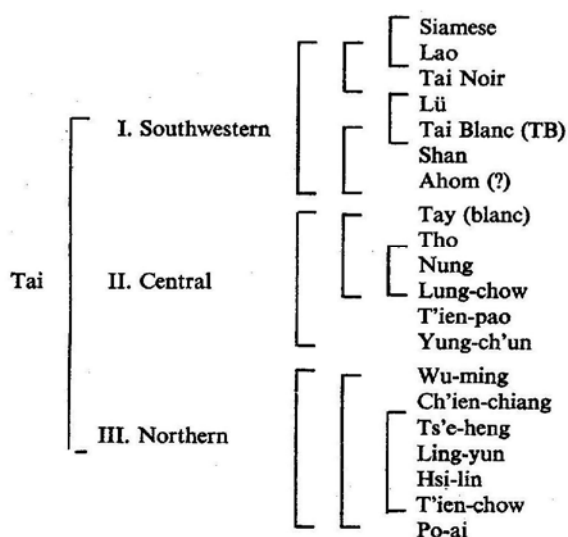
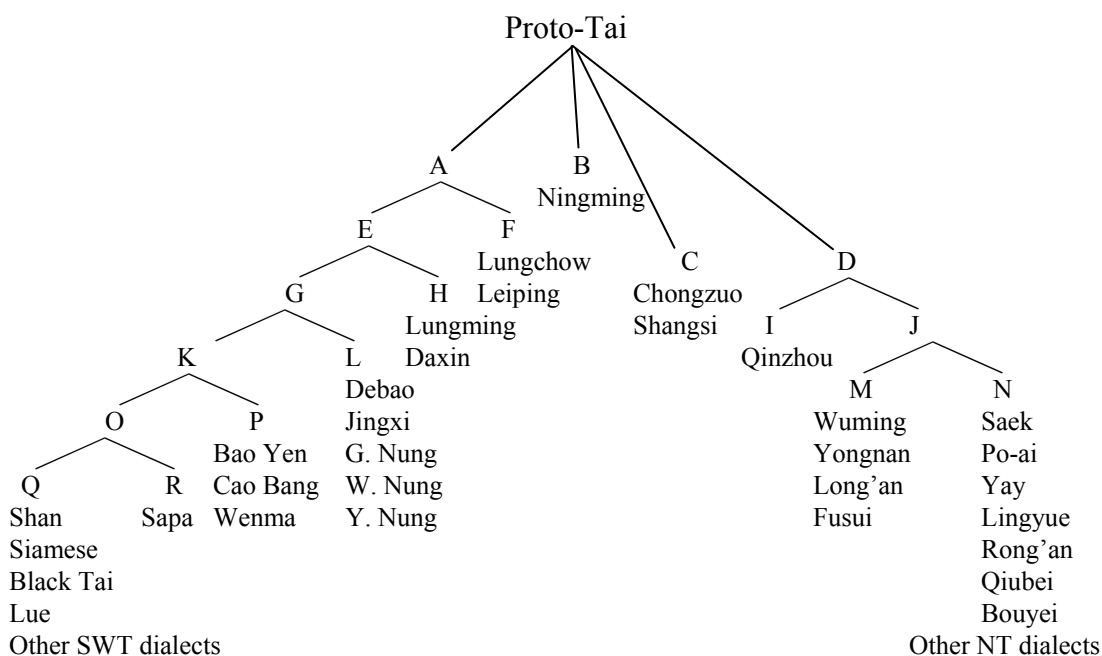


Figure 7-1 Tai subgrouping according to Li (1960).

Although the disagreements among the different proposals are not trivial, it is safe to say that they are all based on the traditional reconstruction of PT, that is, Li (1977). Because the current reconstruction of PT phonology differs markedly from the traditional system, the subgrouping it implies will be significantly different from its

predecessors. Using phonological innovations as criteria for subgrouping, I propose in this section a tentative subgrouping of Tai dialects from the perspective of the current reconstruction without discussing earlier proposals in detail. The proposed subgroup structure of Tai is given in Figure 7-2.

Figure 7-2 **Tentative subgroupings of some Tai dialects studied**



The Tai family tree shown in Figure 7-2 contains altogether 10 terminal nodes: B, C, F, H, I, L, M, N, P, Q, and R. The varieties grouped together under these nodes are combined to form larger subgroups represented in the family tree by nodes A, D, G, I, J, K, and O. Methodologically, I assume that a single sound change does not define a subgroup because most, if not all, innovations can diffuse through contact among dialects. What determines a subgroup is the clustering of several innovations. Furthermore, I assume that innovations that define each subgroup must be exclusively shared only by the members of the subgroup. In fact, many innovations are used as

criteria for defining more than one group. Because the proposed classification is preliminary in nature, the precise sets of phonological innovations that define the different subgroups have not been systematically identified.

The change from *ɯj and *ɯw to *i: and *u: is a good illustration of how innovations are used in this preliminary subgrouping. This sound change is used as a criterion for both group A and C. For group A, the change from *ɯj and *ɯw to *i: and *u: clusters with four other innovations, three of which are not shared by C. Similarly, subgroup C shows three other innovations, two of which are shared with subgroup D rather than C. Therefore, A and C must be recognized as two separate subgroups. In this sense, Lungchow belongs to subgroup A because it shows clustering of the five innovations that define subgroup A. The membership of the terminal subgroups and the phonological criteria that define them are shown in Table 7-5.

Table 7-5 Phonological criteria for subgrouping of Tai languages

Groups	Innovations	Varieties
A	1) *ɯj, *ɯw = *i:, *u: 2) *ɣj, *ɣw, *ɣɯ = *aj, *aw, *aɯ 3) *we, *wo = *e:, *o: 4) *ɟm.r- > *br- 5) *ɯ:k > *u:k	Subgroups F, H, L, P, Q, and R
B	1) *ɣj, *ɣw, *ɣɯ = *i:, *u:, *ɯ: 2) *ɯj, *ɯw = *i:, *u: 3) *we:, *wo: = *e:, *o:	Ningming

Table 7-5 (Continued)

Groups	Innovations	Varieties
C	1) *uj, *uw = *i:, *u: 2) :*we:, *wo: = *e:, *o: 3) *k.t- = *tr- 4) *jṃ.r- > *jr-	Chongzuo, Shangsi, Caolan
D	1) *ɣj, *ɣw, *ɣu = *i:, *u:, *u: 2) *we, *wo = *i:, *u: 3) *k.t- = *tr- 4) *jṃ.r- > *jr-	Subgroups I, M, and N
E	1) *p.t- = *p.r- 2) *um > *ɣm	Subgroups H, L, P, Q, and R
F		Lungchow, Leiping
G	*k.r- = *qr-	Subgroups L, P, Q, and R
H	*qr- = *hr-	Lungming, Daxin
I	*ɣ:# = *a:#	Qinzhou
J	*ɣn, *ɣt, *ɣc = *an, *at, *ac	Subgroups M and N
K	*e:, *o: > *ɛ: and *ɔ:	Subgroups P, Q, and R
L	*qr- = *kr-	Debao, Jingxi, Western Nung, Guangnan Nung, Yanshan Nung
M	*p.t- = *tr-	Wuming, Yongnan, Long'an, Fusui
N	*uj, *uw = *aj, *aw	Saek, Yay, Po-ai, Bouyei, Hengxian, Yongbei, Pingguo, Tiandong, Tianlin, Lingyue, Guangnan Sha, Qiubei, Liujiang, Yishan, Huanjiang, Rong'an, Longsheng, Hechi, Nandan, Dong'an, Dulan, Shanglin, Laibin, Guigang, Lianshan.
P		Bao Yen, Cao Bang, Wenma

Table 7-5 (Continued)

Groups	Innovations	Varieties
O	*ɲn > *on	Subgroups Q, and R
Q	*kr- = *hr-	Black Tai, White Tai, Dehong, Tai Aiton, Tai Phake, Southern Shan, Tai Ya, Lue, Nyo, Yoy, Kaloeng, Phuan, Phu Thai, Lao, Khuen, Tai Yuan, Siamese, Southern Thai, Tak Bai
R	*kr- = *s-	Sapa

In Table 7-5 some subgroups are given only one or two innovations, but this does not mean that the subgroups are defined by those innovations only. Varieties belonging to a given subgroup must show all the innovations that are shared by all nodes that dominate that subgroup. For example, Siamese belongs to Q because it shows all the innovations that characterize A, E, G, K, O, and Q. This means that subgroup Q is defined not by one single innovation *kr- > *hr- but by 11 innovations. Similarly, Wuming belongs to subgroup M because it shows all innovations associated with nodes D, J, and M. In other words, subgroup M is defined by six innovations altogether.

Note that some of the subgroups are defined negatively, In other words, they do not share any innovation that would distinguish them from their sister subgroup(s), but are grouped together due to absence of the innovations that their sisters went through. For example, Bao Yen and Cao Bang are put together in one group because they clearly belong to K but do not show the change from *ɲn to *on, characteristic of subgroup O. To my knowledge, they do not share any innovations that can identify them as members of subgroup O. Alternatively, we might posit a flat structure for

subgroup K, which would consist of three separate branches, namely, Cao Bang, Bao Yen, and subgroup O.

The proposed preliminary subgrouping of Tai dialects shows some notable features. First, the subgroup represented by node Q comprises languages conventionally included in the reconstructions of PSWT (Jonsson 1991; Li 1977; Pittayaporn 2008b, to appear-b). In this sense, SWT is the only one of Li's three groups that is recognized as a valid subgroup. It corresponds to subgroup Q in the family tree in Figure 7-2. Note, however that in addition to the dialects previously recognized as SWT included here in the Q group, other varieties sharing a large number of SWT traits are not included in Q. The only such language included in my sample is Sapa. This variety is extremely similar to subgroup Q, but did not go through the change from *kr- to *hr- characteristic of SWT. In addition, it shows the change from *kr- to *s- not found in the SWT dialects in subgroup Q. Some understudied Tai varieties along the Red River not included in this study, e.g. Payi (Kullavanijaya & L-Thongkum 2000), and Yuanyang (Zhou & Luo 2001), might have a similar relationship to SWT.

Second, none of the subgroups in the proposed subgrouping corresponds exactly to NT in Li's classification. The subgroup that most closely resembles Li's NT is subgroup J, which comprises all of Li's NT dialects plus the CT dialects Yongnan, Long'an, and Fusui in central Guangxi. Of particular relevance is Yongnan, which is almost identical to the NT dialect of Wuming. The only two differences are that in Yongnan *Cr- clusters are reflected as aspirated, and the cluster *tr- that resulted from the merger of PT *tr-, *p.t-, and *p.r-, is reflected as /t^h-. In contrast, Wuming has unaspirated reflexes for the *Cr- clusters, and has /r-/ for the intermediate *tr-. Given their closeness to the NT dialect Wuming, the three CT dialects in subgroup J thus constitute a very strong piece of evidence for considering aspiration to be a recent

areal development. In short, Li's NT can be kept as a subgroup only if the defining criteria are modified to include such dialects as Yongnan, Long'an, and Fusui.

Third, the CT group does not have any genetic status in this subgrouping schema. Varieties classified as CT by Li's criteria are distributed in all four primary branches A, B, C, and D in the proposed subgrouping. This supports Strecker's view (1985) that the phonological innovations shared by CT dialects and used by Li as subgrouping criteria are in fact recent areal innovations. Furthermore, distributing CT dialects among the four primary branches of Tai means that Li's hypothesized CT by itself in fact encompasses the full genetic diversity of Tai languages. As shown in Figure 7-2, all of Li's SWT dialects together make up subgroup Q, which is one low-level subgroup within a single primary branch (A). Similarly, all NT dialects together with Wuming-like CT dialects form subgroup J, which represents only part of the diversity within primary branch D. This situation straightforwardly captures the homogeneity of SWT and NT widely noted in the literature (Gedney 1989b; Pittayaporn 2008b).

Lastly, this proposal does not consider Saek as an outlier of Tai but merely a minor subgroup within one of the primary branches. More specifically, Saek is a member of subgroup N, which comprises all NT dialects in Li's classification except Wuming. As Saek shares a great number of phonological innovations that characterize NT (Gedney 1989c; Haudricourt 1963a), it follows that Saek belongs to subgroup N within the proposed subgrouping. Therefore, the peculiarities of Saek must be considered relatively recent innovations or retentions from the common ancestor of subgroup N. For example, PT *t.n- must have survived until after the separation of Saek from the rest of subgroup N so that the sesquisyllabic cluster did not undergo the implosivization of *t.n- found in all other Tai dialects. After the separation, Saek went its own way and changed *t.n- into /tr-/.

As has become clear, the preliminary subgroup structure proposed here shows characteristics that depart greatly from the conventional three-way classification. It does not recognize either Li's CT or NT as genealogical subgroups, even though subgroup N resembles NT rather closely. Nor does it view SWT as one of the primary branches. Most importantly, it claims that the Tai family tree is rather flat as it consists of more than two primary branches. In short, the reconstruction of PT phonology proposed in this study implies a subgroup structure radically different from all earlier proposals.

7.4 *PT phonology in Kra-Dai context*

Not only is the current reconstruction of PT phonology consistent with advances in the comparative study of Kra-Dai made in the past decades, but it also presents new ideas that further our understanding of Kra-Dai phonology. In this section, I briefly discuss three implications that the current reconstruction of PT has for Kra-Dai phonology: 1) the shape of the prosodic word, 2) the complexity of the vowel system, and 3) the nature of the lexical contrast underlying the *ABCD system.

The first implication is the shape of the prosodic word in Proto-Kra-Dai (PKD). Mounting evidence from other Kra-Dai languages shows that PKD could not have been strictly monosyllabic (see §2.3) but it is not clear exactly what the situation in PKD was. The current reconstruction of PT seems to suggest that the canonical shape of the PKD prosodic word might have been disyllabic. Assuming a genetic relationship between Kra-Dai and Austronesian, Benedict (1982; 1988; 1997a) and Haudricourt (1975) propose that monosyllables in modern Tai languages resulted from a transference of the vowel in the first syllable of disyllables to the second, e.g. PKD **bulan*^A > PT **bluən*^A 'moon' as discussed in (see §2.3.1). Although I showed that

PT was sesquisyllabic rather than disyllabic, such “vocalic transfer”, to use Benedict’s term, is likely to have occurred prior to the formation of PT.

At several points, reconstructions in this dissertation seem to support the position that PKD was a disyllabic language. In particular, it suggests that the sesquisyllabic etyma in PT and their cognates in other Kra-Dai languages followed independent but parallel developments from common disyllabic sources. For example, PT * $\text{C}_\text{.dip}^D$ ‘raw’ corresponds to Sui (Institute of Language and Culture for Rural Development & Thailand and the Kam-Tai Institute 2003) / ʔdju^{DS1} / ‘raw’, suggesting PKD * $\text{C}_\text{.u.dip}^D$. The PKD form may have gone through two parallel vocalic transfers, which resulted in the discrepancy in the vowel reflex between the PT and the Sui forms. In Tai, the transferred vowel *u landed in front of the original vowel of the second syllable *i. It was then lost through simplification, i.e. PKD * $\text{C}_\text{.u.dip}^D > \text{C}_\text{.duip}^D > \text{PT } * \text{C}_\text{.dip}^D$. In contrast, in Sui, the transferred vowel *u was relocated after *i, which subsequently underwent through glide formation to *-j-, i.e. * $\text{C}_\text{.u.dip}^D > \text{C}_\text{.diup}^D > \text{Sui } / \text{ʔdju}^{DS1} /$. If this scenario is correct, PKD must be reconstructed with disyllabic prosodic words.

The second implication is the complexity of the vowel system. The current reconstruction presents a very symmetrical and relatively rich vocalism for PT. One of the most daunting challenges in the study of Kra-Dai phonology is its vocalism. The lack of an understanding of the PKD vowel system is reflected in Strecker (1988), who shows that examining the vowel correspondences among Kam-Sui languages alone has proven to be a daunting task. To my knowledge, there has not been any systematic study of the Kra-Dai vocalic system. The literature on comparative Kra-Dai has focused mainly on the consonants, e.g. Ostapirat (2005), and Edmondson and Yang (1988).

Putting aside complications created by the vocalic transfer discussed above, a clearer picture of the PKD vowel system promises to emerge as a result of the proposed PT vocalism. The best illustration is perhaps the case of PKD long back unrounded vowels *a:, *u:. In Li's system, a diverse set of vocalic nuclei including PT *a, *ua, *(u)u, and *u, all correspond to Sui (Institute of Language and Culture for Rural Development & Thailand and the Kam-Tai Institute 2003) /a:/⁸⁷ as shown in Table 7-6.

Table 7-6 Proto-Tai and Sui reflexes of PKD back unrounded vowel.

Etyma	Li's	PT	Sui	PKD
'stone mountain'	*p ^h ra ^A	*pra: ^A	pra: ^{A1}	*a:
'village'	*ʔban ^C	*ʔa:n ^C	ʔba:n ^{C1}	
'house'	*ruan ^A	*rɯ:n ^A	ɣa:n ^A	*ɤ:
'rope'	*dzwak ^D	*jɯ:k ^D	la:k ^{DL1}	
'ear'	*xrɯu ^A	*k.rwu: ^A	q ^h a: ^{A1}	*u:
'child'	*luɯk ^D	*lu:k ^D	la:k ^{DL2}	
'hand'	*mwu ^A	*mwu: ^A	mja: ^{A2}	

With the current reconstruction of PT phonology, it has now become clear that Sui /a:/ corresponds regularly to back unrounded *a:, *ɤ:, and *u: in PT as shown in Table 7-6. Most likely, PT retains the PKD long back unrounded vowels intact while Sui merged them all into /a:/. From a Kra-Dai perspective, this systematic correspondence seems to indicate that PKD had a relatively rich vowel system with

⁸⁷ Sui /a/ in open syllables is retranscribed as /a:/.

length contrast. In other words, PKD vocalism may have been very similar to the reconstructed PT vowel system.

The last implication to be discussed is the nature of the lexical prosodic contrast in PKD. Although it has never been systematically examined, it is generally assumed that PKD was a tonal language. The basis for this assumption is clearly the typological fact that all attested Kra-Dai languages are tonal. This view is clearly reflected in Ostapirat (2005), who attempts to correlate PAN final consonants to tones *A, *B, *C, and *D in Kra-Dai. Although the current reconstruction of PT phonology does not provide unequivocal evidence for or against this standard view, it highlights the striking similarities among the tonal systems of Kra-Dai languages.

The current reconstruction claims that PT had 3 contrastive tones *A, *B, and *C on non-checked syllables but neutralized them into the so-called *D tone on checked syllables. This is exactly the tonal system found in all of the other branches of Kra-Dai, including Kra (Ostapirat 2000), Hlai (Norquest 2007)⁸⁸, and Kam-Sui (Thurgood 1988). Moreover, both pitch and voice quality are claimed to be crucial to the identity of the lexical tones. PT *A was a modal tone with a mid level contour. *B was a creaky tone with a low rising contour. Lastly, *C was a high falling contour ending with a glottal constriction. Clearly not accidentally, the Proto-Kra tonal system as proposed by Ostapirat (2000) is very similar to the proposed PT system with regards to the number of categories and their phonetic contents. This intriguing resemblance among the tonal systems of the different branches needs to be explained. Perhaps the tonal categories were part of PKD, but it is equally plausible that they represent parallel developments in different branches.

⁸⁸ Norquest (2007) reconstructs Proto-Hlai with an ABCD tonal system but further projects these categories back to final laryngeals in the Pre-Hlai stage.

The implications discussed above illustrate the relevance of the current reconstruction of PT to the study of Kra-Dai. The proposed reconstructed PT phonology is a significant step forward in understanding the Kra-Dai family as a whole.

7.5 *Convergence of Tai varieties*

The reconstruction of PT phonology proposed in this dissertation presents a clear picture of the development of Tai languages. More specifically, it takes an explicit position that the history of Tai varieties is essentially one of convergence. Despite its ability to explain the range of comparative data, its claim that PT was a sesquisyllabic language is at variance with the typological fact that all modern Tai languages are monosyllabic, at least in the native portion of the lexicon (see §2.3). Although cases in which daughter languages are typologically very different from the parent language are common, the fact that all modern Tai varieties agree in exhibiting strong monosyllabicity, despite the sesquisyllabicity of PT, invites discussion. The situation is quite different from the case of Chamic languages, where the modern languages do not conform to one single type (Lee 1966; Thurgood 1999). Although these mainland Austronesian languages have arguably evolved away from the disyllabism of Proto-Chamic toward monosyllabism, they attest various stages of development so that disyllabic, sesquisyllabic, and monosyllabic Chamic languages are all attested. I claim that two factors led to the typological homogeneity of modern Tai dialects: prolonged dialect contact and shared external influence.

The first factor is the prolonged contact among Tai dialects. The preliminary subgroup structure reveals a large number of innovations that cut across subgroup boundaries. Such innovations indicate that the dialects remained in contact with each other after initial diversification so that an innovation that originated in one group can diffuse to others. A clear example of an innovation that spread across subgroup

boundaries is aspiration of PT *Cr- (see §3.5.2). All modern dialects except for Wuming and those in the N subgroup underwent the process of aspiration but, contra Li (1960), the change does not define any of the subgroups. The case of Yongnan and Wuming discussed above shows clearly that aspiration of *Cr- is a case of diffusion. Another example is the neutralization of *C̣.l- and *C̣.r- found commonly in subgroup N. Although the majority of dialects in this subgroup have /r-/ or other consonants that came from earlier *r- (see §3.6.8.1 and §4.4.5) as reflex of this merger, Saek still keeps the two distinct as /l-/ and /r-/ respectively. Similarly, a diverse range of varieties in the A branch show /l-/ for both, but Sapa and the Q subgroup (i.e. subgroup O) still maintain the distinction (see §4.4.5). These two cases are just a few examples of the many phonological innovations that diffuse from one subgroup to others. These innovations are clear attestations of the prolonged dialect contact that led to the observed typological homogeneity among Tai dialects.

The second factor is shared external influence. Obviously, the language that had the greatest influence on Tai languages was Chinese (Benedict 1975; Li 1976; Manomaivibool 1975; Maspéro 1912; Wulff 1934). The Chinese influence is evident both in the amount of Chinese vocabulary found in Tai languages and in the fact that the three-plus-one tonal system of PT was structurally identical to the Middle Chinese tonal system. The Chinese influence on Tai must have persisted for an extended period of time as Manomaivibool (1975; 1976) shows successfully that Siamese, despite being the most Indianized of all Tai dialects, nevertheless consists of at least four layers of Chinese loanwords: 1) pre-Middle Chinese (including OC and LH), 2) (Early) Middle Chinese (MC), 3) Late Middle Chinese (LMC), and 4) post-Middle Chinese.

Siamese is not the only variety that shows these layers. Etyma from the earliest three layers are found widely among Tai dialects. For example, 稼 *jià* ‘sow crops’ (OC

*krâh, LH *ka*^C, MC *ka*^C, LMC *kja*·) is the source for PT *kla:^C ‘rice seedling’. It belongs to the pre-Middle Chinese layer. The presence of the liquid *-l- indicates that the borrowing predated the Later Han dynasty. Similarly, many etyma found widely in Tai belong to the Middle Chinese layer. For instance, Chinese 破 *pò* ‘to smash’ (OC *phâih, LH *phai*^C, MC *p^hwâ*^B, LMC *p^hua*·) is found in Siamese and Lungming as /p^ha:^{B1}/ ‘to split’ and in Yay as /pa:^{B1}/. This post-PT loan must have been incorporated around the Early Middle Chinese period as suggested by the rime /a:/. Lastly, Chinese borrowings from the Late Middle Chinese layer are also commonly found in Tai dialects. For example, 夫 *fū* ‘husband’ (OC *pa, LH *pua*, MC *pju*, LMC *fu*ǎ) is attested in most if not all SWT and CT dialects, e.g. Siamese /p^huǎ^A/, and Western Nung /p^hu:^A/ ‘husband’. The aspirated /p^h-/ in modern dialects suggests that OC and MC *p- had already been spirantized to *f*- in the source language, dating the borrowing of this etymon to Late Middle Chinese. These examples of Chinese loanwords clearly show that Tai languages have been in contact with Chinese since Han times at least, and that contact has been persistent and widespread across Tai varieties, even after differentiation. This prolonged contact with Chinese contributed to the observed typological homogeneity among Tai dialects.

The prolonged contact among Tai dialects and the shared Chinese influence were the most important factors responsible for the typological homogeneity of Tai dialects on the one hand, and the typological discrepancy between reconstructed PT and its daughter languages on the other. Even though the sesquisyllabicity of PT as reconstructed here at first glance seems at odds with the monosyllabicity characteristic of modern Tai varieties, such a discrepancy is in fact expected for dialects that remain in close contact for an extended period of time. In this sense, the typological homogeneity among Tai languages is a product of convergence rather than a shared retention from PT. More specifically, Tai dialects, which remained in contact with

each other, all followed the same path toward monosyllabicity, with Chinese as a model. As shown by Matisoff (2001), monosyllabicization is clearly an areal phenomenon of Mainland Southeast Asia. In Kra-Dai, this tendency to become monosyllabic is particularly strong, to the extent that all but a very few languages have become monosyllabic.

7.6 Conclusion

This dissertation presents a new phonological reconstruction of PT. It takes as its premises a systematic application of the Comparative Method and the force of language contact upon language change. Essentially, the interpretation of the comparative data was constrained by explicit and motivated hypotheses that did not underestimate the role of areal convergence and dialect contact in language change. Moreover, it incorporates a large amount of dialect data, much of which has only become available recently. The resulting reconstruction thus has various advantages over previous proposals.

The first advantage is the success of the current reconstruction in accounting for the various problems in historical studies of Tai languages in a unified way. My reconstruction of PT phonology claims that PT was a sesquisyllabic language despite the monosyllabicity characteristic of modern Tai languages. This view of PT connects several phenomena that are generally regarded as unconnected and gives them a unified account. Such phenomena include the voicing alternation, certain alleged split reflexes in Saek, and many other complex correspondences that are viewed here as results of a single phenomenon of monosyllabification.

The second advantage of the current reconstruction is its simplicity. Although the proposed reconstructed PT sound system is complex in some areas, such as having sesquisyllabic roots and a richer set of contrasts among places of articulation, it

allows us to eliminate many more complexities in other areas of the system. For example, reconstructing a richer place contrast help reduce the number of phonation type contrasts. The establishment of the uvular series and the elimination of the aspiration contrast were independently motivated, but the former got rid of the need to reconstruct aspirated *k^h-, the only aspirated phoneme in Li's system that is supported by a large number of cognate sets.

Finally, the PT reconstruction here clearly distinguishes between shared retentions and shared innovations that resulted from convergence. One of the most serious shortcomings common among earlier proposals is a lack of appreciation of the effect of language contact. As a result, many phonological changes that diffused among Tai languages through contact were taken to be properties retained from PT. One of the best examples is the case of aspiration. Because contrastive aspiration is one of the most common properties shared by a large set of modern Tai varieties, it is assumed in Li's reconstruction to have to been a characteristic of PT. The current reconstruction systematically examined Tai-internal comparative data and showed that aspiration developed secondarily in different dialects that remained in contact after PT broke up into different branches.

In conclusion, this dissertation is an attempt to overcome the challenging task of distinguishing universal tendencies, chance, diffusion, genetic retentions, and parallel development in the Tai language family of Mainland Southeast Asia, one of the world's most intriguing linguistic areas. Equipped with the classic Comparative Method, an abundance of highquality comparative data, and an appreciation of language contact, the current reconstruction of PT phonology presents a vivid picture of the history of Tai languages as characterized by divergent changes overridden by successive waves of convergent trends that radically transformed the ancestral Proto-Tai language into a network of typologically homogenous dialects that differ markedly from their parent.

APPENDIX A
LIST OF LANGUAGES

	Varieties	Alternate names/spellings	Groups	Locations	Sources
1	Aiton	Tai Aiton	SWT	Northeastern India	Morey (2005)
2	Bac Va		CT	Northeastern Vietnam	Hudak (1997) Hudak (2008)
3	Bao Yen		CT	Northwestern Vietnam	Pittayaporn (2007c)
4	Black Thai	Tai Dam, Song	SWT	Northwestern Vietnam, relocated groups Laos, Thailand, and USA)	Gedney (n.d.) Hudak (1994) Hudak (2008)
5	Bouyei	Buyi (布衣)	NT	Southern Guizhou	Chinese Academy of Sciences (1959) Institute of Language and Culture for Rural Development, Mahidol University Kam-Tai Institute, Central University for Nationalities (2001)

	Varieties	Alternate names/spellings	Groups	Locations	Sources
6	Cao Bang		CT	Northwestern Vietnam	Hoang (1997) Hudak (1997) Pittayaporn (2007d)
7	Chongzuo (崇左)		CT	Southwestern Guangxi	Zhang et al. (1999)
8	Daxin (大新)		CT	Southwestern Guangxi	Zhang et al. (1999)
9	Debao (德保)		CT	Southwestern Guangxi	Kullavanijaya (1992a) Zhang et al. (1999)
10	Dehong (德宏)	Tai Nüa, Tai Mao, Dai dialects of Mengshi (芒市) and Menglian (孟连) belong to this group.	SWT	Southwestern Yunnan, northwestern Myanmar	Harris (1975) Kullavanijaya (2001) Luo (1999) Zhou, Fang, and Xi (2001)
11	Donglan (东兰)		NT	Northwestern Guangxi	Zhang et al. (1999)
12	Du'an (都安)		NT	Northwestern Guangxi	Zhang et al. (1999)
13	Fusui (扶绥)		CT	Central Guangxi	Zhang et al. (1999)

	Varieties	Alternate names/spellings	Groups	Locations	Sources
14	Guangnan Nung	Guangnan Nong (广南侬)	CT	Eastern Yunnan	Zhang et al. (1999)
15	Guangnan Sha (广南沙)		NT	Eastern Yunnan	Zhang et al. (1999)
16	Guigang (贵港)		NT	Central Guangxi	Zhang et al. (1999)
17	Hechi (河池)		NT	Northeastern Guangxi	Zhang et al. (1999)
18	Hengxian (横县)		NT	Central Guangxi	Zhang et al. (1999)
19	Huanjiang (环江)		NT	Northeastern Guangxi	Zhang et al. (1999)
20	Jingxi (靖西)		CT	Southwestern Guangxi	Zhang et al. (1999)
21	Kaloeng	Kaloeng	SWT	Southern Laos, relocated groups in central and northeastern Thailand	Thisapong (1985)
22	Khuen	Khün	SWT	Eastern Myanmar, Northwestern Thailand	Hudak (1994) Petsuk (1978)

	Varieties	Alternate names/spellings	Groups	Locations	Sources
23	Laibin (来宾)		NT	Central Guangxi	Zhang et al. (1999)
24	Lao	Northeast Thai, Isan	SWT	Laos, northeastern Thailand	Gedney (n.d.) Hudak (1997) Hudak (2008)
25	Lei Ping	Leiping (雷平)	CT	Southwestern Guangxi	Gedney (n.d.) Hudak (1995) Hudak (2008)
26	Lianshan (连山)		NT	Northwestern Guangdong	Zhang et al. (1999)
27	Lingyue (凌乐)		NT	Northwestern Guangxi	Zhang et al. (1999)
28	Liujiang (柳江)		NT	Northeastern Guangxi	Zhang et al. (1999)
29	Long'an (隆安)		CT	Central Guangxi	Zhang et al. (1999)
30	Longsheng (龙胜)		NT	Northeastern Guangxi	Zhang et al. (1999)
31	Lue	Lü, Tai Lue, Sipsongpanna Tai. Tai Yong is a dialect	SWT	Northern Thailand, northern Laos,	Gedney (n.d.)

	Varieties	Alternate names/spellings	Groups	Locations	Sources
		of Lue, cf. Gedney's Lue dialect of Muang Yong. Dai dialect of Jinghong (景洪 = Chiang Rung) is a Lue dialect.		Northwestern Vietnam, Southern Yunnan	Hudak (1996) Hudak (2008) Zhou and Luo (2001)
32	Lungchow	Longzhou (龙州), Nung Chao	CT	Southwestern Guangxi	Gedney (n.d.) Hudak (2008) Li (1940) Li (1977) Zhang et al. (1999)
33	Lungming	Longming (龙明)	CT	Southwestern Guangxi	Gedney (n.d.) Hudak (1991a) Hudak (1995) Hudak (2008)
34	Nandan (南丹)		NT	Northeastern Guangxi	Zhang et al. (1999)
35	Ningming (宁明)	Ning Ming	CT	Southwestern Guangxi	Hudak (1995) Hudak (2008) Zhang et al. (1999)

	Varieties	Alternate names/spellings	Groups	Locations	Sources
36	Nung Phan Slinh	Nung Fan Slihng	CT	Northeastern Vietnam	Freiberger (1964) Freiberger (1976) Saul and Freiberger Wilson (1980)
37	Nyo	Nyaw	SWT	Southern Laos, relocated groups in central and northeastern Thailand	Boonsner (1984)
38	Phake	Tai Phake	SWT	Northeastern India	Morey (2005)
39	Phu Thai	Thai Kapong is a dialect of Phu Thai.	SWT	Southern Laos, relocated groups in central and northeastern Thailand	Phothiban and Trongdi (1998) Pittayaporn (2005b)
40	Phuan	Thai Phuan	SWT	Eastern Laos, relocated groups in central and northeastern Thailand	Phothiban and Trongdi (1998) Trongrat (1998)
41	Ping Siang	Pingxiang (虎祥)	CT	Southwestern Guangxi	Hudak (1995) Hudak (2008)

	Varieties	Alternate names/spellings	Groups	Locations	Sources
42	Pingguo (平果)		NT	Central Guangxi	Zhang et al. (1999)
43	Po-ai	Bo'ai (剥隘), Jui	NT	Eastern Yunnan	Hudak (2008) Li (1977)
44	Qinzhou (钦州)		CT	Southeastern Guangxi	Zhang et al. (1999)
45	Qiubei (丘北)		NT	Eastern Yunnan	Zhang et al. (1999)
46	Rong'an (融安)		NT	Northeastern Guangxi	Zhang et al. (1999)
47	Saek		NT	Northeastern Thailand, Central Laos	Gedney (n.d.) Gedney (1989c) Hudak (1993) Hudak (2008)
48	Sapa		SWT	Northwestern Vietnam	Pittayaporn (2007e)
49	Shanglin (上林)		NT	Northwestern Guangxi	Zhang et al. (1999)
50	Shangsi (上思)		CT	Southeastern Guangxi	Pittayaporn (2008c)

	Varieties	Alternate names/spellings	Groups	Locations	Sources
51	Siamese	Thai, Standard Thai, Bangkok Thai	SWT	Central Thailand	Zhang et al. (1999) Gedney (n.d.) Hudak (2008) Li (1977) The Royal Institute (2003)
52	Southern Shan	Shan, Tai Long, Tai Luang, Tai Yai	SWT	Eastern Myanmar, Northwestern Thailand	Cushing (1914) Gedney (n.d.) Hudak (1994) Hudak (2008) Sarawit (1979)
53	Southern Thai	Southern Siamese, Pak Tai, Kedah Thai	SWT	Southern Thailand, northern Malaysia	Phothiban and Trongdi (1998) Umar (2003)
54	Tai Ya		SWT	Eastern Yunnan	Xing (2000)
55	Tai Yuan	Northern Tai, Lanna Tai, Kam Muang	SWT	Northern Thailand, relocated groups in Central Thailand	Hudak (1997) Hudak (2008) Rungruengsri (1991)
56	Tak Bai	Kelantan Thai, Jehe	SWT	Southern Thailand, northern Malaysia	Phothiban and Trongdi (1998)

	Varieties	Alternate names/spellings	Groups	Locations	Sources
57	Tiandong (田东)		NT	Northwestern Guangxi	Zhang et al. (1999)
58	Tianlin (田林)		NT	Northwestern Guangxi	Zhang et al. (1999)
59	Wenma (文马)	Wenma Thu (土)	CT	Eastern Yunnan	Zhang et al. (1999)
60	Western Nung	Nung Inh	CT	Northwestern Vietnam	Gedney (n.d.) Hudak (2008)
61	White Tai	Tai Don, Dialects of Jinping (金平) is a dialect of White Tai	SWT	Northwestern Vietnam, Southwestern Yunnan	Gedney (n.d.) Hudak (1994) Hudak (2001) Hudak (2008) Zhou and Luo (2001)
62	Wuming (武鸣)		NT	Central Guangxi	Gedney (n.d.) Hudak (2008) Li (1956) Li (1977) Zhang et al. (1999)

	Varieties	Alternate names/spellings	Groups	Locations	Sources
63	Yanshan Nung	Yanshan Nong (侗)	CT	Eastern Yunnan	Zhang et al. (1999)
64	Yay	Giay	CT	Northwestern Vietnam	Gedney (n.d.) Hudak (1991b) Hudak (2008) Pittayaporn (2006)
65	Yishan (宜山)		NT	Northeastern Guangxi	Zhang et al. (1999)
66	Yongbei (邕北)		NT	Central Guangxi	Zhang et al. (1999)
67	Yongnan (邕南)		CT	Central Guangxi	Zhang et al. (1999)
68	Yooy		SWT	Southern Laos, relocated groups in central and northeastern Thailand	Boonsner (1984)

APPENDIX B

GLOSSARY OF RECONSTRUCTED PT ETYMA

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
	A. Body parts									
1	head (1)	*kraw ^C						law ^{CI}	caw ^{CI}	t ^h raw ^{CI}
2	head (2)	*truə ^A	huə ^{A1}	hu ¹	huə ^{A1}	t ^h uə ^{A1}	hu. ^{A1}			
3	head hair	*prym ^A	p ^h om ^{A1}	p ^h um ^{A1}	p ^h jom ^{A1}	p ^h jom ^{A1}	p ^h jum ^{A1}	p ^h om ^{A1}	piəm ^{A1}	p ^h ram ^{A1}
4	hair knot	*klaw ^C	kla.w ^{CI}		caw ^{CI}	caw ^{CI}	kjaw ^{CI}			
5	body hair, feather	*q.puul ^A	k ^h on ^{A1}		k ^h on ^{A1}	k ^h on ^{A1}	k ^h un ^{A1}		pun ^{A1}	pul ^{A1}
6	gray-haired	* ^h ɣwu:k ^D	ŋɔ:k ^{DLI}	hoʔ ^{DLI}	hoʔ ^{DLI} -v					hu:k ^{DLI}
7	forehead	*pra:k ^D	p ^h a:k ^{DLI}		p ^h ja:k ^{DLI}	p ^h ja:k ^{DLI}	p ^h ja:k ^{DLI}	p ^h ak ^{DLI}	pja:k ^{DLI}	p ^h ra:k ^{DLI}
8	nose	*daj ^A	?daj ^{CI} -t	daj ^{A1}	daj ^{A1}	daj ^{A1}	daj ^{A1}	daj ^{A1}	daj ^{A1}	daj ^{A1}
9	mucus of the nose	*mu:k ^D	mu:k ^{D12}	muʔ ^{D12}	mu:k ^{D12}		muk ^{D12}	muk ^{D12}	muk ^{DS2}	mu:k ^{D12}
10	face	* ^h na: ^C	na: ^{CI}	na: ^{CI}	na: ^{CI}	na: ^{CI}	na: ^{CI}	na: ^{CI}	na: ^{CI}	na: ^{CI}
11	eye	*p.ta: ^A	ta: ^{A1}	ta: ^{A1}	p ^h ja: ^{A1}	t ^h a: ^{A1}	ha: ^{A1}	t ^h a: ^{A1}	ta ^{A1}	pra: ^{A1}
12	mouth	*pa:k ^D	pa:k ^{D1}	paʔ ^{D1}	pa:k ^{D1}	pa:k ^{D1}	pa:k ^{D1}	pa:k ^{D1}	pa:k ^{D1}	pa:k ^{D1}
13	tongue	*li:n ^C	lin ^{C2}	lin ^{C2}	lin ^{C2}	lin ^{C2}	lin ^{C2}	hɪn ^{C2}	lin ^{C2}	li:n ^{C2}
14	tooth	*wan ^A	fan ^{A2}		p ^h an ^{A2}	van ^{A2}	fan ^{A2}		fan ^{A2}	
15	gum; gill	* ^h ɣwək ^D	ɣwək ^{D1}	huʔ ^{D1}	huwək ^{D1}	huwək ^{D1}	hɜ:k ^{D1}		ɣwək ^{DS1}	
16	saliva	*la:j ^A	la:j ^{A2}		la:j ^{A2}	na:j ^{A2} -i	la:j ^{A2}	ma:j ^{A2}	na:j ^{A2}	mla:j ^{A2}

