

Linguistics of The Tibeto-Burman Area; Vol.10,
Matisoff James A.

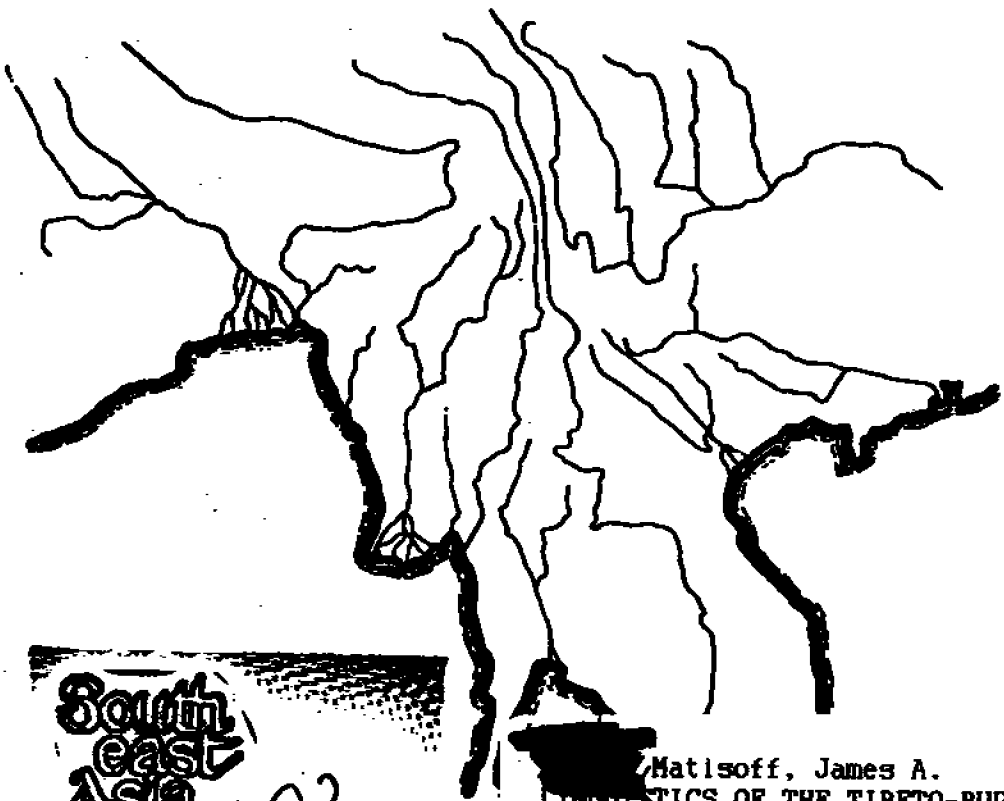
GEN.1872.00

WINTER 2003

TIBETO-BURMAN AREA

Volume 10.2

Fall 1987



South
East
Asia 02
Community Resource Center
Folsom Cordova USD
1460 Cordova Lane
Rancho Cordova, CA 95670

Matisoff, James A.
LINGUISTICS OF THE TIBETO-BURMAN AREA
ENG; periodical

LINGUISTICS OF THE TIBETO-BURMAN AREA

James A. Matisoff, Editor
University of California, Berkeley

EDITORIAL COMMITTEE

Paul K. BENEDICT
Briarcliff Manor, NY

Nicholas C. BODMAN
Cornell University

David BRADLEY
La Trobe University,
Melbourne, Australia

Scott DE LANCEY
University of Oregon

Inga-Lill HANSSON
University of Lund, Sweden

F. K. LEHMAN
University of Illinois

Martine MAZAUDON
Centre National pour la
Recherche Scientifique, Paris

Boyd MICHAILOVSKY
Centre National pour la
Recherche Scientifique, Paris

Graham THURGOOD
California State University,
Fresno

Julian K. WHEATLEY
Cornell University

010 - LCCN
022 - LSSN
050 - Call number (LC)

OCLC number: 4790670

FROM THE EDITOR

This issue of LTBA is devoted entirely to the fascinating and understudied Hmong-Mien (Miao-Yao) language family. Many of the articles date from a panel on Hmong Language and Linguistics chaired by David Strecker during the Southeast Asian Studies Summer Institute (SEASSI) Conference at the University of Michigan in the summer of 1985. Later several papers on Mien (by Caron, Court, Purnell, and Solnit) were added, along with last-minute contributions by Lyman and Jaisser. The end result is a well-rounded set of papers that cover a range of synchronic and diachronic topics in Hmong-Mien phonology, grammar, and orthography.

We would like to thank David Strecker and Brenda Johns for conceiving this idea of a special issue on Hmong-Mien. Tanya Smith was ably assisted in the preparation of the manuscripts by Steve Baron, Amy Dolcourt, John Lowe, and Jean McAneny, to all of whom many thanks.

A cumulative index to the first ten volumes of LTBA appears on pp. 177-180. Also, back issues of LTBA (originals when available, photocopies otherwise) may be purchased according to the price schedule on p.182.

We are about to launch a serious Subscription Drive to further stabilize our financial base. We welcome suggestions from our present readership as to whom we might solicit to become new subscribers!

JAM

* * *

All correspondence and manuscripts should be sent to:

Ms. Tanya Smith, Editorial Secretary, LTBA
Sino-Tibetan Etymological Dictionary and Thesaurus Project
2041 Bancroft Way, Suite 204
Berkeley, CA 94704

We welcome manuscripts submitted on Macintosh disks. Otherwise, manuscripts should be typed single-spaced on standard size paper with at least 3/4 inch (2 centimeters) margins and page numbers pencilled on the reverse of each page. The format for notes and bibliography is left to the author's discretion. Authors will receive 10 free copies of their published article.

The Hmong-Mien Languages

David Strecker

The Hmong-Mien family, which is also called the Miao-Yao family, is one of the major language stocks of Southeast Asia. Most of the languages of the family are spoken in southwestern China, but several languages are spoken in the northern portions of Vietnam, Laos, Thailand, and Burma.

A detailed understanding of the family first became available to linguists as a result of the dialect surveys and descriptive linguistic work undertaken in China after the Liberation. More recently, the Indochinese war forced thousands of speakers of Hmong-Mien languages from Laos to resettle in Australia, Europe, and the Americas, so that a number of linguists outside of China and Southeast Asia have now become active Hmong-Mienists. The present collection includes the work of several of these scholars.

Even today some linguists who are not specialists in these languages tend to think of Hmong-Mien as a small and obscure group, comprising just two languages, "Miao" and "Yao". In fact, Hmong-Mien comprises some two dozen major subgroups which are sufficiently different from one another to be mutually unintelligible. Within several of these subgroups there exist further subdivisions showing marked differences in vocabulary and phonology and considered by their speakers to be separate languages, despite some degree of mutual intelligibility. It is probably closer to the mark to say that there are between 30 and 40 Hmong-Mien languages.

The various subgroups of Hmong-Mien can be grouped in three major branches:

Hmongic
Ho Nte
Mienic

Mienic, also called Yao, is a fairly close-knit group comprising six languages as shown in figure 2. Mun and Mien are widely spoken in China, Vietnam, Laos, and Thailand. The other four languages are confined to a relatively limited region within China.

Ho Nte, also called She, is a single language spoken in four districts near Hong Kong (figure 1).

By contrast, Hmongic, also called Miao, is extremely diverse. I have tried to give a rough idea of the diversity of Hmongic languages and their approximate geographic distribution in Figure 1. In Chinese publications, Hmongic languages are subdivided into Miáoyǔ 'Miao language' and Bùnǚyǔ 'Bunu language' according to whether the speakers are culturally Hmong (Miáozú) or Yao (Yáozú). This distinction is cultural rather than linguistic.

The Na-e language of Vietnam, also called Pateng, seems in most respects to be simply a southern outlier of the Pa Hng subgroup of Hmongic, as was first pointed out by André Haudricourt. Recently, however, Paul Benedict has argued that Na-e actually constitutes a fourth branch of Hmong-Mien. This suggestion needs further study.

A more detailed breakdown of Hmong-Mien languages is as follows:

I. Hmongic (Miáo yǔzhī)

- A. West Hunan group or QoXiong language (Miáoyǔ Xiāngxī fāngyán; Northern Hmongic)
- B. East Guizhou group or Mhu language (Miáoyǔ Qiándōng fāngyán; Eastern Hmongic)
- C. Sichuan-Guizhou-Yunnan group (Miáoyǔ Chuānqiándiān fāngyán plus Bùnǚyǔ Bù-Nǎo fāngyán; Western Hmongic; Purnell's Western and Central)
 1. Sichuan-Guizhou-Yunnan subgroup (Miáoyǔ Chuānqiándiān cìfāngyán; Purnell's West A); Green Mong, White Hmong, etc.
 2. Northeastern Yunnan subgroup (Miáoyǔ Diāndōngběi cìfāngyán) (A-Hmau)
 3. Guiyang subgroup (Miáoyǔ Guìyáng cìfāngyán) (Hmong)
 4. Huishui subgroup (Miáoyǔ Huìshuǐ cìfāngyán) (Mhong)
 5. Mashan subgroup (Miáoyǔ Mǎshān cìfāngyán) (Mang)
 6. Luobo River subgroup (Miáoyǔ Luóbó Hé cìfāngyán) (A-Hmyo)
 7. Eastern or Zhong'an River subgroup (Miáoyǔ Dōngbù cìfāngyán or Miáoyǔ Zhōng'ān Jiāng cìfāngyán) (Mhong)
 8. Pingtang subgroup (Miáoyǔ)
 9. Qianxi-Pingba-Qingzhen-Liuzhi subgroup (Miáoyǔ)
 10. Luodian Moyin subgroup (Miáoyǔ)

I. Hmongic, continued.

C. Sichuan-Guizhou-Yunnan group, continued.

11. Dushan subgroup (Miáoyǔ)
 12. Luodian Pingyan subgroup (Miáoyǔ)
 13. Ziyun-Zhenning subgroup (Miáoyǔ)
 14. Wangmo subgroup (Miáoyǔ)
 15. Wangmo-Luodian subgroup (Miáoyǔ) (Mhang)
 16. Pu-Nao subgroup (Bùnyǔ Bù-Nǎo fāngyán)
 - a. Pu Nu (Tung Nu) (Bùnyǔ Dōngnǔ tǔyǔ)
 - b. Nu Nu (Bùnyǔ Nǔnǔ tǔyǔ)
 - c. Pu No (Bùnyǔ Bùnuò tǔyǔ)
 - d. Nao Klao (Bùnyǔ Nàogéláo tǔyǔ)
 - e. Nu Mhou (Bùnyǔ Nǔmào tǔyǔ)
- D. Pa Hng (Bùnyǔ Bāhēng fāngyán)
- E. Hm Nai (Bùnyǔ Wúnài fāngyán)
- F. Kiong Nai (Bùnyǔ Jiōngnài fāngyán)
- G. Yu Nuo (Bùnyǔ Yōunuò fāngyán)

II. Ho Nte (Shēyǔ)

- A. Western or Lianhua dialect (Liánhuā fāngyán)
- B. Eastern or Luofu dialect (Luófú fāngyán)

III. Mienic (Yáo yǔzhǐ Miǎnyǔ)

- A. Mien-Kim (Miǎn-Jīn fāngyán)
 1. Mien (Iu Mien) (Yōumiǎn tǔyǔ)
 2. Mun (Kim Mun) (Jīnmén tǔyǔ)
 3. Biao Mon (Biāomàn tǔyǔ)
- B. Biao-Chao (Biāo-Jiāo fāngyán)
 1. Biao Min (Biāomīn tǔyǔ)
 2. Chao Kong Meng (Jiāogōngmiǎn tǔyǔ)
- C. Dzao Min (Zǎomīn fāngyán)

IV. Classification uncertain: Na-e (Pateng)

Note that the articles on Hmong in this collection all deal with White Hmong of Thailand and Laos, which belongs to the Sichuan-Guizhou-Yunnan subgroup of the Sichuan-Guizhou-Yunnan group of Hmongic.