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
17	cheek	*ke:m ^C	ke:m ^{C1}	kem ^{C1}	ke:m ^{C1}	kem ^{C1}	ke:m ^{C1}		cem ^{C1}	ke:m ^{C1}
18	ear	*krwu: ^A	hu: ^{A1}	su: ^{A1}	su: ^{A1}	su: ^{A1}	hu: ^{A1}	ly: ^{A2}	ruə: ^{A2} -t	ruə: ^{A2} -t
19	fang	*ɣe:w ^C	k ^h ɛw ^{C1}	xew ^{C1}	k ^h ɛ:w ^{C1}	k ^h ɛw ^{C1}	k ^h ɛ:w ^{C1}	hew ^{C1}	hew ^{C1}	
20	chin, jaw	*ga:ŋ ^A	k ^h a:ŋ ^{A2}	ka:ŋ ^{A2}	ka:ŋ ^{A2}	gä:ŋ ^{A2}	ka:ŋ ^{A2}	k ^h a:ŋ ^{A2} -i	ha:ŋ ^{A2}	ya:ŋ ^{A2}
21	beard	*mom ^B				mum ^{B2}	mum ^{B2}	mum ^{B2}	mum ^{B2}	
22	neck	*ɣo: ^A	k ^h ɔ: ^{A2}	xo: ^{A2}	k ^h ɔ: ^{A2}	gɔ: ^{A2}	ko: ^{C1}	ho: ^{A2}	ho: ^{A2}	ɣo: ^{A2}
23	goiter	* ^h niəŋ ^A	niəŋ ^{A2}	niəŋ ^{A1}	niəŋ ^{A1}	niəŋ ^{A1}	ni:ŋ ^{A1}			
24	shoulder	*Ç.ba: ^B	ba: ^{B1}	ba: ^{B1}	ba: ^{B1}	ba: ^{B1}	ba: ^{B1}	ba: ^{B1}	ba: ^{B1}	va: ^{B1}
25	arm	*qe:n ^A	k ^h ɛ:n ^{A1}	xen ^{A1}	k ^h ɛ:n ^{A1}	k ^h ɛn ^{A1}	k ^h ɛ:n ^{A1}	ken ^{A1}	cen ^{A1}	ke:n ^{A1}
26	elbow	*Ç.swo:k _D	so:k ^{DL1}	soɰ ^{DL1}	ʰɔ:k ^{DL1}	ʰək ^{DL1}			suək ^{DL2}	suək ^{DL2}
27	hand	*mwu: ^A	mu: ^{A2}	mu: ^{A2}	mu: ^{A2}	mu: ^{A2}	mu: ^{A2}	moy: ^{A2}	fuy ^{A2}	mu: ^{A2}
28	finger, toe	*ni:w ^C	niw ^{C2}	niw ^{C2}	niw ^{C2}	niw ^{C2}	ni:w ^{C2}			
29	finger nail, toenail	*Ç.lep ^D	lep ^{DS2}	lip ^{DS2}	lop ^{DS2}	lep ^{DS2}	lip ^{DS2}	lip ^{DS2}	rit ^{DS2}	li:p ^{DL2}
30	joint	*ɣo: ^C	k ^h ɔ: ^{C1}	xo: ^{C1}	k ^h ɔ: ^{C1}	k ^h ɔ: ^{C1}			ho: ^{B2} -t	
31	leg	*p.qa: ^A	k ^h a: ^{A1}	xa: ^{A1}	k ^h a: ^{A1}	k ^h a: ^{A1}	k ^h a: ^{A1}	ha: ^{A1}	ka: ^{A1}	kwa: ^{A1}
32	knee	*ɣow ^B	k ^h aw ^{B1}	xaw ^{B1}	k ^h aw ^{B1}	k ^h aw ^{C1}	k ^h aw ^{B1}	hø: ^{A1}	ho: ^{B1}	ko: ^{C1} -i
33	shin, lower leg	*ye:ŋ ^B	k ^h ɛŋ ^{B2}	xɛŋ ^{C1}	k ^h ɛ:ŋ ^{B2}	gɛŋ ^{B2}	ke:ŋ ^{B2}	hɛŋ ^{B1}	heŋ ^{B2}	ye:ŋ ^{B2}
34	foot	*ti:n ^A	ti:n ^{A1}	fin ^{A1}	ti:n ^{A1}	fin ^{A1}		ɬɪn ^{A1}	tin ^{A1}	ti:n ^{A1}
35	heel	*syn ^C	son ^{C1}	sun ^{C1}	ʰɪn ^{C1}	ɬɪn ^{C1}	ɬɪn ^{C1}	ɬon ^{C1}	θan ^{B1} -t	son ^{C1}
36	chest	*ʔək ^D	ʔok ^{DS1}	ʔuək ^{DS1}	ʰɪn ^{C1}	ʔɪk ^{DS1}	ʔɪk ^{DS1}	ʔak ^{DS1}	ʔak ^{DS1}	
37	back	* ^h lan ^A	lan ^{A1}	lan ^{A1}	lan ^{A1}	lan ^{A1}	lan ^{A1}	lan ^{A1}	lan ^{A1}	lan ^{A1}

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
38	waist (1)	*ʔje:w ^A	ʔew ^{A1}	ʔiw ^{A1}	ʔe:w ^{A1}		ji:w ^{A1}		ʔew ^{A1}	ʔe:w ^{A1}
39	waist (2)	*ʔ _C .wuət ^D						hut ^{DL1}	huət ^{DL1}	vuət ^{DS1}
40	stomach, belly	*dwu:ŋ ^C	tʰɔ:ŋ ^{C2}	toŋ ^{C2}	to:ŋ ^{C2}	doŋ ^{C2}	to:ŋ ^{C2}	toŋ ^{C2}	tuŋ ^{C2}	tʰuŋ ^{C2}
41	navel (1)	*dfwu:ʔ ^A	du:ʔ ^{A1}	duʔ ^{A1}		duʔ ^{A1}			duʔ ^{A1}	duʔ ^{A1}
42	navel (2)	*ʔli:ʔ ^C				di ^{C1}		boy ^{C1} -t		
43	buttock	*kon ^C	kon ^{C1}	kun ^{C1}	kon ^{C1}	kon ^{C1}	kun ^{C1}			
44	pubic hair	* ^h mwu:ʔ ^A	mɔj ^{A1}	mɔj ^{A1}	mɔj ^{A1}	mɔj ^{A1}		mɔy ^{A1}	mi ^{A1}	
45	vagina	*hi:ʔ ^A	hi:ʔ ^{A2}	hi ^{A1}	hi:ʔ ^{A1}	hi ^{A1}	hi:ʔ ^{A1}	hoj ^{A1}		
46	penis	*ywjaj ^A	k ^h uəj ^{A2}	xwəj ^{A2}	vɔj ^{A2}	wɔj ^{A2}		wəj ^{A2}		
47	testicles	*tram	ham ^{A1}	ham ^{A1}	ham ^{A1}	t ^h am ^{A1}		t ^h am ^{A1}	ram ^{A1}	ram ^{A1}
48	pubic mound	* ^h naw ^B	naw ^{B1}						naw ^{B1}	naw ^{B1}
49	urine	*niəw ^B	jiəw ^{B2}	new ^{B2}	nə:w ^{B2}	new ^{B2}		new ^{B2}	nu ^{B2}	nu ^{B2}
50	excrement	*ʔ _C .quj ^C	k ^h i:ʔ ^{C1}		k ^h i:ʔ ^{C1}	k ^h i ^{C1}	k ^h i:ʔ ^{C1}	k ^h oy ^{C1}	haj ^{C2}	yaŋ ^{C2}
51	fart	*k.tɔʔ ^D	toʔ ^{DS1}	tut ^{DS1}	tɔʔ ^{DS1}	tɔʔ ^{DS1}		t ^h oʔ ^{DS1}	rat ^{DS1}	rɛʔ ^{DS1}
52	bone	*ʔ _C .duk ^D	du:k ^{DL1}	duk ^{DS1}	du:k ^{DL1}	duk ^{DL1}	duk ^{DL1}	dok ^{DL1}	dok ^{DL1}	rɔ:k ^{DL1}
53	bone marrow	*ʔwu:k ^D			ʔɔ:k ^{DL1}	ʔɔk ^{DL1}	ʔuk ^{DL1}	ʔuk ^{DS1}	ʔuk ^{DS1}	ʔu:k ^{DL1}
54	tendon	*ʔjen ^A	ʔen ^{A1}	ʔin ^{A1}	ʔən ^{A1}	ʔien ^{A1}	jin ^{A2}	jen ^{A1}	jin ^{A2}	ʔin ^{A1}
55	side, ribs	*k.ra:ŋ ^C	k ^h ra:ŋ ^{C1}	sa:ŋ ^{C1}	sa:ŋ ^{C1}	sa:ŋ ^{C1} -i	ha:ŋ ^{C1}			
56	blood	*luət ^D	luət ^{DL2}	luut ^{DL2}	luət ^{DL2}	luət ^{DL2}	ly:t ^{DL2}	lut ^{DL2}	luət ^{DL2}	luət ^{DL2}
57	meat, flesh	*n.mv:ʔ ^C	nuə ^{C2}		nuə ^{C2}		nu:ʔ ^{C2}		no ^{B2}	mlɔ:ʔ ^{B2}
58	skin	* ^h naŋ ^A	naŋ ^{A1}	naŋ ^{A1}	naŋ ^{A1}	nang ^{A1}	naŋ ^{A1}	naŋ ^{A1}	naŋ ^{A1}	naŋ ^{A1}

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
59	sweat	*r.tuə ^B	huə ^{BI}	hu ^{BI}	huə ^{BI}	t ^h uə ^{BI}	huu: ^{BI}	ty ^{B2}		
60	pus	* ^h no:nɔ ^A	nɔ:nɔ ^{AI}	nɔ:nɔ ^{AI}	nɔ:nɔ ^{AI}	nɔ:nɔ ^{AI}	nɔ:nɔ ^{AI}	nɔ:nɔ ^{AI}	nɔ:nɔ ^{AI}	nɔ:nɔ ^{AI}
61	impetigo	*krit ^D	hit ^{DS1}		k ^h it ^{DS1}	k ^h it ^{DS1}		lit ^{DS1}		hit ^{DS1}
62	wound	* ^h a:t ^D	ba:t ^{DL1}	ba:t ^{DL1}	ba:t ^{DL1}	ba:t ^{DL2}	ba:t5		ba:t ^{DL1}	ba:t ^{DL1}
63	scurf	*yuj ^A			k ^h yj ^{A2}	g̃yj ^{A2}		hoy ^{A2}	hi ^{A2}	yi. ^{A2}
64	brain	*ʔe:k ^D		ʔeʔ ^{DL1}			ʔe:k ^{DL1}			
65	heart, breath	*cɯɯ ^A	caj ^A	cau ^{AI}	cɯɯ ^{AI}	cɯɯ ^{AI}	cau ^{AI}	coy ^{AI}	su ^{AI}	cuu. ^{AI}
66	lung	*pwɯt ^D	pɔ:t ^{DL1}		pɔ:t ^{DL1}	put ^{DS1}	pɯt ^{DS1}	put ^{DL1}	put ^{DS1}	
67	intestine	*saj ^C	saj ^{CI}	saj ^{CI}	t ^h yj ^{CI}	t ^h yj ^{CI}	ɬaj ^{CI}	ɬaj ^{CI}	θaj ^{CI}	saj ^{CI}
68	liver	*tap ^D	tap ^{DS1}	tap ^{DS1}	tap ^{DS1}	tap ^{DS1}	tap ^{DS1}	tap ^{DS1}	tap ^{DS1}	tap ^{DS1}
69	spleen	*ma:m ^C	ma:m ^{C2}		ma:m ^{C2}	ma:m ^{C2}		ma:m ^{C2}	ma:m ^{C2}	ma:m ^{C2}
70	gall bladder, bile	*bli: ^A	di: ^{AI}	bi: ^{AI}	bi: ^{AI}	di: ^{AI} -i	di: ^{AI}	boy ^{AI}	di: ^{AI}	bli: ^{AI}
71	illness, fever	*k.raj	k ^h aj ^I	saj ^{CI}	k ^h yj ^{CI}	k ^h yj ^{CI}	haj ^{CI}		caj ^{CI}	
72	cold (n.)	*q.wat ^D	wat ^{DS1}			wat ^{DS1}				vat ^{DS1}
73	epidemic	*ra: ^B	ha: ^{BI} -i			ra: ^{B2}		la: ^{CI} -t		ha: ^{BI}
74	kidney-stones	* ^h ni:w ^C	niw ^{CI}			niw ^{CI}			niw ^{CI}	ni:w ^{CI}
	B. Animals									
75	dog	* ^h ma: ^A	ma: ^{AI}	ma: ^{AI}	ma: ^{AI}	ma: ^{AI}	ma: ^{AI}	ma: ^{AI}	ma: ^{AI}	ma: ^{AI}
76	pig	* ^h mu: ^A	mu: ^{AI}	mu: ^{AI}	mu: ^{AI}	mu: ^{AI}	mu: ^{AI}	maw ^{AI}	mu: ^{AI}	mu: ^{AI}
77	mouse, rat	* ^h nu: ^A	nu: ^{AI}	nu: ^{AI}	nu: ^{AI}	nu: ^{AI}	nu: ^{AI}	naw ^{AI}	nu: ^{AI}	nu: ^{AI}

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
78	water buffalo	*ywa:j ^A	k ^h wa:j ^{A2}	xwa:j ^{A2}	wa:j ^{A2}	wa:j ^{A2}	va:j ^{A2}	wa:j ^{A2}	va:j ^{A2}	va:j ^{A2}
79	horse	*ma: ^C	ma: ^{C2}	ma: ^{C2}	ma: ^{C2}	ma: ^{C2}	ma: ^{C2}	ma: ^{C2}	ma: ^{C2}	ma: ^{C2}
80	bear	* ^h mwuj ^A	mi: ^{A1}	mi: ^{A1}	mi: ^{A1}	mi: ^{A1}	mi: ^{A1}	muəj ^{A1}	muəj ^{A1}	muəj ^{A1}
81	jackal	*nwaj ^A	naj ^{A2}		nwɔj ^{A2}	nwɔj ^{A2}		naj ^{A1}	naj ^{A2}	naj ^{A2}
82	elephant	*ja:ŋ ^C	c ^h a:ŋ ^{C2}	ca:ŋ ^{C2}	ca:ŋ ^{C2}	za:ŋ ^{C2}	ca:ŋ ^{C2}		sa:ŋ ^{C2}	sa:ŋ ^{C2}
83	deer	*kwuəŋ ^A	kwa:ŋ ^{A1}	kwa:ŋ ^{A1}	kwa:ŋ ^{A1}	kwa:ŋ ^{A1}	kwa:ŋ ^{A1}		vuuəŋ ^{A2}	vuuəŋ ^{A2}
84	antelope	*wa:l ^A	fa:n ^{A2}	fa:n ^{A2}				fa:n ^{A2}	fa:n ^{A2}	va:l ^{A2}
85	monkey	*li:ŋ ^A	liŋ ^{A2}	liŋ ^{A2}	liŋ ^{A2}	liŋ ^{A2}	liŋ ^{A2}		liŋ ^{A2}	li:ŋ ^{A2}
86	ape	*kaŋ ^A	kaŋ ^{A1}	kaŋ ^{A1}	kaŋ ^{A1}	kaŋ ^{A1}	kaŋ ^{A1}	kaŋ ^{A1}	kaŋ ^{A1}	
87	flying squirrel	*ba:ŋ ^B	ba:ŋ ^{B1}	ba:ŋ ^{B1}	ba:ŋ ^{B1}	ba:ŋ ^{B1}			ba:ŋ ^{B1}	
88	squirrel	*ro:k ^D	rɔ:k ^{DL2}	hoʔ ^{DL2}	rɔ:k ^{DL2}				ro:k ^{DL2}	rɔ:k ^{DL2}
89	porcupine	* ^h men ^C	men ^{C1}		mɛn ^{C1}	men ^{C1}	min ^{C1}	mɛn ^{C1}		man ^{C1}
90	civet cat	* ^h ne:l ^A	hen ^{A1}	hin ^{A1}	hɛn ^{A1}	hen ^{A1}	hin ^{A1}		nan ^{A1}	ne:l ^{A1}
91	otter	*na:k ^D	na:k ^{DL2}	naʔ ^{DL2}	na:k ^{DL2}	na:k ^{DL2}		na:k ^{DL2}	na:k ^{DL2}	na:k ^{DL2}
92	pangolin	*li:l ^B				lin ^{B2}		hɛn ^{B2}	lin ^{B2}	li:l ^{B1}
93	chicken	*kaŋ ^B	kaŋ ^{B1}	kaŋ ^{B1}	kaŋ ^{B1}	kaŋ ^{B1}	kaŋ ^{B1}	kaŋ ^{B1}	kaŋ ^{B1}	kaŋ ^{B1}
94	young hen	*Ç.qv:ŋ ^B	k ^h tuəŋ ^{B1}	xuəŋ ^{B1}	k ^h tuəŋ ^{B1}			k ^h ɥŋ ^{B1}		
95	duck	*pit ^D	peŋ ^{DS1}	pit ^{DS1}	peŋ ^{DS1}	peŋ ^{DS1}	pit ^{DS1}	pit ^{DS1}	pit ^{DS1}	pit ^{DS1}
96	goose	*ha:n ^B	ha:n ^{B1}	han ^{B1}	ha:n ^{B1}	ha:n ^{B1}			ha:n ^{B1}	ha:n ^{B1}
97	bird	*Ç.nok ^D	nok ^{DS2}	nuk ^{DS2}	nok ^{DS2}	nok ^{DS2}	nuk ^{DS2}		rok ^{DS2}	nok ^{DS2}
98	crow	*ka: ^A	ka: ^{A1}	ka: ^{A1}	ka: ^{A1}	ka: ^{A1}	ka: ^{A1}	ka: ^{A1}		ka: ^{A1}

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
99	swallow (n.)	*ʔe:n ^B	ʔe:n ^{B1}	ʔe:n ^{C1}	ʔe:n ^{B1}	ʔe:n ^{B1}	ʔe:n ^{B1}	ʔe:n ^{B1}	ʔe:n ^{B1}	ʔe:n ^{B1}
100	owl	*gaw ^C	k ^h aw ^{C2}	kaw ^{C2}	kaw ^{C2}	gaw ^{C2} -v	kaw ^{C2}	k ^h efj ^{A2}		
101	hawk	*lam ^B			lam ^{B2}	lam ^{B2}	lam ^{B2}			lam ^{B2}
102	fish	*pla: ^A	pla: ^{A1}	pa: ^{A1}	pja: ^{A1}	pja: ^{A1}	pja: ^{A1}	pa: ^{A1}	pja: ^{A1}	pla: ^{A1}
103	catfish	*ʔ.dok ^D	duk ^{DS1}	duk ^{DS1}	duk ^{DS1}	duk ^{DS1}		dok ^{DS1}	duk ^{DS1}	rɔk ^{DS1} -v
104	shellfish	*ho:j ^A	ho:j ^{A1}	ho:j ^{A1}	ho:j ^{A1}	ho:j ^{A1}	ho:j ^{A1}			
105	crab	*puw ^A	pu: ^{A1}	pu: ^{A1}	pu: ^{A1}	pu: ^{A1}	pu: ^{A1}	paw ^{A1}	paw ^{A1}	paw ^{A1}
106	shrimp	*kuŋ ^{B/C}	kuŋ ^{C1}	kuŋ ^{C1}	kuŋ ^{C1}	kuŋ ^{C1}	kuŋ ^{C1}	koŋ ^B	kuŋ ^{B1}	kuŋ ^{C2}
107	small shrimp	*ŋjəw ^A					ji:w ^{A2}	ŋj:w ^{A2}		
108	frog	*kɔp ^D	kop ^{DS1}	kup ^{DS1}	kop ^{DS1}	kop ^{DS1}	kup ^{DS1}	kop ^{DS1}	kap ^{DS1}	kap ^{DS1}
109	small frog	*krwe: ^C							kwe: ^{C1}	t ^h rɛ: ^{C1}
110	tree frog	*pa:t ^D	pa:t ^{DL1}	pa:t ^{DL1}	pa:t ^{DL1}	pa:t ^{DL1}			pa:t ^{DL1}	
111	turtle	*taw ^B	taw ^{B1}	taw ^{B1}	tɔw ^{B1}	tɔw ^{B1}		taw ^{B1}		
112	water tortoise	* ^h wuə ^A	fa: ^{A1}	fa: ^{A1}	p ^h a: ^{A1}	p ^h a: ^{A1}		p ^h y ^{A1}		viə ^{C1}
113	snake	*ŋwuu: ^A	ŋu: ^{A2}	ŋu: ^{A2}	ŋu: ^{A2}	ŋu: ^{A2}	ŋu: ^{A2}	ŋy: ^{A2}	ŋuə ^{A2}	ŋuə ^{A2}
114	crocodile	*ŋuək ^D	ŋuək ^{DL2}	ŋuɔ ^{DL2}	ŋuək ^{DL2}		ŋɔ:k ^{DL2}	ŋyɔk ^{DL2}	ŋuək ^{DL2}	ŋuək ^{DL2}
115	land leech	*da:k ^D	t ^h a:k ^{DL2}	taɔ ^{DL2}	ta:k ^{DL2}	da:k ^{DL2}	tek ^{DL2} -v		ta:k ^{DL1}	t ^h a:k ^{DL2}
116	aquatic leech	*pli:ŋ ^A	pliŋ ^{A1}	piŋ ^{A1}	piŋ ^{A1}	piŋ ^{A1}	piŋ ^{A1}	pɔŋ ^{A1}	piŋ ^{A1}	pi:ŋ ^{A1}
117	bedbug	*ʔ.ruəɓ ^D	ruəɓ ^{DL2}		ruəɓ ^{DL2}	luəɓ ^{DL2}	lɔ:t ^{DL2}	lut ^{DL2}		ruəɓ ^{DL2}
118	grasshopper	*p.tak ^D	tak ^{DS2}	tak ^{DS1}		t ^h ak ^{DS1}		t ^h ak ^{DS1}	tak ^{DS1}	
119	mosquito	*nuŋ ^A	juŋ ^{A2}	nuŋ ^{A2}	nuŋ ^{A2}	nuŋ ^{A2}	juŋ ^{A2}	noŋ ^{A2}	nuŋ ^{A2}	nuŋ ^{A2}

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
120	wasp	*b.twi:l ^A	tɛ:n ^{A2}	ten ^{A1}	p ^h ɛ:n ^{A1}	t ^h ɛn ^{A1}	p ^h ɛ:n ^{A1}	tin ^{A2}	tin ^{A2}	t ^h i:l ^{A2}
121	bee	*to: ^B	tɔ: ^{B1}	tɔ: ^{B1}	tɔ: ^{B1}	tɔ: ^{B1}	tɔ: ^{B1}	tø ^{B1}		tɔ: ^{B1}
122	insect	*m.le:ŋ ^A	lɛ:ŋ ^{A2}	mɛŋ ^{A2}	mje:ŋ ^{A2}	mɛŋ ^{A2}	mɛ:ŋ ^{A2}	mɛŋ ^{A2}	nɛŋ ^{A2}	mɛ:ŋ ^{A2}
123	ant	*mɯc ^D	mo:t ^{DS2}	mut ^{DS2}	mo:t ^{DS2}	mɯt ^{DS2}	mɯt ^{DS2}	mo:t ^{DS2}	mat ^{DS2}	mɛk ^{DS2}
124	termite	*mo:t ^D	mɔ:t ^{DL2}		mɔ:t ^{DL2}	mo:t ^{DL2}		mo:t ^{DL2}	mo:t ^{DL2}	mɔ:t ^{DL2}
125	worm	*hno:l ^A	nɔ:n ^{A1}	non ^{A1}	nɔ:n ^{A1}	nɔ:n ^{A1}	no:n ^{A1}	nɔn ^{A1}	non ^{A1}	nɔ:l ^{A1}
126	earthworm	*t.nuəl ^A	duən ^{A1}	duən ^{A1}	duən ^{A1}	duən ^{A1}	dx:n ^{A1}	nun ^{A1}	duən ^{A1}	truəl ^{A1}
127	gadfly	* ^h luək ^D	luəp ^{DL1} -f		luək ^{DL1}	luəp ^{DL1} -f			luək ^{DL1}	
128	head louse	*traw ^A	haw ^{A1}	haw ^{A1}	hɯw ^{A1}	t ^h ɯw ^{A1}	haw ^{A1}	t ^h aw ^{A1}	raw ^{A1}	raw ^{A1}
129	tick	*trep ^D	hep ^{DS1}	hip ^{DS1}		t ^h ep ^{DS1}		t ^h it ^{DS1} -f		rip ^{DS1}
130	flea	* ^h mat ^D	mat ^{DS1}	mat ^{DS1}	mat ^{DS1}	mat ^{DS1}	mat ^{DS1}	mat ^{DS1}	mat ^{DS1}	mat ^{DS1}
131	body louse	*m.lel ^A	len ^{A2}		mɯn ^{A2}	men ^{A2}	min ^{A2}	mɯn ^{A1}	nan ^{A2}	mlel ^{A2}
132	chicken louse	*rwɯŋ ^A	raj ^{A2}			rɯŋ ^{A2}	təj ^{A2}	loy ^{A2}	ri ^{A2}	ri: ^{A2}
133	spider	*krwa:w ^A		saw ^{C1}	sa:w ^{A1}	sa:w ^{A1}	k ^h ja:w ^{A1}	ha:w ^{A1} -i	ca:w ^{A1}	t ^h ra:w ^{A1}
134	coconut grub	*dʉŋ ^C	duəŋ ^{C1}	duəŋ ^{C1}	duəŋ ^{C1}	duəŋ ^{C1}			duŋ ^{C2}	duəŋ ^{C1}
135	stink bug	*ge:ŋl	k ^h ɛ:ŋ ^{A2}		ke:ŋ ^{A2}	gɛŋ ^{A2}	ke:ŋ ^{A2}		keŋ ^{A2}	
136	caterpillar	*C.boŋ ^C	buŋ ^{C1}	buŋ ^{C1}	boŋ ^{C1}			boŋ ^{C1}		vɔŋ ^{C2}
137	silkworm	*mo:n ^C		mon ^{C2}	mɔ:n ^{C2}	mɔn ^{C2}	mu:n ^{C2}			
138	butterfly	*ɬɯ:j		buɯ ^{C1}	buə ^{C1}	buə ^{C1}			ba: ^{C2}	ba: ^{C1}
139	centipede	*q.sip ^D	k ^h ep ^{DS1}	xip ^{DS1}	k ^h op ^{DS1}	k ^h ep ^{DS1}	k ^h ip ^{DS1}	lit ^{DS1} -f	θip ^{DS1}	
140	cockroach	*sa:p ^D	sa:p ^{DL1}			sa:p ^{DL1}	sa:p ^{DL1}	sa:p ^{DL1}		sa:p ^{DL1}

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
141	horn	*qaw ^A	k ^h aw ^{A1}		saw ^{A1}				kaw ^{A1}	kaw ^{A1}
142	tusk, ivory	*ŋa. ^A	ŋa. ^{A2}	ŋa. ^{A2}	ŋa. ^{A2}	ŋa. ^{A2}		ŋa. ^{A2}	ŋa. ^{A2}	ŋa. ^{A2}
143	wing (1)	*wuwet ^D							fuət ^{DL2}	viət ^{DL2}
144	wing (2)	*pi:k ^D	pi:k ^{DL1}		pi:k ^{DL1}	pik ^{DL2}	pik ^{DL1}	pik ^{DL1}		
145	tail	*trwɔ:ŋ ^A	ha:ŋ ^{A1}	ha:ŋ ^{A1}	ha:ŋ ^{A1}	ha:ŋ ^{A1}	ha:ŋ ^{A1}	t ^h yŋ ^{A1}	ruwəŋ ^{A1}	ruwəŋ ^{A1}
146	fish scales	*klec ^C	klet ²	kit ^{A2}	kɔt ^{A2}	kep ^{DS1} -f	kit ^{B1}	kip ^{DS1} -f	cip ^{C1} -f	tlek ^{A2}
147	fishbone	*ka:ŋ ^C	ka:ŋ ^{C1}	ka:ŋ ^{C1}		ka:ŋ ^{C1}		ka:ŋ ^{C1}		ka:ŋ ^{C1}
148	egg	*qraj	k ^h aj ²	saj ^{B1}	k ^h aj ^{B1}	saj ^{B1} -v	k ^h aj ^{B1}	laj ^{B1}	caj ^{B1}	
149	gizzard	*p.tauq ^A	taj ^{A1}	tauq ^{A1}	p ^h yw ^{A1}	t ^h yj ^{A1}		toy ^{A1}	tauq ^{A1}	tɔ: ^{A1}
150	cockscorb	*ho:n ^A	ŋɔ:n ^{A1}	hon ^{A1}	ho:n ^{A1}	hon ^{A1}	ho:n ^{A1}	høn ^{A1}		hɔ:n ^{A1}
151	cockspur	*t.nɔ:j ^A	duwej ^{A1}	duw ^{A1}	duw ^{A1}	duw ^{A1}			da ^{A1}	tra: ^{A1}
152	hump of an ox	* ^h no:k ^D	nɔ:k ^{DL1}			nɔk ^{DL1}		nok ^{DL1}	nok ^{DL1}	nɔ:k ^{DL1}
153	stinger (of a bee)	*ləj ^A	naj ^{A2} -i	ləj ^{A2}		ləj ^{A2}		ləj ^{A2}	ləj ^{A2}	ləj ^{A2}
	C. Plants									
154	tree, wood	*mwaj ^C	ma:j ^{C2}	maj ^{C2}	maj ^{C2} -v	maj ^{C2} -v	maj ^{C2}	maj ^{C2}	maj ^{C2} -i	maj ^{C2}
155	leaf	*ɬauq ^A	baɣ ^{A1}	bauq ^{A1}	bɔuq ^{A1}	bɔuq ^{A1}	bauq ^{A1}	boy ^{A1}	bauq ^{A1}	bɔ: ^{A1}
156	big leaf for wrapping	*k.to:ŋ ^A	tɔ:ŋ ^{A1}	toŋ ^{A1}	tɔ:ŋ ^{A1}	toŋ ^{A1}	to:ŋ ^{A1}	t ^h øŋ ^{A1}	roŋ ^{A1}	rɔ:ŋ ^{A1}
157	flower	*ɬlo:k ^D	do:k ^{DL1}	boɔ ^{DL1}	bjo:k ^{DL1}	bjo:k ^{DL1}	bjo:k5	bok ^{DL1}	dok ^{DL1}	blo:k ^{DL1}
158	banana blossom	*pli: ^A	pli: ^{A1}	pi: ^{A1}	pi: ^{A1}		pi: ^{A1}			pli: ^{A1}
159	stem	*ka:n ^C	ka:n ^{C1}	ka:n ^{C1}					ka:n ^{C1}	ka:n ^{C1}

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
160	peel, bark	*plwək ^D	plwək ^{DL1}	puɿ ^{DL1}		puwək ^{DL1}	pɿ:k ^{DL1}			pla:k ^{DL1} - v
161	husk	*ka:p ^D	ka:p ^{DL1}	ka:p ^{DL1}	ka:p ^{DL1}				ka:p ^{DL1}	ka:p ^{DL1}
162	thorn	* ^h na:m ^A	na:m ^{A1}	na:m ^{A1}	na:m ^{A1}	na:m ^{A1}	na:m ^{A1}	na:m ^{A1}		
163	fruit	* ^h ma:k ^D	ma:k ^{DL1}	maɿ ^{DL1}	ma:k ^{DL1}	ma:k ^{DL1}	ma:k ^{DL1}	ma:k ^{DL1}	ma:k ^{DL1}	ma:k ^{DL1}
164	sheath, pod	*q wak ^D	fak ^{DS1}		p ^h ak ^{DS1}	p ^h ak ^{DS1}		p ^h ak ^{DS1}	hok ^{DS1}	vak ^{DS1}
165	grain	*m.lec ^D	met ^{DS2}	mit ^{DS2}	mjet ^{DS2}				nat ^{DS2}	mlek ^{DS2}
166	stump	*to: ^A	tɔ: ^{A1}		tɔ: ^{A1}	tɔ: ^{A1}	tɔ: ^{A1}			
167	root	*Ç.ra:k ^D	ra:k ^{DL2}	haɿ ^{DL2}	ra:k ^{DL2}	la:k ^{DL2}	la:k ^{DL2}		ra:k ^{DL2}	ra:k ^{DL1}
168	clump (as of bamboo)	*ko: ^A	kɔ: ^{A1}	ko: ^{A1}	ko: ^{A1}	kɔ: ^{A1}	ko: ^{A1}	kɔ: ^{A1}	ko: ^{A1}	
169	sprout, shoot	* ^h no: ^B	no: ^{B1}	no: ^{B1}	no: ^{B1}	nɔ: ^{B1}	no: ^{B1}	nɔ: ^{B1}		
170	bamboo shoot	*r.na:ŋ ^A							ra:ŋ ^{A2}	na:ŋ ^{A2}
171	ear (of rice)	*rwu:ŋ ^A	ruəŋ ^{A2}		ruəŋ ^{A2}	ruəŋ ^{A2}	tu:ŋ ^{A2}	luŋ ^{A2}		ruəŋ ^{A2}
172	rice	*Ç.qaw ^C	k ^h a:w ^{C1}	xaw ^{C1}	k ^h ɿw ^{C1}	k ^h ɿw ^{C1}	k ^h aw ^{C1}	haw ^{C2}	haw ^{C2}	yaw ^{C2}
173	rice seedling	*kla: ^C	kla: ^{C1}	ca: ^{C1}	ca: ^{C1}	ca: ^{C1}	kja: ^{C1}	ka: ^{C1}	ca: ^{C1}	tla: ^{C1}
174	husked rice	*sa:l ^A	sa:n ^{A2}	sa:n ^{A1}	t ^h a:n ^{A1}	la:n ^{A1}	la:n ^{A1}		θa:n ^{A1}	sa:l ^{A1}
175	millet	* ^h wuətəŋ ^C	fa:ŋ ^{C1}		p ^h a:ŋ ^{C2}	p ^h a:ŋ ^{C1}		p ^h ɿŋ ^{A1}		viəŋ ^{C1}
176	resin	* ^h ja:ŋ ^A	ja:ŋ ^{A1}	ja:ŋ ^{A1}	ja:ŋ ^{A1}	ja:ŋ ^{A1}	ja:ŋ ^{A1}	jɿŋ ^{A1}		ja:ŋ ^{A1}
177	banana	*kluəj ^C	kluəj ^{C1}	kuj ^{C1}	kuəj ^{C1}	kuəj ^{C1}	ku:j ^{C1}	koy ^{C1}	cuəj ^{C1}	
178	sugarcane	*ʔo:j ^C	ʔɔj ^{C1}	ʔɔj ^{C1}	ʔɔ:j ^{C1}	ʔɔj ^{C1}	ʔɔ:j ^{C1}	ʔoy ^{C1}	ʔɔj ^{C1}	ʔɔ:j ^{C1}
179	bamboo	*praŋ ^B	p ^h əj ^{B1}		p ^h ɿ:ŋ ^{A1}	p ^h jeŋ ^{C1}	p ^h əj ^{B1}	p ^h əj ^{B1}		p ^h raŋ ^{B1}
180	cucumber, melon	*p.rwe:ŋ ^A	tɛ:ŋ ^{A1}	teŋ ^{A1}		t ^h ɛ:ŋ ^{A1}		t ^h ɛ:ŋ ^{A1}	tiəŋ ^{A1}	priəŋ ^{A1}

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
181	ashgourd	*wak ^D	fak ^{DS2}	fak ^{DS2}		vak ^{DS2}	fak ^{DS2}	fak ^{DS2}		vak ^{DS2}
182	wild olive	*ko:k ^D	ko:k ^{DL1}		ko:k ^{DL1}				kok ^{DL1}	ko:k ^{DL1}
183	starfruit	*gwuəŋ ^A	fuəŋ ^{A2}				fə:ŋ ^{A2}	fɤŋ ^{A2}		viəŋ ^{A1}
184	plum	*man ^C		man ^{C2}	man ^{C2}	man ^{C1}		man ^{C1}		
185	pomelo	*bu:k ^D		puk ^{DL2}	puk ^{DS2} -v	buk ^{DL2}	puk ^{DL2}		puk ^{DS2}	
186	acacia	*ge: ^A	k ^h ɛ: ^{A2}						kɛ ^{A2}	
187	banyan	*raj ^A		haj ^{B2}		rɔj ^{A2}				haj ^{A2}
188	vegetable	*prak ^D	p ^h ak ^{DS1}	fak ^{DS1} -i	p ^h jak ^{DS1}	p ^h jak ^{DS1}	p ^h jak ^{DS1}		piək ^{DS1}	p ^h rak ^{DS1}
189	morning glory	*buŋ ^C	buŋ ^{C1}	buŋ ^{C1}	buŋ ^{C1}	buŋ ^{C1}	buŋ ^{C1}	boŋ ^{C1}		
190	mustard green	*ka:t ^D	ka:t ^{DL1}	ka:t ^{DL1}	ka:t ^{DL1}	ka:t ^{DL1}	ka:t ^{DL1}	ka:t ^{DL1}	ka:t ^{DL1}	
191	yam	*man ^A	man ^{A2}		man ^{A2}	man ^{A2}	man ^{A2}	man ^{A2}	man ^{A2}	man ^{A2}
192	taro	*pruək ^D	p ^h uək ^{DL1}	p ^h ur ^{DL1}	p ^h uək ^{DL1}	p ^h uək ^{DL1}	p ^h ɤ:k ^{DL1}	p ^h ɤ:k ^{DL1}	puək ^{DL1}	
193	ginger	*ɣi:ŋ ^A	k ^h iŋ ^{A1}	xiŋ ^{A1}	k ^h iŋ ^{A1}		k ^h iŋ ^{A1}	hiŋ ^{A1}	hiŋ ^{A1}	hi:ŋ ^{A1}
194	galangal	*xa: ^B	k ^h a: ^{B1}	xa: ^{B1}	k ^h a: ^{B1}	k ^h a: ^{B1}		ha: ^{B1}	ha: ^{B1}	ha: ^{B1}
195	sesame	*ɾ.ŋa: ^A	ŋa: ^{A2}	ŋa: ^{A2}	ŋa: ^{A2}	ŋa: ^{A2}	ŋa: ^{A2}	ŋa: ^{A2}	ra ^{A2}	
196	tea	*ja: ^A	c ^h a: ^{A2}			za: ^{A2}	ca: ^{A2}	sa: ^{A2}	sa: ^{A2}	sa: ^{A2}
197	raisin	*ɕit ^D		ɕit ^{DS1}	ɕit ^{DS1}	ɕit ^{DS1}		ɕit ^{DS1}		ɕit ^{DS1}
198	rattan	*ɕ.wa:j ^A	wa:j ^{A1}	wa:j ^{A1}	wa:j ^{A1}	wa:j ^{A1}	wa:j ^{A1}	wa:j ^{A1}	va:j ^{A1}	va:j ^{A1}
199	reed	*ɕo: ^C	ɕo: ^{C1}	ɕo: ^{C1}	ɕo: ^{C1}	ɕo: ^{C1}	ɕo: ^{C1}	ɕo: ^{C1}		
200	grass	*hjuə ^C	ja: ^{C1}	ja: ^{C1}	ja: ^{C1}	ja: ^{C1}	ja: ^{C1}	ny: ^{C1}	niə ^{C1}	juə ^{C1}
201	thatch grass	*ya: ^A	k ^h a: ^{A2}	xa: ^{A2}	k ^h a: ^{A2}	gä: ^{A2}	ka: ^{A2}	ha: ^{A2}	ha: ^{A2}	ya: ^{A2}

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
202	mushroom	* ^h rwet ^D	het ^{DS1}	hit ^{DS1}	hət ^{DS1}		vit ^{DS2}	lit ^{DS1}	rat ^{DS1}	
203	fern	*ku.t ^D	ku.t ^{DL1}	ku.t ^{DS1}	ku.t ^{DL1}	ku.t ^{DL1}	ku.t ^{DL1}	ku.t ^{DL1}	ku.t ^{DS1}	ku.t ^{DL1}
204	duckweed	* ^h ne. ^A	ne. ^{A1}	ne ^{A1}	ne. ^{A1}	ne ^{A1}	ne. ^{A1}			ne. ^{A1}
205	aquatic moss	*daw ^A	t ^h aw ^{A2}	taw ^{A2}	tw ^{A2}	ɗɔw ^{A2}			taw ^{A2}	
	D. Nature									
206	water	*Ç.nam ^C	na.m ^{C2}	nam ^{C2}	nam ^{C2}	nam ^{C2}	nam ^{C2}	nam ^{C2}	ram ^{C2}	nam ^{C2}
207	fire	*wxj ^A	faj ^{A2}		p ^h xj ^{A2}	vɔj ^{A2}	faj ^{A2}	foy ^{A2}	fi ^{A2}	vi. ^{A2}
208	flame	*ple.w ^A	ple.w ^{A2}	pew ^{A1}	pje:w ^{A1}	pjew ^{A1}		pew ^{A1}		ple:w ^{A1}
209	smoke	*ywan ^A	k ^h wan ^{A2}	van ^{A2}	wan ^{A2}	wan ^{A2}	van ^{A2}	wan ^{A2}		yon ^{A2}
210	soot	* ^h mi. ^C	mi. ^{C1}	mi ^{C1}	mi. ^{C1}	mi. ^{C1}	mi. ^{C1}		mi ^{C1}	mi. ^{C1}
211	hard firewood	*wu.r ^A	fu.r ^{A2}	fun ^{A2}	p ^h u.r ^{A2}	wuun ^{A2}	fɔn ^{A2}		fun ^{A2}	vuul ^{A1}
212	bamboo firewood	* ^h luwəw ^A			luə ^{A1}	luə ^{A1}		liw ^{A1}		
213	leaf ashes	*braw ^B			pjɔw ^{B2}		pjaw ^{B2}			
214	ashes (wood)	*daw ^B	t ^h aw ^{B2}	taw ^{B2}	tw ^{B2}	ɗɔw ^{B2}		taw ^{B2}	taw ^{B2}	t ^h aw ^{B2}
215	iron (1)	* ^h lek ^D	lek ^{DS1}	lik ^{DS1}	lek ^{DS1}	lek ^{DS1}	lik ^{DS1}	lik ^{DS1}		
216	iron (2)	*mwa. ^A							fa ^{A2}	ma. ^{A2}
217	salt	*klwuə ^A	klwə ^{A1}	kuw ^{A1}	kuwə ^{A1}	kuwə ^{A1}	kuw. ^{A1}	ky ^{A1}	kuə ^{A1}	tlwə ^{A1}
218	lye	*ɗaŋ ^B	da.ŋ ^{B1} -v			daŋ ^{B1}	daŋ ^{B1}	daŋ ^{B1}	daŋ ^{B1}	daŋ ^{B1}
219	grease, fat	*man ^A	man ^{A2}			man ^{A2}				man ^{A2}
220	soil	*tuum ^A	tom ^{A1}			tom ^{A1}	tum ^{A1}	tom ^{A1}		

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
221	earth	*dɪn	dɪn ^{A2}	dɪn ^{A1}	dɪn ^{A1}	dɪn ^{A1}				
222	sand	*zwuɛj ^A	sa:j ^{A2}	sa:j ^{A2}	tʰa:j ^{A2}	ra:j ^{A2}	lɑ:j ^{A2}	luy ^{A2}	θiej ^{A2}	jo:j ^{A2}
223	stone	*tri:j ^A	hin ^{A1}	hin ^{A1}	hi:n ^{A1}	tʰin ^{A1}	hin ^{A1}	tʰɻn ^{A1}	rin ^{A1}	ri:j ^{A1}
224	dry land	*ʂok ^D	bok ^{DS1}	buk ^{DS1}	bok ^{DS1}	bok ^{DS1}	buk5		bok ^{DS1}	bok ^{DS1}
225	wild	*C.tuən ^B	tʰuən ^{B1}	tʰuən ^{B1}	tʰuən ^{B1} -v	tʰuən ^{C1} -t	tʰɻ:n ^{A1}	tun ^{B2}	tuən ^{B2}	tʰuəl ^{B2}
226	forest	*dɔŋ ^A	dɔŋ ^{A1}	dɔŋ ^{A1}	dɔŋ ^{A1}	dɔŋ ^{A1}	dɔŋ ^{A1}	dɔŋ ^{A1}	dɔŋ ^{A1}	dɔŋ ^{A1}
227	grove, wood	*pa: ^B	pa: ^{B1}	pa: ^{B1}	pa: ^{B1}	pa: ^{B1}	pa: ^{B1}	pa: ^{B1}		pa: ^{B1}
228	place, ground	*di: ^B	tʰi: ^{B2}	tʰi: ^{B2}	ti: ^{C2}	d̪i: ^{B2}	ti: ^{B2}	toy ^{B2}	ti: ^{B2}	tʰi: ^{B1}
229	mountain	*C.do:j ^A	dɔ:j ^{A2}			dɔ:j ^{A1}		dɔy ^{A1}	dɔj ^{A1}	rɔ:j ^{A1}
230	stone mountain	*pra: ^A	pʰa: ^{A1}	pʰa: ^{A1}	pʰja: ^{A1}	pʰja: ^{A1}	pʰja: ^{A1}	pʰa: ^{A1}		pʰra: ^{A1}
231	cave	*cram ^C	tʰam ^{C1}	tʰam ^{C1}	tʰam ^{C1}	tʰam ^{C1}				tʰam ^{C1}
232	hole	*ru: ^A	ru: ^{A2}	hu: ^{A2}	ru: ^{A2}	ru: ^{A2}	tu: ^{A2}			
233	crack, hole	*jo:ŋ ^B	c ^h ɔŋ ^{B2}	coŋ ^{A2} -t					soŋ ^{B2}	so:ŋ ^{B2}
234	pit	*C.kum ^A	k ^h um ^{A1}		sum ^{A1} -i		k ^h um ^{A1}	k ^h om ^{A2}	kum ^{A2}	k ^h um ^{A2}
235	well, spring	*ʂo: ^B	bo: ^{B1}	bo ^{B1}	bo: ^{B1}	bo ^{B1}	bo: ^{B1}	bo ^{B1}	bo ^{B1}	bo: ^{B1}
236	mountain stream	*qrwɻj ^C	huɛj ^{C1}	huj ^{C2} -t	k ^h uɛj ^{C1}	k ^h uɛj ^{C1}	vu:j ^{C1}		vi ^{C1}	ri: ^{C1}
237	creek	*ro:ŋ ^B	rɔ:ŋ ^{B2}	hoŋ ^{B2}	rɔ:ŋ ^{C2} -t	rɔŋ ^{B2}	ta: ^{B2}	loŋ ^{B1}		
238	river	*da: ^B	tʰa: ^{B2}	ta: ^{B2}	da: ^{B2}	da: ^{B2}	ta: ^{B2}	ta: ^{B2}	ta ^{B2}	tʰa: ^{B2}
239	river bank	* ^h waŋ ^B	faŋ ^{B1}		p ^h aŋ ^{B1}			p ^h aŋ ^{B1}		vaŋ ^{B1}
240	beach, sandbar	*ha:t ^D	ha:t ^{DL1}		ha:t ^{DL1}	ha:t ^{DL1}	ha:t ^{DL1}			ha:t ^{DL2}
241	mud	*lɻm ^B	lom ^{B1}			lum ^{B1}	lum ^{B1}		lam ^{B1}	

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
242	moon, month	*ɬlueŋ ^A	duen ^{A1}	buen ^{A1}	buen ^{A1}	buen ^{A1}	br:n ^{A1}	bun ^{A1}	duen ^{A1}	blien ^{A1}
243	star (general)	*t.na:w ^A	da:w ^{A1}	da:w ^{A1}	da:w ^{A1}	da:w ^{A1}	da:w ^{A1}	da:w ^{A1}	da:w ^{A1}	tra:w ^{A1}
244	star (in the sky)	*dfi: ^B		di: ^{B1}		di: ^{B1}	di: ^{B1}	doy ^{B1}	di: ^{B1}	
245	moonlight	* ^h ŋa:j ^A	ŋa:j ^{A1}		ha:j ^{A1}	ha:j ^{A1}	ha:j ^{A1}	ha:j ^{A1}		
246	sunshine	*Ç.dwi:t ^D	dɛ:t ^{DL1}	dɛt ^{DL1}	dɛ:t ^{DL1}	dɛt ^{DL2}	dɛ:t ^{DL1}	dit ^{DS1}	dit ^{DS1}	ri:t ^{DL1}
247	cloud	* ^h wuə ^C	fa: ^{C1}			p ^h a: ^{C1}	p ^h a: ^{C1}	p ^h y ^{C1}	vuu ^{C1}	viə ^{C1}
248	fog	* ^h mo:k ^D	mɔ:k ^{DL1}	mɔ ^ɔ ^{DL1}	mɔ:k ^{DL1}	mɔk ^{DL1}	mɔ:k ^{DL1}	mɔk ^{DL1}	mɔk ^{DL1}	mɔ:k ^{DL1}
249	wind	*Ç.lum ^A	lom ^{A2}		lom ^{A2}	lom ^{A2}	lum ^{A2}	lom ^{A2}	rum ^{A2}	lum ^{A2}
250	sky, heaven	*ɬum ^A	bon ^{A1}		bon ^{A1}	bɔn ^{A1}		bon ^{A1}	bun ^{A1}	bun ^{A1}
251	sky, weather	*va: ^C	fa: ^{C2}	fa: ^{C2}		fa: ^{C2}	fa: ^{C2}			
252	rain	*Ç.wuŋ ^A	fon ^{A1}		p ^h ɔn ^{A1}	p ^h ɔn ^{A1}	p ^h ɔn ^{A1}	p ^h on ^{A1}	hun ^{A1}	vun ^{A1}
253	gust (of rain)	*kra: ^B	ha: ^{B1}		sa: ^{B1}	sa: ^{B1}	ha: ^{B1}	la: ^{B1}	ra: ^{B1}	ra: ^{B1}
254	lightning	*m.le:p ^D	lɛ:p ^{DL2}			mjɛp ^{A2}	mɛ:p ^{DL2}			
255	thunder	*pra: ^C							pja ^{C1}	p ^h ra: ^{C1}
256	hoarfrost	* ^h muəj ^A		muj ^{B2}	muəj ^{A1}	muəj ^{A1}		muy ^{A1}		
257	dew, mist	*Ç.nwa:j ^A					na:j ^{A2}	na:j ^{A2}		
258	hail	*trep ^D	hɛp ^{DS1}	hip ^{DS1}	hop ^{DS1}	t ^h ɛp ^{DS1}	hat ^{DS1} -f	t ^h it ^{DS1}	rip ^{DS1}	
259	steam, vapor	*s.ʔwɔ:j ^A	ʔaj ^{A2} -v	ʔa:j ^{A1}	ʔa:j ^{A1}	ja:j ^{A1}	ja:j ^{A1}		θuəj ^{A1}	so:j ^{A1}
260	shade	*ɾɪm ^B	rom ^{B2}	hom ^{B2}	rom ^{B2}				ram ^{B2}	
261	shadow, reflection	*ŋaw ^A	ŋaw ^{A2}	hom ^{B2}	ŋɔw ^{A2}	ŋɔw ^{A2}	ŋaw ^{A2}	ŋaw ^{A2}	ŋaw ^{A2}	ŋaw ^{A2}
262	dry season	*Ç.le:ŋ ^C	lɛ:ŋ ^{C2}	leŋ ^{C2}	lɛ:ŋ ^{C2}	lɛŋ ^{C2}	le:ŋ ^{C2}	leŋ ^{C2}	reŋ ^{C2}	

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
	E. Society									
263	father, man	*bo. ^B	p ^h ɔ. ^B	po. ^B	po. ^{B2}	bɔ. ^{B2}	po. ^{B2}	po. ^{C2} -t	po. ^{B2}	p ^h ɔ. ^{B2}
264	mother; woman	*me. ^B	mɛ. ^{B2}	mɛ. ^{B2}	mɛ. ^{B2}	mɛ. ^{B2}	mɛ. ^{B2}	mɛ. ^{B2}	mɛ. ^{B2}	mɛ. ^{B2}
265	older sibling	*bi. ^B	p ^h i. ^{B2}	pi. ^{B2}	pi. ^{C2} -t	bi. ^{B2}	pi. ^{B2}	pi. ^{C2} -t	pi. ^{C2} -t	p ^h i. ^{B2}
266	younger sibling	*nwo. ^C	nɔ. ^{C2}	noŋ. ^{C2}	no. ^{C2}	noŋ. ^{C2}	no. ^{C2}	nøŋ. ^{C2}	nuøŋ. ^{C2}	nuøŋ. ^{C2}
267	paternal grandfather	*puw. ^B	pu. ^{B1}	pu. ^{B1}	pu. ^{B1}	pu. ^{B1}				
268	maternal grandfather	*ta. ^A	ta. ^{A1}	ta. ^{A1}	ta. ^{A1}	ta. ^{A1}	ta. ^{A1}	ta. ^{A1}	ta. ^{A1}	ta. ^{A1}
269	maternal grandmother	*na.j. ^A	ja.j. ^{A2} -i	na.j. ^{A2}	na.j. ^{A2}	ta.j. ^{B1} -i	ta.j. ^{A1}	ta.j. ^{A1}	ta.j. ^{B1} -i	?a.j. ^{A2} -i
270	child (offspring)	*lu.k. ^D	lu.k. ^{DL2}	lu? ^{DL2}	luk. ^{DL2}	luk. ^{DL2}	luk. ^{DL2}	lak. ^{DL1}	luuk. ^{DS2}	luuk. ^{DL1}
271	great-grandparent	*juə. ^C	c ^h uə. ^{C2}	cu. ^{C2}	cuə. ^{C2}	zuə. ^{C2}				
272	nephew, niece, or grandchild	*h _l a.n. ^A	la.n. ^{A1}	la.n. ^{A1}	la.n. ^{A1}	la.n. ^{A1}	la.n. ^{A1}	la.n. ^A	la.n. ^{A1}	la.n. ^{A1}
273	great-grandchild	*h _l en. ^C	le.n. ^{A1} t	lin. ^{A1} -t	h _l n. ^{A1} -t	h _l n. ^{C1}		h _l n. ^{C1}		
274	parent's older brother	*luŋ. ^A	luŋ. ^{A2}		luŋ. ^{A2}	luŋ. ^{A2}	luŋ. ^{A2}			luŋ. ^{A2}
275	parent's older sister	*pa. ^C	pa. ^{C1}	pa. ^{C1}	pa. ^{C1}	pa. ^{C1}	pa. ^{C1}	pa. ^{C1}	pa. ^{C1}	pa. ^{C1}
276	father's younger sister	*ʔa. ^A	?a. ^{A1}	?a. ^{A1}	?a. ^{A1}	?a. ^{A1}	?a. ^{A1}			
277	mother's younger sibling	*na. ^C	na. ^{C2}	na. ^{C2}	na. ^{C2}	na. ^{B1} -t	na. ^{B2} t	na. ^{B2} -t	na. ^{C2}	na. ^{C2}
278	father's younger brother	*ʔa.w. ^A	?a.w. ^{A1}	?a.w. ^{A1}	?a.w. ^{A1}	?a.w. ^{A1}	?a.w. ^{B1}		?a.w. ^{A1}	?a.w. ^{B2}
279	wife of man's younger brother	*luəw. ^A			luə. ^{A2}	luə. ^{A2}	lu. ^{A2}	liw. ^{A2}		luəw. ^{A2}
280	wife	*miə. ^A	miə. ^{A2}	mi. ^{A2}	miə. ^{A2}					

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
281	son-in-law	*Ç.kuəj ^A	k ^h ɔːj ^{A1}		k ^h uəj ^{A1}	k ^h uəj ^{A1}	k ^h ɔːj ^{A1}		kuej ^{A2}	k ^h ɔːj ^{A2}
282	daughter-in-law	*bau ^C	p ^h aj ^{C2}	pau ^{C2}	pɔu ^{C2}	ɔu ^{C2}			pau ^{C2}	k ^h wɔː ^{C2} -i
283	person, human being	*gwuun ^A	k ^h on ^{A2}	kuun ^{A2}	kɔn ^{A2}	gɔn ^{A2}	kɔn ^{A2}	hon ^{A2}	hun ^{A2}	hun ^{A2}
284	child (young person)	*dək ^D	dek ^{DS1}	diŋ ^{A1} -f	diŋ ^{A1} -f	dek ^{DS1}	diŋ ^{DS1}	dik ^{DS1}		
285	man, male	*zaːj ^A	c ^h aːj ^A	caːj ^{A2}	caːj ^{A2}	zaːj ^{A2}	caːj ^{A2}		θaːj ^{A1}	saːj ^{A1}
286	unmarried man	*ɬaːw ^B	baːw ^{B1}	baːw ^{B1}	baːw ^{B1}	baːw ^{B1}	baːw ^{B1}	baːw ^{B1}	baːw ^{B1}	baːw ^{B1}
287	unmarried woman	*saːw ^A	saːw ^{A1}	saːw ^{A1}	ʰaːw ^{A1}	ʰaːw ^{A1}	ʰaːw ^{A1}	ʰaːw ^{A1}	θaːw ^{A1}	saːw ^{A1}
288	girl	*ɬuːk ^D							buuk ^{DS1}	buuk ^{DS1}
289	lady	*naːŋ ^A	naːŋ ^{A2}	naːŋ ^{A2}	naːŋ ^{A2}	naːŋ ^{A2}				naːŋ ^{A2}
290	related by marriage	*t.noːŋ ^A	doːŋ ^{A1}	doŋ ^{A1}	doːŋ ^{A1}	doŋ ^{A1}			doŋ ^{A1}	troːŋ ^{A1}
291	widowed	* ^h maːj ^C	maːj ^{C1}	maːj ^{B2} -t	maːj ^{C1}	maːj ^{C1}	maːj ^{C1}	maːj ^{C1}	maːj ^{B1}	maːj ^{C2}
292	orphan	*gm.raː ^C	p ^h raː ^{C2}	paː ^{C2}	pjaː ^{C2}	ɣjaː ^{C2}	pjaː ^{C2}			ʰraː ^{C2}
293	name	*jɔː ^B	c ^h uː ^{B2}						so ^{B2}	soː ^{B2}
294	master, owner	*ɛw ^C	caːw ^{C1}		cɔw ^{C1}	cɔw ^{C1}	caw ^{C1}		θu ^{C1}	suː ^{C1}
295	slave (1)	*ɣɔːj ^B	k ^h ɔːj ^{C1} -t		k ^h ɔːj ^{C1} -t	k ^h ɔːj ^{C1} -t	k ^h ɔːj ^{C1} -t	hoy ^{B1}	hoj ^{B1}	hoːj ^{B1}
296	slave (2)	*kraː ^C	k ^h aː ^{C1}	xaː ^{C1}	saː ^{C1}					
297	Vietnamese	*keːw ^A	keːw ^{A1}	kew ^{A1}	keːw ^{A1}	kew ^{A1}	keːw ^{A1}	kew ^{A1}	cew ^{A1}	keːw ^{A1}
298	shaman	* ^h moː ^A	moː ^{A1}	mo ^{A1}	moː ^{A1}	mo ^{A1}	moː ^{A1}		mo ^{A1}	moː ^{A1}
299	spirit (1)	*priː ^A	p ^h iː ^{A1}	p ^h iː ^{A1}	p ^h iː ^{A1}	p ^h iː ^{A1}	p ^h iː ^{A1}			
300	spirit (2)	*mwaːŋ ^A				maːŋ ^{A2}		maːŋ ^{A2}	faːŋ ^{A2}	maːŋ ^{A2}
301	soul; whorl in the hair	*qwan ^A	k ^h wan ^{A1}	xwan ^{A1}	k ^h wan ^{A1}	k ^h wan ^{A1}	k ^h wan ^{A1}	wan ^{A1}	van ^{A1}	hon ^{A1}

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
302	village	*ʔa:n ^C	ba:n ^{C1}	ba:n ^{C1}	ba:n ^{C1}	ba:n ^{C1}	ba:n ^{C1}	ba:n ^{C1}	ba:n ^{C2}	ba:n ^{C1}
303	township	*m ^A uəŋ	m ^{A2} uəŋ	m ^{A2} uəŋ	m ^{A2} uəŋ	m ^{A2} uəŋ				
304	debt	* ^h ni: ^C	ni: ^{C1}		ni: ^{C1}	ni: ^{C1}			ni: ^{C1}	
	F. Material culture									
305	liquor	* ^h law ^C	law ^{C1}	law ^{C1}	ɬw ^{C1}	ɬw ^{C1}	law ^{C1}	law ^{C1}	law ^{C1}	law ^{C1}
306	medicine	* ^ʔ juə ^A	ja: ^{A1}	ja: ^{A1}	ja: ^{A1}	ja: ^{A1}	ja: ^{A1}	jy: ^{A1}	jə: ^{A1}	(ja: ^{A1})
307	vinegar	* ^h mi: ^B		m ^{B1} i	m ^{B1} i	m ^{B1} i	m ^{B1} i	moy ^{B1}		
308	flour	* ^ʔ uə ^A		buə ^{A1}	buə ^{A1}	buə ^{A1}	bu: ^{A1}			buə ^{A1}
309	house	* ^ʔ ɿ:n ^A	ruən ^{A2}	huən ^{A2}	ruən ^{A2}	ruən ^{A2}	ɬ: ⁿ A2	lu:n ^{A2}	ra:n ^{A2}	ra:n ^{A2}
310	granary	* ^ʔ juəw ^C	ja:w ^{C1}		ja:w ^{B1}	ja:w ^{C1}			jəw ^{C2}	jəw ^{C1}
311	stake	* ^h lak ^D	lak ^{DS1}		lak ^{DS1}	lak ^{DS1}			lak ^{DS1}	lak ^{DS1}
312	eaves	*za:j ^A	c ^h a:j ^{A2}		ca:j ^{A2}				θa:j ^{A2}	ja:j ^{A2}
313	door	*tu: ^A	tu: ^{A1}	tu: ^{A1}	tu: ^{A1}	tu: ^{A1}	tu: ^{A1}	taw ^{A1}	tu: ^{A1}	tu: ^{A1}
314	stairs, ladder	* ^ʔ fwaj ^A	daj ^{A2}	daj ^{A1}	dwaj ^{A1}	dwaj ^{A1}	daj ^{A1}	daj ^{A1}	ləj ^{A1}	raj ^{A1}
315	pillar	*saw ^A	saw ^{A1}	saw ^{A1}	ɬw ^{A1}	ɬw ^{A1}	ɬaw ^{A1}	ɬaw ^{A1}	θaw ^{A1}	
316	partition, lid	* ^h wa: ^A	fa: ^{A1}	fa: ^{A1}	p ^h a: ^{A1}	p ^h a: ^{A1}	p ^h a: ^{A1}		va: ^{A1}	va: ^{A1}
317	split bamboo flooring	*wa:k ^D	fa:k ^{DL2}		p ^h a:k ^{DL2}	va:k ^{DL2}				va:k ^{DL2}
318	window	*ta:ŋ ^B	ta:ŋ ^{B1}	ta:ŋ ^{B1}	ta:ŋ ^{B1}	ta:ŋ ^{C1}	ta:ŋ ^{B1}			
319	stool	*taŋ ^B	taŋ ^{B1}	taŋ ^{B1}	taŋ ^{B1}	taŋ ^{B1}	taŋ ^{B1}	taŋ ^{B1}	taŋ ^{B1}	taŋ ^{B1}
320	tripod	*giəŋ ^A	k ^h iəŋ ^{A2}		kiəŋ ^{A2}	giəŋ ^{A2}	ki:ŋ ^{A2}	k ^h iŋ ^{A2}	ciəŋ ^{A2}	k ^h iəŋ ^{A2}

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
321	board	*pe:n ^C	pen ^{C2}	pen ^{C1}			pe:n ^{C1}	pen ^{C1}		pe:n ^{C1}
322	thing	*ɣɔ:ŋ ^A	k ^b ɔ:ŋ ^{A1}	xoŋ ^{A1}				høŋ ^{A1}		ho:ŋ ^{A1}
323	knife	*mit ^D	mi:t ^{DL2} -v	mit ^{DS2}	mit ^{DS2}				mit ^{DS2}	
324	machete, big knife	*ɣm.ra: ^C	p ^h ra: ^{C2}	pa: ^{C2}	pja: ^{C2}	bja: ^{C2}	pja: ^{C2}		sa ^{C2}	t ^h a: ^{C2}
325	chopping board	*ɣjəŋ ^A	k ^h əŋ ^{A1}	xiəŋ ^{A1}			k ^h i:ŋ ^{A1}	hiŋ ^{A1}	heŋ ^{A1}	
326	spear	*kro:k ^D	hɔ:k ^{DL1}						rok ^{DL1}	
327	crossbow	*h ₁ nwuə ^C	na: ^{C1}		na: ^{C1}	na: ^{C1}			nuə ^{C1}	nuə ^{C1}
328	axe	*xwa:n ^A	k ^h wa:n ^{A1}	xwa:n ^{A1}	k ^h wa:n ^{A1}	k ^h wa:n ^{A1}			va:n ^{A1}	va:n ^{A1}
329	chisel	*stəw ^B	siw ^{B1}	siw ^{B1}	t ^h i:w ^{B1}	fiw ^{B1}	fi:w ^{B1}	fiw5	θiəw ^{B1}	si:w ^{B1}
330	hook	*xɔ: ^A	k ^b ɔ: ^{A1}	xo ^{A1}	k ^b ɔ: ^{A1}					ho: ^{A1}
331	pliers, thongs	*gi:m ^A	k ^h i:m ^{A2}	kim ^{A2}	ki:m ^{A2}	gim ^{A2}	kim ^{A2}		cim ^{A2}	k ^h i:m ^{A2}
332	handle (of a knife)	*də:m ^C	da:m ^{C1}	dam ^{B1}	da:m ^{C1}	da:m ^{C1}	da:m ^{C1}		da:m ^{C2}	da:m ^{C1}
333	handle, rod	*gal ^A	k ^h an ^{A2}					k ^h an ^{A1}	kan ^{A2}	k ^h aj ^{A2}
334	carrying pole (1)	*tra:p ^D	ha:p ^{DL1}	ha:p ^{DL1}	ha:p ^{DL1}	t ^h a:p ^{DL1}	ha:p ^{DL1}	t ^h a:p ^{DL1}	ra:p ^{DL1}	ra:p ^{DL1}
335	carrying pole (2)	*ga:n ^A	k ^h a:n ^{A2}	ka:n ^{A2}	ka:n ^{A2}		ka:n ^{A2}	ha:n ^{A2}	ha:n ^{A2}	ya:n ^{A2}
336	rope, cord	*jɜ:k	c ^h uək ^{DL2}	cuɾ ^{C2}		zuək ^{DL2}	cɜ:k ^{A2}	sa:k ^{DL2}	sa:k ^{B2}	sa:k ^{B2}
337	top for spinning	*k.ra:ŋ ^B	k ^h a:ŋ ^{B1}		k ^h a:ŋ ^{B1} -i	sa:ŋ ^{B1}	ha:ŋ ^{B1}	la:ŋ ^{B1}	ca:ŋ ^{B1}	
338	basin	*ʔa:ŋ ^B	ʔa:ŋ ^{B1}	ʔa:ŋ ^{B1}	ʔa:ŋ ^{B1}		ʔa:ŋ ^{B1}			ʔa:ŋ ^{B1}
339	jug	*kra:j ^A	haj ^{A1}	haj ^{A1}	k ^h aj ^{A1}				raj ^{A1}	haj ^{A1}
340	pot	*hmo: ^C	mɔ: ^{C1}		mɔ: ^{C1}	mɔ: ^{C1}	mɔ: ^{C1}	mø ^{C1}	mo ^{C1}	mɔ: ^{C1}
341	bamboo tube	*baŋ ^{B/C}			baŋ ^{C1}	baŋ ^{C1}			baŋ ^{B1}	baŋ ^{C2}

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
342	bag	*croŋ ^A	t ^h uŋ ^{A1}	t ^h uŋ ^{A1}	t ^h oŋ ^{A1}	t ^h oŋ ^{A1}	t ^h uŋ ^{A1}			t ^h oŋ ^{A1}
343	walking cane	*daw ^C	t ^h a:w ^{C2}	taw ^{C2}	tɔw ^{C2}	ɔw ^{C2}		taw ^{C6}		
344	comb	* ^h ɔwɔw:j ^A	wi: ^{A1}	vi: ^{A1}	wi: ^{A1}	wi: ^{A1}	vi: ^{A1}	loy ^{A1}	roj ^{A1}	ho:j ^{A1}
345	writing, book	*su: ^A	su: ^{A1}	su: ^{A1}		tu: ^{A1}	tu: ^{A1}	loy ^{A1}	θu: ^{A1}	su: ^{A1}
346	broom	*ju: ^A		ju: ^{A2}		ju: ^{A2}			ŋu: ^{A2}	
347	thread (1)	*Ç.da:j ^C	da:j ^{C1}						da:j ^{C2}	ra:j ^{C1}
348	thread (2)	* ^h maj ^A	maj ^{A1}	maj ^{A1}	mɔj ^{A1}	mɔj ^{A1}	maj ^{A1}	maj ^{A1}	maj ^{A1}	
349	hemp	*pa:n ^B	pa:n ^{B1}	pa:n ^{B1}	pa:n ^{B1}	pa:n ^{B1}	pa:n ^{B1}			
350	loom	*truk ^D	hu:k ^{D1}	hu? ^{D1}	hu:k ^{D1}	t ^h uk ^{D1}	huk ^{DS1}	t ^h ok ^{D1}	rok ^{D1}	hu:k ^{DS2}
351	shuttle of loom	*p.raw ^B			p ^h ɔw ^{B1}	t ^h ɔw ^{B1}			taw ^{B1}	
352	spool	* ^h ɔwɔ:t ^D	ɔ:t ^{D1}		ɔ:t ^{D1}	ɔ:t ^{D1}	ɔ:t ^{D1}			lu:t ^{D1}
353	indigo (1)	*g.ra:m ^A	k ^h ra:m ^{A2}	ca:m ^{A2}	ca:m ^{A2}	za:m ^{A2}	kja:m ^{A2}		sa:m ^{A2} -i	
354	indigo (2)	*krom ^C	hom ^{C1}	hom ^{C1}	so:m ^{C1}	so:m ^{C1}			rom ^{C1}	
355	bamboo strip for tying or weaving	*b.twu:k ^D	tɔ:k ^{D1}	to? ^{D1}	p ^h jo:k ^{D1}	t ^h ɔk ^{D1}	p ^h jo:k ^{D1}	t ^h uk ^{D1}	tuk ^{DS2}	pruk ^{DS2}
356	needle	*qem ^A	k ^h em ^{A1}	xim ^{A1}	k ^h om ^{A1}	k ^h em ^{A1}	k ^h im ^{A1}	hɔm ^{A1}	ci:m ^{A1}	kim ^{A1}
357	strand (of rope)	*kliəw ^A	kliəw ^{A1}						cew ^{A1}	tlɛ:w ^{A1}
358	lacquer	*rak ^D	rak ^{DS2}	hak ^{DS2}	rak ^{DS2}	rak ^{DS2}				
359	bamboo hat	*kɔp ^D			cup ^{DS1}	cup ^{DS1}	kip ^{DS1}		cap ^{DS1}	
360	puppet, marionette	*hun ^B	hun ^{B1}						hun ^{B1}	hun ^{B1}
361	marking, patterns	*Ç.la:j ^A	la:j ^{A2}	la:j ^{A2}	la:j ^{A2}	la:j ^{A2}	la:j ^{A2}		ra:j ^{A2}	la:j ^{A2}
362	paddy field	*na: ^A	na: ^{A2}	na: ^{A2}	na: ^{A2}	na: ^{A2}	na: ^{A2}	na: ^{A2}	na: ^{A2}	na: ^{A2}

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
363	dry field	*rɔj ^B	raj ^{B2}	haj ^{B2}	rɔj ^{B2}	rɔj ^{B2}	ləj ^{B2}	loy ^{B2}	ri ^{B2}	ri. ^{B2}
364	open field	*doŋ ^B	t ^h uŋ ^{B2}	tui ^{B2}	toŋ ^{B2}	doŋ ^{B2}			toŋ ^{B2}	t ^h oŋ ^{B2}
365	garden	*swu:n ^A	suən ^{A1}	suən ^{A1}	t ^h uən ^{A1}	tuən ^{A1}	tu:n ^{A1}	tu:n ^{A1}	θuən ^{A1}	suən ^{A1}
366	ditch	* ^h muəŋ ^A	muəŋ ^{A1}		muəŋ ^{A1}	muəŋ ^{A1}	mɔ:ŋ ^{A1}	myŋ ^{A1}		muəŋ ^{A1}
367	dike between rice fields	*yal ^A	k ^h an ^{A2}		k ^h an ^{A2}	gän ^{A2}	kan ^{A2}	han ^{A2}	han ^{A2}	yal ^{A2}
368	dam	*hwa:j ^A	fa:j ^{A1}	fa:j ^{A1}	pha:j ^{A1}	p ^h a:j ^{A1}	p ^h a:j ^{A1}	p ^h a:j ^{A1}	va:j ^{A1}	
369	water pipe	*lin ^A		lin ^{A2}	lin ^{B1} -t	lin ^{A2}			lin ^{A2}	lin ^{A2}
370	plough	*crwaj ^A	t ^h aj ^{A1}	t ^h aj ^{A1}	t ^h aj ^{A1}	t ^h aj ^{A1}	t ^h aj ^{A1}	t ^h aj ^{A1}	saj ^{A1}	t ^h aj ^{A1}
371	yoke	*ʔe:k ^D	ʔe:k ^{DL1}	ʔeʔ ^{DL1}	ʔe:k ^{DL1}	ʔek ^{DL2}	ʔe:k ^{DL1}	ʔek ^{DL1}	ʔek ^{DL1}	ʔe:k ^{DL1}
372	part of yoke under neck	*ʔo:ŋ ^C		ʔoŋ ^{B2} -t	ʔoŋ ^{C1}	ʔoŋ ^{C1}			ʔoŋ ^{C2}	ʔoŋ ^{C1}
373	mortar (1)	*grok ^D	k ^h rok ^{DS2}		cok ^{DS2}		kjuk ^{DS2}		cok ^{DS2}	
374	mortar (2)	* ^h rum ^A						lom ^{A1}	rum ^{B1}	
375	pestle	*sa:k ^D	sa:k ^{DL1}	saʔ ^{DL1}	t ^h a:k ^{DL1}	ta:k ^{DL2}	ta:k ^{DL1}		θa:k ^{DL1}	sa:k ^{DL1}
376	winnowing basket	*doŋ ^C	doŋ ^{C1}	doŋ ^{C1} -v	doŋ ^{C1}	doŋ ^{C1}	duŋ ^{C1}	doŋ ^{C1}		
377	bran	*ram ^A	ram ^{A2}	ham ^{A2}	ram ^{A2}	ram ^{A2}	lam ^{A2}	lam ^{A2}	ram ^{A2}	ram ^{A2}
378	straw, stubble	*wuəŋ ^A	fa:ŋ ^{A2}	fuəŋ ^{A2} -v	p ^h a:ŋ ^{A2}	va:ŋ ^{A2}	fa:ŋ ^{A2}	fyŋ ^{A2}	fuəŋ ^{A2}	viəŋ ^{A1}
379	fish hook	*bet ^D	bet ^{DS1}	bit ^{DS1}	bx ^{DS1}	bet ^{DS1}	bit5			
380	fish net	*kre: ^A	he: ^{A1}	he: ^{A1}	k ^h ε: ^{A1}	k ^h ε: ^{A1}	he: ^{A1}	le: ^{A1}	re: ^{A1}	re: ^{A1}
381	fishtrap	*zaj ^A	saj ^{A2}	saj ^{A2}	t ^h aj ^{A2}	rɔj ^{A2}	ləj ^{A2}	ləj ^{A2}	θaj ^{A2}	jaj ^{A2}
382	snare	*re:w ^C	re:w ^{C2}	he:w ^{C2}	re:w ^{C2}	re:w ^{C2}				re:w ^{C2}
383	gutter, trough	*rwuəŋ ^A	ra:ŋ ^{A2}	ha:ŋ ^{A2}	ra:ŋ ^{A2}	ra:ŋ ^{A2}	la:ŋ ^{A2}		ruəŋ ^{A2}	

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
384	cooked in bamboo tube	* ^h la:m ^A	la:m ^{A1}	la:m ^{A1}	la:m ^{A1}	la:m ^{A1}	la:m ^{A1}			
385	boat	* ^ç .rwuə ^A	ruə ^{A2}	huə ^{A2}	luə ^{A2} -i	luə ^{A2}	luu ^{A2}	lu ^{A2}	ruə ^{A2}	ruə ^{A1}
386	raft	*be: ^A	p ^h ɛ: ^{A2}	pe ^{A2}	pe: ^{A2}	bɛ: ^{A2}		pe ^{A2}	pe ^{A2}	p ^h ɛ: ^{A2}
387	road	* ^h rwɪn ^A	hon ^{A1}						ran ^{A1}	
388	track	*rwu:j ^A	rɔ:j ^{A2}	hoj ^{A2}		rɔj ^{A2}		loy ^{A2}	rɪ: ^{A2}	rɪ: ^{A2}
389	saddle	*ʔa:n ^A	ʔa:n ^{A1}	ʔa:n ^A	ʔa:n ^A	ʔa:n ^{A1}		ʔa:n ^{A1}	ʔa:n ^{A1}	ʔa:n ^{A1}
390	drum	*klo:ŋ ^A	klo:ŋ ^{A1}	kun ^{B1} -t		coŋ ^{A1}	kjo:ŋ ^{A1}	køŋ ^{A1}	coŋ ^{A1}	flo:ŋ ^{A1}
	G. Adjectives									
391	red	* ^ç .dwi:ŋ ^A	dɛ:ŋ ^{A1}	dɛŋ ^{A1}	dɛ:ŋ ^{A1}	dɛŋ ^{A1}	dɛ:ŋ ^{A1}	dɪŋ ^{A1}	dɪŋ ^{A1}	rɪ:ŋ ^{A1}
392	black	* ^ç .dam ^A	dam ^{A2}	dam ^{A1}	dam ^{A1}	dam ^{A1}	dam ^{A1}	nam ^{A1}		ram ^{A1}
393	white	*xa:w ^A	k ^h a:w ^{A1}	xa:w ^{A1}	k ^h a:w ^{A1}	k ^h a:w ^{A1}	k ^h a:w ^{A1}		ha:w ^{A1}	ha:w ^{A1}
394	green	*xiew ^A	k ^h iəw ^{A1}	xew ^{A1}	k ^h ɛ:w ^{A1}	k ^h ɛ:w ^{A1}	k ^h ɛ:w ^{A1}	hew ^{A1}	hew ^{A1}	hɛ:w ^{A1}
395	yellow	* ^h luəŋ ^A	luəŋ ^{A1}		luəŋ ^{A1}		h: ^{A1}	lyŋ ^{A1}		
396	dark (red)	*klam ^B	klam ^{B1}			kam ^{B1} -v			cam ^{B1}	
397	gray	* ^h mo:ŋ ^A	mɔ:ŋ ^{A1}					møŋ ^{A1}		
398	white-spotted	*ʔla:ŋ ^B	da:ŋ ^{B1}	ba:ŋ ^{B1}	bja:ŋ ^{B1}	da:ŋ ^{B1}			da:ŋ ^{B1}	
399	clear, clean	*sau ^A	saj ^{A1}	sau ^{A1}	t ^h ɪu ^{A1}	tɪu ^{A1}	lau ^{A1}	loy ^{A1}	θau ^{A1}	
400	dark	*mur:t ^D	mur:t ^{DL2}	mut ^{DS2}	mur:t ^{DL2}					
401	bright, light (n.)	*roŋ ^B	ruŋ ^{B2}	huŋ ^{B2}	ruŋ ^{B2}	ruŋ ^{B2}	tuŋ ^{B2}	løŋ ^{B1}	roŋ ^{B2}	rɔ:ŋ ^{B2} -i
402	thin (not fat)	*pro:m ^A	p ^h ɔ:m ^{A1}				p ^h jo:m ^{A1}	p ^h øm ^{A1}	pjom ^{A1}	p ^h ɔ:m ^{A1}