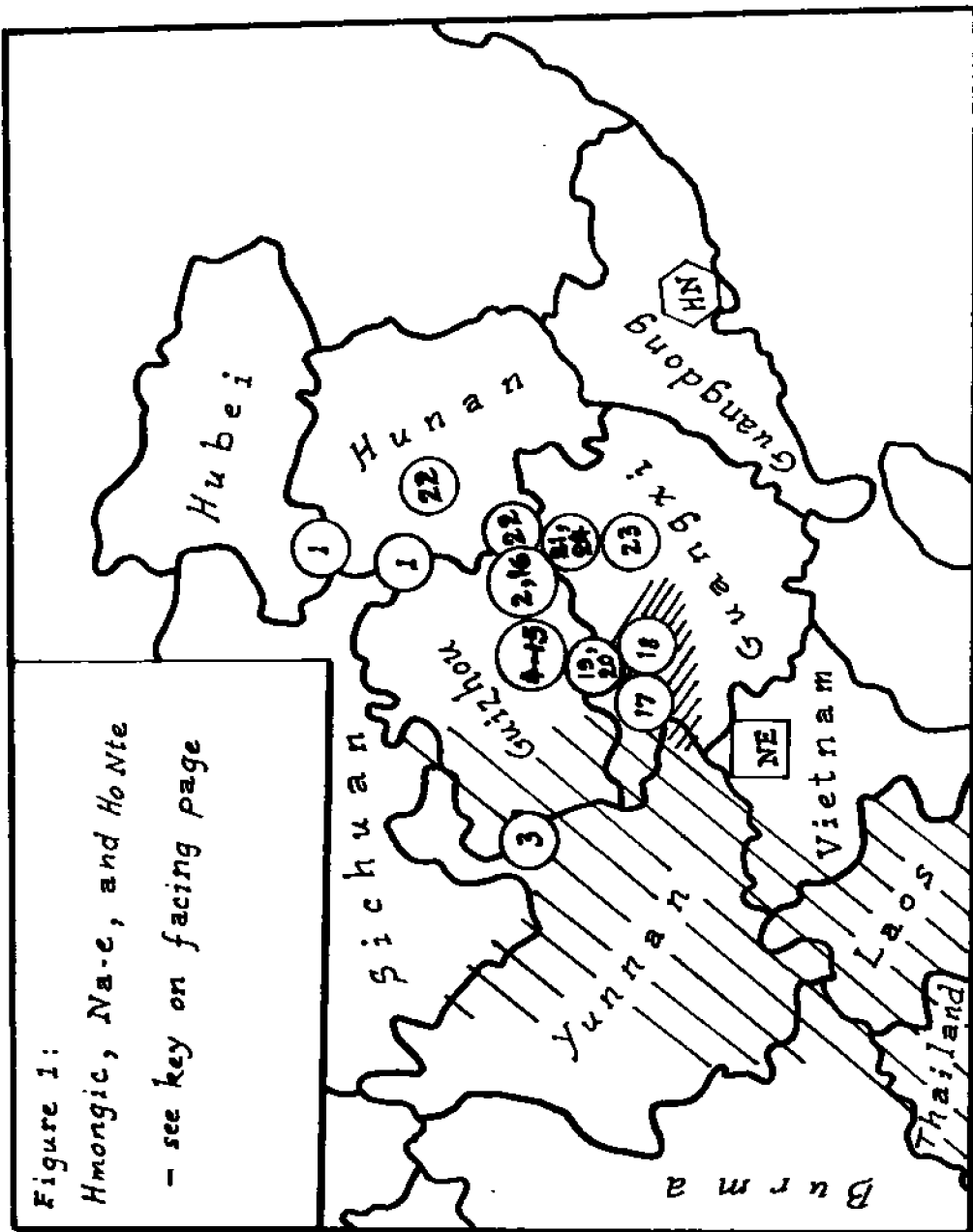
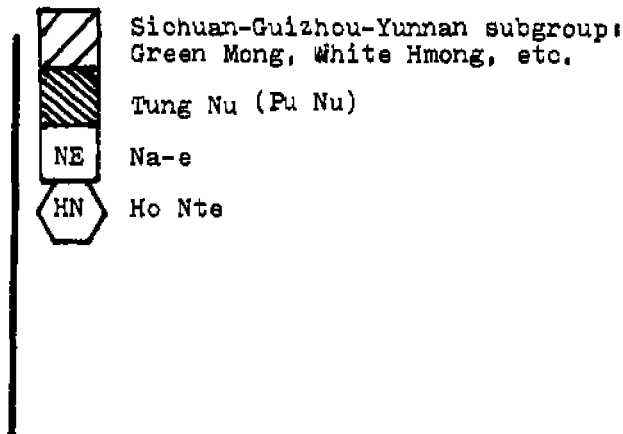
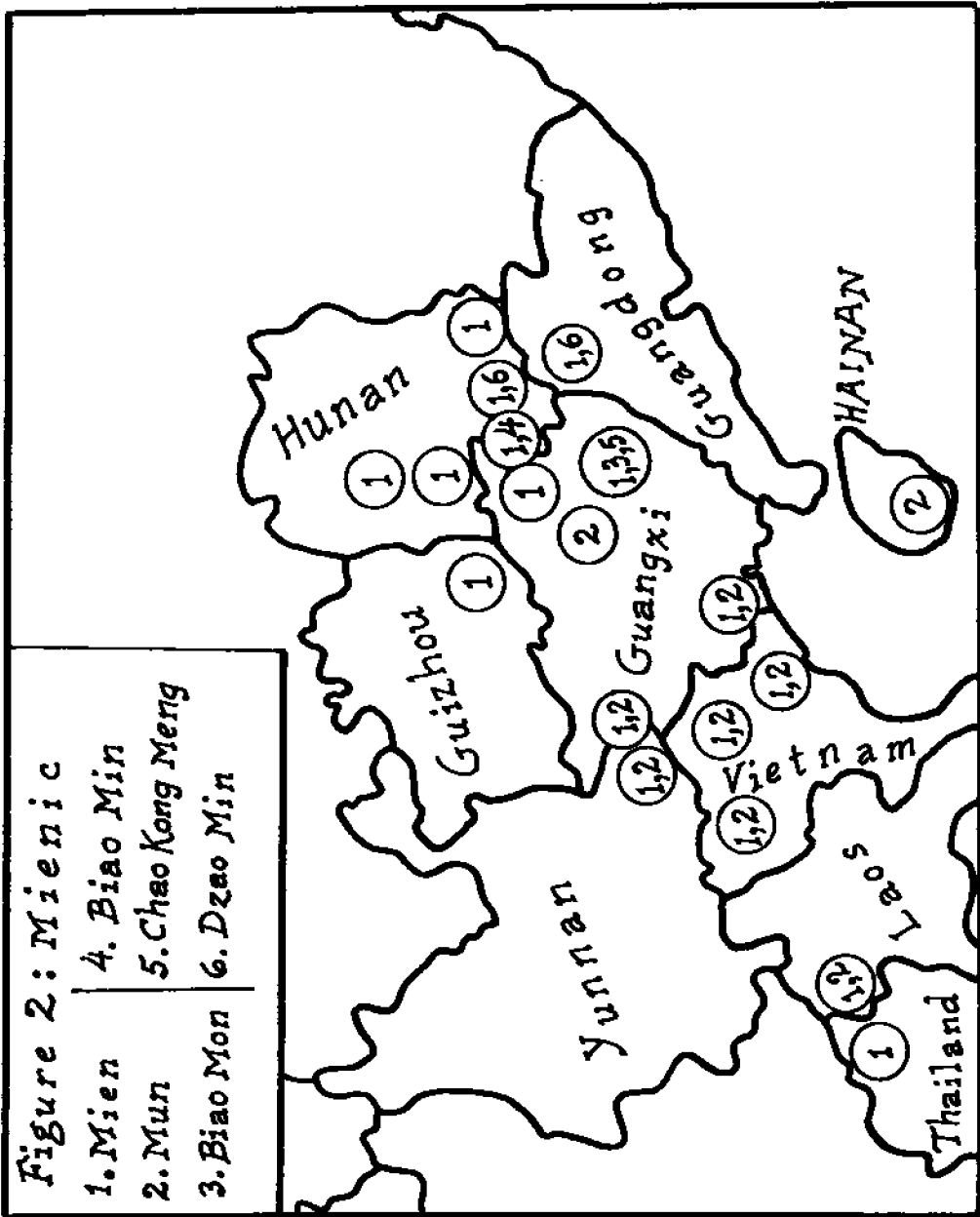


Figure 1: Hmongic, Na-e, and Ho Nte.

- 1 QoXiong or West Hunan group (also speech islands in or near the Tung Nu speaking area)
- 2 Mhu or East Guizhou group (also speech islands in south central and SW Guizhou and in NW Guangxi)
- 3 A-Hmau or Northeastern Yunnan subgroup of Sichuan-Guizhou-Yunnan group
- 4-15 Diverse Hmongic languages of south central Guizhou: Guiyang Hmong, Huishui Mhong, Mang (Mashan subgroup), A-Hmyo (Luobo River subgroup), Pingtang subgroup, Qianxi-Pingba-Qingzhen-Liuzhi subgroup, Luodian Moyin subgroup, Dushan subgroup, Luodian Pingyan subgroup, Ziyun-Zhenning subgroup, Wangmo subgroup, Mhang (Wangmo-Luodian subgroup)
- 16 Zhong'an River Mhong (speech islands within Mhu area)
- 17 Nu Nu
- 18 Pu No
- 19 Nao Klao
- 20 Nu Mhou
- 21 Pa Hng
- 22 Hm Nai
- 23 Kiong Nai
- 24 Yu Nuo





White Hmong Orthography and IPA Equivalents
(Broad Transcription)

1. Final consonant letters denote tones, not consonants.
2. Doubling a vowel denotes final [ŋ].

Thus the name of the language is spelled Hmoob [mɔŋ⁵⁵].

Initials

p	ph	np	nph	v	f	m	hm
[p]	[ph]	[mb]	[mph]	[v]	[f]	[m]	[ɣ]
pl	plh	npl	nplh			ml	hml
[pl]	[phl]	[mbl]	[mphl]			[ml]	[ɣl]
t	th	nt	nth	l	hl	n	hn
[t]	[th]	[nd]	[nth]	[l]	[l̥]	[n]	[ɳ]
tx	txh	ntx	ntxh		x		
[ts]	[tsh]	[ndz]	[ntsh]		[s]		
d	dh						
[d]	[d̥]						
r	rh	nr	nrh				
[ɹ]	[ɹh]	[nd̥]	[nɹh]				
ts	tsh	nts	ntsh	z	s		
[tʂ]	[tʂh]	[nd̥z]	[nʂh]	[z]	[s]		
c	ch	nc	nch	y	xy	ny	hny
[tʃ]	[tʃh]	[nd̥]	[nʃh]	[j]	[ç]	[ɲ]	[hɲ]
k	kh	nk	nkh			ɛ	
[k]	[kh]	[ŋg]	[ŋkh]			[ɛ]	
q	qh	nq	nqh				
[q]	[qh]	[ŋg]	[Nqh]				
∅					h		
[ʔ]					[h]		

For d and dh see Jarkey's paper in this collection.

Finals

i[i]	ia[ia]	ai[ai]	
e[e]			ee[Λη]
a[a]			
o[ɔ]			oo[oη]
u[u]	ua[ua]	au[au]	
w[ɥ]		aw[aw]	

Tones

- b [55] high level
- j [52] high falling
- v [24] rising
- s [22] mid-low level
- ∅ [33] mid level
- g [42] falling, whispery voiced
- m [21] low falling, creaky voiced
- d [13], [213] low rising, or falling-rising (a syntactically determined variant of -m)

Bibliography

Huffman's bibliography provides a very complete listing of sources on Hmong-Mien languages up to 1983 or early 1984. Here we give a small update of Huffman, listing works which do not appear in his bibliography or which he lists as conference presentations and which have since been published.

Benedict, Paul K. 1986a.

Miao-Yao Enigma: the Na-e Language. *Linguistics of the Tibeto-Burman Area* 9.1:89-90.

_____. 1986b.

Early Chinese Dialect 'Processing'. 19th International Conference on Sino-Tibetan Languages and Linguistics.

Chen Qiguang. 1984.

Gu Miao-Yao yu bi guan bise yin shengmu zai xiandai fangyan zhong fanying xingshi-de leixing (The Reflected Patterns in Modern Dialects of the Prenasalized Stop Initials of Proto-Miao-Yao). *Minzu Yuwen* 1984.5:11-22.

_____. 1985.

Sheyu he Kejiayu (She and Hakka). In Zhou Zumo, Xu Guozhang, Yu Min, Xu Shicai, Guo Liangfu, Xu Zhonghua, and Li Bisong, editors, *Yuyan Lunwenji* (Collected Essays on Language), Linguistic Society of Peking, pp. 129-138.