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
403	fat	*bwi: ^A	p ^{hi} : ^{A2}	pi: ^{A2}	pi: ^{A2}	bi: ^{A2}	pi: ^{A2}	poy: ^{A2}	pi: ^{A2}	p ^{hu} :j: ^{A2}
404	long	*ruj: ^A	ri: ^{A2}	hi: ^{A2}	ri: ^{A2}	ri: ^{A2}	hi: ^{A2}	loy: ^{A2}	raj: ^{A2}	raj: ^{A2}
405	short (not long)	*tin ^B		tin ^{C1}	tyŋ ^{C1}	tyŋ ^{C1}	tin ^{C1}	tyŋ ^{C1} -t	tin ^{C1}	tin ^{B1}
406	big (1)	* ^{hi} luəŋ ^A	luəŋ ^{A1}	luəŋ ^{A1}	luəŋ ^{A1}	luəŋ ^{A1}	lu:ŋ ^{A1}			luəŋ ^{A1}
407	big (2)	* ^{bu} u:k ^D							buk ^{DS1}	buk ^{DS1}
408	small	*no:j ^C	no:j ^{C2}	noj ^{C2}	no:j ^{C2}		no:j ^{C1} -t	nøy ^{C2}		
409	low, short (not tall)	*tam ^B	tam ^{B1}	tam ^{B1}	tam ^{B1}	tam ^{B1}	tam ^{B1}	tam ^{B1}	tam ^{B1}	tam ^{B1}
410	heavy	* ^{hi} nak ^D	nak ^{DS1}	nak ^{DS1}	nak ^{DS1}	nak ^{DS1}	nak ^{DS1}	nak ^{DS1}	nak ^{DS1}	nak ^{DS1}
411	light (in weight)	* ^ɕ .baw ^A	baw ^{A1}	baw ^{A1}	bɔw ^{A1}	bɔw ^{A1}	baw ^{A1}	baw ^{A1}	baw ^{A1}	vaw ^{A1}
412	thick	* ^{hi} na: ^A	na: ^{A1}	na: ^{A1}	na: ^{A1}	na: ^{A1}	na: ^{A1}	na: ^{A1}	na: ^{A1}	na: ^{A1}
413	thin (not thick)	* ^ɕ .ba:ŋ ^A	ba:ŋ ^{A1}	ba:ŋ ^{A1}	ba:ŋ ^{A1}	ba:ŋ ^{A1}	ba:ŋ ^{A1}	ba:ŋ ^{A1}	ba:ŋ ^{A1}	va:ŋ ^{A1}
414	new	* ^{hi} mɔ: ^B	maj ^{B1}	mau ^{B1}	mɔu ^{B1}	mɔu ^{B1}	mau ^{B1}	mø ^{B1}	mø ^{B1}	mø: ^{B1}
415	old (of living beings)	*ke: ^B	ke: ^{B1}	ke ^{B1}	ke: ^{B1}	ke ^{B1}	ke: ^{B1}		ce: ^{B1}	ke: ^{B1}
416	old (of things)	*kaw ^B	kaw ^{B1}	kaw ^{B1}		kaw ^{B1} -v	kaw ^{B1}	kaw ^{B1}	kaw ^{B1}	kaw ^{B1}
417	senior	*craw ^C	t ^h aw ^{C1}	t ^h aw ^{C1}		t ^h ɔw ^{C1}	t ^h aw ^{C1}			t ^h aw ^{C1}
418	young, soft	*ʔwu:n ^B	ʔɔ:n ^{B1}	ʔon ^{B1}	ʔɔ:n ^{B1}	ʔon ^{B1}	ʔo:n ^{B1}	ʔøn ^{B1}	ʔun ^{B1}	ʔu:n ^{B1}
419	raw, not ripe	* ^ɕ .dip ^D	dip ^{DS1}	dip ^{DS1}	dip ^{DS1}	dip ^{DS1}	dip ^{DS1}	dip ^{DS1}	dip ^{DS1}	rip ^{DS1}
420	hard	*k.re:ŋ ^A	k ^h ɛŋ ^{A1}	xɛŋ ^{A1}	k ^h ɛ:ŋ ^{A1}	k ^h ɛŋ ^{A1}	k ^h ɛ:ŋ ^{A1}	leŋ ^{A1}		t ^h riəŋ ^{A1}
421	coarse, tough	* ^{hi} na:p ^D	ja:p ^{DL1}	na:p ^{DL1}	na:p ^{DL1}	na:p ^{DL1}	ja:p ^{DL1}		na:p ^{DL1}	niəp ^{DL1} -v
422	deep	*lɔk ^D	luək ^{DS2}	luək ^{DS2}	lɔk ^{DS2}	lɔk ^{DS2}	dak ^{DS1} -i	lak ^{DS2}	lak ^{DS2}	lak ^{DS2}
423	loose	* ^{hi} lo:m ^A	luəm ^{A1} -v	luəm ^{A1} -v					lom ^{A1}	

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
424	tight, narrow	*gap ^D	k'ap ^{DS2}	kap ^{DL2}	kap ^{DS2}	gəp ^{DS2}	kap ^{DS2}	k'ap ^{DS2}		
425	steep	* ^h liŋ ^B		liŋ ^{B1}	liŋ ^{B1}		liŋ ^{B1}	hɔŋ ^{B1}	liŋ ^{B1}	
426	stuck	*ga. ^A	k'ə. ^{A2}	ka. ^{A2}		gə. ^{A2}			ka ^{A2}	k'ə. ^{A2}
427	straight	*zɔ. ^B	suu. ^{B2}	suu ^{B2}			tu. ^{B2}	lɔ ^{B2}	θo ^{B2}	jo. ^{B2}
428	crooked	*got ^{D2}	k'hot ^{DS2}				ku ^{DS2}			
429	hot	*rwu:l ^C	rɔ:n ^{C2}	hon ^{C2}						ru:l ^{C2}
430	warm	* ^h un ^B	ʔun ^{B1}		ʔun ^{B1}	ʔun ^{B1}	ʔun ^{B1}			
431	blind	*fo:t ^D	bɔ:t ^{DL1}	bot ^{DL1}	bɔ:t ^{DL1}	bɔ:t ^{DL1}	bo:t ^{DL1}		bot ^{DL1}	
432	night blind	* ^h wa:ŋ ^A	fa:ŋ ^{A2}	fa:ŋ ^{A2}	p'ə:ŋ ^{A2}	va:ŋ ^{A2}		fa:ŋ ^{A2}		
433	deaf	* ^h nuək ^D	nuək ^{DL1}		nuək ^{DL1}		nu:k ^{DL1}	nuək ^{DL1}	nuk ^{DS1} -v	nu:k ^{DL1}
434	bitter	*C.qɪm ^A	k'om ^{A1}		k'om ^{A1}	k'om ^{A1}	k'um ^{A1}	ham ^{A2}	ham ^{A2}	ɣam ^{A2}
435	astringent in taste	* ^h wuət ^D	fa:t ^{DL1}	fa:t ^{DL1}	p'ə:t ^{DL1}			fa:t ^{DL1}	vuət ^{DL1}	viət ^{DL1}
436	sour	*sɪm ^C	som ^{C1}		t'om ^{C1}	lom ^{C1}	tum ^{C1}	lom ^{C1}	θam ^{C1}	sam ^{C1}
437	sweet, delicious	*C.wa:l ^A	wa:n ^{A1}	wa:n ^{A1}	wa:n ^{A1}	wa:n ^{A1}	va:n ^{A1}	wa:n ^{A1}	va:n ^{A1}	va:l ^{A1}
438	insipid	*cu:t ^D	cu:t ^{DL1}	cut ^{DS1} -v	cu:t ^{DL1}	cut ^{DL1}				
439	core	*ke:l ^B	kɛn ^{B1}	kɛn ^{B1}		kɛn ^{B1}				ke:l ^{B1}
440	fragrant	*ho:m ^A	hɔ:m ^{A1}	hom ^{A1}	hɔ:m ^{A1}	hɔm ^{A1}	ho:m ^{A1}	hɔm ^{A1}	hom ^{A1}	
441	stinky	* ^h men ^A	mɛn ^{A1}	min ^{A1}	mɛn ^{A1}	mɛn ^{A1}				
442	putrid	*xwi:w ^A	k'hiw ^{A1}	xiw ^{A1}	k'hiw ^{A1}	k'hiw ^{A1}	k'hi:w ^{A1}	hiw ^{A1}		k'hi:w ^{A1}
443	fishy	*ya:w ^A	k'ə:w ^{A2}	xa:w ^{A2}	k'ə:w ^{A2}				ha:w ^{A2}	ya:w ^{A2}
444	empty	*plɔw ^B	plə:w ^{B1}	paw ^{B1}	pjɔw ^{B1}	pjɔw ^{B1}	pjaw ^{B1}		pju ^{B1}	plu. ^{B1}

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
445	full	*k.tem ^A	tem ^{A1}	tim ^{A1}	tɔm ^{A1}	tem ^{A1}	tim ^{A1}	hɔm ^{A1} -i	rim ^{A1}	rim ^{A1}
446	deficient	*bro:ɲ ^B	p ^h brɔŋ ^{B2}		pjɔ:ɲ ^{B2}	bjɔŋ ^{B2}	pjɔ:ɲ ^{B2}		pjɔŋ ^{B2}	p ^h ro:ɲ ^{B2}
447	good	*dɔj ^A	di ^{A1} -v	di ^{A1}	dɔj ^{A1}	dɔj ^{A1}	daj ^{A1}	doy ^{A1}	di ^{A1}	di ^{A1}
448	bad	*rɔw:j ^C	ra:j ^{C2}	ha:j ^{C2}	ra:j ^{C2}	ra:j ^{C2}	la:j ^{C2}		ruəj ^{C2}	ra:j ^{C2}
449	long (of time)	*hɔŋ ^A	huŋ ^{A1}	huŋ ^{A1}	hɔŋ ^{A1}	hɔŋ ^{A1}	hɔŋ ^{A1}			
450	slow	*na.n ^A	na.n ^{A2}	na.n ^{A1} -t	na.n ^{A2}	na.n ^{A2}			na.n ^{A2}	
451	late	* ^h la: ^C	la: ^{C1}	la: ^{C1}	la: ^{C1}	la: ^{C1}	la: ^{C1}			
452	near	*k.rau ^C	klaj ^{C1}	sau ^{C1}	svu ^{C1}	svu ^{C1}	k ^h jau ^{C1}	koy ^{C1}	cau ^{C1}	tv: ^{C1}
453	far	*k.laj ^A	klaj ^{A1}		kwɔj ^{A1}	kwɔj ^{A1}	kwaj ^{A1} -i	kaj ^{A1}	caj ^{A1}	flaj ^{A1}
454	face down	*q ^w am ^C	k ^h wam ^{C1}	xam ^{C1} -w	k ^h wam ^{C1}		k ^h um ^{C1}		ham ^{C1}	ŋam ^{C1} -i
455	face up	* ^h ŋa:j ^A	ŋa:j ^{A1}		ha:j ^{A1}				ha:j ^{A1}	ŋa:j ^{A1}
456	sharp-pointed	* ^h le:m ^A	le:m ^{A1}	lem ^{A1}					lem ^{A1}	
457	sharp, sharp edge	*yɔm ^A	k ^h om ^{A2}		k ^h om ^{A2}	gɔm ^{A2}	kum ^{A2}			yam ^{A2}
458	slippery, smooth	*m.lu:ɲ ^B	lu:n ^{B2}	mun ^{B2}	mju:n ^{B2}		h: ^{B1} -t			mlu:ɲ ^{B2}
459	sticky	* ^h niəw ^A	niəw ^{A1}		niəw ^{A1}			niw ^{A1}	niəw ^{A1}	niəw ^{A1}
460	liquid, soft	* ^h le:w ^A	le:w ^{A1}				li:w ^{A1}	lew ^{A1}		le:w ^{A1}
461	rotten	*naw ^B	naw ^{B2}		nɔw ^{A2}	nɔw ^{B2}		naw ^{B2}	naw ^{B2}	naw ^{B2}
462	withered	* ^h riəw ^B	hiəw ^{B1}	hew ^{B1}	he:w ^{B1}		he:w ^{B1}	hoy ^{B1} -v	rew ^{B1}	he:w ^{B1} -i
463	dried up	* ^h re:ɲ ^C	he:ɲ ^{C1}	heɲ ^{C1}					re:ɲ ^{C1}	he:ɲ ^{C1} -i -v
464	dry	*ɣau ^B			k ^h ɣu ^{B1}	k ^h ɣu ^{B1}	k ^h au ^{B1}	hoy ^{B1}		hu: ^{B1} -v
465	wet	*dom ^A					tum ^{A2}	tom ^{A2}		

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
466	mute	*ɲwam ^C			wam ^C		wam ^C	wam ^C	ɲam ^{C2}	
467	dumb	*ɬwɯ ^C	baj ^{C1}			bɯɯ ^{C1}				
468	easy	*ɲa:j ^B	ɲa:j ^{B2}	ɲa:j ^{B2}	ɲa:j ^{B2}	ɲa:j ^{B2}			ɲa:j ^{B2}	ɲa:j ^{B2}
469	drunk (1)	*maw ^A	maw ^{A2}	maw ^{A2}	mɯw ^{A2}		maw ^{A2}	maw ^{A2}		maw ^{A2}
470	drunk (2)	*mwi: ^A						moy ^{A2}		
471	sterile	* ^h man ^A	man ^{A1}	man ^{A1}	man ^{A1}	man ^{A1}			man ^{A1}	man ^{A1}
472	entangled	*ɲuɲ ^{B/C}	juɲ ^{B2}		juɲ ^{B2}		juɲ ^{C1-t}		juɲ ^{C2}	juɲ ^{C2}
473	blistered	*bo:ɲ ^A	p ^h ɔ:ɲ ^{A2}		po:ɲ ^{A2}				poɲ ^{A2}	p ^h ɔ:ɲ ^{A2}
474	expensive	*be:ɲ ^A	p ^h ɛ:ɲ ^{A2}	peɲ ^{A2}		bɛɲ ^{A2}	pe:ɲ ^{A2}	peɲ ^{A2}	peɲ ^{A2}	
475	familiar	*gun ^C	k ^h un ^{C2}						kun ^{C2}	k ^h un ^{C2}
476	asleep	*ɬak ^D			dak ^{DS1}	dak ^{DS1}	dak ^{B1}	dak ^{DS1}	dak ^{DS1}	dak ^{DS1}
477	hungry	*ɣjuək ^D	ja:k ^{DL1}		ja:k ^{DL1}	ja:k ^{DL1}	ja:k ^{DS1}	ɣyk ^{DL1}	ɣiək ^{DL1}	juək ^{DS2}
478	satiated	*ɣi:m ^B	ɣim ^{B1}	ɣim ^{B1}	ɣi:m ^{B1}	ɣim ^{B1}	ɣim ^{B1}	ɣɯm ^{B1}	ɣim ^{B1}	ɣi:m ^{B1}
479	swollen	*gau ^B		kau ^{C1-t}	kɯɯ ^{B2}	gɯɯ ^{B2}	kau ^{B2}	k ^h oy ^{B2}		
480	stiff and tired	*mwəj ^B	mwəj ^{B2}	muj ^{B2}	mwəj ^{B2}				mwəj ^{B2}	
481	tired, exhausted	* ^h mwəj ^B	nwəj ^{B1}	nuj ^{B1}	nwəj ^{B1}	nwəj ^{B1}				
482	lazy	*kli:k ^D						kik ^{DL1}	cik ^{DS1}	tli:k ^{DL1}
483	tired, bored	* ^h na:j ^B	na:j ^{B1}	na:j ^{B1}	na:j ^{B1}	na:j ^{B1}	na:j ^{B1}	na:j ^{B1}	na:j ^{B1}	na:j ^{B1}
484	idle, free	*ɬwɯ:j ^A	da:j ^{A1}	da:j ^{A1}	da:j ^{A1}	da:j ^{A1}	da:j ^{A1}	duy ^{A1}		dx:j ^{A1}
485	disgusted	*ɬwə ^B	buə ^{B1}	buɯ ^{B1}	buə ^{B1}	buə ^{B1}	bu:ɯ ^{B1}	by ^{B1}	buə ^{B1}	
486	crazy	*ɬa: ^C	ba: ^{C1}	ba: ^{C1}	ba: ^{C1}	ba: ^{C1}	ba: ^{C1}	ba: ^{C1}	ba: ^{C1}	

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
487	hurt (1)	*cep ^D	cep ^{DS1}	cip ^{DS1}	cep ^{DS1}	cep ^{DS1}	cip ^{DS1}			
488	hurt (2)	*ke:t ^D						ket ^{DL1}	cet ^{DL1}	ke:t ^{DL1}
489	torn	*qa:t ^D	k'a:t ^{DL1}	xa:t ^{DL1}	k'a:t ^{DL1}	k'a:t ^{DL1}	k'a:t ^{DL1}		ka:t ^{DL1}	ka:t ^{DL1}
490	plugged	*sak ^A				lak ^{DS1}		sak ^{DS1 -i}		
491	bruised (1)	*jam ^C	c ^h am ^{C2}		cam ^{C2}				cam ^{C2}	sam ^{C2}
492	bruised (2)	*wok ^D	fok ^{DS2}		p ^h ok ^{DS2}	vok ^{DS2}		fok ^{DL2 -v}	fok ^{DS2}	vok ^{DS2}
493	lost	* ^h loŋ ^A	loŋ ^{A1}	luŋ ^{A1}	loŋ ^{A1}	loŋ ^{A1}	luŋ ^{B1}		loŋ ^{A1}	loŋ ^{A1}
494	alive	* ⁿ aj ^A	jaŋ ^{A2}		jaŋ ^{A2}	jaŋ ^{A2}				
495	strong, strength	*re:ŋ ^A	re:ŋ ^{A2}	heŋ ^{A2}	re:ŋ ^{A2}	re:ŋ ^{A2}	te:ŋ ^{A2}	leŋ ^{A2}	reŋ ^{A2}	re:ŋ ^{A2}
	H. Verbs									
496	weave, to (cloth)	*tam ^B	tam ^{B1}	tam ^{B1}	tam ^{B1}	tam ^{B1}	tam ^{B1}	tam ^{B1}	tam ^{C1}	tam ^{B1}
497	dye, to	* ⁿ wu:m ^C	jo:m ^{C2}	nom ^{C2}	jo:m ^{C2}	nom ^{C2}	jo:m ^{C2}	nom ^{C2}	num ^{C2}	num ^{C2}
498	sew, to	* ⁿ ep ^D	jep ^{DS2}	nap ^{DS1}	nap ^{DS2 -v}	nap ^{DS2}	jap ^{DS2 -v}	nip ^{DS2}	nip ^{DS2}	nip ^{DS2}
499	embroider, to	*se:w ^B		sew ^{B1}	t ^h ε:w ^{B1}	tεw ^{B1}				
500	weave, to (baskets, mats)	*sa:n	sa:n ^{A1}	sa:n ^{A1}	t ^h a:n ^{A1}	ta:n ^{A1}	ta:n ^{A1}	ta:n ^{A1}	θa:n ^{A1}	sa:n ^{A1}
501	sow, to; scatter, to	* ^ç .wa:l ^B	wa:n ^{B1}	va:n ^{B1}	wa:n ^{B1}	wa:n ^{A1}	va:n ^{B1}	wa:n ^{B1}		va:l ^{B1}
502	transplant, to	*t.nam ^A	dam ^{A1}		dam ^{A1}	dam ^{A1}	dam ^{A1}	dam ^{A1}	dam ^{A1}	tram ^{A1}
503	plant, to	*plwu:k ^D	plu:k ^{DL1}	pu ^h k ^{DL1}		pjuk ^{DL1}	pjɤ:k ^{DL1}	pok ^{DL1}		plɔ:k ^{DL1}
504	sift, to	*qrxŋ ^A	k ^h uŋ ^{A1}	suŋ ^{A1}	svŋ ^{A1}	svŋ ^{A1}	k ^h jŋŋ ^{A1}	laŋ ^{A1}		raŋ ^{A1}
505	imprison, to	*k.raŋ ^A	k ^h aj ^{A1}	saŋ ^{A1}	saŋ ^{A1}	saŋ ^{A1}	haŋ ^{A1}		caŋ ^{A1}	t ^h raŋ ^{A1}

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
506	hunt, to	*p.raw ^B		taw ^{B1}	p ^h ɔw ^{B1}		p ^h ɔw ^{B1}		taw ^{B1}	
507	mark, to; aim, to	* ^h ma:j ^A	ma:j ^{A1}	ma:j ^{A1}	ma:j ^{A1}	ma:j ^{A1}			ma:j ^{A1}	ma:j ^{A1}
508	shoot, to	* ^h ɲuɲ ^A	ɲiɲ ^{A2}				ɲaɲ ^{A2}		ɲiɲ ^{A2}	ɲiɲ ^{A2}
509	poison (fish), to	* ^ʈ buə ^A	buə ^{A1}	buə ^{A1}	buə ^{A1}	buə ^{A1}	buə ^{A1}			viə ^{A1}
510	tend animal, to	* ^ʈ ɲuəɲ ^C		cuəɲ ^{C2}	zuəɲ ^{C2}	zuəɲ ^{C2}	c ^h ɣ:ɲ ^{C2}	syɲ ^{C2}	siəɲ ^{C2}	
511	raise, to	*liəɲ ^C	liəɲ ^{C2}	liəɲ ^{C2}	liəɲ ^{C2}	liəɲ ^{C2}				liəɲ ^{C2}
512	make noise, to	* ^ʈ daɲ ^A			daɲ ^{A1}	daɲ ^{A1}			daɲ ^{A1}	
513	differ, to	* ^ʈ a:ɲ ^B	ta:ɲ ^{B1}		ta:ɲ ^{B1}	ta:ɲ ^{B1}	ta:ɲ ^{B1}		ta:ɲ ^{B1}	ta:ɲ ^{B1}
514	count, to	* ^h nap ^D	nap ^{DS2}		nap ^{DS2}					nap ^{DS2}
515	heap up, to	*ko:ɲ ^A	ko:ɲ ^{A1}		ko:ɲ ^{A1}				koɲ ^{A1}	ko:ɲ ^{A1}
516	weigh, to	* ^h ɲaɲ ^B	c ^h ɲaɲ ^{B2}				saɲ ^{B2}		saɲ ^{B2}	saɲ ^{B2}
517	announce, to	*pa:w ^B	pa:w ^{B1}	pa:w ^{B1}	pa:w ^{B1}	pa:w ^{B1}			pa:w ^{B1}	pa:w ^{B1}
518	answer, to	* ^h ɣa:n ^A	k ^h a:n ^{A1}	xa:n ^{A1}	k ^h a:n ^{A1}	k ^h a:n ^{A1}			ha:n ^{A1}	ɲa:n ^{A1} -i
519	ask, to	*c.ra:m ^A	t ^h a:m ^{A1}	t ^h a:m ^{A1}	t ^h a:m ^{A1}	sa:m ^{A1}	t ^h a:m ^{A1}	t ^h a:m ^{A1}	sa:m ^{A1}	t ^h a:m ^{A1}
520	beg, to	*k.ro: ^A	k ^h ɔ: ^{A1}	so: ^{A1}	so: ^{A1}	so: ^{A1}	ho: ^{A1}	lɔ: ^{A1}		t ^h ɔ: ^{A1}
521	order, to; blow the nose, to	*saɲ ^B	saɲ ^{B1}	saɲ ^{B1}	t ^h ɲaɲ ^{B1}	ɲaɲ ^{B1}	ɲaɲ ^{B2}	ɲaɲ ^{B1}	θaɲ ^{B1}	saɲ ^{B1}
522	scold, to; revile, to	* ^ʈ da: ^B	da: ^{B1}	da: ^{B1}		da: ^{B1}	da: ^{B1}	da: ^{B1}	da: ^{B1}	
523	warn, to	*tuən ^A	tuən ^{A1}						tuən ^{B1}	
524	weep, to	* ^h t.haɲ ^C	haɲ ^{C1}		haɲ ^{C1}	haɲ ^{C1}	haɲ ^{C1}	haɲ ^{C1}	taɲ ^{C1}	
525	moan, to	* ^h gra:ɲ ^A	k ^h ra:ɲ ^{A2}	ca:ɲ ^{A2}	ca:ɲ ^{A2}	za:ɲ ^{A2}				
526	bark, to	* ^h raw ^B	haw ^{B1}	haw ^{B1}	haw ^{B1}	haw ^{B1} -v	haw ^{B1}	haw ^{B1}	raw ^{B1}	haw ^{B1} -i

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
527	crow, to	*ɣal ^A	k ^h an ^{A1}	xan ^{A1}	k ^h an ^{A1}	k ^h an ^{A1}	k ^h an ^{A1}	han ^{A1}	han ^{A1}	ha ^{A1}
528	laugh, to	*kruwəw ^A	huə ^{A1}		k ^h uə ^{A1}	k ^h uə ^{A1}	hu. ^{A1}	liw ^{A1}	riəw ^{A1}	ruəw ^{A1}
529	recite, to	*ʔa:n ^B	ʔa:n ^{B1}	ʔa:n ^{B1}	ʔa:n ^{B1}	ʔa:n ^{B1}	ʔa:n ^{B1}			ʔa:n ^{B1}
530	open (the mouth), to	*ʔa: ^C	ʔa: ^{C1}	ʔa: ^{C1}	ʔa: ^{C1}	ʔa: ^{C1}	ʔa: ^{C1}		ʔa: ^{C2}	ʔa: ^{C1}
531	eat, to	*kuɯŋ ^A	kin ^{A1}	kin ^{A1}	kin ^{A1}	kin ^{A1}	kin ^{A1}	kɯŋ ^{A1}	kuuŋ ^{A1}	kin ^{A1}
532	chew, to	*giəw ^C	k ^h iəw ^{C2}	kew ^{C2}	ke:w ^{C2}	hgew ^{C2}		k ^h ew ^{C2}	cew ^{C2}	
533	lick, to	*Ç.lwiə ^A	liə ^{A2}		li: ^{A2} -v	li: ^{A2} -v	li: ^{A2}	li: ^{A2}	riə ^{A2}	liə ^{A2}
534	put out of the mouth, to	*ya:j ^A	k ^h a:j ^{A2}	xa:j ^{A2}	k ^h a:j ^{A2}	ga:j ^{A2}			ha:j ^{A2}	ya:j ^{A2}
535	vomit, to	*rwuək ^D	ra:k ^{DL2}	ha ^{ʔDL2}	ra:k ^{DL2}	ra:k ^{DL2}	la:k ^{DL2}	lu:k ^{DL2}	ruək ^{DL2}	ruək ^{DL2}
536	take a bite, to	*kat ^D	kat ^{DS1}	kat ^{DS1}			kat ^{DS1}		kat ^{DS1}	kat ^{DS1}
537	hold in the mouth, to	*ʔɯm ^A	ʔom ^{A1}	ʔum ^{A1}	ʔom ^{A1}	ʔom ^{A1}				ʔam ^{A1}
538	bite, to	*Ç.qɯp ^D	k ^h op ^{DS1}	xup ^{DS1}	k ^h op ^{DS1}	k ^h op ^{DS1}	k ^h o:p ^{DS1}	hap ^{DS2}	hap ^{DS2}	ɣap ^{DS2}
539	nibble, to; peck, to	*to:t ^D	to:t ^{DL1}	tot ^{DL1}	to:t ^{DL1}	tot ^{DL2}	to:t ^{DL1}	tot ^{DL1}	tot ^{DL1}	to:t ^{DL1}
540	kiss, to	*cu:p ^D	cu:p ^{DL1}	cup ^{DS1}	cu:p ^{DL1}	cup ^{DL1}	cup ^{DL1}	cup ^{DL1}	sup ^{DS1}	cu:p ^{DL1}
541	yawn, to	* ^h ra:w ^A	ha:w ^{A1}	ha:w ^{A1}	ha:w ^{A1}	ha:w ^{A1}	ha:w ^{A1}	ha:w ^{A1}	ra:w ^{A1}	ra:w ^{A1}
542	blow, to	*pow ^B	paw ^{B1}	paw ^{B1}	pyw ^{B1}	pyw ^{B1}	paw ^{B1}	pø ^{B1}	po ^{B1}	p ^h ɔ: ^{B2}
543	snore, to	*kly ^A	kron ^{A1} -i	kon ^{A1}	cɯŋ ^{A1}	can ^{A1} -v		kan ^{A1}		tlej ^{A1}
544	cough, to	*ʔaj ^A	ʔaj ^{A1}	ʔaj ^{A1}	ʔaj ^{A1}	ʔaj ^{A1}	ʔaj ^{A1}	ʔaj ^{A1}	ʔaj ^{A1}	ʔaj ^{A1}
545	hold the breath, to	*klan ^C	klan ^{C1}		can ^{C1}				can ^{C1}	
546	hug, to; embrace, to	*ko:t ^D	ko:t ^{DL1}	kot ^{DL1}	ko:t ^{DL1}	ko:t ^{DL1}	ko:t ^{DL1}	ko:t ^{DL1}	ko:t ^{DL1}	ko:t ^{DL1}
547	build, to	*ko: ^B	ko: ^{B1}		ko: ^{B1}	ko: ^{B1}			ko: ^{B1}	ko: ^{B1}

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
548	close (the eyes), to	* ^h lap ^D	lap ^{DS1}	lap ^{DS1}	lap ^{DS1}	lap ^{DS1}	lap ^{DS1}		lap ^{DS1}	
549	lie down, to, sleep, to	*nwu:n ^A	non ^{A2}	non ^{A2}	no:n ^{A2}	non ^{A2}	no:n ^{A2}	nøn ^{A2}	nin ^{A2} -v	nu:n ^{A2}
550	incubate, to	*wak ^D	fak ^{DS2}		p ^h ak ^{DS2}	vak ^{DS2}	fak ^{DS2}	fak ^{DS2}	fak ^{DS2}	
551	hold in the jaws, to	*ga:p ^D	k ^h a:p ^{DL2}	ka:p ^{DL2}		gä:p ^{DL2}			ka:p ^{DL2}	k ^h a:p ^{DL2}
552	hold in the closed hand, to	*kam ^A	kam ^{A1}	kam ^{A1}	kam ^{A1}	kam ^{A1}	kam ^{A1}	kam ^{A1}	kam ^{A1}	kam ^{A1}
553	take up in the two cupped hands, to	*ko:p ^D	ko:p ^{DL1}	kop ^{DL1}		kop ^{DL1}		kop ^{DL1}		
554	pick up, to	* ^ʔ jip ^D	jip ^{DS1}	jip ^{DS1}		ʔep ^{DS1}	ʔip ^{DS1}	jip ^{DS1}		
555	grasp, to; perch, to	*cap ^D	cap ^{DS1}	cap ^{DS1}	cap ^{DS1}	cap ^{DS1}	cap ^{DS1}			cap ^{DS1}
556	squeeze, to	*gal ^C	k ^h an ^{C2}		kan ^{C2}	gan ^{C2}				k ^h al ^{C2}
557	pinch, to	* ^h ni:p ^D	ni:p ^{DL1}				nip ^{DL1}	nip ^{DL1}	nip ^{DS1}	ni:p ^{DL1}
558	raise, to	* ^h rk ^D	jok ^{DS2}		jak ^{DS2} -v				jak ^{DS2}	ruk ^{DS2}
559	lift, to	* ^h o: ^A	jo: ^{A2}	jo ^{A2}	jo: ^{A2}	jo: ^{A2} -i			jo ^{A1}	
560	carry in the arms, to	* ^h u:m ^C	ʔum ^{C1}	ʔum ^{C1}	ʔom ^{C1}	ʔum ^{C1}	ʔum ^{C1}	ʔom ^{C1}	ʔum ^{C1}	ʔu:m ^{C1}
561	carry (a bag), to	*tri:w ^C	hiw ^{C1}	hiw ^{C1}	hiw ^{C1}	t ^h iw ^{C1}		t ^h iw ^{C1}	riəw ^{C1}	ri:w ^{C1}
562	carry on the back, to	*trwa:m ^A	ha:m ^{A1}	ha:m ^{A1}	ha:m ^{A1}	t ^h a:m ^{A1}	ha:m ^{A1}	t ^h a:m ^{A1}	ra:m ^{A1}	ra:m ^{A1}
563	pull, to; drag, to	*Ç.la:k ^P	la:k ^{DL2}	la ^ʔ DL2	la:k ^{DL2}	la:k ^{DL2}	la:k ^{DL2}		ra:k ^{DL2}	
564	bend over, to	*kxm ^C	kom ^{C1}	kum ^{C1}	kom ^{C1}	kom ^{C1}	kum ^{C1}	kom ^{C1}		kam ^{C1}
565	turn upside down or inside out, to	*pli:n ^C	plin ^{C1}		pi:n ^{C1}				pin ^{C1}	pli:n ^{C1}
566	lean back, to	* ^h i:ŋ ^A	ʔiŋ ^{A1}	ʔiŋ ^{A1}	ʔiŋ ^{A1}	ʔiŋ ^{A1}			ʔiŋ ^{A1}	ʔiŋ ^{A1} -v
567	point, to	* ^h i: ^C	ci: ^{C2}	ci: ^{C2}	ci: ^{C2}	zi: ^{C2}				

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
568	stroke, to; carress, to	*Ç.lu.p ^D	lu:p ^{DL2}	lup ^{DS2}	lu:p ^{DL2}	lup ^{DL2}		lup ^{DL2}	rup ^{DS2}	lu:p ^{DL2}
569	scratch, to	*kaw ^A	kaw ^{A1}				kaw ^{A1}	kaw ^{A1}	kaw ^{A1}	kaw ^{A1}
570	put, to	*s.cɜ:l ^B	saj ^{B1}	sauj ^{B1}	tʰɜtʃ ^{B1}				so ^{B1}	co: ^{B1}
571	snap, to	*ɬli:t ^D	di:t ^{DL1}	bit ^{DS1}						bli:t ^{DL1}
572	stamp (on), to	*duu:p ^D	tʰu:r:p ^{DL2}							tʰu:r:p ^{DL2}
573	step on, to (1)	*nam ^B	jam ^{B2}		nam ^{B2}	nam ^{B2}	jam ^{B2}		nam ^{B2}	
574	step on, to (2)	*ʔjiəp ^D	jiəp ^{DS1}							jiəp ^{DS2}
575	repeat, to	*zam ^C	sam ^{C2}	sam ^{C2}						jam ^{C2}
576	tremble, to	*sal ^B	san ^{B1}	san ^{B1}		lan ^{B1}			θan ^{B1}	seɿ ^{B1}
577	shake, to	*Ç.waj ^A	waj ^{A1}		vɜj ^{A1}	wɜj ^{A1}				
578	wrestle, to	*plam ^C	plam ^{C1}		pjam ^{C1}					plam ^{C1}
579	twist, to; wring, to	*bit ^D	bit ^{DS1}			bit ^{DS1}	bit ^{DS1}	bit ^{DS1}		bit ^{DS1}
580	stretch out, to	*ʔjiət ^D	jiət ^{DL1}		jet ^{DS1}	jiət ^{DL2}	ji:t ^{DL1}	ji:t ^{DL1}		
581	sit, to	*naj ^B	naɲ ^{B2}	naɲ ^{B2}	naɲ ^{B2}	naɲ ^{B2}	naɲ ^{B2}	naɲ ^{B1}	naɲ ^{B2}	naɲ ^{B2}
582	come, to	*h ^{ma} . ^A	ma: ^{A2} -i	ma: ^{A2} -i	ma: ^{A2} -i	ma: ^{A2} -i	ma: ^{A2} -i	ma: ^{A2} -i	ma ^{A1}	ma: ^{A1}
583	go, to	*paj ^A	paj ^{A1}	paj ^{A1}	pɜj ^{A1}	pɜj ^{A1}	paj ^{A1}	poy ^{A1}	paj ^{A1}	paj ^{A1}
584	arrive, to	*Ç.tɜŋ ^A	tʰuŋ ^{A1}	tʰuŋ ^{A1}	tʰɜŋ ^{A1}	tʰɜŋ ^{A1}	tʰɜŋ ^{A1}		taŋ ^{A2}	tʰaŋ ^{A2}
585	go upward, to	*muə ^A	muə ^{A2}		muə ^{A2}	muə ^{A2}	mu: ^{A2}			muə ^{A2}
586	ascend to	*χun ^C	k ^h uun ^{C1}		k ^h uun ^{C1}	k ^h uun ^{C1}	k ^h ɜn ^{C1}		huun ^{C1}	huun ^{C1}
587	descend, to	*n.loŋ ^A	loŋ ^{A2}		loŋ ^{A2}	loŋ ^{A2}	nuŋ ^{A2}	noŋ ^{A2}	roŋ ^{A2}	loŋ ^{A2}
588	enter, to	*χaw ^C	k ^h aw ^{C1}	xaw ^{C1}	k ^h aw ^{C1}	k ^h ɜw ^{C1}	k ^h aw ^{C1}	haw ^{C1}	haw ^{C1}	haw ^{C1}

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
589	exit, to	*ʔo:k ^D	ʔo:k ^{DL1}	ʔoʔ ^{DL1}	ʔo:k ^{DL1}	ʔok ^{DL1}	ʔo:k ^{DL1}	ʔok ^{DL1}	ʔok ^{DL1}	ʔo:k ^{DL1}
590	cross, to	*ʔa:m ^C	k ^h a:m ^{C1}	xa:m ^{C1}	k ^h a:m ^{C1}	k ^h a:m ^{C1}	k ^h a:m ^{C1}	ha:m ^{C1}	ha:m ^{C1}	ha:m ^{C1}
591	cross over, to	*kwa: ^B	kwa: ^{B1}	kwa: ^{B1}	kwa: ^{B1}	kwa: ^{B1}	kwa: ^{B1}	kwa: ^{B1}	kwa: ^{B1}	kwa: ^{B1}
592	drop, to	*tok ^D	tok ^{DS1}	tuk ^{DS1}	tok ^{DS1}	tok ^{DS1}	tuk ^{DS1}	tok ^{DS1}	tok ^{DS1}	tok ^{DS1}
593	fall off, to	*h ^h ɣn ^B	lon ^{B1}	lon ^{C1} -t	lon ^{B1}	ɣn ^{B1}		lon ^{B1}	lan ^{B1}	
594	fall down, to	*ɣm ^C	lom ^{C2}	lum ^{C2}	lom ^{C2}	lom ^{C2}		lom ^{C2}	lam ^{C2}	lam ^{C2}
595	slip and fall, to	*bla:t ^D	p ^h la:t ^{DL2}	p ^h a:t ^{DL2} -i	pja:t ^{DL2}	bja:t ^{DL2}			pja:t ^{DL2}	
596	stand, to	*ʔju:n ^A	ju:n ^{A1}	jin ^{A1}					dum ^{A1}	
597	walk, to	*pra:j ^C	p ^h a:j ^{C1}	fa:j ^{C1} -i	p ^h ja:j ^{C1}	p ^h ja:j ^{C1}	p ^h ja:j ^{C1}	p ^h a:j ^{C1}	pja:j ^{C1}	
598	fly, to	*ʔil ^A	bin ^{A1}	bin ^{A1}	byn ^{A1}	ben ^{A1} -v	bin ^{A1}	byn ^{A1}	bin ^{A1}	buɿ ^{A1} -v
599	crawl, to	*g.lwɿ:n ^A	k ^h la:n ^{A2}	ca:n ^{A2}	za:n ^{A2}	za:n ^{A2}	kja:n ^{A2}	lu:n ^{A2}	ruən ^{A2}	luən ^{A2}
600	crawl over, to	*twaj ^B	ta:j ^{B1}		twj ^{B1}	twɿj ^{B1}				ta:j ^{B1}
601	climb, to	*pi:n ^A	pi:n ^{A1}	pin ^{A1}	pi:n ^{A1}	pin ^{A1}			pin ^{A1}	pi:n ^{A1}
602	flow, to	*h ^h waj ^A	ləj ^{A1}	ləj ^{A1}	lwɿj ^{A1}	lwɿj ^{A1}	ləj ^{A1}	ləj ^{A1}	ləj ^{A1}	
603	flood, to	*ʔ.tuəm ^B	t ^h uəm ^{B2}	t ^h um ^{C1}	t ^h uəm ^{C1}	t ^h uəm ^{C1}	t ^h u:m ^{C1}	tom ^{C2}	tum ^{C2}	t ^h um ^{B1}
604	spill, to	*ʔa: ^B	ba: ^{B1}	ba: ^{B1}	ba: ^{B1}	ba: ^{B1}	ba: ^{B1}			
605	pour, to	*rwa:t ^D	ra:t ^{DL2}		rwa:t ^{DL2}					rɔ:t ^{DL1}
606	swim, to	*lo:j ^A	lo:j ^{A2}	loj ^{A2}	lo:j ^{A2}	loj ^{A2}				
607	float, to	*wu: ^A	fu: ^{A2}		p ^h u: ^{A2}	vu: ^{A2}	fu: ^{A2}	fo: ^{A2} -v	fu: ^{A2}	vu: ^{A1}
608	sink, to	*cɿm ^A	com ^{A1}	cum ^{A1}					sam ^{A1}	cam ^{A1}
609	dive, to	*dam ^A	dam ^{A2}	dam ^{A1}	dam ^{A1}	dam ^{A1}				