Fuller, Judith Wheaton. 1985.

Topic and Comment in Hmong. PhD dissertation, University of Minnesota.

_____. 1986a.

Zero Anaphora and Topic Prominence in Hmong. In Hendricks, Downing, and Deinard 1986:261-277.

_____. 1986b.

Chinese Le and Hmong Lawn. 19th International Conference on Sino-Tibetan Languages and Linguistics.

Heimbach, Ernest E. 1979.

White Hmong - English Dictionary. Revised edition. Data Paper 75, Southeast Asia Program, Department of Asian Studies, Cornell University.

Hendricks, Glenn L., Bruce T. Downing and Amos S. Deinard. 1986.

The Hmong in Transition. Center for Migration Studies of New York, Inc. The Southeast Asian Refugee Studies Project of the University of Minnesota.

- Huffman, Franklin E. 1986.
Bibliography and Index of Mainland Southeast Asian Languages and Linguistics. Yale University Press, New Haven and London. See index under Hmong, Man, Miao, Mien, She, Yao, Hmongic, Miao-Yao.
- Jaisser, Annie. 1986.
The Morpheme Kom: A First Analysis and Look at Embedding in Hmong. In Hendricks, Downing, and Deinard 1986:245-260.
- Li, Charles N., Bettina Harriehausen and Don Litton. forthcoming.
Iconicity: A View from Green Hmong Serial Verbs. To appear in Eric Schiller, editor, Proceeding of the Conference on Southeast Asia as a Linguistic Area. University of Chicago, Chicago Linguistic Society.
- Lu Yichang. 1985.
Yaozu Mianyu Biaomin fangyan-de gouci biandiao yu gouxing biandiao (The Tonal Change in Word-Formation and the Tonal Change in Morphology in Biaomin Dialect of Mian Language Spoken by Yao Nationality). Minzu Yuwen 1985. 6:16-19.
- Mao Zongwu and Meng Chaoji. 1986.
Sheyu Jianzhi (A Sketch of the She Language). Nationalities Press, Peking.
- Mills, Carl and David Strecker. 1982.
Diphthongization, Syllable Structure, and the Feature [high] in Hmu. Conference on Phonological Distinctive Features, State University of New York at Stony Brook.
- Owensby, Laurel. 1986.
Verb Serialization in Hmong. In Hendricks, Downing and Deinard 1986:237-243.
- Ratliff, Martha. 1986a.
An Analysis of Some Tonally-Differentiated Doublets in White Hmong (Miao). Linguistics of the Tibeto-Burman Area 9.2:1-35.
- _____. 1986b.
Two-Word Expressives in White Hmong. In Hendricks, Downing and Deinard 1986:219-236.
- _____. 1986c.
The Morphological Functions of Tone in White Hmong. PhD dissertation, The University of Chicago.

- Solnit, David B. 1985.
Introduction to the Biao Min Yao Language. Cahiers de Linguistique Asie Orientale 14.2:175-191.
- _____. 1986.
Some evidence from Biao Min on the initials of proto-Mienic (Yao) and proto-Hmong-Mien (Miao-Yao). 19th International Conference on Sino-Tibetan Languages and Linguistics.
- Strecker, David. 1986.
Proto-Hmongic Finals. 19th International Conference on Sino-Tibetan Languages and Linguistics.
- _____. forthcoming.
Vocalization of Syllable-Final Stops in Tai and Hmong-Mien. To appear in Eric Schiller, editor, Proceedings of the Conference on Southeast Asia as a Linguistic Area. University of Chicago, Chicago Linguistic Society.
- _____. ms.
Some Comments on Benedict's 'Miao-Yao Enigma: the Na-e Language'.
- Wang Chunde. 1986.
Miaoyu Yufa; Qiandong Fangyan (Mhu Grammar). Guangming Ribao Press.
- Wang Deguang. 1986.
Weining Miaoyu huayu cailiao (Text of Miao Language in Weining, Guizhou Province). Minzu Yuwen 1986.3:69-80. (Text in A-Hmau with Chinese morpheme for morpheme and free translation.)
- Wang Fushi. 1985.
Miaoyu Jianzhi (A Sketch of the Miao Language). Nationalities Press, Peking.
- _____. 1986.
Miao-Yaoyu-de xi shu wenti chu tan (A Preliminary Investigation of the Genetic Affiliation of the Miao-Yao Languages). Minzu Yuwen 1986.1:1-18, 79.
- Yu Weiwen and Chao Zongqi. 1984.
Youling Yao hua gai shu (A brief description of the Yao dialect in Youling). Journal of Sun Yat-sen University, social sciences edition, 1984.3:115-126, 62. (Phonological and grammatical sketch of a Dzao Min dialect.)
- Zhao Bingxuan. 1986.
Taiyuan fangyan li-de fufuyin yiji (Residues of Consonant Clusters in the Taiyuan Dialect). 19th International Conference on Sino-Tibetan Languages and Linguistics. (Includes a list of 21 words in Biao Min.)

EARLY MY/TB LOAN RELATIONSHIPS

Paul K. Benedict

The Miao-Yao (MY) languages appear to have split off from the mainland bloc of Austro-Tai (AT) languages at an early period, moving to the north and west and there coming into contact with the Sino-Tibetan (ST) languages. To make use of an analogy from geomorphology, they came to be positioned at the 'grinding edges' of the vast AT and ST tectonic plates, along a line of maximum anticipated activity. In this case the activity involved the transformation of MY into a monosyllabic, fully tonal language family, as earlier described by the writer (1975 - hereafter ATLC), with many of the details yet to be explored. As regards the monosyllabism, this is a feature of both Tibeto-Burman (TB) and Sinitic (Chinese, Bai) and must be set up for the parental proto-language (PST), hence no clues as to the precise source of the influences are provided. The tones are another matter, however; at an early period (2nd/1st mill. B. C.) Chinese converted the two-tone (or two-accent) *A ~ *B system of PST into a three-tone (or three-accent) system with the addition of a third, sandhi tone (or accent) *C and a system of precisely this kind must be reconstructed for the parent MY language (PMY), with one-to-one tonal correspondences in the early loanwords for such cultural items as 'horse' and 'charcoal'. This is an example of 'direct' diffusion, in contrast to the 'stimulus' diffusion found in Mon-Khmer (e.g. Riang, Khamu) and even in Austronesian (Hulhui, the Chamic language spoken on Hainan); the tonal system itself was borrowed, not simply the 'idea of tones', with the loanwords serving as the bearers of tone. The process undoubtedly began in a highly selective manner, with later spread throughout the language; cf. the situation in T'in, a Mon-Khmer language now on its way to becoming fully tonal, which has borrowed Thai (Siamese) numerals *along with the tones* (see Benedict 1984: 67).

This sort of linguistic happening cannot occur at a distance, inasmuch as it requires a virtual cultural 'flooding', with extensive diglossia on the part of the 'submerged' population (as in the case of T'in). The early (Archaic level) Chinese/MY loans, as currently analyzed (Benedict 1986), point to exactly this kind of prolonged, intimate relationship, thus dovetailing with the tonal evidence. The picture is complicated, however, by the fact that the bulk of Chinese loans into MY reflect early dialect(s) that are distinct from the 'standard' Archaic, notably in the retention of PST final *-a, where the latter shows a regular shift to -o. These loans also commonly show evidence of Chinese 'processing', reflecting manifold prefix + initial shifts of the type described in an earlier paper (1976) by the writer. Thus, the picture is exceedingly complex, even under the best of circumstances, yet all would make good sense historically were it not for the following: a TB rather than Chinese source must be recognized for the basic cardinal numerals of MY as well as for the 'core' (basic) lexical items: 'sun/day' - 'moon/month', along with a number of other items rather less basic in nature. This paper presents the relevant linguistic data, attempts to delineate (if not identify) the donor source or sources (DMY or DMY's) and,

finally, offers some preliminary speculation as to how this strange linguistic state of affairs is likely to have come about.

The numerals of MY, like those of the Kadai languages and even Austronesian (Huihui), nicely exemplify selective lexical retention (SLR - see Benedict 1983b), with a primary 'rule' that cardinals are better retained than ordinals and a secondary 'rule' that the lowest numerals are better retained than the higher (it should be noted here that the Swadesh basic 100-word list includes only 'one' and 'two'). In the case of the MY numeral system, only one loanword managed to break into the native 'one' through 'three' alignment and the intrusion was relatively late: Middle Chinese (MC) — ʔ₁ēt 'one' > Yao (Mien only) *yet¹. The MY numerals above 'ten' are also of Chinese origin but those from 'four' through 'ten' appear to have been borrowed as a set from a TB source (or sources). Both Shafer (1964) and Downer (1971) took special note of the resemblances shown by these numerals while the writer (1976) has analyzed them at some length, along with the early Chinese loans to Tai. Contra the view expressed there, it now seems evident (see fn. 1) that 'four' also belongs in this set of early loans from TB, with the further strong likelihood that 'five' must be included as well, i.e. these numerals were borrowed as a set. As shown in the following listing, the corresponding Chinese numerals developed in different ways (see the above-cited paper), precluding the possibility of their having served as sources for the MY numerals (see Benedict 1972 - hereafter *STC* - for the TB numerals):

'four': PTB *b-ləy; PMY *plei^A; MC 四 st-. PMY regularly has -ei for earlier (PAT-level) final *-i, with *-i as an alternative reconstruction, hence the indicated DMY form is *pli^A, with *-i for final *-ay (as in most TB languages) along with unvoicing of the prefix.

'five': PTB *r-ŋa ~ (secondary) *b-ŋa; PMY *pra^A; MC 五 ŋuo: PTB prefixed *r- is maintained in Old Kuki but replaced in Written Tibetan (WT) by l-, apparently through influence from PTB *lak 'arm/hand'²; general replacement by prefixed *b- through influence from the preceding numeral (see above), with occasional preemption of initial *ŋ-. The indicated DMY form is *pra^A, with the secondary *b- (> p-) preceding the *r- rather than replacing it, followed by preemption of the initial and unvoicing of the prefix, as in 'four'.

'six': PTB *d-ruk ~ *k-ruk; PMY *tru[?]; MC 六 liuk. The PMY initial *tr- has been reconstructed on the basis of the Na-e evidence (see Benedict: forthcoming); the final *-ʔ for earlier *-k is a MY feature, hence the indicated DMY form is *truk, with unvoicing of the prefix as in 'four' and 'five'.

'seven': PTB *snis; PMY *zńia^C; MC 七 ts'iet. Contra *STC*, the PTB *s- stands for the first part of the *sn- cluster rather than for prefixed *s-, with Stau exhibiting a distinction between /zńi/ 'seven' (secondary voicing and palatalization with loss of final *-s) and /sńi/ 'day', from *s-nəy^A (see below). The nasal element of the MY root is maintained only in Yao: Mun (ńi) but it yielded typical secondary nasalization of the final in Miao, with some Western dialects reflecting secondary unvoicing (*zńia^C > *sńia^C). The initial of the DMY form can be reconstructed as *z[n,ń]-, with secondary voicing as in Stau; the ambiguity results from the fact that

secondary palatalization is characteristic of MY, tying in with the same feature in Chinese, probably reflecting an ancient Sprachbund. The final of the DMY form could hardly have been *-i, with loss of *-s (as in Stau), since this would have yielded PMY *-ei (cf. 'four'); PMY does have medial *-ia- for medial *-i-, however (see 'ten' and 'year', below), hence the indicated reconstruction of the DMY final is *-is, with the vocalic shift preceding the eventual loss of *-s (lacking in PMY).

'eight': PTB *(-)ryát; PMY *hyat; MC /\ pwät. The earliest ST prefixation pattern, reflected in WT (brgyad < *b-g-ryat) and Jingpho (mətsát) as well as in Chinese (apparently unvoiced by an original prefixed *-s-) can be set up as *b-g- but within TB the Kuki-Naga forms reflect prefixed *d- (< *d-rát) while those of Burmese-Lolo reflect prefixed *s- > Written Burmese (WB) hrac (< *hryat), with /hyat/ appearing in the inscriptions, exactly matching the PMY root. The indicated DMY form is *[hry,ylat, with *-ry- a possibility in view of the fact that PMY lacked this cluster and may well have substituted *-y-, paralleling the Burmese development.