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
610	close, to	*hap ^D	hap ^{DS1}	hap ^{DS1}		hap ^{DS1}		hap ^{DS1}	hap ^{DS1}	
611	revolve, to; spin, to	*pan ^B	pan ^{B1}	pan ^{B1}	pan ^{B1}		pan ^{B1}	pan ^{B1}	pan ^{B1}	
612	go hungry, to	*ʔɔt ^D	ʔot ^{DS1}		ʔot ^{DS1}			ʔot4		
613	hate, to	*jaŋ ^A	c ^h afj ^{A2}	caŋ ^{A2}	caŋ ^{A2}	zaŋ ^{A2}	caŋ ^{A2}	saŋ ^{A2}	saŋ ^{A2}	saŋ ^{A2}
614	quit, to	*ʔja. ^B	ja. ^{B1}			ja. ^{B1}				
615	see, to	*tran ^A	hen ^{A1} -v	hin ^{A1} -v	han ^{A1}	han ^{A1}	han ^{A1}	th ^h ɤn ^{A1} -v	ran ^{A1}	ren ^{A1}
616	hear, to	*ŋin ^A	jin ^{A2}	ŋin ^{A2}	ŋin ^{A2}	jin ^{A2}	hin ^{C2}			
617	taste, to	*jim ^A	c ^h im ^{A2}	cim ^{A2}	ci:m ^{A2}	zim ^{A2}	cim ^{A2}	sɤm ^{A2}	sim ^{A2}	
618	smell (tr.), to	*ɬɤm ^A	dom ^{A1}	dum ^{A1}	dom ^{A1}					dam ^{A1}
619	forget, to	*lu:m ^A	lu:m ^{A2}	lum ^{A2}	lu:m ^{A2}	lum ^{A2}	lum ^{A2}	lom ^{A2}	lum ^{A2}	
620	dream, to	*hwan ^A	fan ^{A1}		p ^h an ^{A1}	p ^h an ^{A1}		p ^h an ^{A1}		van ^{A2} -t
621	believe, to	*juə ^B	c ^h uə ^{B2}							sue ^{B2}
622	recognize, to	*cak ^D	cak ^{DS1}	cak ^{DS1}		cak ^{DS1}				cak ^{DS1}
623	know, to	*ru:w ^C	ru. ^{C2}	hu ^{C2}	ru. ^{C2}		lu. ^{C2}	lø ^{C2}	ro ^{C2}	rɔ. ^{C2}
624	be, to; become, to	*ŋ.pel ^A	pen ^{A2}	pin ^{A1}	pɤn ^{A1}	bən ^{A2}	pin ^{A2}		pan ^{A2}	p ^h al ^{A2}
625	be (in a place), to	*ʔju. ^B	ju. ^{B1}	ju ^{B1}	ju. ^{B1}	ju ^{B1}	ju. ^{B1}	jaw ^{B1}	ju ^{B1}	ju. ^{B1}
626	fear, to (1)	*hja:w ^A			la:w ^{A1}	la:w ^{A1}	la:w ^{A1}		la:w ^{A1}	la:w ^{A1}
627	fear, to (2)	*ja:n ^C	jan ^{C1} -v	ja:n ^{C1}	ja:n ^{C1}	ja:n ^{C2}	ja:n ^{C1}			
628	itch, to (1)	*yal ^A	k ^h an ^{A2}	xan ^{A2}	k ^h an ^{A2}					yal ^{A2}
629	itch, to (2)	*yom ^A						hom ^{A2}	hum ^{A2}	yum ^{A2}
630	seek, to	*kra ^A	ha. ^{A1}		sa. ^{A1}	sa. ^{A1}	ha. ^{A1}		ra ^{A1}	ra. ^{A1}

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
631	join, to	*to. ^B	tɔː ^{BI}		tɔː ^{BI}	tɔː ^{BI}			tɔː ^{BI}	tɔː ^{BI}
632	divide, to	*pan ^A	pan ^{A1}	pan ^{A1}			pan ^{A1}			pan ^{A1}
633	separate, to	*bra:k ^D	p ^h ra:k ^{DL2}		pja:k ^{DL2}	bja:k ^{DL2}	pja:k ^{DL2}		pja:k ^{DL2}	
634	castrate, to	*to:n ^A	tɔːn ^{A1}	ton ^{A1}	tɔːn ^{A1}	ton ^{A1}	to:n ^{A1}		ton ^{A1}	tɔːn ^{A1}
635	cut, to	*tac ^D	tat ^{DS1}	tat ^{DS1}		tat ^{DS1}	tat ^{DS1}	tat ^{DS1}	tat ^{DS1}	tɛk ^{DS1}
636	slash, to	*wal ^A	fan ^{A2}	fan ^{A2}	p ^h an ^{A2}	van ^{A2}			fan ^{A2}	val ^{A1} -t
637	chop, to	*tram ^C	ham ^{C1}	ham ^{C1}	ham ^{C1}	t ^h am ^{C1}	t ^h am ^{C1}	t ^h am ^{A1} -t	ram ^{C1}	
638	weed, to	*ɬla:j ^A	da:j ^{A1}	ba:j ^{A1}	bja:j ^{A1}	bja:j ^{A1}	bja:j ^{A1}	ba:j ^{A1}	da:j ^{A1}	
639	break, to	*t.rak ^D	hak ^{DS1}	hak ^{DL2}	hak ^{DS1}	tak ^{DS1}	tak ^{DS1}	t ^h ak ^{DS1}	rak ^{DS1}	rak ^{DS1}
640	burst, to	*p.re:k ^D	tɛ:k ^{DL1}	teʔ ^{DL1}	p ^h ɛ:k ^{DL1}	t ^h ɛɛk ^{DL1}	p ^h ɛ:k ^{DL1}	t ^h ɛk ^{DL1}	tek ^{DL1}	pre:k ^{DL1}
641	pluck, to	*Ç.dec ^D	det ^{DS1}							rɛk ^{DS1}
642	take down, to; put down, to	*ploŋ ^A	ploŋ ^{A1}				pjuŋ ^{B1}		pjoŋ ^{A1}	ploŋ ^{A1}
643	peel, to	*po:k ^D	po:k ^{DL1}	poʔ ^{DL1}	po:k ^{DL1}	po:k ^{DL1}	po:k ^{DL1}		pok ^{DL1}	
644	hammer, to	*yo:l ^C	k ^h ɔ:n ^{C2}		k ^h ɔ:n ^{C2}		ko:n ^{C2}		hon ^{C2}	yo:l ^{C2}
645	pound, to	*tam ^A	tam ^{A1}	tam ^{A1}	tam ^{A1}	tam ^{A1}		tam ^{A1}	tam ^{A1}	tam ^{A1}
646	pound (rice), to	*zo:m ^C	so:m ^{C2}	som ^{C2}	t ^h ɔ:m ^{C2}		to:m ^{C2}			
647	slap, to	*tɥp ^D	top ^{DS1}	tup ^{DS1}	tɥp ^{DS1}	top ^{DS1}	tup ^{DS1}			tap ^{DS1}
648	pound, to (2)	*dup ^A	t ^h up ⁴			dup ^{DS2}			tup ^{DS2}	
649	pound, to (1)	*to:k ^D	tɔ:k ^{DL1}	toʔ ^{DL1}	tɔ:k ^{DL1}	tɔ:k ^{DL1}			tok ^{DL1}	tɔ:k ^{DL1}
650	come into contact, to	*Ç.tu:k ^D	t ^h u:k ^{DL1}	t ^h uʔ ^{DL1}	t ^h u:k ^{DL1}	t ^h uk ^{DL1}	t ^h uk ^{DL1}		tuuk ^{DS2}	t ^h uuk ^{DL1}
651	tie, to	*çm.ruk ^D	p ^h u:k ^{DL1}	fɥʔ ^{DL1} -i	p ^h ju:k ^{DL1}	p ^h u:k ^{DL1}	p ^h uk ^{DS1}	suk ^{DL2}	suk ^{DS2}	

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
652	tie up, to	*[la:m ^B	la:m ^{B2}		la:m ^{B2}	la:m ^{B2}	la:m ^{B2}	la:m ^{B2}		la:m ^{B2}
653	undo, to; untie, to	*ke:C	ke:C ^{C1}	ke:C ^{C1}	ke:C ^{C1}	ke:C ^{C1}	ke:C ^{C1}			ke:C ^{C1}
654	fold, to	*bap ^D	p ^h ap ^{DS2}			bap ^{DS2}			pap ^{DS2}	p ^h ap ^{DS2}
655	fold double, to	*dop ^D	t ^h op ^{DS2}	tup ^{DS2}	top ^{DS2}	dop ^{DS2}		top ^{DS2}		t ^h op ^{DS2}
656	stretch, to	*juu:t ^D	juu:t ^{DL2}		juu:t ^{DL2-i}	juut ^{DL2}				
657	erect, to	*taj ^C	taj ^{C1}	taj ^{C1}	taj ^{C1}	taj ^{C1}	taj ^{C1}		taj ^{C1}	taj ^{C1}
658	plug (a hole), to	*ʔwuut ^D	ʔut ^{DS1}	ʔut ^{DS1}	ʔot ^{DS1}	ʔot ^{DS1}	ʔut ^{DS1}			ʔot ^{DS1}
659	salt, to	*ʔuəp ^D			ʔuəp ^{DL1}	ʔuəp ^{DL1}	ʔx:p ^{DS1}			ʔuəp ^{DS2}
660	pickle, to	*do:ŋ ^A	do:ŋ ^{A1}						doŋ ^{A1}	do:ŋ ^{A1}
661	soak, to (2)	* ^h ma:B	ma:B ^{B1}	ma:B ^{B1}	ma:B ^{B1}	ma:B ^{B1}			ma ^{B1}	ma:B ^{B1}
662	scoop, to	*tak ^D	tak ^{DS1}	tak ^{DS1}	tak ^{DS1}	tak ^{DS1}	tak ^{DS1}	tak ^{DS1}	tak ^{DS1}	tak ^{DS1}
663	soak, to (1)	*je:B	c ^h ε:B2		ce:B2	zε ^{B2}			se ^{B2}	se:B2
664	burn (intr.), to	* ^h maj ^C	maj ^{C1}	maj ^{C1}	mɔj ^{C1}	mɔj ^{C1}	maj ^{C1}			
665	burn (tr.), to	*praw ^A	p ^h aw ^{A1}		p ^h jɔw ^{A1}	p ^h jɔw ^{A1}			piəw ^{A1}	p ^h raw ^{A1}
666	roast, to	* [?] juəŋ ^C	ja:ŋ ^{C1}	ja:ŋ ^{B2}	ja:ŋ ^{B2}	ja:ŋ ^{C1-t}			jiəŋ ^{C2}	juəŋ ^{C1}
667	boil (tr.), to	*tom ^C	tom ^{C1}	tum ^{C1}	tom ^{C1}	tom ^{C1}	tum ^{C1}	tom ^{C1}	tum ^{B1}	
668	boil (intr.), to	*ɬx:t ^D	duəɬ ²			duəɬ ^{DL1}	ɬx:t ^{DL1}	ɬut ^{DL1}	da:t ^{DL1}	da:t ^{DL1}
669	scald, to	*luək ^D	luək ^{DL2}	luɔ ^{DL2}	luək ^{DL2}		lu:k ^{DL2}		luək ^{DS1}	luək ^{DL2}
670	cook (rice), to	*truŋ ^A	huŋ ^{A1}		huŋ ^{A1}		huŋ ^{A1}		ruŋ ^{A1}	ruŋ ^{A1}
671	steam, to	* ^h nuŋ ^C	nuŋ ^{C1}	nuŋ ^{A2-t}	nuŋ ^{C1}	nuŋ ^{C1}			naŋ ^{C1}	nuŋ ^{C1}
672	make a fire, to	*ɬaŋ ^A		ɬaŋ ^{C1-t}	ɬaŋ ^{A1}	ɬaŋ ^{A1}	ɬaŋ ^{A1}	ɬaŋ ^{A1}		

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
673	extinguish, to	*dap ^D	dap ^{DS1}	dap ^{DS1}	dap ^{DS1}	dap ^{DS1}	dap ^{DS1}	dap ^{DS1}	dap ^{DS1}	dap ^{DS1}
674	roof, to	*mwuɣ ^A	muɣ ^{A2}	muɣ ^{A2}	muɣ ^{A2}	muɣ ^{A2}	fuj ^{A2}		foɣ ^{A2}	
675	cover (with cloth), to	*hɣm ^B	hom ^{B1}		hom ^{B1}	hom ^{B1}	hum ^{B1}	hom ^{B1}		ham ^{B1}
676	deceive, to	*bra:ɲ ^A	p ^h ra:ɲ ^{A2}	pa:ɲ ^{A2}	pja:ɲ ^{A2}	pja:ɲ ^{A2}	pja:ɲ ^{A2}			
677	buy, to	*z.fu:. ^C	su:. ^{C2}	su:. ^{C2}	ʰu:. ^{C2}	ru:. ^{C2}	ʰu:. ^{C2}	soy ^{C2}	su:. ^{C2}	su:. ^{C2}
678	sell, to	*p.qa:j	k ^h a:j ^{A1}	xa:j ^{A1}	k ^h a:j ^{A1}	k ^h a:j ^{A1}	k ^h a:j ^{A1}	ha:j ^{A1}	ka:j ^{A1}	kwa:j ^{A1}
679	borrow, to	*ʔju:m ^A	ju:m ^{A1}			ju:m ^{A1}	jim ^{A1}	jom ^{A1} -v		
680	defeat, to; be defeated, to	*be:. ^C	p ^h ɛ:. ^{C2}		pe:. ^{C2}				pe ^{B2}	p ^h ɛ:. ^{B2}
681	have, to	*mi:. ^A	mi:. ^{A2}	mi:. ^{A2}	mi:. ^{A2}	mi:. ^{A2}	mi:. ^{A2}	moy ^{A2}	mi:. ^{A2}	mi:. ^{A2}
682	take, to	*ʔaw ^A	ʔaw ^{A1}	ʔaw ^{A1}	ʔaw ^{A1}	ʔaw ^{A1}	ʔaw ^{A1}	ʔaw ^{A1}	ʔaw ^{A1}	ʔaw ^{A1}
683	obtain, to	*daj ^C	da:j ^{C1}	daj ^{C1}	daj ^{C1} -v	daj ^{C1} -v	daj ^{C1}	ʔaj ^{C1} -i	daj ^{C2}	daj ^{C1}
684	pick, to	*kep ^D	kep ^{DS1}		kop ^{DS1}	kep ^{DS1}	kip ^{DS1}		cip ^{DS1}	kip ^{DS1}
685	steal, to	*C.lak ^D	lak ^{DS2}	lak ^{DS2}	lak ^{DS2}	lak ^{DS2}	lak ^{DS2}		rak ^{DS2}	lak ^{DS2}
686	give, to	*hau ^C	haj ^{C1}	hau ^{C1}	hvu ^{C1}	hvu ^{C1}	hu:. ^{C1}	hoy ^{C1}	hau ^{C1}	hy:. ^{C1}
687	dust off, to	*pac ^D	pat ^{DS1}			pat ^{DS1}	pat ^{DS1}		pat ^{DS1}	pek ^{DS1}
688	sweep, to	*kwa:t ^D	kwa:t ^{DL1}						kwa:t ^{DL1}	kwa:t ^{DL1}
689	wipe, to	*jet ^D	c ^h et ^{DS2}		ce:. ^{DS2}					se:. ^{DS2}
690	rinse, to	*C.lwɣ:ɲ ^C	la:ɲ ^{C2}	la:ɲ ^{C2}	la:ɲ ^{C2}		la:ɲ ^{C2}		ruwɣ ^{C2}	
691	wash (clothes), to	*zak ^D	sak ^{DS2}	sak ^{DL2}	ʰak ^{DS2}	rak ^{DS2}	ʰak ^{DS2}		θak ^{DS2}	jak ^{DS2}
692	wash, to	*za:w ^A	sa:w ^{A2}			ra:w ^{A2}	ʰa:w ^{A2}	ʰa:w ^{A2}		
693	bathe, to	*ʔa:p ^D	ʔa:p ^{DL1}	ʔa:p ^{DL1}	ʔa:p ^{DL1}	ʔa:p ^{DL1}	ʔa:p ^{DL1}		ʔa:p ^{DL1}	ʔa:p ^{DL1}

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
694	teach, to	*so:ɿ ^A	sɔ:n ^{A1}		tʰɔ:n ^{A1}		tɔ:n ^{A1}	tɔn ^{A1}		sɔ:l ^{A1}
695	rest, to	*bak ^D	pʰak ^{DS2}			bak ^{DS2}	pak ^{DS2}	pak ^{DS2}		
696	wait, to	*cra: ^C	tʰa: ^{C1}	tʰa: ^{C1}	tʰa: ^{C1}	tʰa: ^{C1}			sa ^{C1}	tʰa: ^{C1}
697	be left over, to	* ^h luuə ^A	luuə ^{A1}		luuə ^{A1}	luuə ^{A1}	luu: ^{B1}		luuə ^{A1}	luə ^{A1}
698	leak, to	*rwo: ^B	ruə ^{B2}	hu ^{B2}	ruə ^{B2}	ruə ^{B2}	tʰu: ^{B2}	lu ^{B1}	ro ^{B2}	rɔ: ^{B2}
699	awaken (someone), to	*plok ^D	pluk ^{DS1 -v}	puk ^{DS1}	pjok ^{DS1}	pjok ^{DS2}	pjuk ^{DS1}		pjok ^{DS1}	pluk ^{DS1 -v}
700	wake up, to	*k.tu:n ^B	tu:n ^{B1}	tum ^{B1}	tu:n ^{B1}	tum ^{B1}			rum ^{B1}	
701	grow, to; rise, to	* ^h ma: ^C		ma: ^{C1}		ma: ^{C1}	ma: ^{C1}		ma ^{C1}	
702	shrink, to	* ^h rot ^D	hot ^{DS1}	hut ^{DS1}	hot ^{DS1}	hot ^{DS1}			rut ^{DS1}	
703	disappear, to	* ^h rwɔ: ^{jA}	hai: ^{A1}						ruəj ^{A1}	rɔ: ^{A1}
704	die, to	*p.ta:j ^A	ta:j ^{A1}	ta:j ^{A1}	pʰa:j ^{A1}	tʰa:j ^{A1}	ha:j ^{A1}	tʰa:j ^{A1}	ta:j ^{A1}	pra:j ^{A1}
705	kill, to	*qa: ^C	kʰa: ^{C1}	xa: ^{C1}	kʰa: ^{C1}	kʰa: ^{C1}	kʰa: ^{C1}	ka: ^{C1}	ka ^{C1}	ka: ^{C1}
706	expose to the sun, to	*p.ra:k ^D	ta:k ^{DL1}	taʔ ^{DL1}	pʰja:k ^{DL1}	tʰa:k ^{DL1}	pʰja:k ^{DL1}	tʰa:k ^{DL1}	ta:k ^{DL1}	pra:k ^{DL1}
707	hang up, to	*qwe:n ^A	kʰwe:n ^{A1}	xwen ^{A1}	kʰe:n ^{A1 -i}	kʰwen ^{A1}			ven ^{A1}	ve:n ^{A1}
708	hang down, to	*ho: ^C	hoj ^{C1}		ho: ^{C1}	hoj ^{C1}				
709	smear, to	*da: ^A	tʰa: ^{A2}			da: ^{A2}	ta: ^{A2}			tʰa: ^{A2}
710	lay (a cloth, etc.) across, to	*ba:t ^D	pʰa:t ^{DL2}		pa:t ^{DL2}	ba:t ^{DL2}			pa:t ^{DL2}	
711	drive away, to	*k.rap ^D	kʰap ^{DS1}	sap ^{DS1}	sap ^{DS1}					
712	chase, to	* ^h aj ^B	aj ^{B2}	aj ^{B2}	aj ^{B2}		aj ^{B2}		aj ^{B2}	aj ^{B2}
713	be finished, to	* ^h e:w ^C	le:w ^{C2}		le:w ^{C2}	lew ^{C2}		lew ^{C1 -t}	lew ^{C2}	
714	to commission	*zau ^C	cʰaj ^{C2}				cau ^{C2}		θau ^{C2}	

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
715	entrust, to	* ^h wa:k ^D	fa:k ^{DL1}	faʔ ^{DL1}	p ^h a:k ^{DL1}	p ^h a:k ^{DL1}	p ^h a:k ^{DL1}			
	I. Miscellaneous									
716	1s pronoun (weak)	*ku. ^A	ku. ^{A1}	ku ^{A1}				kaw ^{A1}	ku ^{A1}	ku. ^{A1}
717	1s pronoun (strong)	*kaw ^A			kɔw ^{A1}	kɔw ^{A1}	kaw ^{A1}		ku ^{A1}	ku. ^{A1}
718	2s pronoun (weak)	*muŋ ^A	muŋ ^{A2}	muŋ ^{A2}				maŋ ^{A1} -t	muŋ ^{A2}	muŋ ^{A2}
719	2s pronoun (strong)	*mau ^A			mɔu ^{A2}		mau ^{A2}			
720	3s pronoun (weak)	*mum ^A		min ^{A2}			mɔn ^{A2}	man ^{A1} -t		
721	3s pronoun (strong)	*man ^A	man ^{A2}		man ^{A2}				man ^{A2}	man ^{A2}
722	1p pronoun	*raw ^A	raw ^{A2}	haw ^{A2}	rɔw ^{A2}		law ^{A2}		raw ^{A2}	ro. ^{A2} -v
723	one	*nu:ŋ ^B	nuŋ ^{B1} -t	nuŋ ^{B2}	mɔŋ ^{A1} -t		nɔŋ ^{A1/B2}			nuŋ ^{B2}
724	two	*so:ŋ ^A	so:ŋ ^{A1}	soŋ ^{A1}	t ^h ɔ:ŋ ^{A1}	tɔŋ ^{A1}	tɔ:ŋ ^{A1}	tɔŋ ^{A1}	θoŋ ^{A1}	so:ŋ ^{A1}
725	three	*sa:m	sa:m ^{A1}	sa.m ^{A1}	t ^h a:m ^{A1}	tə.m ^{A1}	tə:m ^{A1}	tə.m ^{A1}	θa.m ^{A1}	sa.m ^{A1}
726	four	*si. ^B	si. ^{B1}	si ^{B1}	t ^h i: ^{B1}	t ^h i: ^{B1}	t ^h i: ^{B1}	tɔy ^{B1}	θi ^{B1}	si. ^{B1}
727	five	*ha. ^C	ha. ^{C1}	ha. ^{C1}	ha. ^{C1}	ha. ^{C1}	ha. ^{C1}	ha. ^{C1}	ha. ^{C1}	ha. ^{C1}
728	six	*krok ^D	hok ^{DS1}	huk ^{DS1}	sok ^{DS1}	sok ^{DS1}	huk ^{DS1}	lok ^{DS2}	rok ^{DS1}	rɔk ^{DS1}
729	seven	*cet ^D	cet ^{DS1}	cit ^{DS1}	cet ^{DS1}	cet ^{DS1}	cit ^{DS1}	cit ^{DS1}	sat ^{DS1}	cet ^{DS1}
730	eight	*pe:t ^D	pe:t ^{DL1}	pet ^{DL1}	pe:t ^{DL1}	peʔ ^{DL1}	pe:t ^{DL1}	pet ^{DL1}	pet ^{DL1}	pe:t ^{DL1}
731	nine	*kɔw ^C	ka:w ^{C1}	kaw ^{C1}	kɔw ^{C1}	kaw ^{C1}	kaw ^{C1}	kaw ^{C1}	ku ^{C1}	ku. ^{C1}
732	hundred; string, to	*ro:j ^C	rɔ:j ^{C2}	hoj ^{C2}	rɔ:j ^{C2}		lɔ:j ^{C2}		roj ^{C2}	rɔ:j ^{C2}
733	single, only one	*dɛw ^A	dɛw ^{A1}		dɛw ^{A1}	dɛw ^{A1}	de:w ^{B1}		dew ^{A1}	dɛw ^{A1}

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
734	pair	*gu. ^B	k ^h u. ^{B2}			g ^u . ^{B1}	ku. ^{B2}		ku. ^{B2}	
735	how many, several	*ki. ^C	ki. ^{B1} -t	ki. ^{C1}	ki. ^{B2} -t	ki. ^{C1}	ki. ^{C1}	koy. ^{C1}	ci. ^{C1}	ki. ^{C1}
736	all	*daŋ ^A	t ^h aŋ ⁴	tiŋ ^{A2} -v	taŋ ^{A2}	daŋ ^{A2}	taŋ ^{A2}			t ^h uŋ ^{A2} -v
737	many, much	* ^h la:j ^A	la:j ^{A1}	la:j ^{A1}	la:j ^{A1}	la:j ^{A1}	la:j ^{A1}	la:j ^{A1}	la:j ^{A1}	la:j ^{A1}
738	little, few	* ^h no:j ^B	no:j ^{B1}						no:j ^{B2} -t	no:j ^{B2} -t
739	one or so, any	*sak ^D	sak ^{DS1}		t ^h a:k ^{DS1}	lak ^{DS1}	lak ^{DS1}			sak ^{DS1}
740	half	*grɔŋ ^B	k ^h ruŋ ^{B2}							t ^h ruŋ ^{B2}
741	each other, together	*kan ^A	kan ^{A1}	kan ^{A1}	kan ^{A1}	kan ^{A1}	kan ^{A1}			kin ^{A1} -v
742	other	*ʔu.n ^B	ʔu.n ^{B1}		ʔu.n ^{B1}		ʔu.n ^{B1}			ʔu.n ^{B1}
743	measure from thumb to fingertip	*yuu.p ^D	k ^h u:p ^{DL2}	xu:p ^{DL2}					hup ^{DS2}	(k ^h u:p ^{DL2} -i -v)
744	bite (n.), speech	*gam ^A	kham ^{A2}	kam ^{A2}	kam ^{A2}	gam ^{A2}				
745	classifier for things	*ʔal ^A	ʔan ^{A1}		ʔan ^{A1}	ʔan ^{A1}	ʔan ^{A1}	ʔan ^{A1}		ʔal ^{A1}
746	cord, string	*sa:j ^A	sa:j ^{A1}	sa:j ^{A1}	t ^h a:j ^{A1}	la:j ^{A1}	la:j ^{A1}	la:j ^{A1}	θa:j ^{A1}	sa:j ^{A1}
747	classifier for long, thin, object	*sel ^C	sen ^{C1}							sel ^{C1}
748	classifier for tools	*mwa:k ^D			ma:k ^{DL2}	ma:k ^{DL2}	ma:k ^{DL2}		fa:k ^{DL2}	ma:k ^{DL2}
749	litter (of young)	*gro:k ^D	k ^h ro:k ^{DL2}				kjo:k ^{DL2}		cok ^{DL2}	ro:k ^{DL2}
750	time (classifier)	*baj ^A			paj ^{A2}		paj ^{A2}	paj ^{A2}		
751	name of first month	*ciəŋ ^A		ciəŋ ^{A1}	ciəŋ ^{A1}	ciəŋ ^{A1}	ci:ŋ ^{A1}	ciŋ ^{A1}		ciəŋ ^{A1}
752	last (year)	*kla:j ^A	kla:j ^{A1}	ka:j ^{A1}	ca:j ^{A1}	ca:j ^{A1}	kja:j ^{A1}	ka:j ^{A1}		fla:j ^{A1}
753	daytime	*ŋwan ^A	wan ^{A2}	vin ^{A2} -v	wan ^{A2}	wan ^{A2}	van ^{A2}	ŋm ^{A2} -v	van ^{A2}	ŋen ^{A2}

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
754	night	*yuu.n ^A	k ^h u:n ^{A2}	xu:n ^{A2}	k ^h u:n ^{A2}	g ^h u:n ^{A2}	k ^h u:n ^{A2}	han ^{A2}	hu:n ^{A2}	y ^h u:n ^{A2}
755	day	*mwu: ^C	mu: ^{C2}	mu: ^{C2}	mu: ^{C2}	mu: ^{A2}			fu: ^{C2}	mu: ^{C2}
756	early morning	*jaw ^C	c ^h a:w ^{C2}	caw ^{C2}	caw ^{C2}		caw ^{C2}	saw ^{C2}	saw ^{C2}	
757	night	*yam ^B	k ^h am ^{B2}	xam ^{B2}	k ^h am ^{B2}	g ^h am ^{B2}	kam ^{B2}	ham ^{B2}	ham ^{B2}	yam ^{B2}
758	late at night	*dfu:k ^D	du:k ^{DS1}	du:k ^{DS1}		h ^h ak ^{DS1} -i		da:k ^{DS1}	da:k ^{DS1}	du:k ^{DS1}
759	early meal	*ŋa:j ^A	ŋa:j ^{A2}	ŋa:j ^{A2}	ŋa:j ^{A2}	ŋa:j ^{A2}	ŋa:j ^{A2}	ŋa:j ^{A2}	ŋa:j ^{A2}	ŋa:j ^{A2}
760	late meal	*Ç.lwi:ŋ ^A	lɛ:ŋ ^{A2}	lɛŋ ^{A2}	lɛ:ŋ ^{A2}	lɛŋ ^{A2}	lɛ:ŋ ^{A2}	hɛŋ ^{A2}	riŋ ^{A2}	lɛ:ŋ ^{A2}
761	evening meal	*j ^h m.raw ^A	p ^h raw ^{A2}		lɛ:ŋ ^{A2}	b ^h jw ^{A2}	p ^h aw ^{C2}		saw ^{A2}	
762	day before afterday	*zu:n ^A	su:n ^{A2}	su:n ^{A2}	t ^h u:n ^{A2}	ru:n ^{A2}	fin ^{A2}	lon ^{A1}		
763	yesterday	*ŋwa. ^A	wɑ:n ^{A2}	ŋwa. ^{A2}	ŋwa. ^{A2}	ŋwa. ^{A2}	va. ^{A2}			
764	tomorrow	*j ^h m.ruk ^D	p ^h ruŋ ^{B2} -f	pu: ^{DL1}	p ^h ruk ^{DL2}		p ^h ruk ^{DS2}	tok ^{ADL2}	sok ^{DL2}	t ^h o:k5
765	day after tomorrow	*Ç.ru: ^A	ru:n ^{A2} -f	hu: ^{A2}	ru: ^{A2}	lu: ^{A2}	lu: ^{A2}	loy ^{A2}	ru: ^{A2}	ru: ^{A1}
766	year	*pi. ^A	pi: ^{A1}	pi: ^{A1}	pi: ^{A1}		pi: ^{A1}	poy ^{A1}	pi: ^{A1}	pi: ^{A1}
767	upstream, above	* ^h nuə ^A	nuə ^{A1}		nuə ^{A1}	nuə ^{A1}	nu: ^{A1}			
768	downstream, below	*tau ^C	ta: ^{C1}	tau ^{C1}	tau ^{C1}	t ^h u: ^{C1}	tau ^{C1}	taj ^{C1} -v		tɛ: ^{C1}
769	below	* wɔ:ŋ ^B	la:ŋ ^{B2}	la:ŋ ^{B2}	la:ŋ ^{B2}		la:ŋ ^{B2}			luəŋ ^{B2}
770	inside	*Ç.dau ^A	naj ^{A2} -i			d ^h u: ^{A1}	dau ^{A1}	doy ^{A1}	dau ^{A1}	rɛ: ^{A1}
771	outside	* .no:k ^D	no:k ^{DL2}	no: ^{DL2}	no:k ^{DL2}	nok ^{DL2}	no:k ^{DL2}	nok ^{ADL2}	rok ^{DL2}	lo:k ^{DL2}
772	middle	*kla:ŋ ^A	kla:ŋ ^{A1}	ka:ŋ ^{A1}	ca:ŋ ^{A1}		kja:ŋ ^{A1}	ka:ŋ ^{A1}	ca:ŋ ^{A1}	fla:ŋ ^{A1}
773	side	*Ç.buəŋ ^C	buəŋ ^{C1}		buəŋ ^{C1}	buəŋ ^{C1}	bɛ:ŋ ^{C1}			viəŋ ^{C1}
774	this	*naj ^C	ni: ^{C2} -v	ni: ^{C2}	naj ^{C2}	nɛ: ^{C2}	naj ^{B2}	noy ^{C2}	ni: ^{C2} -v	ni: ^{C2} -v

	Gloss	PT	Siamese	Sapa	Bao Yen	Cao Bang	Lungchow	Shangsi	Yay	Saek
775	before	*ko:n ^B	ko:n ^{B1}	kon ^{B1}	ko:n ^{B1}	kon ^{B1}	ko:n ^{B1}	kon ^{B1}	kon ^{B1}	ko:n ^{B1}
776	tip, end	*pla:j ^A	pla:j ^{A1}		pla:j ^{A1}	pla:j ^{A1}	pla:j ^{A1}		pla:j ^{A1}	pla:j ^{A1}
777	tip, highest point	*no:t ^D	jo:t ^{DL2}	not ^{DL2}	no:t ^{DL2}	not ^{DL2}	jo:t6			nu:t ^{DL2} -v
778	edge	*ri:m ^A	rim ^{A2}			rim ^{A2}				rim ^{A2}
779	gradually, slowly	*go:j ^B	k ^h o:j ^{B2}	koj ^{B2}	ko:j ^{B2}	gɔj ^{B2}	ko:j ^{B2}			
780	not (strong 1)	*ɓaw ^B			ɓaw ^{B1}	ɓaw ^{B1}	ɓaw ^{B1}		ɓaw ^{B1}	bo: ^{B1}
781	not (strong 2)	*bo: ^B	bo: ^{B1}	bo ^{B1}		bo: ^{B1}	bo: ^{B1}		bo: ^{B1}	bo: ^{B1}
782	not (weak)	*mi: ^A	mi: ^{A2}		mi: ^{A2}	mi: ^{C1} -t	mi: ^{B1} -t	mi: ^{C1} -t	mi: ^{A2}	
783	not yet	*paj ^B	paj ^{B1}		paj ^{BA1}	paj ^{B1}				
784	do not	* ^h pa: ^B	ja: ^{B1}				ja: ^{B1}		ja: ^{B1} , ja: ^{B1}	ja: ^{B1}
785	which	*dau ^A	daj ^{A1}	dau ^{A1}	dau ^{A1}	dau ^{A1}	nau ^{A2} -i			da: ^{A1}
786	also	*ko: ^C	ko: ^{C1}	ko ^{C1}					ko ^{C2} -t	
787	with, and	*kap ^D	kap ^{DS1}		kap ^{DS1}	kap ^{DS1}				kap ^{DS1}
788	matter	*ywa:m ^A	k ^h wa:m ^{A2}			wa:m ^{A2}	va:m ^{A2}			

Notes to glossary

1. The reflexes of this etymon in NT dialects point to PT *-r. It is generally thought to be the same etymon as 'hair knot' found only in CT and SWT with *-l-. The Siamese form generally refers to 'hair knot' but means 'head' in the expression /puət^{DL1} siən^{A1} wiən^{A2} kla:w^{CL}/ 'to have a headache, to be confused'. See 'hair knot'.
2. cf. 頭 *tóu* 'head' (Sagart 1999) from MC *dau*, LH *do*, and OC *dó. This etymon is not found in NT, and might in fact be a post-PT loanword.
5. Reflexes in some dialects, e.g. Ningming /hon^{A1}/, point to an immediate stage *χ-. See 'leg'.
6. The Siamese form means 'nose bridge'.
11. cf. PAN *matá 'eye'.
14. The Cao Bang form is found only in compounds, e.g. /kʰi:CL van^{A2}/ 'food stuck between teeth'.
15. cf. 顎 è 'palate' (Schuessler 2007: 224) from MC *ŋák*, LH *ŋak*, and OC *ŋák.
16. NT varieties point to earlier *ml-. The cluster is probably due to influence of the final *-m in PT *Ç.nam^C 'water', with which the etymon 'saliva' usually occurs, cf. Saek /nam^{C2} mla:j^{A2}/ 'saliva'.
18. The initial *k.- may have been a prefix that was lost early on in dialects that have reflexes with tone A2.
19. This etymon independently came to be the general word for 'tooth' in most dialects, e.g. Lao, White Tai, Lungchow, Cao Bang, etc. Siamese still preserves the original meaning. See 'tooth'.
20. Shangsi shows /kʰ-/ instead of the expected /h-/.
22. cf. 喉 *hóu* 'throat' (MC *ȳau*; LH *go*; OC *gô).
24. cf. PAN *qabáya 'shoulder'.
25. cf. 肩 *jiān* 'shoulder' (MC *kien*; LH *ken*; OC *kên).
27. The /-ŋ/ in Yay and other NT languages resulted from a sporadic nasalization, i.e. *mwuu.^À > *mwüt.^À > *mwuu.ŋ.^À.
28. Not found in NT.
29. The long vowel in Saek seems irregular but the Lakkja cognate /pli:p^{DL1}/ also has a long vowel.

30. The Yay form means 'a piece or section of a tree cut' and might not be related.
31. Some dialects including Wuming, Yongbei, Shangsi, and Ningming point to an immediate stage *ɣ-.
32. Saek /k-/ is aberrant. It might be a result of contamination from 'leg'.
33. cf. 脛 *jing* 'shin' (MC *yieŋ*^C; LH *gef*^C; OC *gêŋh). Lungming unexpectedly has /k-/.
38. cf. 腰 *yāo* 'waist' (MC *?jäu*, LH *?iau*, OC *?iaw*).
39. PT *C.wuət^P > *hwuət^P.
41. Sapa /du^{A1}/ is found in /du^{A1} maj^{C2}/ 'wood knuckle' only. See 'gall bladder, bile'.
49. cf. 尿 *niào* 'urine' (Schuessler 2007: 401) from MC *nieu*^C, LH *neu*^C, and OC *niâu(k)h. Independently in SWT and NT dialects, *n- became palatalized in front of *-i-, yielding *niəw. NT further simplify the intermediate form to due *ɲw, i.e. *niəw > *ɲw. CT /ne:w/ and /ne:w/ most likely resulted from simplification of *iəw to *e:w, cf. 'single, one', but it is also possible that they are late Chinese loans, cf. Mandarin *niao*.
53. In many dialects including Shan, Yay, Saek, etc., this etymon means 'brain'. In other dialects, it also refers to 'bone marrow', cf. Bao Yēn /ʔɔ:k^{DL1} du:k^{DL1}/ 'brain'. See 'brain'.
54. The onset reflexes in Yay and other NT dialects show sporadic nasalization under the influence of *-n.
57. Contrast PT *n.m- > *mn- > *ml- in NT with PT *n.m- > *nu- in SWT and CT.
59. Cao Bang shows PT *r.t- > *tr- > /th-^{B1}/ but Shangsi indicates *r.t- > *d- > /t-^{B2}/. Standard Siamese now shows an irregular /ŋ-/, possibly due to spontaneous nasalization.
60. cf. 膿 *nóng* 'pus' (MC *nuoŋ*; LH *nouŋ*; OC *nũŋ).
63. The /-l-/ in Siamese /khlaj^{A2}/ is unexpected. CT forms /laj^{A2}/ or /jaj^{A2}/ are not related to this etymon.
64. In many dialects including White Tai, Black Tai, Lue, Leiping, etc., this etymon means 'brain'. In other dialects, it occurs in a compound with *ʔwu:k^D 'bone marrow', cf. Bao Yēn /ʔɔ:k^{DL1} ʔɛ:k^{DL1}/ 'brain'. See 'bone marrow'.
66. cf. 肺 *fèi* 'lung' (MC *p'jwət*^C, LH *p'uas*, OC *phats).
67. Lakkja /kja:j^{C1}/ 'intestine' suggests PT *k.saj^C, but Tai-internal evidence is lacking.

70. Cao Bang /d-/ might have been due to contamination from 'navel' (2)!. The confusion between 'navel' and 'gall bladder, bile' is common among Tai dialects, cf. Sapa /saj^{C1} bi^{A1}/ 'navel, umbilical cord' from PT *saj^C 'intestine' + *bi^A 'gall bladder, bile'.
73. The irregular tonal reflexes might be related to the taboo nature of this etymon.
75. It is conceivable that this etymon was in fact *q.ma.^A, cf. Lakkja /k^hwǝ^{A1}/, but Tai evidence is lacking.
76. It is possible to reconstruct *q.mu.^A, for this etymon, cf. Lakkja /k^hũ^{A1}/, but Tai-internal evidence is lacking.
79. cf. 馬 *mǎ* 'horse' (MC *mǎ*^B; LH *mǎ*^B; OC *mrǎʔ).
80. cf. PAN *tumáy 'bear'.
82. This etymon is reconstructible at the PT, though ultimately from 象 xiàng 'elephant' (MC *xiang*^B; LH *ziang*^B; OC *s-jaŋʔ). Wuming /ɕiəŋ^{C2}/, Yongbei /tsu:ŋ^{C2}/, Lianshan /θe:ŋ^{C2}/ among others are later Mandarin loan from the same Chinese etymon.
83. PT *kwuəŋ^A > *wuəŋ^A > /vuəŋ^{A2}/ in Yay and Saek.
87. The Saek form /pa:ŋ^{B1}/ is unrelated.
90. PT *^hɲ- became *h- in front of front vowels in all SWT and some CT dialects.
92. Thai form /nim^{B2}/ is not related
93. See 'tree, wood'.
95. The vowel irregular reflexes in SWT and many CT dialects suggest that this etymon might in fact be a post-PT loan. The unaspirated /p-/ indicates an early date of borrowing (see §3.5)
96. cf. 雁 *yàn* 'wild goose' (MC *yan*^C; LH *yan*^C; OC *ɲans).
97. cf. PAN *manúk 'bird'.
99. cf. 燕 *yàn* 'swallow' (MC *ɣien*^C; LH *ɣen*^C; OC *ɣêns).
100. See 'tree, wood'.
104. The Yay form /θaj^{A1}/ is not related and might have come from Chinese 蝸 *wā* 'snail'. Chongzuo has both /tai/ and /ho:i/.
106. NT dialects point to *B. It is hard to determine what the original tone was.