'nine': PTB *d-kəw ~ *d-gəw; PMY *gy[ou]A; MC 九 kiəu. The initial *g- form, represented notably in WT (dgu), shows secondary voicing after the prefix, which was replaced by *-s- in some languages (Jingpho, Bodo-Garo). The PMY palatalized initial is generally reflected by palatals or dentals (but Na-e kó); the reconstruction of the final is provisional, with the Yao forms apparently influenced by the similar ordinal forms (< Chinese). PMY regularly has final *-ou for earlier (PAT-level) *-u, paralleling *-ei for earlier *-i (see 'four', above), hence the indicated DMY form here is *[prefix+] [g.gy]u, with ambiguity arising from the palatalization (see the discussion under 'seven', above). The initial voicing points to earlier prefixed *d-, as in WT, with later shift to *t- (cf. 'six', above).

'ten': PTB *gip; PMY *g(y)lap; MC 十 z̄iəp. The PMY initial palatalization appears to have been variable, with Yao forms perhaps influenced by competing forms of Chinese origin. As in 'seven', the PMY medial *-ia- is to be considered secondary, with *gip indicated for the DMY form.

Reconstructed tones (*A or *B) have not been cited for the three relevant PTB numeral roots ('four', 'five' and 'nine') in view of the widespread variation shown by the modern forms, including tonal 'form classes', e.g. WB all < tone *B; Trung (Nungish) all < tone *A, for these three numerals as well as for 'three' (PTB *g-sum). PMY tone *A for the trio indicates that DMY had a 'form class' like that of Trung. The fact that PMY has tone *C for 'seven' supports the final *-s reconstruction since this tone has a sandhi origin, reflecting syllable-final features.

To sum up, the DMY numerals were probably as follows:

	DMY		DMY
four	*pli ^A	eight	*hryat
five	*pra ^A	nine	*t-gu ^A
six	*truk	ten	*gip
seven	*znis		

The DMY phonology is featured by the unvoicing of stop prefixes (the *t- of 'nine' only by inference), along with the *s-r- > *hr- shift, contrasting

with secondary voicing of the *sn- cluster; also the parallel final *-əy > *-i and *-əw > *-u shifts. Although only one numeral root (*gip 'ten') remained without change, none of the DMY forms appears very unusual for TB, with the exception of that for 'five', and even here there are parallels in the modern languages to the initial-preemption involved.

Three other lexical items, all at least partially of calendrical type, have long attracted the attention of comparativists because of the obvious MY/TB resemblances. As reported in the above-cited article (Benedict 1976), there is a competing AT etymology in the case of one of these items ('moon/month') but this must now be discarded, especially in view of the additional evidence available here (see below). As in the case of the numerals, the Chinese cognates exhibit different lines of development, effectively precluding them as possible loan sources:

'sun/day': PTB *nəyA 'sun' ~ *s-nəyA 'day'; PMY *hnɔiA (often prefixed) 'sun/day'; MC 日 *n̄z̄īt̄* 'id.'. The 'sun' vs. 'day' distinction can be established at the PTB level; cf. Stau *sni* 'day' (above); Jingpho *sənf* 'id.'; WB *ne* (< *nəyA) 'sun' - *ne'* (< *s-nəyA) 'day', but has been lost in some TB languages, e.g. Lushai has /ni/ for both 'sun' and 'day'. Chinese also has lost the distinction: *n̄z̄īt̄* < *s-n̄z̄īt̄ (Min evidence for the prefixation), with secondary final -t after the prefix (paralleling the WB glottalization; see Benedict 1983a). In MY the 'sun' vs. 'day' distinction is often made through prefixation but basically the PMY root shows a loss here, of the Chinese rather than Lushai type, i.e. it is the *s- prefixed rather than the unprefixed form that has prevailed. On the basis of this MY loan evidence one can posit a similar loss of distinction in the donor language, although it is possible that MY innovated here. The indicated *hn- for DMY, from an earlier (PTB) *s-n-, contrasts with *zn- from a *sn- cluster in 'seven', precisely as in Stau (see above). The final presents a problem, however, inasmuch as in 'four' PTB final *-əy is represented by PMY *-ei, regularly from an earlier *-i, whereas in this root it is represented by PMY *-ɔi, very likely standing for *-əy itself. At least three possible solutions present themselves: (a) the final reflex was conditioned by the initial (*pl- vs. *hn-) (b) PMY lacked the specific *hnei combination (Mien, at least, appears to lack it) and substituted *hnɔi (3) a somewhat different DMY, perhaps a dialect of the 'standard' DMY, which retained final *-əy, was the source in the case of this item. Of these three possible solutions (a) seems the least likely and (c) the most likely.

'moon/month': PTB *(s-)(g-)laA; PMY *hlaC; MC 月 *ʃək* [GSR 769a] 'new moon', from *s-glāk (see Benedict 1976:189 for the phonetic element here), with prefixed *s- yielding secondary -k after the low vowel, in contrast to the final -t after high front vowel that appears in the preceding root (see Benedict 1983a). The Garo cognate: *ja* (< *sglaA) is also glossed as 'spirit/apparition/phantom'; WT has the doublet *zla* (< *sglaA) 'moon' - *hla* (< *sla) 'the gods'; Jingpho has only *ʃətā* (< *s-glaA). Chang-Tangsa (STC: Konyak group) only *glaA and WB only *la'* (< *s-laA), all in the meaning 'moon' (~ 'month'), but Lushai parallels WT in showing a doublet, in this case the product of regular tonal changes (see Benedict 1983a): *thlā* (< *s-glaA) 'moon' - *thlā'* (< *sglaA) 'spirit/soul/one's double'. This all points to an

underlying cult of the moon, with WT hla 'the gods' playing a key semantic role. An identical form but with tone added (*hla^A) must be set up for the donor language (DMY), in this case with the basic meaning of 'moon/month' retained, as in WB la' (< *s-la^A). The indicated *s-1 > *hl- shift here parallels the similar shift before *r- in 'eight' (above). The tone *A reconstruction is conjectural; the PMY (sandhi) tone *C points to an earlier suffix (cf. WT zla-ba 'moon').

This is all straightforward enough, with 'moon/month' joining 'sun/day' to form an interesting pair of 'heavenly body' terms that were borrowed by MY from an early TB source, either the same DMY that donated the numerals or a language very similar to it. There is no evidence in the 'moon' loan of the cult associations of the TB root but, curiously enough, this may simply reflect the fact that the replaced native term had already undergone a parallel shift, thus making 'semantic space' available for the loan; cf. PMY (Miao only) *b|a^{A/C} 'spirit/soul', apparently the direct cognate (PMY lacks final *-1) of PAT *(m)bu|al 'moon/month', represented in Kadai as well as in Austronesian. And the case gets 'curiouser and curiouser': a similar root can be reconstructed for PTB and it shows the same range of meanings as *(s-)(g-)la^A, cited above: PTB *b-la (tone variable): Proto-Tamang (Nepal) *bla^B 'spirit, soul'; Rawang (Nunglish) phəla (< *bəla) 'soul; demon'; Southern Loloish *bəla^C 'moon' (cited under No. 318 in Bradley 1979). The loan here must have been in the reverse (MY > TB) direction, as confirmed by the phonology: if the donor language had been Tibeto-Burman the yield in PMY would have been medial *-1-, as in 'four' (above), while a prototypical *-l- could have yielded only *-l- in TB, which lacks the contrast; further, prefixed *b- is known to have been unvoiced to *p- in DMY (see 'four' and 'five' above), excluding this language as a possible loan source for PMY *b|a^{A/C}.

year: PTB *(s-)ni-ŋ^B; PMY *śńian^C MC 年 nien, from *s-nien^A
 < *s-ni-ŋ^A [GSR 364a] (Min evidence for the prefix), with regular *B > *A tonal shift after *s- and assimilative -n for PST-level final *-ŋ. The PMY medial *-ia- for *-i- is anticipated (cf. 'seven' and 'ten' above) but the initial *śń- (the sibilant is retained only in Western Miao) can hardly be reconciled with the DMY *hn- for *s-n- indicated by the loan for 'sun/day' (above). It is certain, from the Min evidence, that Chinese retained prefixed *s- (variable within TB) and the palatalization of the initial is typically Chinese (although not present in /nien/). It is probable, therefore, that the source of this early MY loan is not Tibeto-Burman but Chinese, specifically an early dialect (other than 'standard' Archaic) that maintained final *-ŋ, with the (sandhi) tone *C pointing to an earlier suffix. Valuable support for this view is supplied by a parallel loan; cf. MC 新 sī^A 'new' [GSR 382k], representing a PST-level *sī^A as shown by PMY (Yao only) *sia^A. The MY loan clarifies another matter here: in STC the /sī^A/ is compared with the well-represented PTB *(s-)sar^A 'new', setting up PST *sar^A with *-ar > -i^A a regular development; the early MY loan now shows that the PST root here is *sī^A, represented in TB by WB sac (< *sik < *sī^A), precisely paralleling the nearly homophonous root for 'tree/wood': *sī^B; PTB *sī^B > WB sac; MC

薪 sɿɛ̃n 'firewood' [GSR 382 n], with *B > *A tonal shift after *s- (as in 'year' above).

The numerals, along with the 'sun' ~ 'moon' pair, stand out in any over-view of early MY loan relationships simply because they must be assigned a Tibeto-Burman, rather than a Chinese, origin, thus differing from the bulk of MY lexical items. The occasional MY/TB lexical resemblances that have been noted to date rarely involve 'core' (basic) vocabulary or even widespread roots, for that matter, hence one can scarcely rule out the chance factor. One of the 'better' comparisons in this group involving both basic vocabulary and well-represented roots, is the following:

'horn': PTB *rɯŋ^A ~ *rwaŋ^A, yielding Garo grɔŋ (g- is a pronominal prefix); PMY krɔŋ^A. If these forms indeed reflect an early loan, the DMY form would be *krɔŋ^A, from *k-rwaŋ^A, with unvoicing of an earlier *g- prefix, as in the *b- and *d- prefixes encountered in the numerals (above).

As indicated above, the cultural items such as 'horse' and 'charcoal' in MY routinely exhibit early loan connections with Chinese, with the direction of the borrowing in doubt at times. In one item of this kind MY has a curious doublet involving inter alia medial -i- ~ -ia- (see above), with TB (limited) as well as Chinese connections:

'field': PTB *liŋ (limited distribution, with indeterminate tone): WT zɿŋ < l(y)liŋ (regular shift) 'field, ground, soil, arable land'; Lepcha lyǎŋ < *s-liŋ (regular shifts) 'land, field'; MC (probably an early loan from a Tibetan group source); MC 田 d'ien 'field', from *s-lien^A < *s-li-ŋ^A (with final -n < *-ŋ shift as in 'year', 'new' and 'firewood' above); PMY *liŋ^A 'paddy field (lowland irrigated rice field)' ~ *ʔliŋ^C 'mountain field' (in Chiangrai Yao [Mien] specified as 'not yet under cultivation'). The WT/Lepcha form is isolated in TB (cf. STC : fn. 246) and appears to be intrusive (borrowed). This is hardly surprising in view of the mountainous region inhabited by the speakers of these two neighboring languages and, in fact, the 'core' agricultural AT root for '(inundated) lowlands/(wet)field/mud': *buna (> *bənə), represented even in Japanese (hena 'earth/mud/clay'), somehow found its way into Tibetan for the alpine equivalent: WT na 'meadow' (cf. the Kadal root: *naa^A 'paddy field'). The Chinese form is *s- prefixed³ and this also is hardly surprising, given the ubiquity of that element in Chinese. Perhaps PMY *liŋ^A stands for the root itself, whatever its ultimate origin (no AT etymology has yet been uncovered), with the original meaning ('lowland field') maintained, serving as a basis for the early loans into TB as well as into Chinese (note the tonal agreement); PMY *ʔliŋ^C, on the other hand, might then represent a back-loan from an *ʔa- prefixed TB form: *ʔa-liŋ, with the meaning accommodated to the terrain (cf. WT na 'meadow', cited above) and the tone modified by the glottal prefix, also with *-l for *-l- through influence from the 'native' doublet. This would point to an underlying AT > ST direction for the loan, which is in keeping with the overall evidence re south > north cultural movements at an early (PAT-level) period in SEA (see ATLC). This all sounds, and is, complicated but it is the kind of historical scenario that has come to be expected in this 'language-crowded' corner of the Asian mainland.

To complete this review of lexical categories, kinship terminologies tend to include more than their fair share of loan words, especially for older-than-Ego and affinal kin. The MY languages present a bewildering variety of kinship terms, far more than one would have anticipated for this fairly compact language family, and it seems evident that extensive borrowing of various kinds has taken place. Chinese loans are much in evidence of course, and even apparent loans from Tai, complicated in this case by the fact that both MY and Tai are of AT descent. As for possible loans from TB, the consanguineal terminology reveals little that can definitely be 'tagged' as borrowed from this source. The Mien branch of Yao on occasion exhibits what appear to be isolated loans from TB, in various lexical categories, and one of these appears in the consanguineal terminology: Mien *naw^B 'younger brother (female sp.)' > 'uncle (mother's y. br.)'; cf. PTB *naw^A 'younger sibling' > (Byangsi, Mru) 'y. br.' ~ (Chinbok) 'y. br. (male sp.)'. The comparison is weakened by the fact that younger, rather than older-than-Ego kin, are involved as well as by the tonal discrepancy.

The MY affinal terminology is quite another matter, however, inasmuch as it seems very likely that both of the key younger-than-Ego terms in this terminology are of TB origin:

'daughter-in-law': PTB *(s-)nam^A 'daughter-in-law'; also (WB: archaic) '(comp.) wife'; (Gyarung) 'sister'; also (Nunglish: Rawang) *ʔa-nam^A 'cousin [female]; (comp.) sister'; also (Magari)*ʔa-r-nam^A 'maiden'; PMY *ʔnaam^A 'daughter-in-law'; also (Yao:Mien) 'wife of older brother'; in Mien (Chiangrai Yao) used only in address or in 'personalized reference' ('She is my /naam/'), in contrast to the general referential term (< *bwaŋ^B). The indicated DMY form is *ʔ-nam^A, with the widespread (< 3rd person) pronominal *ʔa- prefix, as in Nunglish; the secondary palatalization is probably of MY origin, as in the numerals (see above), and the length also appears to be secondary, apparently through influence exerted by the prefixed *ʔa-, although for this PTB root medial length (*-naam^A) cannot be ruled out; note the tonal agreement.

'son-in-law': PTB *krwəy^B: Jingpho khri (< *khrwi) 'son-in-law; paternal aunt's children; sister's children'; WB khrwê:-ma' 'daughter-in-law', from *khrwê: 'son-in-law' + -ma' 'fem. suffix'; PMY *ʔwei^B 'son-in-law'; also (Yao: Mien) 'husband of older sister': in Mien (Chiangrai Yao) used only in address or 'personalized reference' ('He is my /wei/'), in contrast to the general referential term (< *laaŋ^A). Here the indicated DMY form is *ʔa-wi^B, with the final reflexes as in 'four' (above) and prefixed by *ʔa-, as in the preceding affinal term. As for the initial, it appears that the *k- of the root was metanalyzed as common *g- (> k- in DMY) pronominal element (see 'horn', above), often employed with kinship terms in Jingpho, with simplification of the medial *-rw- to -w- rather than to the -r- found in Jingpho; note the tonal agreement.

It is possible, of course, that still other key TB > MY loans (and/or the reverse) will in time be uncovered but for the moment we are left with three groups of such loans, all crying out for an explanation:

1. the numerals from 'four' through 'ten'. Loans of this kind, sparing only the very lowest numerals, are generally explained through a need to communicate in trading. Apparently this must suffice for an explanation here as well,

but one might wonder why the terms for the things exchanged, i.e. the cultural items, did not at times get borrowed along with the numerals. Why would a language borrow, say, the word for 'silver' from Chinese but its higher numerals (through '10') from TB, as the early PM-speakers clearly did? By contrast, the early Tai- (and other Kadai-) speakers borrowed both from a single source (Chinese), as one would have expected. Clearly there was 'something special' about the early MY relationships vis-à-vis TB and Chinese, but what? And can the answer here in any way be connected with the answers for the following two points?

2. the 'sun/day' ~ 'moon/month' pair. The writer originally emphasized the calendrical ('day' ~ 'month') aspects of these two roots and attempted to explain them along the lines of trading, e.g. agreements to return for a given exchange of goods in so many days or months. This was all radically altered, however, by the uncovering of the second 'moon' loan, this in the reverse (MY > TB) direction, with its powerful confirmation of an underlying cult of the moon. It is now apparent that it was this cult trait, with overlapping distribution at the TB/MY border, that shaped the two loans, with PMY supplementing with a loan from TB for 'moon' after having undergone specialization of its own (< AT) root in the cult sense of 'spirit/soul'. The other ('sun') member of this pair of 'heavenly body' roots can be explained along similar lines (cult of the sun) although in this case the linguistic evidence (and by inference the culture trait itself) lies at a rather deeper level. Matisoff (1983) has pointed out that the Mikir cognate for PTB *nəy^A 'sun' (see above) is ar-ni, glossed not only as 'sun/day' but also as 'god, deity' and suggests a relationship to WB ne 'be, dwell'. Although not cited in STC, a homophonous *nəy^A 'be/dwell' (> 'sit') can be reconstructed at the PTB level, to include the above-cited WB ne (- 'sit' in Loloish) as well as the following: Lepcha nyi(-m) 'to be; exist'; West Himalayish *ni 'dwell'; Bahing ni(-so), Sunwari ni 'sit'; Lushai ni (< *s-ni^A) 'to be; become'. Thus, with the help of the Mikir form, one can set up a single basic etymon, widespread in both senses, for PTB: *nəy^A ~ *s-nəy^A (~ *r-nəy^A) 'to be/dwell' > '(heavenly being =) '(sun) god' > 'sun' ~ 'day'.

3. the 'daughter-in-law' ~ 'son-in-law' pair. As pointed out above, affinal terms are rather more susceptible to replacement through borrowing than are consanguineal terms but the writer, who has made extensive/intensive analyses of Tibeto-Burman as well as Chinese, Tai and Vietnamese kinship nomenclatures, knows of nothing even remotely like the apparent early MY borrowing (from TB) of *just this key pair of younger-than-Ego in-law terms*. Where one or (more rarely) both of these terms have been borrowed elsewhere, as commonly in northern TB languages from Indic and also in northern Tai languages from Chinese, they invariably represent only a single aspect of a larger pattern of loans, at times on a very grand scale indeed. This clearly did not occur in the case of these early MY loans from TB and, again, the fact in itself cries out for an answer. The existence of the two sets of in-law terms in Yao (see above) suggests that one of the sets may well have been intrusive (borrowed), in this case the set employed in address and for 'personalized reference', a most unusual distinction. One is tempted to speculate about the existence of an early TB/MY 'bride exchange' marital program or the like, with some transference of the terms involved as well, but in this event influences upon other sectors of the nomenclatures would be likely.

As for the precise source of the early loans from TB it is possible or even likely that a single donor language (DMY) was involved, perhaps with some dialectical variation (see the discussion of 'sun/day', above). Despite its early date (see below), DMY displays prefixial unvoicing and other secondary changes, presenting little of significance to the TB comparativist. In addition, it shows 'a little of this and a little of that', making it impossible to place within any existing TB group or division, hence one must regard it as another 'extinct' language, on the order of Pyu (the precursor of Burmese). As for location, the very fact of the loans into MY places DMY well to the east, certainly within the borders of China, but this is of no great help. Finally, it does present one significant bit of data, viz. *pra^A 'five', providing much needed evidence for PTB prefixed *r- in this numeral root.

What is one to make of these findings, some of a most unusual nature, re early MY/TB borrowings? The cultural flow, as evidenced by these loanwords, was almost entirely *towards* MY, even more clearly so than in the case of the early MY/Chinese borrowings. These loans also reflect an entirely different 'borrowing milieu', so to speak, than the MY/Chinese loans: much less 'diffuse', with the loans sparse and rigidly confined to specific categories. The early MY-speakers made good use of the higher numerals of the TB-speakers on their west and even shared in their heavenly body (sun, moon) cults, perhaps also entered into certain marital alliances with them, but they kept their distance; with their Chinese neighbors, on the other hand, they shared a community existence of sorts as a 'substratumized' population, the two groups exchanging cultural items of various kinds. To put it somewhat differently, they had the DMY-speakers as neighbors; they lived with the Chinese.

Finally, how is one to date these early happenings, at least on a relative basis? The one solid bit of dating here is that for the MY community's co-existence with the ancestral Chinese, dated as early as the Archaic (Dobson's 'Early Archaic') period, from before 1,000 B.C. to ca. 750 B.C., as attested by the extensive loans in MY that 'tie in' phonologically (see Benedict 1986) with AC (and AC dialect) forms. It would seem, surely, that the TB loanwords had already become part of the parent MY language by that time in view of the continuous nature of the subsequent MY/Chinese relationship. One must suppose, therefore, that the early MY-speakers first made contact with TB-speaking populations as they moved west, probably during the 2nd millennium B.C., borrowing the higher numerals and a few other, very 'special' items from one of their languages (DMY); further, that their next move was towards the north, bringing them face to face with the ancestral Chinese, with whom they settled down into an arrangement of community living. This historical scenario is conjectural, to be sure, but it does serve to explain the linguistic findings. It also agrees well with the view held by a number of scholars (e. g. Pulleyblank 1983), that the ancient state of Chu 楚, that arose during the 1st millennium B.C. in the middle Yangtze region, was of MY origin. Actually, by placing the ancestral MY in northwest China as early as the Archaic period, the scenario readily lends itself to the conclusion that the early population of much of China, particularly in the west, was MY-speaking, with the continuing post-Archaic Chinese expansion to the south and east overlaying this autochthonous stratum.

Footnotes

1. See Benedict 1975 (cited as ATLC) for references to MY and other AT sources, vastly strengthened now for Miao dialects by Wang 1979. ATLC (pp. 211 ff.) derives the MY numerals for 'one', 'three' and 'four' from the corresponding AT roots and suggests AT etymologies for both 'two' and 'five'. It now appears (Benedict: 1986b) that the various MY forms for 'two' reflect, rather, the corresponding AT numeral root. As for 'one' and 'three', both present problems in reconstruction of the PMY rhymes but a good case can be made out for 'one', at least, as connected to the PAT root, giving the proto-language the 'one' - 'two' pair. Finally, the ATLC analysis for 'four' involved handling PMY medial *-l- as an infix, with an apparent parallel in Atayal (Formosan), but Li (1983) has recently shown that in the latter case simply a feature of men's speech is involved, virtually invalidating the ATLC analysis as a whole. With 'four' thus restored to the TB loanword roster, so to speak, and with 'six' and higher numerals already there, the likelihood of a straight numeral run: 'four', 'five', 'six' . . . becomes very great, indeed, with a TB source for 'five' to be preferred over the ATLC etymology (< 'palm' via 'hand').
2. Shafer (1964) cites Abor pilŋa, which can be used in support of a PTB reconstruction of *b-l-ŋa rather than *b-r-ŋa type, with the doublet as an alternative. In any event, the Abor form does provide a likely parallel for the DMY double-prefixed form.
3. Contra Benedict 1976, prefixed Archaic *s-l- yielded MC d' (-*s-l̥- > d̥-), as in this form for 'field', whereas cluster sl̥- yielded s̥-, with initial *l- maintained as a lateral.

References

- Benedict, P. K. 1975. *Sino-Tibetan: a Conspectus*. Contributing editor: James A. Matisoff. Cambridge University Press.
- . 1975. *Austro-Thai: Language and Culture, with a Glossary of Roots*. New Haven: HRAF Press.
- . 1983a. "TB/Karen cluster vs. prefix *s-." In *Papers from the Fourteenth International Conference on Sino-Tibetan Languages and Linguistics*. Edited by Chauncey Chu, W. South Coblin and Feng-fu Tsao. Taipei: Student Book Company, Ltd.
- . 1983b. "Selective lexical retention (SLR) in Southeast Asia." *Sino-Tibetan Conference XVI*, Seattle.
- . 1984. "Comment on W. L. Ballard: 'Mother soup - a south Chinese recipe for tonometamorphogenesis.'" *CAAAL* 22:65-70.
- . 1986a. "Early Chinese dialect 'processing'." *Sino-Tibetan Conference XIX*, Columbus.
- . 1986b. "Miao-Yao enigma: the Na-e language." *LTBA* 9,1:89-90.
- Bradley, D. 1979. *Proto-Loloish*. Scandinavian Institute of Asian Studies Monograph Series, No. 39. London and Malmö: Curzon Press, Ltd.
- Downer, G. B. 1971. "The further relationships of the Miao-Yao languages." *Sino-Tibetan Conference IV*, Bloomington.
- Li, P. J.-k. 1983. "Types of lexical derivation of men's speech in Mayrinax." *Academia Sinica: BIHP* 54.3:1-18.
- Matisoff, J. A. 1983. "God and the Sino-Tibetan copula, with some good news concerning selected Tibeto-Burman rhymes." *Sino-Tibetan Conference XVI*, Seattle. Published 1985. *Journal of Asian and African Studies* (Tokyo), #29, 1-81.
- Pulleyblank, E. G. 1983. "The Chinese and their ancestors in prehistoric and early historic times." In *The Origins of Chinese Civilization*. Edited by D. N. Keightley. Berkeley: University of California Press.
- Shafer, R. 1964. "Miao-Yao". *Monumenta Serica* 23:398-411.
- Wang, F.-s. 1979. "The Comparison of Initials and Finals of Miao Dialects (in Chinese)." Beijing.