107. PT **ŋiəw*^A > **ŋiəw*^A > Lungchow /*ji:w*^{A2}/.
109. Wuming has /*klwe*^{C1}/. For CT, see Jingxi and Debao /*kwe*^{C1}/.
111. Saek /*rɔː*^{A2}/ is not related. It is a borrowing from Vietnamese *rùa* 'turtle'.
113. Lingyue /*ŋwa*^{A2}/ attests PT*-w-.
114. cf. 鱷 'crocodile' (Schuessler 2007: 223) from from MC *ŋāk*, LH *ŋak*, and OC **ŋāk*. This etymon came to mean 'mythical water creature' in SWT and many CT dialects.
115. This word might have in fact been **l.ta:k*^D in PT, as suggested by Lakkja /*la:k*^{PL2}/ 'land leech', but Tai-internal evidence is lacking. This etymon is likely to be related to PAN **qatimaték* and **qatimatéq*, both meaning 'leech'.
122. cf. 螟蛉 *míng-líng* 'insect'. See §4.4.5 and Pittayaporn (to appear-b) for discussion of the Siamese reflex.
123. cf. Proto-Semai **smx:c*. See discussion in §5.5.
127. Saek /*luəp6*/ is loan from Siamese or Lao.
130. This etymon might have in fact been **q.matD* in PT, **q*- being an animal suffix, cf. Lakkja /*k^bwət*^{DS1}/ 'flea'. Tai evidence is lacking. See 'dog' and 'pig'.
132. The medial -w- is preserved in some NT dialects including Huanjiang /*rwiA2*/ and Nandan /*lwejA2*/.
133. Wuming /*klwa:w*^{A1}/ and Liujiang /*kjwa:w*^{A1}/, among others, preserve the medial *-w-.
141. cf. 角 *jiǎo* 'horn' (MC *kāk*; LH *kək*; OC **krók*). Bao Yen /*s-*/ suggests PT **qr-* but other dialects all point to **q-*.
142. cf. 牙 *yá* 'tusk' (MC *ŋa*; LH *ŋa*; OC **ŋrâ*).
146. Aberrant final /-p/ in some CT and NT dialects, e.g. Yay /*cip3*/, Leiping /*kipDS1*/, and Ningming /*kipDS1*/. Also see 'mushroom', 'tick', and 'hail'.
148. See 'tree, wood'.

149. Bao Yen, Cao Bang and, and a few others show reflex of *p.t-, but Leiping, Lungming, and Saek, among others, show reflex of PT *t-. In this latter set of dialects, the onset might be due to contamination from *tap^D 'liver', cf. Siamese expression /tap^{DS1} taj^{A1}/ 'internal organs'.
150. Sporadic nasalization in many varieties including Siamese.
151. PT *uəj simplified to /uəʃ/ in many dialects including Siamese.
153. The Siamese form /naj^{A2}/ is aberrant, possibly due to dissimilation from /lek^{DS2} laj^{A2} > /lek^{DS} naj^{A2}/.
154. Several high-frequency words in Bao Yen and Cao Bang have /aj/ and /aw/ instead of the expected /ɔj/ and /ɔw/.
162. Not found in NT
164. The *q-, is tentatively reconstructed to account for the /h-/ in Yay.
170. Not found in SWT. For CT, Debao, Jingxi, Yanshan Nung, and Guangnan Nung have /ra.ŋ^{A2}/.
173. cf. 稼 *jià* 'sow grain' (MC *ka^C*; LH *ka^C*; OC **kräh*).
183. I tentative reconstruct *gw- to distinguish it from *w-, which would give /fa.ŋA2 / in Siamese and Lungchow.
187. Siamese and Lao /saj^{A2}/ is a Mon-Khmer borrow, cf. Nyakur /chɾɛy/ (Shorto 2006).
190. cf. 芥 *jiè* 'mustard plant' (MC *kái^C*; LH *kɛs*; OC **kréts*)
191. Proto-Kra has *mɔl^A (Ostapirat 2000) suggesting PT *mal^A but Saek has /-n/, possibly an influence from Lao or Siamese.
195. cf. PAN *leŋa 'sesame'.
196. Possibly a late borrowing from Chinese 茶 'tea' (MC *ɖa*; LH *ɖai*; OC **dra* < **r-îa*). The form /cɛ^{B2}/ in Black Tai, and White Tai, as well as /cɛ:^{A2}/ in Bao Yen is from Vietnamese *trè* 'tea'.
198. cf. PAN *quwai 'rattan'.
202. The Saek form /het^{DS1}/ is a Siamese/Lao loan because both the onset and the vowels are unexpected. Ningming has /lap^{DS1}/ with an aberrant /-p/.
203. cf. 蕨 *jué* 'fern, bracken' (MC *kjwət*; LH *kyat*; OC **kot*).
206. cf. PAN *datum 'water' (Blust's **daNum*).

207. cf. Proto-Malayo-Polynesian *apuy < PAN *sehapyú ‘fire’.
208. cf. 燹 *biāo* ‘leaping flame’ (Schuessler 2009) from MC *pjiāu*, LH *piau*, and OC *piauw. PT *-e:w became PSWT *-ε:w, which is now reflected as /-e:w/. See ‘liquid’.
210. cf. 煤 *méi* ‘soot, coal’ (author’s etymology) from MC *mwāi*, LH *mā*, and OC *mê.
212. Saek /luə^{A1}/ has an irregular rime, and is likely a borrowing from Lao.
214. Li (1977: 95) cites Lungchow /pjaw^{B2}/ as reflex of this etyma but this form in fact came from a different etymon—*braw^B ‘leaf ashes’. Debao have both /taw^{B2}/ from *daw^B and /pjaw^{B2}/ from *braw^B.
216. Not found in SWT and CT.
217. Rong’an has /kwɤA1/ attesting the medial *-w-.
219. Proto-Kra has *(m-)nɛl^A ‘fat’ (Ostapirat 2000) suggesting PT *mal^A but Saek has /-n/, possibly an influence from Lao or Siamese.
220. The Siamese. Cao Bang, and Lungchow forms mean ‘mud’.
221. Not found in CT.
222. This is most likely a relative early Chinese loan from before the Later Han period, cf. 沙 *shā* ‘sand’ (MC *ʃa*; LH *ʃa* < *ʃai*; OC *srâi).
225. Some NT and CT dialects have /uə/.
228. Comparing the vowel with Chinese forms, this etymon may have been borrowed from Chinese relatively recently, cf. 地 *dì* ‘earth, ground’ (MC *dī*^C; LH *dī*^C; OC *r-laih)
231. The reconstruction of the onset is tentative because crucial evidence from NT is lacking.
234. Bao Yen /s-/ suggests PT *-r- but all other dialects indicates that PT did not have a liquid in this etymon.
236. cf. Vietnamese *suói* ‘mountain stream’. Longsheng has /jwi/ from earlier *rwi^C.
243. This etymon was the general PT word for ‘star’.
244. This etymon is normally used together with *t.na:w^A ‘star’ to refer to stars in the sky, in contrast with falling stars, cf. Bao Yen /da:w^{A1} di:^{B1}/ ‘star in the sky’ vs. /da:w^{A1} ha:ŋ^{A1}/ ‘falling star’, /ha:ŋ^{A1}/ being Bao Yen’s reflex for ‘tail’.
248. cf. 霧 *wù* ‘fog’ (MC *mju*^C; LH *muo*^C; OC *moh).

250. Cao Bang has /bɿn^{A1}/ for 'heaven', and /va:^{C2}/ for 'sky' and 'weather'. The Siamese form means 'on, above'.
251. See 'sky, heaven'.
254. Tianlin has /mja:p^{DL2}/.
257. Among NT dialects, Longsheng and Wuming have /rwa:j^{A2}/ and /ra:i^{A2}/ respectively.
259. The short vowel in Siamese is unexpected. The NT and CT/SWT may in fact be unrelated.
260. cf. 陰 *yīn* 'overcast, shade' (author's etymology) from MC *ʔjam*, LH *ʔim*, and OC **ʔəm* (Baxter's **ʔ(r)jum*). Saek has /jumA1/ but this is probably not related.
262. Saek /lɛ:ŋ^{C2}/ is probably a borrowing from Siamese or Lao. The expected vowel reflex is /e:/.
265. NT dialects have *C due to contamination from *nwo:ŋ^C 'younger sibling'.
267. For NT, Tianlin, Lingyue, Pingguo, Hechi, etc., all have /paw^{B1}/.
269. The SWT form is contaminated by the post-PT etymon *ja:^B 'maternal grandmother'. See Li (1971) and Matisoff (1992) for discussion.
273. SWT languages have *A probably due to contamination from *hla:nA 'grandchild'. Saek /le:nA1/ is a Siamese borrowing.
277. Lungchow and Shangsi show reflexes of *B2, perhaps due to influence of *me:^B or *bo:^B.
278. Siamese has an archaic form /ʔa:w^{A1}/.
282. Kosaka (2007) claims that the irregular Saek onset reflex results from contact with *lur:k^D bauŋ^C > (*lur:k^D bauŋ^C > (*lur:k^D gbwauŋ^C) > *lur:k^D gwauŋ^C > Saek /lu:k^{DS2} khwɿ:^{C2}/.
284. The final /-ŋ/ in Sapa and Bao Yen results from an assimilation to the onset of *no:j^C 'small', cf. Sapa /dij^{A1} noj^{C2}/ 'child'. The Saek form is /dɛk^{DS1}/ is loan from Siamese or Lao.
286. The Siamese form now means 'servant'.
289. The simple onset suggests that this etymon might in fact be a relatively recent loan, cf. 娘 *niáng* 'young woman, mother' from MC *nriang* (Pulleyblank 1991)
291. NT points to *B, probably due to contamination from PT *me:^B 'mother' or post-PT *ja:^B 'lady', cf. Siamese /mɛ:^{B2} ma:j^{Cl}/ 'widow' and Yay /ja:^{B2} ma:j^{B1}/ 'widow'.

293. cf. 字 *zì* 'polite name, written character' (MC *dzɿ*^C; LH *dzia*^C; OC **dzəh*).
294. cf. 主 *zhǔ* 'host, master, owner' (author's etymology) from *tʃju*^B, LH *tʃjo*^B, OC to? Saek also has /ca:w^C/, a borrowing from Siamese or Lao.
295. SWT and CT dialects point to *C, probably due to contamination from /kra:^C/ 'slave'.
298. cf. 巫 *wū* 'shaman' (MC *mju*, LH *mua*, OC **ma*).
299. The Saek form /p^hi:^{A1}/ is a borrowing from Thai or Lao.
300. Cao Bang /ma:ŋ^{A2}/ means 'coffin for dead body'.
301. cf. 魂 *hún* 'spiritual soul' (MC *ɣwəŋ*, LH *yuən*, OC **wən*).
306. Saek has /ja:^{A1}/ but this is most likely a Simese/Lao loan.
312. The fricative *z is tentative.
315. Wuming has two forms for 'pillar': native Tai /θaw^{A1}/, and a recent loan /cau^{A2}/ from Chinese 柱 *zhù* 'post, pillar'.
319. cf. 凳 *dèng* 'stool, bench' from MC *təŋ*^h. This etymon might in fact be a relatively late borrowing.
327. This etymon is related to PAN **panaŋ* 'crossbow', Chinese 弩 *nǚ* 'crossbow (MC *nuo*^B, LH *nd'*, OC **nâ*?), as well as Mon-Khmer **sna*? (Shorto 2006). The medial *-w- perhaps suggests earlier *p.na:^C.
329. PT *i:w became /iəw/ in Yay. See 'to carry (a bag)'.
330. For NT, Lingyue /ho^{A1}/.
331. cf. 鉗 *qián* 'pincers, tongs' (MC *gjam*; LH *giam*; OC **gam*). This etymon might in fact be a post-PT borrowing.
334. This etyma may have functioned both as a noun 'carrying pole' and a verb 'to carry with a carrying pole', cf. Siamese /ha:p^{DL1}/.
337. Wuming has /ra:ŋ^{B1}/ instead of the expected /kla:ŋ^B/.
338. cf. 盎 *àng* 'bowl' (MC *ʔaŋ*^C; LH *ʔaŋ*^C; OC **ʔaŋh*).
342. The reconstruction of the onset is tentative because crucial evidence from NT is lacking.
344. The medial *-w- is attested as /-u-/ in Lingyue /luej^{A1}/ and Nandan /ruej^{A1}/.

345. cf. 書 *shū* ‘written document’ (MC *šjwo*; LH *śa*; OC **lha*).
348. cf. 徽 *huī* ‘rope’ (Schuessler 2007: 286) from MC *xjwei*, LH *hui*, OC **hməi*. Saek /mi:^{A1}/ is not related.
353. cf. 藍 ‘indigo, blue’. This etymon refers to *Indigofera tinctoria*. Saek form /kha:m^{A2}/ is a loan from Lao or Thai. NT points to **ja:m^A*.
354. *Baphicacanthus cusia* (Ness) *Bremek*
356. cf. 鍼 *zhēn* ‘needle’ (Schuessler 2007: 610) from MC *tšjam*, LH *tšim* < *kim*, OC **kim*.
359. cf. 笠 *lì* ‘bamboo hat’ (Schuessler 2007: 351) from MC *ljəp*, LH *lip*, and OC **rəp* (Baxter’s **g-rjəp*).
365. Possibly related to 園 *yúan* ‘garden’ (author’s etymology) from MC *jwän*, LH *wan*, and OC **wen*.
370. The *-w- is preserved in Tianlin Rong’an /ɛwaj^{A1}/ and Rong’an /kɿwaj^{A1}/, among others.
371. cf. 軛 *è* ‘yoke’ (MC *ʔək*; LH *ʔək*; OC *ʔrək*).
378. Sapa, Phu Thai, and many SWT dialects have /tuəŋ/ instead of the expected /a:ŋ/.
379. Not found in NT
380. Shan dialects have /k^h-/ instead of the expected /h-/.
383. Saek /ha:ŋ^{A2}/ is a borrowing from Lao.
386. cf. 簰 ‘raft’ (MC *bat*, LH, *bɛ*, OC **brè*).
387. Liujiang /hjwan^{A1}/ and Huanjiang /wan^{A1}/ attests the medial *-w-.
388. The medial *-w- is attested in Tianlin and Donglan /lwi^{A2}/.
389. cf. 鞍 *ān* ‘saddle’ (MC *ʔân*; LH *ʔan*; OC **ʔân*).
392. cf. PAN **tudem* ‘dark, black’
397. The Siamese form means ‘not bright, dusty’.
398. Black Tai and Lue show reflexes of **d-*.
403. cf. 肥 *fēi* ‘fat’ (MC *bjwei*; LH *bui*; OC **bəi*). PT *-w- is preserved in Huangjiang and Rong’an /pwi^{A2}/, among others.

407. Not found in CT and SWT
408. The tonal irregularity in certain dialects due to confusion with **hno:j^B* 'few, little'. Siamese and Lungming still maintain the distinction.
412. This form might in fact be reconstructed as **c.na:^A*, as suggested by Lakkja /cā:^{A1}/, but Tai-internal evidence is not available.
417. The reconstruction of the onset is tentative because crucial evidence from NT is lacking.
418. Guigang /ɰwun^{B1}/ attests the medial **-w-*
419. cf. PAN **qudip* 'alive'.
422. Some CT languages, e.g. Lungchow, Leiping, and Pingsiang, point to PT **d̥*. These forms might not be related.
423. SWT varieties and Sapa aberrantly show reflexes of PT **uəm*.
427. Qinzhou has /la:^{C2}/.
436. cf. PAN **leqacém* 'sour'.
442. Yongnan has /he:w^{A1}/. The vowel in Siamese is irregular probably due to false etymological connection to 'green'.
447. This etymon shows a sporadic change from **-xj > PSWT *-i-*.
452. See Gedney (1997) for discussion.
453. It is not clear why many dialects show /-w-/. See Gedney (1997) for discussion. Perhaps, this etymon goes back to PT **k|.wajA*.
454. Saek form is irregular. This might be a case of spontaneous nasalization, or contamination by **hja:jA*
455. This might not be reconstructible (see §3.6.7.4)
456. The expected vowel reflex in Saek is /e:/. The forms /ɛm^{A1}/ and /le:m^{A1}/ are most likely loans from Siamese or Lao.
460. Although **e:* usually gives /ɛ:/ in Siamese, but **e:w* gives /e:w/. See 'flame' and 'waist'.
462. The Saek form is from a case of dialect mixing with Siamese or Lao.
463. The Saek form is a Siamese/Lao borrowing.
465. Wuming has /tum^{A2}/.
470. It is possible to reconstruct **r.mi:^A* for this etymon, cf. Proto-Kra (Ostapirat 2000) **r-me^A*, but Tai-internal evidence is lacking.

472. It is not clear whether the original PT tone was *B or *C.
474. The Saek form /pɛ:ŋ^{A2}/ is a borrowing from Siamese or Lao as suggested by the vowel /ɛ:/.
486. Not found in NT.
487. Not found in NT.
488. Not found in CT and SWT.
490. This etymon is found in Siamese as petrified part of the noun /sa^{DS1}rʉk^{DS1}/ 'hiccups'. See Kullavanijaya (1992).
494. In Siamese, this etymon has become grammaticalized to an adverb 'not yet'.
497. cf. Chinese 染 *rǎn* 'to dye' (Schuessler 2007: 239) from MC *ńzjäm^{B/C}*, LH *ńam^{B/C}*, and OC **nam?* / **nam?* and Vietnamese *nhuôm* 'to dye'.
499. cf. 繡 'to embroider' (MC *sjau^C*; LH *siu^C*; OC **siu(k)h*). For NT, see /tɛ:wB/ in Tiandong, and /tɛ:wB/ in Tianlin, Du'an, etc.
501. cf. Chinese 播 *bō* 'to sow, broadcast' (Maspéro 1912) from MC *pwá^C*, LH *pai^C*, and OC **pâih*. Vietnamese *vãi* 'to broadcast (rice)'.
502. cf. PAN *talem* (Blust's **taNem*)
510. cf. 庠 *xiáng* 'to provide for' (Schuessler 2007: 532) from MC *zjan*, LH *zian*, and OC **s-jan* or **s-lan?*.
511. cf. 養 *yǎng* 'to nourish' (Schuessler 2007: 558-559) from MC *jiaŋ^B*, LH *jaŋ^B*, and OC **jan?* or **lan?*.
512. This etymon means 'lightning' in CT dialects.
516. cf. 稱 *chēng* 'to weigh' (MC *tš^hjəŋ*; LH *tš^hŋ*, OC **k-hlŋ*), or 稱 'equal to' (MC *tš^hjəŋ^C*; LH *tš^hŋ^C*, OC **k-hlŋh*).
517. cf. Chinese 報 *bào* 'to announce' (MC *páu^C*; LH *pou^C*; OC **pūh*), and Vietnamese *báo* 'to announce, to inform'.
518. Saek /ŋ-/ is perhaps due to spontaneous nasalization.
520. Leiping /ho:^{A1}/ and Lungming /ho:^{A1}/ are irregular.
524. cf. PAN **tájic* 'to weep'.
525. For NT, Laibin /kya:ŋ^{A2}/, Liujiang /kja:ŋ^{A2}/, etc.
526. The Saek form /haw^{B1}/ is probably a loan from Lao or Thai that replaced the expected **raw^{B1}*.

531. The reconstruction of *-p is tentative. See §5.5.
533. The medial *-w- is attested in many NT dialects, i.e. Liujiang /hjwe^{A2}/, Hechi /tʰwi^{A2}/ etc.
541. Western Nung has /ŋa:w^{A1}/ in /ŋa:w^{A1}ŋa:p^{DL2}/ 'to yawn'. The nasal is probably due to contamination from /ŋa:p^{DL2}/. The dialect also has /c^{ba}:w^{A1}/.
542. The Saek reflex points to PT *b- rather than *p-.
543. The Siamse aberrantly have /-r-/ instead of the expected /-l-/.
545. The Saek form /kan^{C1}/ is a Lao loan.
549. NT dialects including Wuming and Po-ai have /nuum^{A2}/ point to PT *uu.
553. Most NT dialects have this etymon, even though Yay and Saek do not, e.g. Wuming, Lingyue, Laibin /ko:pD/.
561. PT *i:w became /iəw/ in Yay. See 'chisel'.
562. The *-w- is preserved in Tianlin Rong'an /lwa:m^{A1}/ and /jwa:m^{A1}/ among others.
563. Lakkja has /kja:k^{D2}/, possibly suggesting *g-.
567. Saek has /ni:^{C2}/, and Yay has /ji:^{C2}/ and /vi:^{C2}/, pointing to *ŋwi:^C.
574. Leiping and Lungming have /ji:p^{DL1}/ and /ʔi:p^{DL1}/ respectively.
582. SWT and CT dialects point to earlier voiceless onset probably due to contamination from *muə 'to return'. Kam-Sui languages point to voiceless initial *^hm- (Thurgood 1988).
583. Wuming and Shangsi unexpectedly show reflexes of *i:.
584. cf. PAN *datej 'to arrive'.
588. See 'tree, wood'.
589. Etymologically related to *r.no:k^D 'outside'.
591. cf. 邁 guò 'to transgress' (MC kwā^B; LH kwai^C; OC *kôih).
593. Saek /lon^{B1}/ is most likely a borrowing from Thai/Lao

595. The Saek form /tʰlɿ:t^{DL2}/ is not related.
598. Saek show /tu/ instead of /i/ perhaps due to the influence of both /b-/ and /-l/.
599. Wuming has both /wuen^{A2}/ attesting the *-w-.
601. Wuming /plen^{A1}/ is not related.
605. Phu Thai has /hwa:t^{DL2}/.
607. This etymon might in fact be a late Chinese borrowing, cf. 浮 *fú* 'to float' (author's etymology) from MC *bjəu*, LH *bu*, and OC *bu.
610. cf. 闔 *hé* 'to shut' (author's etymology) from MC *ɣáp*, LH *gap*, and OC *gáp.
615. Some SWT and CT including Siamese, Black Tai, White Tai, Sapa, and Shansi, show reflexes of *-e-.
624. SWT and some CT varieties point to *p- but NT and some other CT point to *b-.
626. cf. PAN *talaw 'to fear'.
643. cf. 剝 *bō* 'to peel' (MC *pāk*; LH *pək*; OC *prók). The correspondence is regular but if this is indeed a Chinese loan, the absence of *-l- suggests that it is a relatively recent loan.
653. The lack of *-r- suggests that it may in fact be a post-PT loan from Chinese, cf. 解 *jiě* 'to untie, undo' (MC *kat^B*; LH *kɛ^B*; OC *krê?).
658. Shanglin and Guangnan /ʔwat^{DS1}/ attests the PT *-w-.
660. Some SWT languages, e.g. Lao, Phu Thai, etc., have forms that would point to earlier *bo:ŋ^A, which is probably not related to this etymon, cf. Kaping /do:ŋ^{A1}/ 'to pickle' vs. /bo:ŋ^{A1}/ 'a kind of dipping sauce (not involving picking)'.
668. Shangsi /dut^{DL1}/ means 'fever'.
669. cf. Vietnamese *luộc* 'to boil (tr.)':
671. cf. 蒸 *zhēng* 'to steam' (author's etymology) from MC *tsjəŋ*, LH *tsiŋ*, and OC *təŋ.
678. Some dialects including Wuming, Yongbei, Shangsi, and Ningming point to an immediate stage *χ-. See 'leg'.
679. For NT, see Qiubei /dzəm^{A1}/.
680. cf. 敗 *bài* 'to be defeated' (MC *bat^C*; LH *bas*; OC *brats).

683. See 'tree, wood'.
685. Lakkja has /lak^{DS2}/. Some NT dialects have reflexes that would point to *jak^P, e.g. Wuming /ɛak^{DS2}/, but this etymon is a post-PT loan, and is related to Siamese /stuk^{DS1}/, cf. 賊 *zéi* 'thief' (author's etymology) from MC *dzək*, LH *dzək*, and OC *dzək.
686. Many dialects including Lungchow, Yuan etc. points to *u:. This monophthongal reflex may have to do with the fact that the verb 'to give' occurs frequently and is extremely susceptible to phonological reduction and grammaticalization.
688. Bao Yen has /kwɛ:t^{PL1}/ from Vietnamese *quét* 'to sweep'.
695. In many NT and CT languages including Debao, Lungchow, Wuming, Liujiang etc., this etymon has come to mean 'tired'.
698. cf. 漏 *lòu* 'to leak' (MC *ləu^C*; LH *lo^C*; OC *rô(k)h). The unique vowel correspondence may indicate a relatively recent date of borrowing.
699. Siamese irregularly raised *o to /u/ before /-k/ in this etymon. The Saek form may be a Siamese borrowing.
702. Saek /hot4/ is a borrowing from Siamese or Lao.
704. cf. PAN *matáy 'to die'
709. cf. 塗 *tú* 'to smear' (MC *duo*; LH *da*; OC *lâ).
711. Guigang, Long'an, and Wuming have /krap^{DS1}/, /k^hlap^{DS1}/, /klap^{DS1}/, respectively.
713. Wuming /liaw^{C2}/ is a later borrowing from Mandarin. The Saek form /ɛ:w^{C2}/ is a borrowing from Lao or Siamese, as suggested by the vowel /ɛ:/.
716. See '1s pronoun (strong)'.
717. PT had both *ku:^A and *kaw^A. See Strecker (1984) for discussion. The distribution of the two may have been prosodically conditioned, i.e. *kaw^A occurred in prosodically prominent positions, while *ku:^A occurred in prosodically weak positions.
718. See Strecker (1984) for discussion. This etymon probably occurred in prosodically prominent position, cf. the reduced form *mau^A.
720. See Strecker (1984) for discussion. This etymon probably occurred in prosodically prominent position, cf. the strong form *muun^A.
722. The Saek reflex /o:/ may have resulted from a reduction of earlier *-aw.
723. This form might not be reconstructible at the PT level.
724. cf. 雙 *shuāng* 'pair' (MC *ṣāŋ*; LH *ṣaŋ*; OC *srōŋ).

725. cf. 三 *sām* ‘three’ (MC *sâm*; LH *sam/ sam*; OC **sâm*).
726. cf. 四 *sì* ‘four’ (MC *sí*; LH *sís*; OC **sis/ slis*).
727. Cf. 五 *wǔ* ‘five’ (MC *ŋuo^B*; LH *ŋɑ^B*; OC **ŋɑʔ*).
728. cf. 六 *liù* ‘six’ (MC *ljuk*; LH *liuk*; OC **ruk*; Baxter’s OC **C-rjuk*).
729. cf. 七 *qī* ‘seven’ (MC *ts^hjet*; LH *ts^hit*; OC **ts^hit < *snhit*).
730. cf. 八 *bā* ‘five’ (MC *pāt*; LH *pət*; OC **prêt*).
731. cf. 九 *jiǔ* ‘nine’ (MC *kjəu^B*; LH *ku^B*; OC **kuʔ/ kwəʔ*).
732. cf. 幾 *jǐ* ‘how many’ (MC *kjəi^B*; LH *kii^B*; OC **kəjʔ*). The tonal reflex is rather irregular because this is a grammatical item.
736. The vowel reflexes are not completely regular because this etymon is a grammatical word.
737. cf. 多 *duō* ‘many’ (Schuessler 2009) from MC *tá*, LH *ta < tai*, and OC **tái < tlai*.
738. See ‘small’.
740. Wuming has /kɿaŋ^{B2}/ ‘half’
743. The Saek form is a borrowing from Lao/Siamese.
747. cf. 綠 *xiàn* ‘thread’ (MC *sjan^C*; LH *sian^C*; OC **sans*).
750. This etyma is attested widely in NT, e.g. /paj^{A2}/ in Nandan, Hengxian, Guigang, etc.
751. cf. 正 *zhēng* ‘five’ (MC *tšjāŋ*; LH *tšəŋ*; OC **teŋ*). This might be in fact be a post-PT borrowing (Manomaivibool 1976).
753. Many dialects in Northwestern Vietnam unexpectedly have /-e-/; e.g. Lao, Black Tai, White Tai, and Sapa.
756. cf. 早 *zǎo* ‘morning’ (MC *tsáu^B*; LH *tsou^B*; OC **tsûʔ*).
759. In NT, this etymon has been generalized and now means that ‘cooked rice’.
763. The final /-n/ in Siamese is unetymological and was acquired under the influence of the deictic /ni:C2/ ‘this’ in the expression /muə^{B2} wa:n^{A2} ni:C2/ ‘this past yesterday’. Also see Li (1977: 128-129).

764. The final /-ŋ/ in Siamese is aberrant and must have been due to the influence of /ni:C2/, cf. the expression /wanA2 phruŋB2 ni:C2/.
765. See 'yesterday'.
766. Wuming /pi:^{A1}/ shows an unexpected vowel reflex (see §5.4).
770. SWT dialects show reflexes of *n-. This is an innovation due to the influence of *l.no:k 'outside', which is reflected as *no:kD in PSWT.
771. This cannot go back to *C.n- because Saek has /l-/ rather than /n-/.
774. SWT and some CT dialects show reflexes *i:. This is probably because this etymon is a grammatical word. It is also possible that two forms *naj^C and *ni:^C co-existed. See *kaw^A and *ku:^A both '1s pronoun'.
775. Some languages unexpectedly have /x/, e.g. Sapa.
780. Many dialects such as Lao, Yuan, Western Nung have /bo:^{B1}/ rather than /baw^{B1}/. These monophthongal reflexes are reduced forms developed independently in individual varieties. Lungchow and Yay have both the full and reduced forms. See 'not (strong 2)'.
781. See 'not strong (1)'.
782. The tonal irregularity is not surprising because this etymon is a grammatical item.
786. Wuming has /kx^{C1}/. Some dialects have irregular tonal reflexes because this etymon is a grammatical item.

APPENDIX C

INDEX OF GLOSSARY

The numbers in the rightmost column are the identifying numbers in Appendix B.

acacia	*ge: ^A	186	astringent in taste	* ^h wuət ^D	435	be (in a place), to	*ju: ^B	625
alive	*naŋ ^A	494	awaken (someone), to	*plok ^D	699	be finished, to	*le:w ^C	713
all	*daŋ ^A	736	axe	*xwa:n ^A	328	be left over, to	* ^h luə ^A	697
also	*ko: ^C	786	back	* ^h laŋ ^A	37	be, to; become, to	*m.pel ^A	624
announce, to	*pa:w ^B	517	bad	*rwx:j ^C	448	beach, sandbar	*ha:t ^D	240
answer, to	*ɣa:n ^A	518	bag	*croŋ ^A	342	bear	* ^h mwuj ^A	80
ant	*mɣc ^D	123	bamboo	*praŋ ^B	179	beard	*mom ^B	21
antelope	*wa:l ^A	84	bamboo hat	*klɔp ^D	359	bedbug	*Ç.ruət ^D	117
ape	*kaŋ ^A	86	bamboo shoot	*r.na:ŋ ^A	170	bee	*to: ^B	121
arm	*qe:n ^A	25	bamboo strip for tying or weaving	*b.twu:k ^D	355	before	*ko:n ^B	775
arrive, to	*Ç.tɔŋ ^A	584	banana	*kluej ^C	177	beg, to	*k.ro: ^A	520
ascend to	*ɣun ^C	586	banana blossom	*pli: ^A	158	believe, to	*juə ^B	621
ashes (wood)	*daw ^B	214	bank, river	* ^h waj ^B	239	below	*lwɔ:ŋ ^B	769
ashes, leaf	*braw ^B	213	banyan	*raj ^A	187	bend over, to	*kɣm ^C	564
ashgourd	*wak ^D	181	bark, to	* ^h raw ^B	526	big (1)	* ^h luəŋ ^A	406
ask, to	*c.ra:m ^A	519	basin	*ʔa:ŋ ^B	338	big (2)	*bu:k ^D	407
asleep	*dak ^D	476	bathe, to	*ʔa:p ^D	693	bird	*Ç.nok ^D	97

bite (n.), speech	*gam ^A	744	build, to	*ko. ^B	547	chew, to	*giəw ^C	532
bite, to	*Ç.qɿp ^D	538	burn (intr.), to	* ^h maj ^C	664	chicken	*kaj ^B	93
bitter	*Ç.qɿm ^A	434	burn (tr.), to	*praw ^A	665	child (offspring)	*lu:k ^D	270
black	*Ç.dam ^A	392	burst, to	*p.re:k ^D	640	child (young person)	*dɛk ^D	284
blind	*bo:t ^D	431	butterfly	*bɿ:j	138	chin, jaw	*ca:ŋ ^A	20
blistered	*bo:ŋ ^A	473	buttock	*kon ^C	43	chisel	*siəw ^B	329
blood	*luət ^D	56	buy, to	*z.ju: ^C	677	chop, to	*tram ^C	637
blow, to	*pow ^B	542	cane for walking	*daw ^C	343	chopping board	*ɣiəŋ ^A	325
board	*pe:n ^C	321	carry (a bag), to	*tri:w ^C	561	civet cat	*hjel ^A	90
boat	*Ç.rwuə ^A	385	carry in the arms, to	*ʔu:m ^C	560	classifier for long, thin, object	*sel ^C	747
boil (intr.), to	*dɿ:t ^D	668	carry on the back, to	*trwa:m ^A	562	classifier for things	*ʔal ^A	745
boil (tr.), to	*tom ^C	667	carrying pole (1)	*tra:p ^D	334	classifier for tools	*mwa:k ^D	748
bone	*Ç.duk ^D	52	carrying pole (2)	*ca:n ^A	335	clear, clean	*sau ^A	399
borrow, to	*ʔju:m ^A	679	castrate, to	*to:n ^A	634	climb, to	*pi:n ^A	601
brain	*ʔe:k ^D	64	caterpillar	*Ç.boŋ ^C	136	close (the eyes), to	*hlap ^D	548
bran	*ram ^A	377	catfish	*Ç.dok ^D	103	close, to	*hap ^D	610
break, to	*t.rak ^D	639	cave	*cram ^C	231	cloud	*hwuə ^C	247
bright, light (n.)	*roj ^B	401	centipede	*q.sip ^D	139	clump (as of bamboo)	*ko: ^A	168
broom	*ju: ^A	346	chase, to	*laj ^B	712	coarse, tough	*hna:p ^D	421
bruised (1)	*jam ^C	491	cheek	*ke:m ^C	17	cockroach	*sa:p ^D	140
bruised (2)	*wok ^D	492	chest	*ʔɿk ^D	36	cockscorn	*ho:n ^A	150

cockspur	*t.nɿːj ^A	151	cross, to	*ɣa:m ^C	590	descend, to	*p.lɔŋ ^A	587
coconut grub	*d̪uəŋ ^C	134	crossbow	*h̪nwuə ^C	327	dew, mist	*ɕ.nwa:j ^A	257
cold (n.)	*q.wat ^D	72	crow	*kaː ^A	98	die, to	*p.ta:j ^A	704
comb	*h̪rwu:j ^A	344	crow, to	*ɣal ^A	527	differ, to	*ta:ŋ ^B	513
come into contact, to	*ɕ.tur:k ^D	650	cucumber, melon	*p.rwe:ŋ ^A	180	dike between rice fields	*ɣal ^A	367
come, to	*h̪maː ^A	582	cut, to	*tac ^D	635	disappear, to	*h̪rwɿ:j ^A	703
cook (rice), to	*truŋ ^A	670	dam	*hwa:j ^A	368	disgusted	*buə ^B	485
cooked in bamboo tube	*h̪la:m ^A	384	dark	*mur:t ^D	400	ditch	*h̪muəŋ ^A	366
cord, string	*sa:j ^A	746	dark (red)	*klam ^B	396	dive, to	*dam ^A	609
core	*ke:l ^B	439	daughter-in-law	*bau ^C	282	divide, to	*pan ^A	632
cough, to	*ʔaj ^A	544	day	*mwu: ^C	755	do not	*jaː ^B	784
count, to	*nap ^D	514	day after tomorrow	*ɕ.rurː ^A	765	dog	*h̪maː ^A	75
cover (with cloth), to	*h̪m ^B	675	day before afterday	*zu:n ^A	762	door	*tuː ^A	313
crab	*puw ^A	105	daytime	*ŋwan ^A	753	downsteam, below	*tau ^C	768
crawl over, to	*twaj ^B	600	deaf	*h̪nuək ^D	433	dream, to	*h̪wan ^A	620
crawl, to	*g.lwɿ:n ^A	599	debt	*h̪niː ^C	304	dried up	*re:ŋ ^C	463
crazy	*ʔaː ^C	486	deceive, to	*bra:ŋ ^A	676	drive away, to	*k.rap ^D	711
creek	*ro:ŋ ^B	237	deep	*ɬk ^D	422	drop, to	*tok ^D	592
crocodile	*ŋuək ^D	114	deer	*kwuəŋ ^A	83	drum	*klo:ŋ ^A	390
crooked	*got ^{D2}	428	defeat, to; be defeated, to	*beː ^C	680	drunk (1)	*maw ^A	469
cross over, to	*kwaː ^B	591	deficient	*bro:ŋ ^B	446	drunk (2)	*mwiː ^A	470

dry	*ɣau ^B	464	embroider, to	*se:w ^B	499	fat	*k.tɔt ^D	51
dry land	*ɓok ^D	224	empty	*plɔw ^B	444	fat	*bwi: ^A	403
dry season	*ɕ.le:ŋ ^C	262	entangled	*nuŋ ^{B/C}	472	father, man	*bo: ^B	263
duck	*pit ^D	95	enter, to	*ɣaw ^C	588	father's younger brother	*ʔa:w ^A	278
duckweed	* ^h ne: ^A	204	entrust, to	* ^h wa:k ^D	715	father's younger sister	*ʔa: ^A	276
dumb	*ɓɔɔ ^C	467	epidemic	*ra: ^B	73	fear, to (1)	* ^h la:w ^A	626
dust off, to	*pac ^D	687	erect, to	*taŋ ^C	657	fear, to (2)	*ja:n ^C	627
dye, to	*ɲwu:m ^C	497	excrement	*ɕ.quj ^C	50	fern	*ku:t ^D	203
each other, together	*kan ^A	741	exit, to	*ʔo:k ^D	589	field, dry	*ɾj ^B	363
ear	*krwu: ^A	18	expensive	*be:ŋ ^A	474	field, open	*doŋ ^B	364
ear (of rice)	*rwu:ŋ ^A	171	expose to the sun, to	*p.ra:k ^D	706	field, paddy	*na: ^A	362
earth	*ɗin	221	extinguish, to	*ɗap ^D	673	finger, toe	*ni:w ^C	28
earthworm	*t.nuəl ^A	126	eye	*p.ta: ^A	11	finger nail, toenail	*ɕ.lep ^D	29
easy	*ɲa:j ^B	468	face	* ^h na: ^C	10	fire	*wɔj ^A	207
eat, to	*kuɲ ^A	531	face down	*qwam ^C	454	firewood, bamboo	* ^h luəw ^A	212
eaves	*za:j ^A	312	face up	* ^h ɲa:j ^A	455	firewood, hard	*wu:ɲ ^A	211
edge	*ri:m ^A	778	fall down, to	*ɪm ^C	594	fish	*pla: ^A	102
egg	*qraj	148	fall off, to	* ^h ɪn ^B	593	fish hook	*bet ^D	379
eight	*pe:t ^D	730	familiar	*gun ^C	475	fish net	*kre: ^A	380
elbow	*ɕ.swo:k ^D	26	fang	*ɣe:w ^C	19	fishbone	*ka:ŋ ^C	147
elephant	*ja:ŋ ^C	82	far	*k.laj ^A	453	fishtrap	*zaj ^A	381

fishy	*ya:w ^A	443	frog, small	*krwe: ^C	109	grandfather, maternal	*ta: ^A	268
five	*ha: ^C	727	frog, tree	*pa:t ^D	110	grandfather, paternal	*puw ^B	267
flame	*ple:w ^A	208	fruit	* ^h ma:k ^D	163	grandmother, maternal	* ^h na:j ^A	269
flea	* ^h mat ^D	130	full	*k.tem ^A	445	grasp, to; perch, to	*cap ^D	555
float, to	*wu: ^A	607	gadfly	* ^h luək ^D	127	grass	* ^h juuə ^C	200
flood, to	*Ç.tuəm ^B	603	galangal	*xa: ^B	194	grass, thatch	*ya: ^A	201
flour	* ^h uə ^A	308	gall bladder, bile	* ^h li: ^A	70	grasshopper	*p.tak ^D	118
flow, to	* ^h lwaj ^A	602	garden	*swu:n ^A	365	gray	* ^h mo:ŋ ^A	397
flower	* ^h lo:k ^D	157	ginger	* ^h i:ŋ ^A	193	gray-haired	* ^h ŋwu:k ^D	6
fly, to	* ^h il ^A	598	girl	* ^h ur:k ^D	288	grease, fat	*man ^A	219
flying squirrel	*ba:ŋ ^B	87	give, to	* ^h au ^C	686	great-grandchild	* ^h len ^C	273
fog	* ^h mo:k ^D	248	gizzard	*p.tau ^A	149	great-grandparent	*juə ^C	271
fold double, to	*dop ^D	655	go hungry, to	*ʔɿt ^D	612	green	*xiəw ^A	394
fold, to	*bap ^D	654	go upward, to	*muə ^A	585	grove, wood	*pa: ^B	227
foot	*ti:n ^A	34	go, to	*paj ^A	583	grow, to; rise, to	* ^h ma: ^C	701
forehead	*pra:k ^D	7	goiter	* ^h niŋ ^A	23	gum; gill	* ^h juək ^D	15
forest	* ^h oŋ ^A	226	good	* ^h ɸj ^A	447	gust (of rain)	*kra: ^B	253
forget, to	*lu:m ^A	619	goose	*ha:n ^B	96	gutter, trough	* ^h ruwəŋ ^A	383
four	*si: ^B	726	gradually, slowly	*gə:j ^B	779	hail	*trep ^D	258
fragrant	*ho:m ^A	440	grain	*m.lec ^D	165	hair, body; feather	*q.puu ^A	5
frog	*kɿp ^D	108	granary	* ^h juəw ^C	310	hair, head	*prɿm ^A	3