Some Comments on Benedict's
"Miao-Yao Enigma: the Na-e Language"

David Strecker

1. Preface

Up to now we have thought that the Hmong-Mien (Miao-Yao) family comprised three branches:

- Hmongic (Miao and Bunu Yao)
- Ho Nte (She)
- Mienic (Yao proper)

Now Benedict has presented evidence for a possible fourth branch: Na-e.

Benedict's remarks have inspired me to review the current classification schemes for Hmong-Mien. I would like to propose the following system of subgrouping:

- I. Hmongic
 - A. Eastern Guizhou (Purnell's "Eastern")
 - B. West Hunan (Purnell's "Northern")
 - C. Sichuan-Guizhou-Yunnan (Purnell's "Central" and "Western")
 - 1. Sichuan-Guizhou-Yunnan (Purnell's "West A")
 - 2. Northeast Yunnan
 - 3. Guiyang
 - 4. Huishui
 - 5. Mashan
 - 6. Luobo River
 - 7. Eastern
 - 8. Pu Nu (Tung Nu)
 - 9. Nu Nu
 - 10. Pu No
 - 11. Nao Klao
 - 12. Nu Mhou
 - D. Unclassified: nine major groups.
- II. Pa Hng (including Na-e)
- III. Hm Nai
- IV. Kiong Nai
- V. Yu Nuo
- VI. Ho Nte
- VII. Mienic
 - A. Mien-Kim
 - 1. Mien (Iu Mien)
 - 2. Mun (Kim Mun)

VII. Mienic, continued.

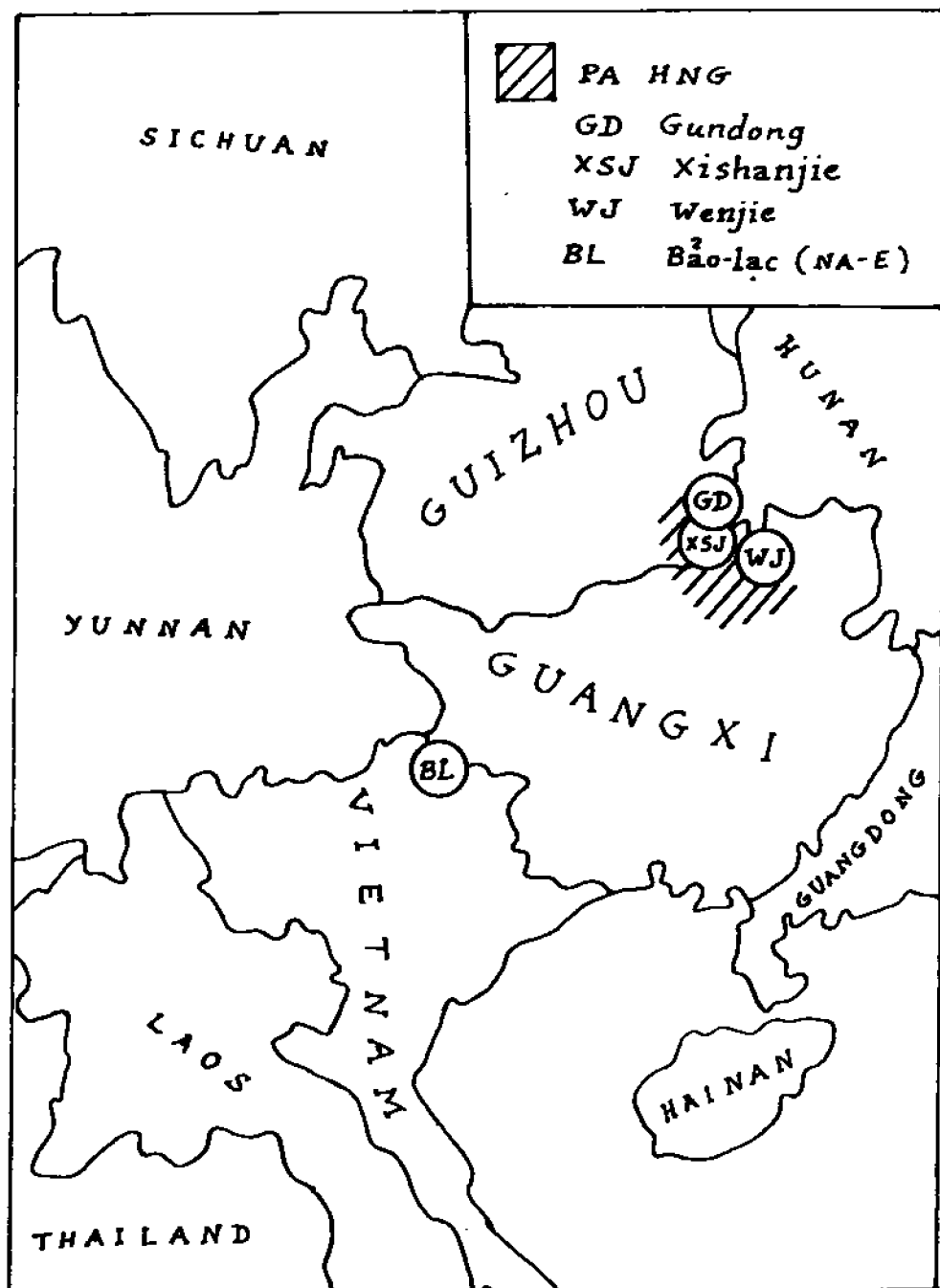
- A. Mien-Kim, continued.
 - 3. Biao Mon
- B. Biao-Chao
 - 1. Biao Min
 - 2. Chao Kong Meng
- C. Dzao Min

This schema is taken directly from recent Chinese and Soviet publications except that I have split up the Punic (Bunu Yao) subdivision of Hmongic. On the one hand, I have put Pu Nu (Tung Nu), Nu Nu, Pu No, Nao Klao, and Nu Mhou into the Sichuan-Guizhou-Yunnan branch of Hmongic, a possibility which has already been considered by such people as Kun Chang, Wang Fushi, and Martha Ratliff. On the other hand, I have provisionally elevated Pa Hng, Hm Nai, Kiong Nai, and Yu Nuo to the status of independent branches of Hmong-Mien, carrying one step further the suggestion made by Mao, Meng, and Zheng (1982:117) that these languages "have almost reached the status of separate yü", that is separate major subdivisions of a language family.

Na-e turns out to be simply a southern outlier of the Pa Hng group: see the map. This was first pointed out by Haudricourt (1954:564/1972:197; 1971:38, 43). See also Benedict (1975:xxi). In addition to Bonifacy's article, we have the following sources for Pa Hng:

1. Mao, Meng, and Zheng (1982:118, 121-123): dialect of Wenjie, Sanjiang County, Guangxi Zhuang Autonomous Region.
2. Institute for the Study of Minority Languages (1959), cited in Moskalev (1978:15). Locality not specified, but appears to be the same dialect as that described by Mao, Meng, and Zheng, or one very close to it.
3. Chang (1947, 1953, 1972, 1976): dialect of Xishanjie, near Yongcong, in the southeastern corner of Guizhou Province. In the literature this dialect is variously known as "Yung-ts'ung", "Hsi-shan-chieh", or "Tahua Yao".
4. Chen (1984:17, dialect #13): dialect of Gundong, Liping County, Guizhou Province.

Benedict has expressed fears that Na-e may have died out since the early recording by Bonifacy. But we have reason to hope that the language may still be spoken. Nguyen (1985:2) cites a recent publication by the Institute of Ethnology in Hanoi which lists a Hmong-Mien language called Pàthen. I think this must be the same name as Bonifacy's "Pà-ten", that is, Na-e. We will need to go through Vietnamese linguistic journals such as Ngôn-ngữ and Dân-tộc-học to see



whether more material on Na-e has become available.

Pa Hng splits into two groups, Northern and Southern, which differ in the development of certain initials and finals:

Northern: Xishanjie, Gundong

Southern: Wenjie, Institute wordlist, Na-e

In this paper I will outline very briefly some of the evidence for Haudricourt's assignment of Na-e to the Pa Hng group and I will outline very briefly some additional evidence for Benedict's proposal that Na-e -- or, rather, the Pa Hng group -- separated from Hmongic at an early date.

2. Transcription

All forms in this paper are written in IPA, except for Na-e, for which I retain Bonifacy's transcription. Note the following equivalents:

Bonifacy	IPA
t'	[tʰ]
ɬ	[tʃ]
ɳ	[ŋ]
y	[j]
ɣ	[ɣ]

Bonifacy transcribes four tones in Na-e:

- 1: level. Occurs in all historical tone categories and appears to represent cases where Bonifacy simply failed to hear the tone.
- 2: descending, like Vietnamese huyền (low falling). Historically A2, B1, B2, C2, and D2, probably representing several different tones which Bonifacy failed to distinguish.
- 3: acute, like Vietnamese sắc (high rising). Historically C1 and D1.
- 4: interrogative, like Vietnamese hỏi (low rising). Historically A1.

3. Evidence that Na-e belongs to the Pa Hng group

Characteristics shared by Na-e and Pa Hng include the following:

(1) *r becomes yod, e.g. Na-e yò 1 'stone'; Xishanjie [jō 33] 'dragon'.

(2) Velar and uvular initials merge into a single series, realized phonetically as uvular in Wenjie and the Institute wordlist and as velar in the other dialects. For

example, in Na-e, kɔ̃ʌ 'horn' (< *k-) is homophonous with the second syllable of pin 1 kɔ̃ʌ 1 'star' (< *q-); in Wenjie, [qɔ̃ 1] 'warm' (< *k-) is homophonous with the second syllable of [a 6 qɔ̃ 1] 'star' (< *q-); in Xishanjie, [kã 34] 'horn' (< *k-) has the same initial as [kai 34] 'sweet' (< *q-).

(3) *q1, *q1w, etc. become [kw], e.g. 'white', Na-e kuã 4, Xishanjie [kuo 34]; 'yellow', Na-e kuɔ̃ 1, Wenjie [khwɔ̃ 2].

(4) Allowing for the vagaries of Bonifacy's transcription, the development of the finals is virtually identical in Na-e and Southern Pa Hng, sometimes slightly different in Northern Pa Hng. For example, corresponding to Wang's Proto-Hmongic final -5, we get [o] or [ɔ]:

	Na-e	Wenjie	Xishanjie	Gundong
house	piɕu ²	pjo ³	pjo ²²	--
ear	-biɕ ¹	-mphjo ²	mjo ³³	--
fish	biɕ ¹	--	mjo ¹¹ , mje ¹¹ (!)	njo ⁴
stone	yɔ̃ ¹	--	--	--

Corresponding to Wang's Proto-Hmongic final -6, we get [e] or [ɛ]:

	Na-e	Wenjie	Xishanjie	Gundong
female	-tiɕ ¹	--	--	--
chicken	-kɕ ¹	--	--	--
meat	gɛ ¹	Nqhɛ ²	ŋee ³³	--
afraid	--	--	ɬe ⁵⁵	ɬhe ⁵

Corresponding to Wang's Proto-Hmongic final -23 we get [ɔ̃] or [ɔ̃ŋ] in Na-e and Southern Pa Hng versus [ã] or [aŋ] in Northern Pa Hng:

	Na-e	Wenjie	Xishanjie	Gundong
(classifier) luã ¹	luã ¹	lɔ̃ ¹	--	--
star	-kɔ̃ʌ ¹	-qɔ̃ ¹	--	--
horn	kɔ̃ʌ	--	kã ³⁴	--
son/male	tɔ̃ʌ ¹	--	tã ³⁴	--
snow	--	--	--	maŋ ⁵

One complication is that Na-e appears to have undergone two dissimilatory changes:

1. [wo] > [wa]
2. Loss of final nasal in syllables which begin with a nasal.

Thus we have

Xishanjie [kwo 34] 'white', but Na-e *[kwo 13] > kuà 4 [kwa 13] (final -18)

Na-e *[kwəŋ] > kuàn 1 [kwaŋ] 'black' (final -23)

Gundong *[ɣwəŋ 2] > [wəŋ 2] 'sky', but Na-e *[ɣwəŋ] > *[ɣwəŋ] > hà 1 [ɣwa] (final -23).

(5) 'dog': Na-e yàn, Xishanjie [1jǎ 54] (tone D1).

Compare Hm Nai [liang 6] (tone C2) (Mao and Meng 1982:78) and 18th century "Yao" of northwest Hunan liang (Lombard-Salmon 1972:316). I suspect that the "Yao" vocabulary reproduced by Lombard-Salmon belongs to the Hm Nai group rather than to the West Hunan branch of Hmongic as Benedict suggests. Benedict's Chinese etymology for these words for 'dog' is very plausible, although there is the problem of the tonal disagreement: Chinese A2 versus Hm Nai C2 versus Pa Hng D1.

(6) 'finger': Na-e tuf 1 vā 1 pu 1 (pu 1 = 'hand'), Wenjie [tɰv 2 va 2].

Hmongic has two variants for 'finger', something like *nta B and *Nqla B (or *Nqlwa B). The latter variant is attested in Green Mong ndliiv [ndli 35] (Xiong, Xiong, and Xiong 1983:480). Compare also Hm Nai [ɣei 1 ɣkwa 3] 'finger'.

The forms *Nqla B and *Nqlwa B would normally correspond to Na-e *guà 2 (or, perhaps, *hà 2 [ɣwa]), Wenjie *[ɣkwa 3]. Is [va 2] a variant of this arising in weak stress position in the phrase *[tɰv 2 ɣkwa 3 phu 4] 'finger of the hand'?

Compare 'sky', below.

(7) 'head': Na-e pi 1 (in 'turban'), Wenjie [a 6 tɰu 5 phi 4], Institute wordlist [a 6 tɰu 5 phi 4].

Compare Dzao Min [pɛ 24], Biao Min [pli 35], Mun [pɛi 44] 'head', Mien (Chiang Rai) [pɛi 452] 'leader, headman', but note the difference in tone: Mienic 3l versus Pa Hng B2.

(8) Noun prefixes:

1. Na-e ā, Wenjie [a 6], Institute wordlist [a 6], e.g. 'ear', Na-e ā 1 kuā 1 biō 1, Wenjie [a 6 qhō 3 mphjo 2], Institute wordlist [a 6 qhō 4 mphio 2].

2. Na-e tā 1, tō 1, Wenjie [tɰv 2], Institute wordlist [tɰv 2], e.g. 'hand, arm', Na-e tā 1 pu 1, Institute wordlist [tɰv 2 pfu 4]; 'tooth', Na-e tō 1 mhō 1; 'skirt', Wenjie

[thv 2 tē 1].

(9) 'red': Na-e tɔ̃², Wenjie [thv 6] (tone C2). Compare possibly Hm Nai [tai 7] (tone D1), but notice the difference in tone.

(10) 'sky': Na-e -fǝǝ 1 [ɲwa] < *ɲwaŋ, Gundong [waŋ 2].

Hmongic has two variants for 'sky', perhaps something like *ndɔŋ A and *NGLwɔŋ A. The latter is suggested by Gaopo [Nqǝ 55] (irregular, for *ɲkǝ 55), Fuyuan [Nqwaŋ 31], Fengxiang [Nquŋ 24] (irregular, for *Nqwuŋ 24), Pu Nu (Meizhu) [ta 33 ɲkuŋ 24], and probably Shiban [Nqwaŋ A]. Compare also Huangluo [ɲo 2] (Chen 1984:17, dialect #12; probably a Yu Nuo dialect). The Pa Hng form appears to correspond to *NGLwɔŋ A.

The Mienic words for 'sky' also look as if they may correspond to the *NGLw- form: Dzao Min [vaŋ 53], Mun [gu:ŋ 31], Mien (Dapingjiang) [luŋ 21], Mien (Chiang Rai) [luŋ 31].

Compare 'finger', above.

(11) 'two': Na-e vǝ 1, Wenjie [va 7] (tone D1), Xishanjie [va 34] (tone A1).

Compare Hm Nai [wa 1'] (Mao and Meng 1982:79).

This is a Proto-Hmong-Mien word, but it shows great irregularity from one language to another in initial, final, and tone. Benedict may now have found the solution to this puzzle.

(12) 'urine': Na-e vi 1, Xishanjie [ve 11].

Most Hmongic dialects have words for 'urine' reflecting a Proto-Hmongic form something like *ra B (Wang 1979, initial 67, final -ʒ). However, in certain Hmongic dialects, in Pa Hng, and in Mienic we have a puzzling array of forms all involving some sort of labial or rounding element:

	Expected form corresponding to *ra B	Actual form
HMONGIC:		
Yanghao	*y ^h ka ¹¹	v ^h ka ¹¹
PA HNG:		
Na-e	*yi ²	vi ¹
Xishanjie	*ji ¹¹	ve ¹¹
MIENIC:		
Dzao Min	*ja ⁴⁴	via ⁴⁴
Biao Min	*la ⁴²	lɔ ⁴²
Mun	*ga ⁵³	wa ⁵³
Mien(Dapingj.)	*la ²³¹	jwe ²³¹

Additional examples of the same phenomenon are Na-e và 1 (pà 2) 'to see' in place of the expected *và 2 (Hmongic 'to watch') and Xishanjie [vi 34] 'vegetable' in place of the expected *[ji 34]. The corresponding Hmongic forms have *ɹ-.

4. Some further remarks on initials

Prenasalized stops develop in various ways in Na-e and Pa Hng. The pattern seems to be as follows:

	Na-e	Wenjie	Xishanjie	Gundong
*mph-	ph	--	ṃ	ṃh
*mp-	b	mp	m	m
*mb-	b	mph, mp	m	m

We see, then, that a characteristic feature of Northern Pa Hng is that prenasalized stops change to nasals, e.g.

	Na-e	Wenjie	Xishanjie	Gundong	
*mphj-	thɕ ¹	--	--	--	'ant'
*ɲtsh-	ʧɕ ²	--	ʧɕ ²²	ʧhei ³	'blood'
*mp-	-bɕ ¹	mpe ⁵	me ⁵⁵	--	'pig'
*mp-	--	--	--	maŋ ⁵	'snow'
*nt-	dɕ ²	--	--	--	'paper'
*mbr-	-biɕ ¹	-mphjo ²	mjo ³³	--	'ear'
*mbr-	biɕ ¹	--	mjo ¹¹	njo ⁴	'fish'
*mbr-	-biɕ ¹	--	mjo ⁴⁴	--	'nose'
*NG-	gɕ ¹	Nqhɕ ²	ŋæ ³³	--	'meat'

Na-e is a Southern Pa Hng language, but in 'leaf' and 'sky' it shows what would seem to be the Northern development of the initial:

	Na-e	Wenjie	Xishanjie	Gundong	
*mbl-	mɔ ¹	--	mā ³³	--	'leaf'
*NGlw-	-ʌɔā ¹	--	--	waŋ ²	'sky'

If Benedict's etymology for the first syllable of Na-e nuh 1 bè 1 'pig' is correct, then this is another example of the same thing. These exceptions might be the result of some sort of borrowing or dialect mixture.

In Southern Pa Hng initials are regularly aspirated in syllables with lower register tones. In Wenjie and the Institute wordlist this occurs quite consistently with tones A2, B2, and C2, e.g. Wenjie [mhiɕ 2] 'person', [vhɛ 4] 'pot', [mhi 6] 'soft'. It does not occur with D2: Wenjie [-mpi 8]

'fan'. In Na-e Bonifacy usually failed to hear the aspiration but he did record it in six words:

	Na-e	Proto-Hmongic	Wang 1979 (initial)
flower	-phā ¹	*b- A2	3
person	mhē ¹	*n- A2	49
earth	lhē ¹	*l- A2	74, 'paddy field'
mountain	-phā ²	*b- B2	3
horse	mhi ¹	*mɸ- B2	17
bird	mhu ¹	*n- C2	49

Benedict reanalyzes Bonifacy's mhu¹ 'bird' as [hmu] and suggests that

The initial h-, with a parallel in Yao (Biao Min hno), probably represents an earlier *s-, for the ubiquitous Sino-Tibetan 'animal prefix' *s-, ...

In this case I disagree with Benedict. I have just shown that the aspiration in Na-e is apparently regular so that we do not need to posit an *s- prefix. As for Biao Min, it appears that in this language voiced nasals regularly become voiceless in words with Entering Tone, that is words that originally ended in a stop:

	Biao Min	Mien (Dapingj.)	Middle Chinese	GSR
bird	ŋoʔ ⁴²	noʔ ¹²	---	---
ink	ŋoʔ ⁴²	ma:t ¹²	*mæk	904 e
socks	munʔ ⁴²	ma:t ¹²	*miwət	311 l
wheat	mɛʔ ⁴²	mɛʔ ¹²	*mwɛk	932 a-c

As it happens, most of our Biao Min examples may be loans from Chinese, but in 'bird' we see the same process affecting a native word. Again, I see no need to posit an *s- prefix.

Finally, with regard to Benedict's remarks on Na-e ku } 'six', we may note that Na-e preserves certain distinctions that have been lost in Xishanjie, for example in the following words (ǰ = IPA [tʃ], ṭ = IPA [t̚]):

	Na-e	Xishanjie	Proto-Hmongic
seven	čá ¹	tɕá ⁴⁴	*dɕ-
to laugh	t ^h á ¹	tɕá ⁵⁴	*tɕ _o -
door	t ^h á ¹	tɕá ³³	*dɕ-
the wind	t ^h i ³	tɕi ⁵⁵	*tɕ-
wine	--	tɕɿ ²²	*tɕ-
husband	t ^h á ²	--	*dɕ- 'man' (vir)
year	t ^h á ³	--	*ɕ-
six	ku ³	tɕu ⁵⁵	*tɕ _o -

Another word in which Na-e č reflects a retroflex initial is 'blood', Proto-Hmongic *ɕtɕh-, Na-e č². Again this changes to a prepalatal initial in Northern Pa Hng: Xishanjie [ɕ²²], [ɕ²²], Gundong [ɕ³].

5. Evidence for separating Pa Hng from Hmongic

The Pa Hng languages show several phonological peculiarities which suggest that they may have split off from the Hmongic stock prior to the period represented by Wang's reconstruction of Proto-Hmongic.

(1) Generally, as one would expect, Proto-Hmongic dentals correspond to Pa Hng dentals, for example:

	Proto-Hmongic	Na-e	Wenjie	Insti-tute	Xishan-jie	Gundong
son/male	*t-	tá ¹	--	--	tá ³⁴	--
skirt	*t-	--	-t ^h e ¹	--	--	--
to come	*d-	tá ¹	--	--	tɕ ³³	--
fire	*d-	tá ¹	--	-tɕɿ ⁴	tɕ ¹¹	--
snake	*ɕn-	ná ¹	--	--	nɕ ³⁴	--
to eat	*n-	noá ¹	--	--	ná ³³	--
cloth	*nt-	--	ntI ¹	--	--	--
paper	*nt-	dá ²	--	--	--	--
to weave	*nt-	--	--	--	ná ⁵⁴	na ⁷

In a number of cases, however, Pa Hng has a labial initial:

	Proto-Hmongic	Na-e	Wenjie	Institute	Xishan-jie	Gundong
hemp	*nd-	bɔ̃ ³	--	--	no ⁴⁴	--
rain	*n-	mu ¹	--	-mō ⁵	--	--
person	*n-	mhɛ ¹	mhie ²	--	--	--
bird	*n-	mhu ¹	--	--	--	--

Benedict's suggestion is that these words originally began with clusters, and that these clusters merged with dentals in Hmongic after Hmongic had separated from Pa Hng. Benedict has already commented on the initials of 'rain' and 'bird'. 'Person' likewise fits very well with the Austro-Tai etymology that has been offered (Benedict 1975:336; 1985).