hair, pubic	* ^h mwu:j ^A	44	hoarfrost	* ^h muəj ^A	256	idle, free	* ^h dwɔ:j ^A	484
half	* ^g ɾɿŋ ^B	740	hold in the closed hand, to	* ^k am ^A	552	illness, fever	* ^k raŋ	71
hammer, to	* ^ɣ o:l ^C	644	hold in the jaws, to	* ^g a:p ^D	551	impetigo	* ^k rit ^D	61
hand	* ^m wu: ^A	27	hold in the mouth, to	* ^ʔ m ^A	537	imprison, to	* ^k raŋ ^A	505
handle (of a knife)	* ^d a:m ^C	332	hold the breath, to	* ^k lan ^C	545	incubate, to	* ^w ak ^D	550
handle, rod	* ^g al ^A	333	hole	* ^{ru} : ^A	232	indigo (1)	* ^g ra:m ^A	353
hang down, to	* ^{ho} :j ^C	708	hole, crack	* ^{jo} :ŋ ^B	233	indigo (2)	* ^k rom ^C	354
hang up, to	* ^q we:n ^A	707	hook	* ^x ɔ: ^A	330	insect	* ^m le:ŋ ^A	122
hard	* ^k re:ŋ ^A	420	horn	* ^q aw ^A	141	inside	* ^ʃ ɔ: ^A	770
hate, to	* ^{ja} ŋ ^A	613	horse	* ^{ma} : ^C	79	insipid	* ^{cu} :t ^D	438
have, to	* ^{mi} : ^A	681	hot	* ^{rwu} :l ^C	429	intestine	* ^s aj ^C	67
hawk	* ^{lam} ^B	101	house	* ^ɾ :n ^A	309	iron (1)	* ^h lek ^D	215
head (1)	* ^k raw ^C	1	how many, several	* ^{ki} : ^C	735	iron (2)	* ^m wa: ^A	216
head (2)	* ^{tru} ə ^A	2	hug, to; embrace, to	* ^{ko} :t ^D	546	itch, to (1)	* ^ɣ al ^A	628
heap up, to	* ^{ko} :ŋ ^A	515	hump of an ox	* ^{no} :k ^D	152	itch, to (2)	* ^ɣ om ^A	629
hear, to	* ^ŋ in ^A	616	hundred; string, to	* ^{ro} :j ^C	732	jackal	* ⁿ waj ^A	81
heart, breath	* ^{cu} ŋ ^A	65	hungry	* ^{ju} ək ^D	477	join, to	* ^{to} : ^B	631
heavy	* ^h nak ^D	410	hunt, to	* ^p raw ^B	506	joint	* ^ɣ o: ^C	30
heel	* ^s ɿn ^C	35	hurt (1)	* ^{cep} ^D	487	jug	* ^k raŋ ^A	339
hemp	* ^{pa} :n ^B	349	hurt (2)	* ^{ke} :t ^D	488	kidney-stones	* ^h ni:w ^C	74
hen, young	* ^ʃ qɔ: ^ŋ ^B	94	husk	* ^{ka} :p ^D	161	kill, to	* ^{qa} : ^C	705

kiss, to	*cu:p ^D	540	lie down, to, sleep, to	*nwu:n ^A	549	make a fire, to	*ɖaŋ ^A	672
knee	*ɣow ^B	32	lift, to	*no: ^A	559	make noise, to	*ɖaŋ ^A	512
knife	*mit ^D	323	light (in weight)	*ɕ.baw ^A	411	man, male	*zai:j ^A	285
knot, hair	*klaw ^C	4	lightning	*m.le:p ^D	254	many, much	*hla:j ^A	737
know, to	*ru:w ^C	623	liquid, soft	*hle:w ^A	460	mark, to; aim, to	*hma:j ^A	507
lacquer	*rak ^D	358	liquor	*hlaw ^C	305	marking, patterns	*ɕ.la:j ^A	361
lady	*na:ŋ ^A	289	litter (of young)	*gro:k ^D	749	marrow, bone	*ʔwu:k ^D	53
last (year)	*kla:j ^A	752	little, few	*hno:j ^B	738	master, owner	*ɕw ^C	294
late	*hla: ^C	451	liver	*tap ^D	68	matter	*ywa:m ^A	788
late at night	*ɖu:k ^D	758	long	*ruj ^A	404	meal, early	*ŋa:j ^A	759
laugh, to	*kruwəw ^A	528	long (of time)	*hŋŋ ^A	449	meal, evening	*jm.raw ^A	761
lay (a cloth, etc.) across, to	*ba:t ^D	710	loom	*truk ^D	350	meal, late	*ɕ.lwi:ŋ ^A	760
lazy	*kli:k ^D	482	loose	*ho:m ^A	423	measure from thumb to fingertip	*yu:p ^D	743
leaf	*bau ^A	155	lost	*hloŋ ^A	493	meat, flesh	*n.mɿ: ^C	57
leaf for wrapping (big)	*k.to:ŋ ^A	156	louse, body	*m.lel ^A	131	medicine	*juwə ^A	306
leak, to	*rwo: ^B	698	louse, chicken	*rwɿj ^A	132	middle	*kla:ŋ ^A	772
lean back, to	*ʔi:ŋ ^A	566	louse, head	*traw ^A	128	millet	*hwuəŋ ^C	175
leech, aquatic	*pli:ŋ ^A	116	low, short (not tall)	*tam ^B	409	moan, to	*gra:ŋ ^A	525
leech, land	*da:k ^D	115	lung	*pwɿt ^D	66	monkey	*li:ŋ ^A	85
leg	*p.qa: ^A	31	lye	*ɖaŋ ^B	218	moon, month	*ɖlutən ^A	242
lick, to	*ɕ.lwiə ^A	533	machete, big knife	*jm.ra: ^C	324	moonlight	*hŋa:j ^A	245

morning glory	*ɸun ^C	189	near	*k.rau ^C	452	order, to; blow the nose, to	*saŋ ^B	521
morning, early	*ɸaw ^C	756	neck	*yo. ^A	22	orphan	*gm.ra. ^C	292
mortar (1)	*grok ^D	373	needle	*qem ^A	356	other	*ʔu:n ^B	742
mortar (2)	* ^h rum ^A	374	nephew, niece, or grandchild	* ^h la:n ^A	272	otter	*na:k ^D	91
mosquito	* ^h nuj ^A	119	new	* ^h mx:l ^B	414	outside	*l.no:k ^D	771
moss, aquatic	*daw ^A	205	nibble, to; peck, to	*to:t ^D	539	owl	*gaw ^C	100
mother; woman	*me. ^B	264	night	*yuu:n ^A	754	pair	*gu. ^B	734
mother's younger sibling	*na. ^C	277	night	*yam ^B	757	pangolin	*li ^B	92
mountain	* ^ç .do:j ^A	229	night blind	* ^h wa:ŋ ^A	432	parent's older brother	*luŋ ^A	274
mountain, stone	*pra. ^A	230	nine	*kɔw ^C	731	parent's older sister	*pa. ^C	275
mouse, rat	* ^h nu. ^A	77	nose	*daj ^A	8	partition, lid	*hwa. ^A	316
mouth	*pa:k ^D	12	not (strong 1)	*ɸaw ^B	780	peel, bark	*pluək ^D	160
mucus of the nose	*mu:k ^D	9	not (strong 2)	*bo. ^B	781	peel, to	*po:k ^D	643
mud	*lɔm ^B	241	not (weak)	*mi. ^A	782	penis	*ywaŋ ^A	46
mushroom	* ^h rwet ^D	202	not yet	*paj ^B	783	person, human being	*ɸwun ^A	283
mustard green	*ka:t ^D	190	obtain, to	*daj ^C	683	pestle	*sa:k ^D	375
mute	*ŋwam ^C	466	old (of living beings)	*ke. ^B	415	pick up, to	*ɸjip ^D	554
name	*ɸɔ. ^B	293	old (of things)	*kaw ^B	416	pick, to	*kep ^D	684
name of first month	*ciəŋ ^A	751	one	*nu:ŋ ^B	723	pickle, to	*do:ŋ ^A	660
navel (1)	*ɸwu:. ^A	41	one or so, any	*sak ^D	739	pig	* ^h mu. ^A	76
navel (2)	*ɸli. ^C	42	open (the mouth), to	*pa. ^C	530	pillar	*saw ^A	315

pinch, to	* ^h ni:p ^D	557	pronoun, 1s (strong)	*kaw ^A	717	raw, not ripe	*Ç.dip ^D	419
pit	*Ç.kum ^A	234	pronoun, 1s (weak)	*ku: ^A	716	recite, to	*ʔa:n ^B	529
place, ground	*di: ^B	228	pronoun, 2s (strong)	*mau ^A	719	recognize, to	*cak ^D	622
plant, to	*plwu:k ^D	503	pronoun, 2s (weak)	*muu ^A	718	red	*Ç.dwi:ŋ ^A	391
pliers, thongs	*gi:m ^A	331	pronoun, 3s (strong)	*man ^A	721	reed	*ʔo: ^C	199
plough	*crwaj ^A	370	pronoun, 3s (weak)	*muun ^A	720	repeat, to	*zam ^C	575
pluck, to	*Ç.dec ^D	641	pubic mound	* ^h naw ^B	48	resin	*ja:ŋ ^A	176
plug (a hole), to	*ʔwut ^D	658	pull, to; drag, to	*Ç.la:k ^D	563	rest, to	*bak ^D	695
plugged	*sak ^A	490	puppet, marionette	*hun ^B	360	revolve, to; spin, to	*pan ^B	611
plum	*man ^C	184	pus	* ^h no:ŋ ^A	60	rice	*Ç.qaw ^C	172
point, to	*ji: ^C	567	put out of the mouth, to	*ya:j ^A	534	rice, husked	*sa:l ^A	174
poison (fish), to	*Ç.buə ^A	509	put, to	*s.cv:l ^B	570	rinse, to	*C.lwɔ:ŋ ^C	690
pomelo	*bu:k ^D	185	putrid	*xwi:w ^A	442	river	*da: ^B	238
porcupine	* ^h men ^C	89	quit, to	*ja: ^B	614	road	* ^h rwɔm ^A	387
pot	* ^h mo: ^C	340	raft	*be: ^A	386	roast, to	*juəŋ ^C	666
pound (rice), to	*zo:m ^C	646	rain	*Ç.wuun ^A	252	roof, to	*mwuŋ ^A	674
pound, to	*tam ^A	645	raise, to	*liəŋ ^C	511	root	*Ç.ra.k ^D	167
pound, to (1)	*to:k ^D	649	raise, to	*ŋɔk ^D	558	rope, cord	*jɔ:k	336
pound, to (2)	*dup ^A	648	raisin	*ʔit ^D	197	rotten	*naw ^B	461
pour, to	*rwa:t ^D	605	related by marriage	*t.no:ŋ ^A	290	saddle	*ʔa:n ^A	389
pronoun, 1p	*raw ^A	722	rattan	*Ç.wa:j ^A	198	saliva	*la:j ^A	16

salt	*klwuuə ^A	217	shake, to	*Ç.waj ^A	577	sink, to	*cym ^A	608
salt, to	*ʔuəp ^D	659	shaman	*hmo. ^A	298	sit, to	*naŋ ^B	581
sand	*zwuəj ^A	222	sharp, sharp edge	*ɣm ^A	457	six	*krok ^D	728
satiated	*ʔi:m ^B	478	sharp-pointed	*hle:m ^A	456	skin	*hnaŋ ^A	58
scald, to	*luək ^D	669	sheath, pod	*q.wak ^D	164	sky, heaven	*ɬun ^A	250
scales, fish	*klec ^C	146	shellfish	*ho:j ^A	104	sky, weather	*va. ^C	251
scold, to; revile, to	*da. ^B	522	shin, lower leg	*ye:ŋ ^B	33	slap, to	*txp ^D	647
scoop, to	*tak ^D	662	shoot, to	*njuŋ ^A	508	slash, to	*wal ^A	636
scratch, to	*kaw ^A	569	short (not long)	*tin ^B	405	slave (1)	*ɣɔ:j ^B	295
scurf	*yuj ^A	63	shoulder	*Ç.ba. ^B	24	slave (2)	*kra. ^C	296
see, to	*tran ^A	615	shrimp	*kuŋ ^{B/C}	106	slip and fall, to	*bla:t ^D	595
seedling, rice	*kla. ^C	173	shrimp, small	*ŋjəw ^A	107	slippery, smooth	*m.lu:ɿ ^B	458
seek, to	*kra ^A	630	shrink, to	*hrot ^D	702	slow	*na:n ^A	450
sell, to	*p.qa:j	678	shuttle of loom	*p.raw ^B	351	small	*no:j ^C	408
senior	*craw ^C	417	sibling, older	*bi. ^B	265	smear, to	*da. ^A	709
separate, to	*bra:k ^D	633	sibling, younger	*nwo:ŋ ^C	266	smell (tr.), to	*ɬsm ^A	618
sesame	*r.ŋa. ^A	195	side	*Ç.buəŋ ^C	773	smoke	*ɣwan ^A	209
seven	*ceɿ ^D	729	side, ribs	*k.ra:ŋ ^C	55	snake	*ŋwu:. ^A	113
sew, to	*nep ^D	498	sift, to	*qrxŋ ^A	504	snap, to	*ɬli:t ^D	571
shade	*rym ^B	260	silkworm	*mo:n ^C	137	snare	*re:w ^C	382
shadow, reflection	*ŋaw ^A	261	single, only one	*ɬfəw ^A	733	snore, to	*klɿɿ ^A	543

soak, to (1)	*je: ^B	663	stamp (on), to	*du: p ^D	572	straight	*zɿ: ^B	427
soak, to (2)	* ^h ma: ^B	661	stand, to	*Ç jur:n ^A	596	strand (of rope)	*klɿw ^A	357
soil	*tum ^A	220	star (general)	*t.na:w ^A	243	straw, stubble	*wuwəŋ ^A	378
son-in-law	*Ç.kuəj ^A	281	star (in the sky)	*dfi: ^B	244	stream, mountain	*qrwɿj ^C	236
soot	* ^h mi: ^C	210	starfruit	*gwuwəŋ ^A	183	stretch out, to	* ^h jiəl ^D	580
soul; whorl in the hair	*qwan ^A	301	steal, to	*Ç.lak ^D	685	stretch, to	*ju: t ^D	656
sour	*sɿm ^C	436	steam, to	* ^h nuuŋ ^C	671	stroke, to; carress, to	*Ç.lu:p ^D	568
sow, to; scatter, to	*Ç.wa:l ^B	501	steam, vapor	*s.ʔwɿ:j ^A	259	strong, strength	*re:ŋ ^A	495
spear	*kro:k ^D	326	steep	* ^h liŋ ^B	425	stuck	*ga: ^A	426
spider	*krwa:w ^A	133	stem	*ka:n ^C	159	stump	*to: ^A	166
spill, to	* ^h a: ^B	604	step on, to (1)	*jam ^B	573	sugarcane	*ʔo:j ^C	178
spirit (1)	*pri: ^A	299	step on, to (2)	* ^h jiəp ^D	574	sunshine	*Ç.dwi:t ^D	246
spirit (2)	*mwa:ŋ ^A	300	sterile	* ^h man ^A	471	swallow (n.)	*ʔe:n ^B	99
spleen	*ma:m ^C	69	sticky	* ^h niəw ^A	459	sweat	*r.tuə ^B	59
split bamboo flooring	*wa:k ^D	317	stiff and tired	*muwəj ^B	480	sweep, to	*kwa:t ^D	688
spool	* ^h lwu:t ^D	352	stinger (of a bee)	*laj ^A	153	sweet, delicious	*Ç.wa:l ^A	437
sprout, shoot	* ^h no: ^B	169	stink bug	*ge:ŋl	135	swim, to	*lo:j ^A	606
squeeze, to	*gal ^C	556	stinky	* ^h men ^A	441	swollen	*gau ^B	479
squirrel	*ro:k ^D	88	stomach, belly	*dwu:ŋ ^C	40	tail	*trwɿ:ŋ ^A	145
stairs, ladder	*dfwaj ^A	314	stone	*tri:l ^A	223	take a bite, to	*kat ^D	536
stake	* ^h lak ^D	311	stool	*taj ^B	319	take down, to; put down, to	*ploŋ ^A	642

take up in the two cupped hands, to	*ko:p ^D	553	tie up, to	*la:m ^B	652	tube, bamboo	*baŋ ^{B/C}	341
take, to	*ʔaw ^A	682	tie, to	*ɕm.ruk ^D	651	turn upside down or inside out, to	*pli:n ^C	565
faro	*pruək ^D	192	tight, narrow	*gap ^D	424	turtle	*taw ^B	111
taste, to	*jim ^A	617	time (classifier)	*baj ^A	750	tusk, ivory	*ŋa: ^A	142
tea	*ja: ^A	196	tip, end	*pla:j ^A	776	twist, to; wring, to	*ɕit ^D	579
teach, to	*so:l ^A	694	tip, highest point	*jo:t ^D	777	two	*so:ŋ ^A	724
tend animal, to	*juwəŋ ^C	510	tired, bored	*hna:j ^B	483	undo, to; untie, to	*ke: ^C	653
tendon	*ʔjen ^A	54	tired, exhausted	*hnuwəj ^B	481	unmarried man	*ba:w ^B	286
termite	*mo:t ^D	124	to commission	*zau ^C	714	unmarried woman	*sa:w ^A	287
testicles	*tram	47	tomorrow	*jm.ruk ^D	764	upstream, above	*hnuwə ^A	767
thick	*hna: ^A	412	tongue	*li:n ^C	13	urine	*niəw ^B	49
thin (not fat)	*pro:m ^A	402	tooth	*wan ^A	14	vagina	*hi: ^A	45
thin (not thick)	*ɕ.ba:ŋ ^A	413	top for spinning	*k.ra:ŋ ^B	337	vegetable	*prak ^D	188
thing	*ɣo:ŋ ^A	322	torn	*qa:t ^D	489	Vietnamese	*ke:w ^A	297
this	*naj ^C	774	tortoise, water	*hwuə ^A	112	village	*ba:n ^C	302
thorn	*hna:m ^A	162	township	*muwəŋ ^A	303	vinegar	*hmi: ^B	307
thread (1)	*ɕ.da:j ^C	347	track	*rwu:j ^A	388	vomit, to	*rwuək ^D	535
thread (2)	*hmaj ^A	348	transplant, to	*t.nam ^A	502	waist (1)	*ʔje:w ^A	38
three	*sa:m	725	tree, wood	*mwaj ^C	154	waist (2)	*ɕ.wuəl ^D	39
thunder	*pra: ^C	255	tremble, to	*sal ^B	576	wait, to	*cra: ^C	696
tick	*trep ^D	129	tripod	*giəŋ ^A	320	wake up, to	*k.tu:n ^B	700

walk, to	*pra:j ^C	597	wife of man's younger brother	*luəw ^A	279	yoke, part of	*ʔo:ŋ ^C	372
warm	*ʔun ^B	430	wild	*ʧ.tuən ^B	225	young, soft	*ʔwu.n ^B	418
warn, to	*tuən ^A	523	wild olive	*ko:k ^D	182			
wash (clothes), to	*zak ^D	691	wind	*ʧ.lum ^A	249			
wash, to	*za:w ^A	692	window	*ta:ŋ ^B	318			
wasp	*b.twi:l ^A	120	wing (1)	*wuət ^D	143			
water	*ʧ.nam ^C	206	wing (2)	*pi:k ^D	144			
water buffalo	*ywa:j ^A	78	winnowing basket	*doŋ ^C	376			
water pipe	*lin ^A	369	wipe, to	*jet ^D	689			
weave, to (baskets, mats)	*sa:n	500	with, and	*kap ^D	787			
weave, to (cloth)	*tam ^B	496	withered	*hriəw ^B	462			
weed, to	*ʃla:j ^A	638	worm	*no:l ^A	125			
weep, to	*t.haj ^C	524	wound	*ʃa:t ^D	62			
weigh, to	*jaŋ ^B	516	wrestle, to	*plam ^C	578			
well, spring	*ʃo: ^B	235	writing, book	*su: ^A	345			
wet	*dom ^A	465	yam	*man ^A	191			
which	*ɗau ^A	785	yawn, to	*hra:w ^A	541			
white	*xa:w ^A	393	year	*pi: ^A	766			
white-spotted	*ʃla:ŋ ^B	398	yellow	*hluəŋ ^A	395			
widowed	*hma:j ^C	291	yesterday	*ŋwa: ^A	763			
wife	*miə ^A	280	yoke	*ʔe:k ^D	371			

REFERENCES

- Abramson, Arthur S. (1978). Static and dynamic acoustic cues in distinctive tones. *Language and Speech*, 21, 319-325.
- Abramson, Arthur S. (1962). The vowels and tones in Standard Thai: Acoustical measurements and experiments. *International Journal of American Linguistics*, 28(2), Part II.
- Abramson, Arthur S., & Erickson, Donna M. (1978). Diachronic tone splits and voicing shifts in Thai: Some perceptual data. *Status report on speech research, Haskins Laboratories SR-53(2)*, 85-96.
- Aikhenvald, Alexandra Y., & Dixon, R. M. W. (2001). Introduction. In A. Y. Aikhenvald & R. M. W. Dixon (Eds.), *Areal diffusion and genetic inheritance: problems in comparative linguistics* (pp. 1-23). Oxford; New York: Oxford University Press.
- Alves, Mark J. (2007). A look at North-Central Vietnamese. In R. Wayland, J. Hartmann & P. Sidwell (Eds.), *Papers from the 12th Annual Meeting of the Southeast Asian Linguistics Society 2002* (pp. 1-7). Canberra: Pacific Linguistics.
- Awoi-hathe, Aviong, A-Tý, A-Ly, Cobbey, Maxwell, & Cobbey, Vurnell. (1977). *Northern Roglai Vocabulary = Suraq Vungã Sanãp Radlai = Ngũ-Vưmg Roglai*. Manila, Philippines: Summer Institute of Linguistics.
- Ball, Douglas. (2003). Insights into the Patterns of Sounds: Stress in Tondano. *Journal of Undergraduate Research, University of Rochester*, 2, 15-19.
- Barrack, Charles M. (2002). The Glottalic Theory revisited: A negative appraisal. *Indogermanische Forschungen*, 107, 76-95.
- Baxter, William. (1992). *A handbook of Old Chinese phonology*. Berlin; New York: Mouton de Gruyter.
- Belotel-Grenié, Agnès, & Grenié, Michel (1994, September 18-22, 1994). *Phonation types analysis in Standard Chinese*. Paper presented at Third International Conference on Spoken Language Processing (ICSLP 94), Yokohama, Japan.
- Benedict, Paul K. (1942). Thai, Kadai and Indonesian: A new alignment in Southeastern Asia. *American Anthropologist*, 44, 576-601.

- Benedict, Paul K. (1975). *Austro-Thai language and culture, with a glossary of roots*. New Haven: HRAF Press.
- Benedict, Paul K. (1982). Qiang monosyllabization: a third phase in the cycle. *Linguistics of the Tibeto-Burman Area*, 7(2), 113-114.
- Benedict, Paul K. (1988). Kadai linguistics: the rule of engagement. In J. A. Edmondson & D. B. Solnit (Eds.), *Comparative Kadai: Linguistic studies beyond Tai* (pp. 323-340). Dallas: Summer Institute of Linguistics and the University of Texas at Arlington Publications in Linguistics.
- Benedict, Paul K. (1997a). Linguistic prediction: The case of Saek. In J. A. Edmondson & D. B. Solnit (Eds.), *Comparative Kadai: The Tai branch* (pp. 161-189). Arlington, TX: Summer Institute of Linguistics and the University of Texas at Arlington.
- Benedict, Paul K. (1997b). Tai Languages: Key to Austro-Tai. In A. Abramson (Ed.), *Southeast Asian Linguistics Studies in honor of Vichin Panupong* (pp. 25-28). Bangkok: Chulalongkorn University Press.
- Bennett, J. Fraser. (1994). Iambicity in Thai. *Studies in the linguistic sciences*, 24(1), 39-57.
- Bennett, J. Fraser. (1995). *Metrical foot structure in Thai and Kayah Li: Optimality-theoretic studies in the prosody of two Southeast Asian languages*. Unpublished Ph.D. dissertation, University of Illinois at Urbana-Champaign, Urbana-Champaign.
- Booij, Geert. (1997). The role of the prosodic word in phonotactic generalizations. In T. A. Hall & U. Kleinhenz (Eds.), *Studies on the phonological words* (pp. 47-72). Amsterdam; Philadelphia: John Benjamins Publishing Company.
- Boonsner, Thepbangon. (1984). *The phonological description of Yooy*. Unpublished M.A. thesis, Mahidol University, Nakhorn Pathom.
- Bradley, David. (1982). Register in Burmese. In *Papers in Southeast Asian Linguistics No.8, Tonation* (pp. 117-132). Canberra: Pacific Linguistics.
- Brown, J. Marvin (1965). From Ancient Thai to modern dialects. In *From Ancient Thai to modern dialects, and other writings on historical Thai linguistics* (pp. 69-254). Bangkok: White Lotus.

- Brugman, Johanna C. (2009). *Segments, tone and alignment in Khoekhoe prosody*. Unpublished Ph.D. dissertation, Cornell University, Ithaca, NY.
- Brunelle, Marc. (2008). Monosyllabicization in Eastern Cham. In W. Khanittanan & P. Sidwell (Eds.), *SEALS XIV, Vol I: Papers from the 14 th annual meeting of the Southeast Asian Linguistics Society 2004* (pp. 43-58). Canberra: Pacific Linguistics.
- Brunelle, Marc. (2009). Tone perception in Northern and Southern Vietnamese. *Journal of Phonetics*, 37(1), 79-96.
- Burenhult, Niclas. (2005). A grammar of Jahai. Canberra: Research School of Pacific and Asian Studies, Australian National University.
- Chamberlain, James R. (1975). A new look at the history and classification of the Tai dialects. In J. G. Harris & J. R. Chamberlain (Eds.), *Studies in Tai linguistics in honor of William J. Gedney* (pp. 49-60). Bangkok: Central Institute of English Language, Office of State Universities.
- Chamberlain, James R. (1984). The Tai dialects of Khammouane Province, Laos: Their diversity and origins. *The Science of Language*, 4, 62-95.
- Chamberlain, James R. (1991). The efficacy of the P/PH distinction for Tai languages. In J. R. Chamberlain (Ed.), *The Ram Khamhaeng controversy* (pp. 453-486). Bangkok: Siam Society.
- Chamberlain, James R. (2000). The origin of the Sek implications for Tai and Vietnamese history. In S. Burusphat (Ed.), *Proceedings: the International Conference on Tai Studies, July 29-31, 1998*. (pp. 97-127). Bangkok: Institute of Language and Culture for Rural Development, Mahidol University.
- Chang, Kun. (1975). The reconstruction of Proto-Miao-Yao tones. *Bulletin of the Institute of History and Philology, Academia Sinica*(44), 541-628.
- Chantanakomes, Veena. (1980). *A description of Moken: A Malayo-Polynesian language*. Unpublished M.A. thesis, Mahidol University, Nakhon Pathom.
- Chao, Yuen-Ren. (1930). A system of tone-letters. *Le maître phonétique*, 45, 24-27.
- Chinese Academy of Sciences. (1959). *Bùyīyǔ diàochá bàogào [Research reports on Bouyei languages]*. Běijīng: Science Press.

- Cho, Young-mee Y., & King, Tracy H. (2003). Semisyllables and universal syllabification. In C. Féry & R. van de Vijver (Eds.), *The syllable in Optimality Theory*. Cambridge; New York: Cambridge University Press.
- Clements, G. N. (1990). The role of syllabification cycle in core syllabification. In J. Kingston & M. E. Beckman (Eds.), *Papers in Laboratory Phonology I: Between the grammar and physics of speech* (pp. 283-333). Cambridge; New York, Port Chester, Melbourne, Sydney: Cambridge University Press.
- Clements, G. N., & Osu, Sylvester. (2002). Explosives, implosives, and nonexplosives: the linguistic function of air pressure differences in stops. In C. Gussenhoven & N. Warner (Eds.), *Laboratory phonology 7* (pp. 299-350). Berlin; New York: Mouton de Gruyter.
- Cohn, Abigail. (in progress). *The phonological structure of Indonesian truncations and abbreviation*. Department of Linguistics, Cornell University.
- Cohn, Abigail, & McCarthy, John. (1998). Alignment and parallelism in Indonesian phonology. *Working Papers of the Cornell Phonetics Laboratory*, 12, 53-137.
- Conrady, August. (1896). *Eine indochinesische Causative-Denominativ-Bildung und ihr Zusammenhang mit den Tonaccenten*. Leipzig: Otto Harrassowitz.
- Cooper, Adam. (to appear). Similarity avoidance in the Proto-Indo-European root. *Proceedings of the 32nd Annual Penn Linguistics Colloquium, Penn Working Papers in Linguistics*, 15(1).
- Croft, William. (1990). *Typology and universals*. Cambridge; New York: Cambridge University Press.
- Cushing, Josiah Nelson. (1914). *A Shan and English Dictionary* (2nd ed.). Rangoon: American Baptist Mission Press.
- Davison, Deborah S. (1991). An acoustic study of so-called creaky voice in Tianjin Mandarin. *UCLA Working Papers in Phonetics*, 78, 50-77.
- de Lacouperie, Terrien. (1886). *The languages of China before the Chinese: Researches on the languages spoken by the pre-Chinese races of China proper previously to the Chinese occupation*. London: D. Nutt.

- de Lacy, Paul. (2004). Maximal words and the Māori passive. In J. McCarthy (Ed.), *Optimality Theory in Phonology: A reader* (pp. 495-512). Malden, MA: Blackwell.
- Dench, Alan. (2001). Descent and diffusion: the complexity of the Pilbara situation. In A. Y. Aikhenvald & R. M. W. Dixon (Eds.), *Areal diffusion and genetic inheritance: problems in comparative linguistics* (pp. 105-132). Oxford; New York: Oxford University Press.
- Dhananjayananda, Puttachart. (1997). The emergence of the length distinction in mid-front vowels *e-ee* in Thai. In J. A. Edmondson & D. B. Solnit (Eds.), *Comparative Kadai: The Tai branch* (pp. 327-336). Arlington, TX: The Summer Institute of Linguistics and the University of Texas at Arlington.
- Diffloth, Gérard. (1976). Minor-syllable vocalism in Senoic languages. In P. N. Jenner, L. C. Thompson & S. Starosta (Eds.), *Austroasiatic Studies* (pp. 229-248). Honolulu: The University Press of Hawai'i.
- Diffloth, Gérard. (1977). Towards a history of Mon-Khmer: proto-Semai vowels. *Tōnan Azia Kenkyū*, 14(4), 463-495.
- Diffloth, Gérard. (1984). *The Dvaravati Old Mon language and Nyah Kur*. Bangkok: Chulalongkorn University Printing House.
- Diffloth, Gérard. (1991). Palaungic vowels in Mon-Khmer perspective. In J. H. C. S. Davidson (Ed.), *Austroasiatic languages: Essays in honour of H. L. Shorto* (pp. 13-27). London: School of Oriental and African Studies.
- Diller, Anthony. (2000). The Tai language family and the Comparative Method. In S. Burusphat (Ed.), *Proceedings : the International Conference on Tai Studies, July 29-31, 1998*. (pp. 1-32). Nakhon Pathom: Institute of Language and Culture for Rural Development, Mahidol University.
- Dixon, R. M. W., & Aikhenvald, Alexandra Y. (2002). Word: a typological framework. In R. M. W. Dixon & A. Y. Aikhenvald (Eds.), *Word: A cross-linguistic typology* (pp. 1-41). Cambridge: Cambridge University Press.
- Đoàn, Thiin Thuut. (1996). *Tay-Nung language in the north Vietnam*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.

- Đoàn, Thiện Thuật. (1992). *Ngữ âm tiếng Việt [Vietnamese phonology]*. Hà Nội: Nhà xuất bản Đại học Quốc gia Hà Nội [Hanoi National University Press].
- Downing, Laura J. (2006). *Canonical forms in prosodic morphology*. Oxford; New York: Oxford University Press.
- Durie, Mark. (1985). *A grammar of Acehnese on the basis of a dialect of north Aceh*. Dordrecht, Holland; Cinnaminson, USA: Foris Publications.
- Edmondson, Jerold A., & Yang, Quan. (1988). Word-initial preconsonants and the history of Kam-Sui resonant initials and tones. In J. A. Edmondson & D. B. Solnit (Eds.), *Comparative Kadai: Linguistic studies beyond Tai* (pp. 143-166). Arlington, TX: The Summer Institute of Linguistics and the University of Texas at Arlington.
- Edmondson, Jerry A., & Solnit, David B. (1988). Introduction. In J. A. Edmondson & D. B. Solnit (Eds.), *Comparative Kadai: Linguistic studies beyond Tai* (pp. 1-17). Arlington, TX: The Summer Institute of Linguistics and the University of Texas at Arlington.
- Edmondson, Jerry A., & Solnit, David B. (1997). Introduction. In J. A. Edmondson & D. B. Solnit (Eds.), *Comparative Kadai: The Tai Branch* (pp. 1-26). Arlington, TX: The Summer Institute of Linguistics and the University of Texas at Arlington.
- Ferlus, Michel. (1975). Vietnamien et proto-Viet-Muong. *Asie du Sud-Est et monde insulindien*, 6(4), 21-55.
- Ferlus, Michel. (1989). *Remarques sur le consonantisme des langues kadai*. Paper presented at the 22nd International Conference on Sino-Tibetan Languages and Linguistics. University of Hawai'i, Manoa. October 5-8, 1989.
- Ferlus, Michel. (1990). Remarques sur le consonantisme de proto thai-yay (révision du proto-tai de Li Fangkuei). Paper circulated at the 23rd International Conference on Sino-Tibetan Languages and Linguistics, University of Texas at Arlington.
- Ferlus, Michel. (1992). Histoire abrégée de l'évolution des consonnes initiales du vietnamien. *Mon-Khmer Studies*, 20, 111-127.
- Ferlus, Michel. (1996a). Evolution vers le monosyllabisme dans quelques langues de l'Asie du Sud-Est. Séance du 23 novembre 1996, la Société de Linguistique de Paris.

- Ferlus, Michel. (1996b). Remarques sur le consonantisme du proto kam-sui. *Cahiers de linguistique Asie orientale*, 25(2), 235-278.
- Ferlus, Michel. (1998). *Nhahuen-French-English lexicon (edited and annotated by Pascale Jacq & Paul Sidwell)*. München: Lincom Europa.
- Ferlus, Michel. (1999). *Les dialectes et les écritures des Tai (Thai) de Nghê An (Vietnam)*. Communication de la Treizièmes Journée de Linguistique d'Asie Orientale, Paris (EHESS-CNRS). 10-11 juin 1999. .
- Ferlus, Michel. (2008). The dialects of Nghê An, Vietnam (Tay Daeng, Tay Yo, and Tai Muong). In A. Diller, J. A. Edmondson & Y. Luo (Eds.), *The Tai-Kadai languages*. London; New York: Routledge.
- Fine Arts Department. (1986). *Charuek Nai Prathet Thai [Inscriptions in Thailand]* (Vol. 5). Bangkok: National Library.
- Flemming, Edward. (2004). Contrast and perceptual distinctiveness. In B. Hayes, R. Kirchner & D. Steriade (Eds.), *Phonetically-based Phonology* (pp. 232-276). Cambridge ; New York: Cambridge University Press.
- Freiberger, Nancy. (1964). A phonemic description of Nong (Nung). *Te Reo*, 7, 15-22.
- Freiberger, Nancy, & Bé, Vy Thị. (1976). *Sọc mahn slú nohng fan slihng = Ngũ-vung nùng phan slinh = Nung fan slihng vocabulary: Nùng-Viit-English* Summer Institute of Linguistics: Viin Chuyên-Khảo Ngữ-Học [Institute of Linguistics].
- Gamkrelidze, Thomas V., & Ivanov, Vjacheslav V. . (1972). Lingvističeskaja tipologija i rekonstrukcija sistemy indoevropejskix smyčnyx [Linguistic Typology and the Reconstruction of the Indo-European Occlusives]. In *Konferencija po sravnitel'no-istoričeskoj grammatike indoevropejskix jazykov: Predvaritel'nye materialy, 15-18 [Working Papers of the Conference on the Comparative-Historical Grammar of the Indo-European Languages (12—14 December 1972)]* (Vol. 15-18). Moscow: Nauka.
- Gandour, J., Potisuk, S., & Dechongkit, S. (1994). Tonal coarticulation in Thai. *Journal of Phonetics*, 22, 477-492.
- Gandour, Jack. (1977). On the interaction between tone and vowel length: Evidence from Thai dialects. *Phonetica*, 34, 54-65.

- Gedney, William J. (1965). *Indic loanwords in spoken Thai*. Unpublished Ph.D. dissertation, Yale University, New Haven.
- Gedney, William J. (1972a). A checklist for determining tones in Tai dialects. In S. M. Estellie (Ed.), *Studies in Linguistics in honor of George L. Trager* (pp. 423-438). The Hague: Mouton.
- Gedney, William J. (1972b). A puzzle in Comparative Tai Phonology. In J. G. Harris & R. B. Noss (Eds.), *Tai phonetics and phonology*. (pp. 52-57). Bangkok: Central Institute of English Language, Office of State Universities.
- Gedney, William J. (1989a). Evidence for another series of voiced initials in Proto-Tai. In R. J. Bickner, J. Hartmann, T. J. Hudak & P. Peyasantiwong (Eds.), *Selected papers on Comparative Tai Studies*. Ann Arbor: Center for Southeast Asian Studies, the University of Michigan.
- Gedney, William J. (1989b). Future directions in Comparative Tai Linguistics. In R. J. Bickner, J. Hartmann, T. J. Hudak & P. Peyasantiwong (Eds.), *Selected papers on Comparative Tai Studies* (pp. 7-116). Ann Arbor: Center for Southeast Asian Studies, the University of Michigan.
- Gedney, William J. (1989c). The Saek language of Nakhon Phanom province. In R. J. Bickner, J. Hartmann, T. J. Hudak & P. Peyasantiwong (Eds.), *Selected papers on Comparative Tai Studies* (pp. 373-400). Ann Arbor: Center for Southeast Asian Studies, the University of Michigan. .
- Gedney, William J. (1989d). Speculations on early Tai tones. In R. J. Bickner, J. Hartmann, T. J. Hudak & P. Peyasantiwong (Eds.), *Selected papers on Comparative Tai Studies* (pp. 207-228). Ann Arbor: Center for South and Southeast Asian Studies, University of Michigan.
- Gedney, William J. (1991). Comments on linguistic arguments relating to Inscription One. In J. R. Chamberlain (Ed.), *The Ram Khamhaeng controversy* (pp. 193-226). Bangkok: Siam Society.
- Gedney, William J. (1993). Saek final *-l*: archaism or innovation? In T. J. Hudak (Ed.), *William J. Gedney's Saek language: Glossaries, Texts, and Translations* (pp. 917-974). Ann Arbor: Center for South and Southeast Asian Studies, University of Michigan.
- Gedney, William J. (1995). Linguistic diversity among Tai dialects of southern Guangxi. In T. J. Hudak (Ed.), *William J. Gedney's Central Tai dialects* :

glossaries, texts, and translations (pp. 803-822). Ann Arbor: Center for South and Southeast Asian Studies, University of Michigan.

Gedney, William J. (n.d.). *A comparative Tai wordlist*. manuscript.

Goldsmith, John A. (1995). Phonological Theory. In J. A. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 1-23). Cambridge, MA: Blackwell.

Gordon, Matthews. (2001). A typology of contour tone restriction. *Studies in Language*, 25(3), 423-462.

Gordon, Matthews, & Ladefoged, Peter. (2001). Phonation types: a cross-linguistic overview. *Journal of Phonetics*, 29, 383-406.

Green, Anthony Dubach. (2005). Word, foot, and syllable structure in Burmese. In J. Watkins (Ed.), *Studies in Burmese linguistics* (pp. 1-25). Canberra: Canberra: Pacific Linguistics.

Grierson, G. A. (1903). *Linguistic survey of India*. Calcutta: Office of the Superintendent of Government Printing House.

Guion, Susan G., & Wayland, Rtree. (2004). Aerodynamic coarticulation in sound change or how onset trills can condition a falling tone. In A. Agwuele, W. Warren & S.-H. Park (Eds.), *Proceedings of the 2003 Texas Linguistics Society Conference*. Somerville, MA: Cascadilla Proceedings Project.

Gussmann, Edmond. (1992). Resyllabification and delinking: The case of Polish voicing. *Linguistic Inquiry*, 23(1), 29-56.

Hall, T. Alan. (1997). Phonotactics and the prosodic structure of German. In T. A. Hall & U. Kleinhenz (Eds.), *Studies on the phonological words* (pp. 47-72). Amsterdam; Philadelphia: John Benjamins Publishing Company.

Halle, Moris, & Vergnaud, Jean-Roger. 1980. (1980). Three Dimensional Phonology. *Journal of Linguistic Research*, 1(1), 83-105.

Handel, Zev. (1998). *The medial systems of Old Chinese and Proto-Sino-Tibetan*. Unpublished Ph.D. dissertation, University of California at Berkeley, Berkeley.

Handel, Zev. (2003). Northern Min tone values and the reconstruction of "soften initials". *Language and Linguistics*, 4, 47-84.

- Harris, James W. (1983). *Syllable structure and stress in Spanish: a nonlinear analysis*. Cambridge, MA: MIT Press.
- Harris, Jimmy G. (1975). A comparative word list of three Tai Nüa dialects. In J. G. Harris & J. R. Chamberlain (Eds.), *Studies in Tai linguistics in honor of William J. Gedney* (pp. 202-230). Bangkok: Central Institute of the English Language and Office of State Universities.
- Haudricourt, André-Georges. (1948). Les phonèmes et le vocabulaire du thai commun. *Journal Asiatique*, 236, 197-238.
- Haudricourt, André-Georges. (1952). Les occlusives velaires en thai. *Bulletin de la Société de Linguistique de Paris*, 48, 86-89.
- Haudricourt, André-Georges. (1954a). Comment reconstruire le chinois archaïque. *Word*, 10, 351-364.
- Haudricourt, André-Georges. (1954b). De l'origine des tons en vietnamien. *Journal Asiatique*, 242, 69-82.
- Haudricourt, André-Georges. (1956). De la restitution des initiales dans les langues monosyllabiques: le problème du thai commun. *Bulletin de la Société de Linguistique de Paris*, 52, 307-322.
- Haudricourt, André-Georges. (1958). Les Sek de la province du Cammon (Laos), migration thai ou déportation chinoise? *Journal Asiatique*, 246, 107-108.
- Haudricourt, André-Georges. (1960). Notes sur les dialectes des régions de Moncay. *Bulletin de l'École Française d'Extrême-Orient*, 50, 167-177.
- Haudricourt, André-Georges. (1961). Bipartition et tripartition des systèmes de tons dans quelques langages d'extrême orient. *Bulletin de la Société de Linguistique de Paris*, 56, 163-180.
- Haudricourt, André-Georges. (1963a). La langue Sek. In *Proceedings of the 25th Congress of Orientalists* (Vol. 4, pp. 414-415).
- Haudricourt, André-Georges. (1963b). Remarques sur les initiales complexes de la langue Sek. *Bulletin de la Société de Linguistique de Paris*, 58(1), 156-163.
- Haudricourt, André-Georges. (1966-1974). Daïque (daic). In R. Shafer (Ed.), *Introduction to Sino-Tibetan* (pp. 453-525). Wiesbaden: O. Harrassowitz.

- Haudricourt, André-Georges. (1975). A propos du puzzle de W.J. Gedney. In J. G. Harris & J. R. Chamberlain (Eds.), *Studies in Tai linguistics in honor of William J. Gedney* (Vol. 252-258). Bangkok: Central Institute of English Language.
- Hayes, Bruce. (1995). *Metrical stress theory : principles and case studies*. Chicago: University of Chicago Press.
- Henderson, Eugénie J.A. (1952). The main features of Cambodian pronunciation. *Bulletin of the School of Oriental and African Studies, University of London*, 14, 149-174.
- Hoang, Van Ma. (1997). The sound system of the Tày language of Cao Bằng Province , Vietnam. In J. A. Edmondson & D. B. Solnit (Eds.), *Comparative Kadai: The Tai branch* (pp. 221-235). Arlington, TX: The Summer Institute of Linguistics and the University of Texas at Arlington.
- Hock, Hans Henrich. (1991). *Principles of Historical Linguistics*. Berlin; New York: Mouton de Gruyter.
- Hombert, Jean-Marie. (1977). A model of tonal systems. *UCLA Working Papers in Phonetics*, 36, 20-32.
- Hombert, Jean-Marie, Ohala, John J., & Ewan, William G. (1979). Phonetic explanations for the development of tones. *Language*, 55(1), 37-58.
- Hooper, Joan B. (1976). *An introduction to Natural Generative Phonology*. New York: Academic Press.
- Hualde, José Ignacio. (1991). On Spanish syllabification. In H. Campos & F. Martínez-Gil (Eds.), *Current studies in Spanish linguistics* (pp. 475-494). Washington, D.C.: Georgetown University Press.
- Hudak, Thomas J. (Ed.). (1991a). *William J. Gedney's the Tai dialect of Lungming : glossaries, texts, and translations*. Ann Arbor: Center for South and Southeast Asian Studies, University of Michigan.
- Hudak, Thomas J. (Ed.). (1991b). *William J. Gedney's the Yay language : glossaries, texts, and translations*. Ann Arbor: Center for South and Southeast Asian Studies, University of Michigan.

- Hudak, Thomas J. (Ed.). (1993). *William J. Gedney's the Saek language : glossaries, texts, and translations*. Ann Arbor: Center for South and Southeast Asian Studies, University of Michigan.
- Hudak, Thomas J. (Ed.). (1994). *William J. Gedney's Southwestern Tai dialects : glossaries, texts, and translations*. Ann Arbor: Center for South and Southeast Asian Studies, University of Michigan.
- Hudak, Thomas J. (Ed.). (1995). *William J. Gedney's Central Tai dialects : glossaries, texts, and translations*. Ann Arbor: Center for South and Southeast Asian Studies, University of Michigan.
- Hudak, Thomas J. (Ed.). (1996). *William J. Gedney's Lue dialects : glossaries, texts, and translations*. Ann Arbor: Center for South and Southeast Asian Studies, University of Michigan.
- Hudak, Thomas J. (Ed.). (1997). *William J. Gedney's Tai dialects : glossaries, texts, and translations*. Ann Arbor: Center for South and Southeast Asian Studies, University of Michigan.
- Hudak, Thomas J. (Ed.). (2001). *Minot's White Tai: a phonetic dictionary*. Tempe, AZ: Program for South and Southeast Asian Studies, Arizona State University.
- Hudak, Thomas J. (Ed.). (2008). *William J. Gedney's Comparative Tai source book*. Honolulu: University of Hawaii Press.
- Huffman, Franklin E. (1972). The boundary between monosyllable and disyllable in Cambodian. *Lingua*, 29, 54-66.
- Inkelas, Sharon. (1989). *Prosodic constituency in the lexicon*. Unpublished Ph.D. dissertation, Stanford University, Palo Alto.
- Institute of Language and Culture for Rural Development, Mahidol University, & Kam-Tai Institute, Central University for Nationalities. (2001). *Bouyei-Chinese-English-Thai dictionary*. Nakhon Pathom, Thailand: Institute of Language and Culture for Rural Development, Mahidol University.
- Institute of Language and Culture for Rural Development, Mahidol University, & Thailand and the Kam-Tai Institute, Central University for Nationalities, People's Republic of China. (2003). *Sui-Chinese-Thai-English Dictionary*. Nakhorn Pathom, Thailand: Institute of Language and Culture for Rural Development, Mahidol University.

- International Phonetic Association (Ed.). (1999). *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge; New York: Cambridge University Press.
- Iverson, Gregory K., & Salmons, Joseph C. (1992). The phonology of Proto-Indo-European *DVD root structure constraint. *Lingua*, 87, 290-320.
- Jonsson, Nanna L. (1991). *Proto-Southwestern Tai*. Unpublished Ph.D. dissertation, State University of New York at Albany, Albany, NY.
- Jrang, Ja Kuang, Ja Wi, Ja Dai, Ja Ngai, & Fuller, Eugene. (1983). *Chru vocabulary*. Manila: Summer Institute of Linguistics.
- Karlgren, Bernhard. (1964). *Grammata serica recensa*. Stockholm: Museum of Far Eastern Antiquities.
- Kasisopa, Benjawan. (2003). *Phonological study and genetic classification of Daraang (Palaung) spoken at Nawlae Village, Fang District, Chiangmai Province*. Unpublished M.A. thesis, Chulalongkorn University, Bangkok.
- Kingston, John. (2004). Mechanisms of tone reversal. In S. Kaji (Ed.), *Proceedings of the Third Symposium on Cross-linguistic Studies of Tone* (pp. 57-120). Tokyo: Institute of the Study of the Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.
- Kingston, John, & Diehl, Randy L. (1994). Phonetic knowledge. *Language*, 70, 419-454.
- Kingston, John, & Solnit, David B. (1989). The inadequacy of underspecification. In J. Carter & R.-M. Déchaine (Eds.), *Proceedings of the 19th Meeting of the Northeast Linguistic Society*, 264-278. (pp. 264-278). Amherst, MA: University of Massachusetts, Amherst.
- Kong, Qingming. (1987). Influence of tones upon vowel duration in Cantonese. *Language and Speech*, 30, 387-399.
- Kosaka, Ryuichi. (1992). *Tentative de reconstruction d'un proto-saek-chuang et comparaison de son vocabulaire avec l'ancien siamois*. Unpublished M.A. thesis, Tokyo University of Foreign Studies, Tokyo.
- Kosaka, Ryuichi. (1997). A phonology of the Bu Dai (Thu Lao) language. *Journal of Asian and African Studies*, 53, 137-153.

- Kosaka, Ryuichi. (2002). On the affiliation of Miao-Yao and Kadai –possibility of forming Miao-Dai linguistic family. *Mon-Khmer Studies*, 32, 71-100.
- Kosaka, Ryuichi. (2007). Consonantal lenition and vocalic transfer in Kadai languages: with disyllabic proto-Be-Tai as supporting evidence. In I. Shoichi, A. Simpson, K. L. Adams & P. Sidwell (Eds.), *Papers from the 13th Annual Meeting of the Southeast Asian Linguistics Society 2003* (pp. 93-103). Canberra: Pacific Linguistics.
- Kreitman, Rina. (2008). *Phonetics and phonology of onset clusters: the case of Modern Hebrew*. Unpublished Ph.D. dissertation, Cornell University, Ithaca, NY.
- Kullavanijaya, Pranee. (1992a). *Photchananukrom Chuang Tai-Thai [Southern Zhuang-Thai dictionary]*. Bangkok: Faculty of Arts, Chulalongkorn University.
- Kullavanijaya, Pranee. (1992b). A study of some two-syllabled words in Thai. In *Pan-Asiatic linguistics: Proceedings of the Third International Symposium on Language and Linguistics* (pp. 651-656). Chulalongkorn University, Bangkok, Thailand.
- Kullavanijaya, Pranee. (2001). *Phasa Thai Nuea [The Tai Nua language]*. Bangkok: Faculty of Arts, Chulalongkorn University.
- Kullavanijaya, Pranee, & L-Thongkum, Theraphan. (2000). Linguistic criteria for determining Tai ethnic groups: case studies on Central and South-Western Tais. In S. Burusphat (Ed.), *Proceedings : the International Conference on Tai Studies, July 29-31, 1998*. (pp. 273-298). Bangkok: Institute of Language and Culture for Rural Development, Mahidol University.
- L'École Française d'Extrême-Orient. (1938). *EFEO Questionnaire linguistique*. Unpublished manuscript, Hanoi.
- L-Thongkum, Theraphan. (1989). An acoustic study of the register complex in Kui (Suai). *Mon-Khmer Studies*, 15, 1-19.
- L-Thongkum, Theraphan. (1991). An instrumental study of Chong registers. In J. H. C. S. Davidson (Ed.), *Austroasiatic languages: Essays in honour of H. L. Shorto* (pp. 141-160). London: School of Oriental and African Studies.

- L-Thongkum, Theraphan. (1992). A preliminary reconstruction of Proto-Lakkja (Cha Shan Yao). *Mon-Khmer Studies*, 20, 57-90.
- L-Thongkum, Theraphan. (1997). Implications of the relation of proto-voiced plosives and fricatives in the Dai Tho language of Yunnan Province for a theory of tonal development and Tai language classification. In J. A. Edmondson & D. B. Solnit (Eds.), *Comparative Kadai: The Tai branch* (pp. 191-220). Arlington, TX: The Summer Institute of Linguistics and the University of Texas at Arlington.
- L-Thongkum, Theraphan. (2001). *Languages of the tribes in Xekong Province Southern Laos : a foundation for research and development*. Bangkok: Thailand Research Fund.
- L-Thongkum, Theraphan. (2002). *Old Tai Dam (Black Tai) and the meaning of ambiguous words in Modern Thai elaborate expressions*. Paper presented at the Conference Name|. Retrieved Access Date|. from URL|.
- L-Thongkum, Theraphan (Ed.). (1978). *Siang Lae Rabop Siang Nai Phasa Thai Thin Tai Changwat Sura Sot Thani 19 Amphoe [Sounds and sound systems of Southern Thai spoken in the 16 amphoes of Suratthani province]*. Bangkok: Indegenous Languages of Thailand Project, Chulalongkorn University.
- Ladefoged, Peter, & Maddieson, Ian. (1996). *The sounds of the world's languages*. Oxford, UK ; Cambridge, MA: Blackwell Publishing.
- Lee, Ernest. (1966). *Proto-Chamic phonologic word and vocabulary*. Unpublished Ph.D. dissertation, Indiana University, Bloomington.
- Lehiste, Ilse. (1970). *Suprasegmentals*. Cambridge, MA: MIT Press.
- Levin, Juliette. (1985). *A Metrical Theory of Syllabicity*. Unpublished Ph.D. dissertation, MIT, Cambridge, MA.
- Li, Fang-kuei. (1940). *Lóngzhōu tǔyǔ [The Tai dialect of Lungchow]*. Shāngwù yìnshū guǎn.
- Li, Fang-kuei. (1960). A tentative classification of Tai dialects. In S. Diamond (Ed.), *Culture in history: Essays in honor of Paul Radin* (pp. 951-959). New York: Columbia University Press.

- Li, Fang-kuei. (1964). The phonemic system of the Tai Lü language. *Bulletin of the Institute of History and Philology, Academia Sinica*, 35, 7-14.
- Li, Fang-kuei. (1965). The Tai and the Kam-Sui languages. In G. B. Milner & E. i. J. A. Henderson (Eds.), *Indo-Pacific Linguistic Studies* (Vol. Vol. I, pp. 148-179). Amsterdam: North Holland Publishing Co.
- Li, Fang-kuei. (1971). On Siamese *jaai*. *Bulletin of the Institute of History and Philology, Academia Sinica*, 42(3), 337-340.
- Li, Fang-kuei. (1973). Some dental clusters in Tai. *Bulletin of the School of Oriental and African Studies, University of London*, 36(2), 334-339.
- Li, Fang-kuei. (1976). Sino-Tai. *Computational Analysis of Asian and African Languages*, 3, 39-48.
- Li, Fang-kuei. (1977). *Handbook of Comparative Tai*. Honolulu: University of Hawai'i Press.
- Li, Fang-kuei. (1980). *Shànggǔ yīn yánjiū [Researches on Old Chinese phonology]*. Běijīng: Shāngwù yìnshū guǎn.
- Li, Fang-kuei. (1989). Proto-Tai *kh and *x-. In J. H. C. S. Davidson (Ed.), *South-East Asian Linguistics: Essays in honour of Eugénie J.A. Henderson* (pp. 143-146). London: School of Oriental and African Studies, University of London.
- Li, Jinfang. (1999). *Bùyāngyǔ yánjiū [Studies in the Buyang Language]*. Běijīng: Central University for Nationalities Press.
- Liang, Min, & Zhang, Junru. (1993). Dòng-Tái yǔzú sòngqì qīngsè yīn shēngmǔ de chǎnshēng hé fāzhǎn [The rise and development of voiceless aspirated initials in Kam-Tai languages]. *Mínzú Yǔwén*(5), 10-15.
- Liang, Min, & Zhang, Junru. (1996). *Dòng-Tái yǔzú gàilùn [Outline of Tai-Kadai languages]*. Běijīng: China Social Sciences Press.
- Liberman, Mark. (1975). *The intonational system of English*. New York: Garland Publishing.
- Liu, Baoyuan. (1999). *Hàn-Yáo Cídiǎn (Lājīā yǔ) [Chinese-Yao dictionary (Lakkja language)]*. Chéngdū: Sichuan Nationalities Press.

- Luo, Yongxian. (1996). Tonal irregularities in Tai revisited. *Mon-Khmer Studies*, 25, 69-102.
- Luo, Yongxian. (1997). *The subgroup structure of the Tai languages: A historical-comparative study*. Berkeley: Project on Linguistic Analysis, University of California, Berkeley.
- Luo, Yongxian. (1999). *A dictionary of Dehong, Southwest China*. Canberra: Pacific Linguistics.
- Maddieson, Ian. (1984a). The effects of a voicing distinction in sonorants and their implication for a theory of tonogenesis. *Journal of Phonetics*, 12, 9-15.
- Maddieson, Ian. (1984b). *Patterns of sounds*. Cambridge; New York: Cambridge University Press.
- Manomaivibool, Prapin. (1975). *Study of Sino-Thai lexical correspondences*. Unpublished Ph.D. dissertation, University of Washington, Seattle.
- Manomaivibool, Prapin. (1976). Layers of Chinese loanwords in Thai. In T. W. Gething, J. G. Harris & P. Kullavanijaya (Eds.), *Tai linguistics in honor of Fang-Kuei Li* (pp. 179-184). Bangkok: Central Institute of English Language, Office of State Universities.
- Maran, La Raw. (1978). *Burmese and Jingpho; a study of tonal linguistic processes*. Urbana-Champaign: Center for Asian studies, University of Illinois.
- Maspéro, Henri. (1912). Études sur la phonétique historique de la langue annamite. Les initiales. *Bulletin de l'École Française d'Extrême-Orient*, 12, 1-126.
- Matisoff, James A. (1973). Tonogenesis in Southeast Asia. In L. M. Hyman (Ed.), *Consonant types & tones* (pp. 71-95). Los Angeles: The Linguistic Program, University of Southern California.
- Matisoff, James A. (1992). Siamese *jaai* revisited, or ¡Ay, madre!: a case study in multiple etymological possibilities. In C. J. Compton & J. F. Hartmann (Eds.), *Papers on Tai languages, linguistics and literatures* (pp. 111-117). DeKalb: Northern Illinois Center for Southeast Asian Studies
- Matisoff, James A. (2001). Genetic versus contact relationship: Prosodic diffusibility in South-East Asian languages. In A. Y. Aikhenvald & R. M. W. Dixon (Eds.),

Areal diffusion and genetic inheritance: problems in comparative linguistics (pp. 291-327). Oxford; New York: Oxford University Press.

- Matisoff, James A. (2006). *English-Lahu lexicon*. Berkeley: University of California Press.
- Mayrhofer, Manfred. (1986). *Lautlehre (Segmentale Phonologie des Indogermanischen)* (Vol. I). Heidelberg: Carl Winter.
- Mazaudon, Martine. (1977). Tibeto-Burman tonogenetics. *Linguistics of the Tibeto-Burman area*, 3(2), 1-123.
- McCarthy, John, & Prince, Alan. (1986). *Prosodic morphology*. manuscript. University of Massachusetts, Amherst, and Brandeis University.
- Michaud, Alexis. (2004). Final consonants and glottalization: new perspectives from Hanoi Vietnamese *Phonetica*, 61(2-3)(2-3), 119-146.
- Morén, Bruce, & Zsiga, Elizabeth. (2006). The lexical and post-lexical phonology of Thai tones. *Natural Language and Linguistic Theory*, 24(1), 113—178.
- Morey, Stephen. (2004). Tonal change in the Tai languages of Northeast India. *Linguistics of the Tibeto-Burman Area*, 28(2).
- Morey, Stephen. (2005). *The Tai languages of Assam: A grammar and texts*. Canberra: Pacific Linguistics.
- Nacaskul, Kanchana. (1978). The syllabic and morphological structure of Cambodian words. *Mon-Khmer Studies*, 7, 183-200.
- Nespor, Marina, & Vogel, Irene. (1986). *Prosodic Phonology*. Dordrecht, Holland; Riverton, USA: Foris Publications.
- Nguyễn, Văn Lợi, & Edmondson, Jerold A. . (1998). Tones and voice quality in modern northern Vietnamese: Instrumental case studies. *Mon-Khmer Studies*, 28, 1-18.
- Nishida, Tatsuo. (1954). Tonēmu ni yoru Tai shogo hikaku gengogakuteki kenkyū [Comparative linguistic studies of the Tai languages based on tonemes]. *Gengo Kenkyū*, 25, 19-46.

- Nishida, Tatsuo. (1955). Makku-Sui go to kyōtsu go hikaku gengogakuteki kenkyū [Mak-Sui and Common Tai]. *Gengo Kenkyū*, 28, 30-62.
- Nishida, Tatsuo. (1975a). Common Tai and Archaic Chinese. *Studia Phonologica* (2-12).
- Nishida, Tatsuo. (1975b). On the developments of tones in Tibetan. *Acta Asiatica*, 25, 43-55.
- Norquest, Peter. (2007). *A phonological reconstruction of Proto-Hlai*. Unpublished Ph.D. dissertation, University of Barcelona, Tucson, AZ.
- Ohala, John J. (2003). Phonetics and historical phonology. In B. D. Joseph & R. D. Janda (Eds.), *The handbook of Historical Linguistics* (pp. 669-686). Malden, MA: Blackwell Publishing.
- Ostapirat, Weera. (1995). Notes on Laha final -l. *Linguistics of the Tibeto-Burman Area*, 18(1), 173-181.
- Ostapirat, Weera. (2000). Proto-Kra. in *Linguistics of the Tibeto-Burman Area*, 23(1), 1-251.
- Ostapirat, Weera. (2004). Proto-Hlai sound system and lexicons. In Y.-c. Lin, F.-m. Hsu, C.-c. Lee, J. T.-S. Sun, H.-f. Yang & D.-a. Ho (Eds.), *Studies on Sino-Tibetan languages: Papers in honor of Professor Hwang-cherng Gong on his seventieth birthday* (pp. 121-175). Taipei: Institute of Linguistics. Academia Sinica.
- Ostapirat, Weera. (2005). Kra-Dai and Austronesian: Notes on phonological correspondences and vocabulary distribution. In L. Sagart, R. Blench & A. Sanchez-Mazas (Eds.), *The peopling of East Asia: putting together archaeology, linguistics and genetics* (pp. 107-131). London; New York: RoutledgeCurzon.
- Ostapirat, Weera. (2007). Proto-Tai velars and postvelars. In J. G. Harris, S. Burusphat & J. E. Harris (Eds.), *Studies in Tai and Southeast Asian Linguistics* (pp. 293-300). Bangkok: Ekphimthai Ltd.
- Parker, Stephen. (2002). *Quantifying the Sonority Hierarchy*. Unpublished Ph.D. dissertation, University of Massachusetts Amherst., Amherst.
- Petsuk, Rasi. (1978). *General Characteristics of the Khün language*. Unpublished M.A. thesis, Mahidol University, Nakhorn Pathom.

- Pham, Andrea Hoa. (2003). *Vietnamese tone : a new analysis*. New York: Routledge.
- Phothiban, Phuthachart, & Trongdi, Thananan. (1998). *Sathana Khong Phasa Tak Bai Nai Phasa Thai Thin [The Position of Tak Bai Tai in Tai Dialects (Phase 2)]* (Research report). Pattani: Faculty of Humanities, Prince of Songkhla Univeristy.
- Pittayaporn, Pittayawat. (2005a). Moken as a Mainland Southeast Asian language. In A. Grant & P. Sidwell (Eds.), *Chamic and Beyond* (pp. 189-210). Canberra: Pacific Linguistics.
- Pittayaporn, Pittayawat. (2005b). *Phu Thai Kapong*. fieldnote. Cornell University.
- Pittayaporn, Pittayawat. (2005c). *Tone, minor syllable, and infixation in Kammu*. Paper presented at the 11th annual Midcontinental Workshop on Phonology, University of Michigan, Ann Arbor, Michigan. November 4-6, 2005.
- Pittayaporn, Pittayawat. (2006). *Giay of Sapa*. fieldnote. Cornell University.
- Pittayaporn, Pittayawat. (2007a). *A chronology-sensitive approach to subgrouping: The case of Southwestern Tai*. Paper presented at the Conference Name|. Retrieved Access Date|. from URL|.
- Pittayaporn, Pittayawat. (2007b, August 6-10, 2007). *Directionality of tone change*. Paper presented at the 16th International Congress of Phonetic Sciences, Saarbrücken, Germany.
- Pittayaporn, Pittayawat. (2007c). *Tay of Bao Yen*. fieldnote. Cornell University.
- Pittayaporn, Pittayawat. (2007d). *Tay of Cao Bang*. fieldnote. Cornell University.
- Pittayaporn, Pittayawat. (2007e). *Tay of Sapa*. fieldnote. Cornell University.
- Pittayaporn, Pittayawat. (2008a). *Lakkja*. fieldnote. Cornell University.
- Pittayaporn, Pittayawat. (2008b). *Proto-Southwestern Tai: A New Reconstruction*. Paper presented at 18th Annual Meeting of the Southeast Asian Linguistics Society, Universiti Kebangsaan Malaysia, Bangi. May 21-22, 2008.
- Pittayaporn, Pittayawat. (2008c). *Zhuang dialect of Shangsi*. fieldnote. Cornell University.

- Pittayaporn, Pittayawat. (to appear-a). Establishing relative chronology of Palaung sound changes using Tai loanwords. *Mon-Khmer Studies*, 39.
- Pittayaporn, Pittayawat. (to appear-b). Proto-Southwestern-Tai revised: A new reconstruction. *Journal of the Southeast Asian Linguistics Society*.
- Pulleyblank, Edwin G. (1962). The consonantal system of Old Chinese. *Asia Major*, 9, 58-144, 206-165.
- Pulleyblank, Edwin G. (1984). *Middle Chinese : a study in historical phonology*. Vancouver: University of British Columbia Press.
- Pulleyblank, Edwin G. (1991). *Lexicon of reconstructed pronunciation in Early Middle Chinese, Late Middle Chinese, and Early Mandarin*. Vancouver: University of British Columbia Press.
- Qin, Xiaohang. (1992). Zhuàngyǔ liàngcí pǒu4 de láiyuán [Origin of Zhuang classifier pǒu4]. *Mínzú Yǔwén*(4).
- Ratliff, Martha S. (1992). *Meaningful tone : a study of tonal morphology in compounds, form classes and expressive phrases in White Hmong*. DeKalb, IL: Northern Illinois University, Center for Southeast Asian Studies.
- Rhee, Seok-Chae. (2003). Onset-coda asymmetries in Mishmi and other Southeast Asian languages. In K. L. Adams, T. J. Hudak & F. K. Lehman (Eds.), *Papers from the Seventh Annual Meeting of the Southeast Asian Linguistics Society* (pp. 137-152). Tempe, AZ: Program for Southeast Asian Studies, Arizona State University.
- Ross, Peter. (1996). Dao Ngan Tay: a B-language in Vietnam. *Mon-Khmer Studies*, 25, 133-139.
- Rungrueangsri, Udom. (1991). *Photchananukrom Lanna - Thai Chabap Mae Fah Luang [Mae Fah Luang's Lanna-Thai Dictionary]*. Bangkok: มูลนิธิแม่ฟ้าหลวง [Mae Fah Luang Foundation].
- Sagart, Laurent. (1986). On the departing tone. *Journal of Chinese Linguistics*, 14(1), 90-113.
- Sagart, Laurent. (1988). Glottalized tones in China and South-East Asia. In D. Bradley, E. J. A. Henderson & M. Mazaudon (Eds.), *Prosodic analysis and*

- Asian linguistics: to honour R. K. Sprigg* (pp. 83-93). Canberra: Pacific Linguistics.
- Sagart, Laurent. (1993). Austronesian finals and the origin of Chinese tones. In J. A. Edmondson & K. Gregerson (Eds.), *Tonality in Austronesian languages* (pp. 41-60). Honolulu: University of Hawai'i Press.
- Sagart, Laurent. (1999). *The roots of Old Chinese*. Amsterdam; Philadelphia: John Benjamins Pub. Co.
- Sagart, Laurent. (2004). The higher phylogeny of Austronesian and the position of Tai-Kadai. *Oceanic Linguistics*, 43(2), 411-444.
- Sagart, Laurent. (2005). Tai-Kadai as a subgroup of Austronesian. In L. Sagart, R. Blench & A. Sanchez-Mazas (Eds.), *The peopling of East Asia: putting together archaeology, linguistics and genetics* (pp. 177-181). London; New York: RoutledgeCurzon.
- Sarawit, Mary E. (1973). *The Proto-Tai vowel system*. Unpublished Ph.D. dissertation, University of Michigan, Ann Arbor.
- Sarawit, Mary E. (1979). A sketch of a dialect of Mae Sot. In T. Gething & D. L. Nguyen (Eds.), *Papers in Southeast Asian Linguistics No.6, Tai Studies in honor of William J. Gedney* (pp. 75-84). Canberra: Pacific Linguistics.
- Sastry, G. Devi Prasada. (1984). *Mishmi grammar*. Mysore: Central Institute of Indian Languages.
- Saul, Janice E., & Freiberger Wilson, Nancy. (1980). *Nùng grammar*. Dallas, TX: Summer Institute of Linguistics ; Arlington : University of Texas at Arlington.
- Schlegel, Gustav. (1902). Siamese studies. *T'oung Pao, ser. 2: vol. 2*(supplement), 1-128.
- Schmidt, Wilhelm. (1926). *Die Sprachfamilien und Sprachenkreise der Erde*. Heidelberg.
- Schuessler, Axel. (2007). *ABC etymological dictionary of Old Chinese*. Honolulu: University of Hawai'i Press.
- Schuessler, Axel. (2009). *Minimal Old Chinese and Later Han Chinese: A Companion to Grammatica Serica Recensa*. Honolulu: University of Hawai'i Press.

- Selkirk, Elizabeth. (1981). On the nature of phonological representation. In J. Anderson, J. Laver & T. Myers (Eds.), *The cognitive representation of speech* (pp. 379-388). Amsterdam: North Holland.
- Selkirk, Elizabeth O. (1984). On the major class features and syllable theory. In M. Aronoff & R. T. Oehrle (Eds.), *Language sound structure* (Vol. 107-136). Cambridge, MA: MIT Press.
- Shaw, Patricia. (1994). The prosodic constituency of minor syllable. In E. Duncan, D. Farkas & P. Spaelti (Eds.), *Proceedings of West Coast Conference on Formal Linguistics 12* (pp. 117-132). Stanford: Stanford Linguistics Association.
- Shorto, Harry. (1960). Word and syllable pattern in Palaung. *Bulletin of the School of Oriental and African Studies, University of London*, 23(3), 544-557.
- Shorto, Harry. (2006). A Mon-Khmer comparative dictionary. In P. Sidwell, D. Cooper & C. Bauer (Eds.). Canberra: Pacific Linguistics.
- Silverman, Daniel, Blackenship, Barbara, Kirk, Paul, & Ladefoged, Peter. (1995). Phonetic structures in Jalapa Mazatec. *Anthropological Linguistics*, 37, 70-88.
- Simmonds, E. H. S. (1965). Notes on some Tai dialects of Laos and neighbouring regions. *Lingua*, 14, 133-147.
- Sloan, Kelly. (1988). Bare-consonant reduplication: Implications for a prosodic theory of reduplication. In H. Borer (Ed.), *Proceedings of West Coast Conference on Formal Linguistics 7* (Vol. 319-330). Stanford: Stanford Linguistics Association.
- Steriade, Donca. (1982). *Greek prosodies and the nature of syllabification*. Unpublished Ph.D. dissertation, MIT, Cambridge, MA.
- Stevens, Kenneth N., Keyser, S. Jay, & Kawasaki, Haruko (1986). Toward a phonetic and phonological theory of redundant features. In J. S. Perkell & D. H. Klatt (Eds.), *Invariance and variability in speech processes* (pp. 426-449). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Strecker, David. (1979). A preliminary typology of tone shapes and tonal sound changes in Tai: the Lān Nā A-tones. In T. L-Thongkum, V. Panupong, P. Kullavanijaya & K. M. R. Tingsabadh (Eds.), *Studies in Tai and Mon-Khmer phonetics and phonology in honour of Eugénie J.A. Henderson*. Bangkok: Chulalongkorn University Press.

- Strecker, David. (1983). Proto-Tai vowels revisited: a comparison and critique of the work of Sarawit and Li. *Linguistics of the Tibeto-Burman area*, 7(2).
- Strecker, David. (1984). *Proto-Tai personal pronouns*. Unpublished Ph.D. dissertation, University of Michigan, Ann Arbor, MI.
- Strecker, David. (1985). The classification of the Caolan languages. In S. Ratanakul, D. Thomas & S. Premrirat (Eds.), *Southeast Asian Linguistic Studies presented to André-G. Haudricourt* (pp. 479-492). Nakhorn Pathom: Institute of Language and Culture for Rural Development, Mahidol University.
- Strecker, David. (1988). Gedney's Puzzle in Kam-Sui. In J. A. Edmondson & D. B. Solnit (Eds.), *Comparative Kadai: Linguistic studies beyond Tai* (pp. 107-128). Arlington, TX: The Summer Institute of Linguistics and the University of Texas at Arlington.
- Svantesson, Jan-Olof. (1983). *Kammu phonology and morphology*. Lund: CWK Gleerup.
- Svantesson, Jan-Olof. (1991). Hu – a language with unorthodox tonogenesis. In J. H. C. S. Davidson (Ed.), *Austroasiatic languages: Essays in honour of H. L. Shorto* (pp. 67-79). London: School of Oriental and African Studies.
- t'Hart, Johan, Collier, René, & Cohen, Antoine. (1990). *A perceptual study of intonation: An experiment-phonetic approach to speech melody*. Cambridge: Cambridge University Press.
- Tài, Nguyyn Văn. (1982). *Ngữ âm tiếng Mường qua các phương ngôn [Muong phonology across dialects]*. Hà Nội: Nguyyn Văn Tài.
- The Royal Institute. (2003). *Photchananukrom Chabap Ratchabandit Sathan Pho So 2542 [The Royal Institute's Dictionary 2002]*. Bangkok: Nanmee Book Publications.
- Theranon, Phaninthara. (2002a). Changes in phonetic characteristics of Thai Tone 4. *Warasarn Aksorasat, Silpakorn University*, 24(1-2), 188-209.
- Theranon, Phaninthara. (2002b). Thai tones in the reigns of King Rama VI and King Rama IX. *Journal of Language and Linguistics*, 21(1), 32-46.
- Thisaphong, Arilak. (1985). *Rabop Siang Phasa Ka Loeng Thi Ban Dong Mafai Changwat Sakon Nakhon [The Phonological description of Kaloeng at Ban*

- Dong Mafai, Sakonnakhorn Province*. Unpublished M.A. thesis, Mahidol University, Nakhorn Pathom.
- Thompson, Laurence C. (1965). *A Vietnamese reference grammar* Seattle: University of Washington Press.
- Thurgood, Graham. (1988). Notes on the reconstruction of Proto-Kam-Sui. In J. A. Edmondson & D. B. Solnit (Eds.), *Comparative Kadai: Linguistic studies beyond Tai* (pp. 179-218). Dallas: Summer Institute of Linguistics and the University of Texas at Arlington Publications in Linguistics.
- Thurgood, Graham. (1996). Language contact and the directionality of internal drift: the development of tones and registers in Chamic. *Language*, 71(1), 1-31.
- Thurgood, Graham. (1999). *From ancient Cham to modern dialects : two thousand years of language contact and change*. Honolulu: University of Hawai'i Press.
- Thurgood, Graham. (2002a). A comment on Gedney's proposal for another series of voiced initials in proto-Tai. . In R. B. Bauer (Ed.), *Studies in Southeast Asian Languages* (pp. 169-183). Canberra: Pacific Linguistics.
- Thurgood, Graham. (2002b). Vietnamese and tonogenesis: revising the model and the analysis. *Diachronica*, 19(2).
- Thurgood, Graham. (2005). The tones from Proto-Chamic to Tsat [Hainan Cham]: Insights from Zheng 1997 and from summer 2004 fieldwork. In A. Grant & P. Sidwell (Eds.), *Chamic and beyond: Studies in Mainland Austronesian languages* (pp. 247-271). Canberra: Pacific Linguistics.
- Thurgood, Graham. (2007). Comments on Gedney's proposal for another series of initials in Proto-Tai revisited. In J. G. Harris, S. Burusphat & J. E. Harris (Eds.), *Studies in Tai and Southeast Asian Linguistics* (pp. 241-262). Bangkok: Ekphimthai Ltd.
- Trask, R. L. (2000). *The dictionary of historical and comparative linguistics*. Chicago: Fitzroy Dearborn.
- Trongrat, Phutphong. (1998). *A Comparative Phonological Study of Thai Phuan at Suphanburi and Sukhothai Provinces*. Unpublished M.A. Thesis, Mahidol, Nakhorn Pathom.

- Umar, Umaiyah Haji. (2003). *A classification of Thai-Kedah dialects using phonological characteristics*. Bangkok: Nutch Printing Limited Partnership.
- Ussishkin, Adam. (2005). A fixed prosodic hierarchy of nonconcatenative templatic morphology. *Natural Language and Linguistic Theory*, 23, 169-218.
- van der Hulst, Harry, & van de Weijer, Jeroen. (1995). Vowel harmony. In J. A. Goldsmith (Ed.), *The handbook of Phonological Theory* (pp. 495-534). Cambridge, MA: Blackwell.
- Varasarin, Uraisi. (1984). *Les éléments khmers dans la formation de la langue siamoise*. Paris: Société d'études linguistiques et anthropologiques de France.
- Venneman, Theo (1972). On the Theory of Syllabic Phonology. *Linguistische Berichte*, 18, 1-18.
- Wang, Stephen S. (1966). *Phonology of Chinese loanwords in a northern Tai dialect*. Unpublished Ph.D. dissertation, University of Washington, Seattle.
- Wayland, Ratre, & Guion, Susan G. (2005). Sound changes following the loss of /r/ in Khmer: a new tonogenetic mechanism. *Mon-Khmer Studies*, 35, 55-82.
- Wichiankhiaw, Arunrat et al. (1996). *The Northern Thai Dictionary of Palm-Leaf Manuscripts* (2nd ed.). Chiangmai, Thailand: Samnakphim Tratsawin.
- Wolff, John U. (in press). *Proto-Austronesian Phonology and glossary*.
- Wulff, K. (1934). *Chinesisch und Tai : sprachvergleichende Untersuchungen*. København [Copenhagen]: Levin und Munksgaard, Ejnar Munksgaard.
- Wulff, K. (1942). *Über das Verhältnis des Malayo-Polynesischen zum Indochinesischen*. København [Copenhagen]: Levin und Munksgaard.
- Wurm, S. A., & Wilson, B. (1975). *English finderlist of reconstructions in Austronesian languages (Post-Brandstetter)*. Canberra: Department of Linguistics, Research School of Pacific Studies, Australian National University.
- Xing, Gongwan. (2000). *Hónghé Shàngyóu Dǎiyǎ yǔ [Tai Ya of the Upper Reach of the Red River]*. Běijīng: Language and Literature Publishing.

- Xu, Ching X, & Xu, Yi. (2003). Effects of consonant aspiration on mandarin Tones. *Journal of the International Phonetic Association*, 33, 165-181.
- Xu, Yi. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics*, 25, 61-83.
- Xu, Yi. (2001). Fundamental frequency peak delay in Mandarin. *Phonetica*, 58, 26-52.
- Xu, Yi. (2004). Understanding tone from the perspective of production and perception. *Language and Linguistics*, 5, 757-797.
- Yip, Moira. (2001). Tonal features, tonal inventories and phonetic targets. In C. Iten & A. Neeleman (Eds.), *UCL Working Papers in Linguistics 13* (pp. 303-329). London: Department of Phonetics and Linguistics, University College London.
- Yu, Alan. (to appear). Tonal effects on perceived vowel duration. In C. Fougeron & M. D'Imperio (Eds.), *Laboratory Phonology 10*. Berlin; New York: Mouton de Gruyter.
- Zec, Draga. (2005). Prosodic differences among function words. *Phonology*, 22, 77-112.
- Zec, Draga. (2007). The syllable. In P. de Lacy (Ed.), *The Cambridge handbook of Phonology* (pp. 161-194). Cambridge: Cambridge University Press.
- Zhang, Jie. (2002). *The effects of duration and sonority on contour tone distribution*. New York & London: Routledge.
- Zhang, Junru. (1995). Dòng-Tái yǔzú qīngchún yīn de chǎnshēng hé fāzhǎn [The rise and development of labiodental sounds in Kam-Tai languages]. *Mínzú Yǔwén*(1), 28-33.
- Zhang, Junru, Liang, Min, Ouyang, Jueya, Zheng, Yiqing, Li, Xulian, & Xie, Jianyou (1999). *Zhuàngyǔ fāngyán yánjiū [Zhuang dialect research]*. Chéngdū: Sichuan Nationalities Press.
- Zhou, Yaowen, & Luo, Meizhen. (2001). *Dǎiyǔ fāngyán yánjiū [Dai dialect studies]*. Běijīng: Minorities Press.