With regard to 'person', note that in 1975 Benedict took the Hmongic reflex to be the form *prouj A (or *smouj A) 'Hmong' (Wang 1979, initial 5, final -31; Wang 1983:4). But in 1985 Benedict revised his reconstruction and took the Hmongic reflex to be, rather, *nan A 'person' (Wang 1979, initial 49, final -20), which agrees much better in initial and final. Benedict now thinks that *smouj A is a back loan from Chinese.

'Hemp' is like 'bird' in that it has a dental initial in both Hmongic and Mienic, with only Na-e showing a labial initial. Even Xishanjie has a dental initial in 'hemp'. I should point out that 'hemp' has a historically unexpected tone in Na-e and that 'rain' has a historically unexpected tone in the Institute wordlist. But these may be transcription errors or misprints in both cases. Wang Fushi has told me that Institute (1959) contains many misprints.

(2) Generally, Proto-Hmongic prepalatals correspond to prepalatals in Pa Hng:

	Proto-Hmongic	Na-e	Wenjie	Institute	Xishanjie
wine	*tɛ-	--	--	--	tɛr ²²
the wind	*tɛ-	t'i ³	--	-tɛi ⁵	tɛi ⁵⁵
husband	*dʒ-	t'ɔ̃ ²	--	--	--
heavy	*pɛ-	--	--	--	pɛ ²²
mushroom	*ntɛ-	--	--	--	no ³⁴
to climb	*ntɛ-	--	--	--	no ⁵⁵
mouth	*ndʒ-	--	--	--	nu ³³

In two words, however, Pa Hng has the velar/uvular series:

	Proto-Hmongic	Na-e	Wenjie	Institute	Xishanjie
nine	*dɜ-	kɔ ¹	--	--	ko ³³
rice bread	*ptɔ-	--	Nqo ³	--	ŋo ²²

Differences of this kind, including the word for 'nine', were noted by Chang (1976:173, 176), who proposed an alternation between plain and yodized velars to account for them. Thus in 'nine' we would have Pa Hng *g- versus Hmongic *gɜ- > *dɜ-.

'Nine', like 'six', has been suggested to be an early loan from Sino-Tibetan. An initial of the back series also occurs in Hm Nai [ko 2] 'nine' (contrast [-tɕi 5] 'the wind') and in Dzao Min [ku 53] 'nine' (contrast [jau 31] 'the wind'). Note that the Hmong-Mien words for 'nine' all have tone A2, whereas the Chinese form has B1.

(3) Benedict has already discussed the word for 'six' where the Na-e form appears to preserve a contrast between *tɿ- and *tr- which has been lost in Hmongic.

(4) Pa Hng appears to have preserved contrasts involving finals which have been lost in Hmongic. Thus Wang's final -ɿ³ corresponds to Pa Hng [a] or [æ] in some words versus [o] in others:

	Na-e	Wenjie	Xishanjie	Gundong
to see ¹	vã ¹ -	--	--	--
mouth ²	-lhã ³	--	la ⁵⁵	--
old ³	--	qa ⁵	--	--
strength	--	--	jæ ⁴⁴	--
to weave	--	--	nae ⁵⁴	na ⁷
to laugh	tã ¹	--	tɔæ ⁵⁴	--
duck	--	--	ɬæ ⁵⁴	--
to see	-pã ²	--	pa ³¹	--
thunder	bɔ ²	--	--	--
deep	--	--	to ³⁴	--
early	--	--	--	ŋo ³
to come	--	--	lo ¹¹	--
buffalo ⁴	-hɔ ¹	--	--	--

(1) Hmongic 'to watch'. (2) Hmongic 'mouthful'. (3) Not in Wang; see Furnell 1970, Appendix, #628; Chen and Li 1981:1'7, 2'1, 3'3. (4) Hmongic 'ox, cow'.

Similarly, Wang's final -l8 corresponds to Pa Hng [ɿ] in some words versus [o] or [ɔ] in others:

	Na-e	Institute	Xishanjie
wine	--	--	tɿ ²²
fire	tɿ ¹	-tɿ ⁴	tɿ ¹¹
correct ¹	--	--	ɿ ⁵⁴
white	kuə ⁴	--	kuo ³⁴
paper	dɔ ²	--	--
to open	--	--	pɔ ⁵⁴

Note the dissimilatory change [wo] > [wa] in the Na-e word for 'white'.

(5) 'fruit': The Proto-Hmong-Mien word for 'fruit' was probably something like *pjiu B or *pjeu B. Mienic languages preserve the final [-u]: Dzaio Min [beu 11], Mun [pjəu 44], Mien [pjəu 452]. In Ho Nte and Hmongic the final [-u] disappears so that 'fruit' comes to rhyme with 'you':

	Hmong (Qing- yan)	Ho Nte (Boluo)	Dzaio Min	Mien (Da- ping- jiang)	Mien (Chiang Rai)	Mun	Proto-Hmong- Mien
fruit	pi ¹³	pi ⁵⁵	beu ¹¹	pjou ⁵²	pjeu ⁴⁵²	pjəu ⁴⁴	*pjiu ^B
you	ɿ ⁵⁴	mpi ⁴²	mui ⁵²	mwei ²¹	mei ³¹	mvi ³¹	*mwi ^A

In Pa Hng the word for 'fruit' is Xishanjie [pjɛ 22], Na-e pɔ̄ 1. The rounded vowel of the Na-e form is intriguing. Is the rounding a remnant of the final [-u]? If it is, this would be another feature preserved in Pa Hng but lost in Hmongic. But it is difficult to be certain. The rounding might also be conditioned by the initial labial.

If the Na-e word for 'you' does not rhyme with the Na-e word for 'fruit', this would be evidence that Na-e has indeed preserved a distinction lost in Hmongic. Unfortunately, the Na-e form for 'you' was not included in Bonifacy's paper.

(6) 'sun/day': One of the characteristics which appears to have set Proto-Hmongic apart from its sister dialects is that the word for 'sun' or 'day' was something like *ɿɔ̄ A, ending in a nasal, whereas in the other dialects it seems to have been something like *ŋɔ̄i A. Thus in Hmongic languages

¹Not in Wang; see Chang 1947:108.

'sun' rhymes with such words as 'son' and 'horn', whereas in other branches of Hmong-Mien it does not. In Pa Hng, 'sun' does not rhyme with 'son', 'horn', etc.:

	Hmong (Qing- yan)	Na-e	Xi- shan- jie pã ³⁴	Hm Nai paŋ ¹	Ho Nte (Boluo) paŋ ³³	Ho Nte (Zeng- cheng) --	Mien (Chiang Rai) pun ⁴⁴
give/ share	--	--	pã ³⁴	paŋ ¹	paŋ ³³	--	pun ⁴⁴
son/ male ¹ (clf.)	tɔŋ ⁵⁵	-tɔŋ ¹	tã ³⁴	taŋ ¹	taŋ ³³	taŋ ²²	tɔŋ ⁴⁴
steam (vb.)	tɔŋ ⁵⁵	--	--	naŋ ¹	ntaŋ ³³	naŋ ²²	nɔŋ ⁴⁴
hot	--	--	--	--	tjaŋ ³³	tjaŋ ²²	tɕa:ŋ ⁴⁴
young plant	zɔŋ ⁵⁵	--	ja ³⁴	--	--	jaŋ ²²	ja:ŋ ⁴⁴
horn	kuŋ ⁵⁵	kɔŋ ¹	kã ³⁴	--	--	kjaŋ ²²	coŋ ⁴⁴
star	quŋ ⁵⁵	-kɔŋ ¹	--	--	--	--	(lhei ²⁵)
black	ʔluŋ ⁵⁵	kuãŋ ¹	--	--	--	kjaŋ ²²	cia ²⁵
sky	ntuŋ ⁵⁴	ãdã ¹	--	--	--	--	luŋ ³¹
full	puŋ ¹³	--	pã ²²	paŋ ³	paŋ ⁵⁵	--	pwãŋ ⁴⁵²
body	tuŋ ¹³	--	--	ntaŋ ³	taŋ ⁵⁵	taŋ ⁵⁵	tã ⁴⁵²
louse	luŋ ¹³	--	lã ²²	laŋ ³	ntaŋ ⁵⁵	--	naŋ ⁴⁵²
short (clf.)	tɕuŋ ³²	--	--	n ⁴ (:)	thaŋ ⁵⁴	thaŋ ⁵³	(tau ³¹)
snow	mpuŋ ⁴³	--	--	mpaŋ ⁵	--	paŋ ³³	(2)
break	tuŋ ⁴³	--	tã ⁵⁵	taŋ ⁵	taŋ ²¹	taŋ ³³	taŋ ²⁵
sun/ day	nuŋ ⁵⁵	-nhɛ ¹	ne ³⁴	-nhɛ ¹	ntɔ ³³	no ²²	noɔi ⁴⁴

In Na-e, 'black' and 'sky' have undergone dissimilatory changes as explained earlier, so that we have [kwaŋ] and [ɲwa] instead of *[kwɔŋ] and *[ɲwɔŋ]. Although 'sky' loses its final nasal, the vowel of 'sky' remains quite different from that of 'sun'. Please note also that in the case of the Na-e word for 'sun' I have assumed that Bonifacy's ã n l hɛ l is actually a false division for ã l nhɛ l.

¹Hm Nai classifier for people.

²Dapingjiang Mien [bwan²⁴].

6. Mienic versus non-Mienic

The various non-Mienic branches of Hmong-Mien share certain characteristics that set them apart from Mienic, for example:

1. Distinctive lexical items. e.g. 'pig': non-Mienic *mpai C versus Mienic *duŋ B.
2. Distinctive forms of particular lexical items, e.g. 'black': non-Mienic final *-ŋ versus Mienic final *-ʔ.
3. In the Mienic languages *-ʔ after *voiceless initials remains distinct from tone C whereas in the non-Mienic languages it merges with tone C.

Pa Hng has non-Mienic characteristics, for example:

1. 'pig': Na-e -bè 1, Wenjie [mpe 5], Xishanjie [mɛ 55].

Compare Hmong (Qingyan) [mpo 43], Hm Nai [mpe 5], Kiong Nai [mpei 5], Yu Nuo (Liutian) [mɛi 4], Ho Nte (Boluo) [pui 21], Ho Nte (Zengcheng) [pi 33]. Contrast Dzao Min [tiŋ 44], Biao Min [twəʔ 42], Mien [tuŋ 231], Mun [tuŋʔ 53].

But Na-e also has nuŋ 1 -, which Benedict suggests may be a loan from Mienic.

2. 'black': Na-e kuən 1. Compare Hmong (Qingyan) [ʔluŋ 55], Ho Nte (Zengcheng) [kjaŋ 22]. Contrast Dzao Min [kia 44] < *kjaʔ, Biao Min [kjaʔ 54], Mien (Dapingjiang) [tɕeʔ 55], Mien (Chiang Rai) [ciaʔ 55], Mun [kja 4232] < *kjaʔ.

3. Final glottal stop merges with tone C:

	Xishan- jie	Hmong (Xian- jin)	Ho Nte (Bo- luo)	Dzao Min	Mien (Chiang Rai)	Mun
sleep (C1)	pai ⁵⁵	pu ⁴⁴	pɔ ²¹	bui ⁴¹	pwei ²⁵	fvi ²⁴
six (*-ʔ)	təu ⁵⁵	tou ⁴⁴	ko ²¹	to ⁴⁴	cu ^{ʔ55}	kjo ⁴²³²
hundred (*-ʔ)	pɛ ⁵⁵	po ⁴⁴	pa ²¹	ba ⁴⁴	pɛ ^{ʔ55}	pɛ ⁴²³²
laugh (D1)	tɕæ ⁵⁴	tɕo ³³	ku ³³⁵	tut ⁴⁴	cet ⁵⁵	kjet ²²

Did Proto-Hmong-Mien begin by splitting into Proto-Mienic and Proto-non-Mienic? Were the non-Mienic characteristics innovations in Proto-non-Mienic? 'Pig' and 'black' perhaps are not. Benedict has presented evidence that both 'pig' and the nasal-final variant of 'black' go back to Proto-Austro-Tai (Benedict 1975:354, 266). Benedict's theory that Hmong-Mien

is a branch of Austro-Tai is controversial; see especially Wang (1985b). But if Benedict's etymologies for 'pig' and the nasal-final variant of 'black' are correct, then these forms already existed in Proto-Hmong-Mien and their presence in Hmongic, Pa Hng, Ho Nte, and so on, is simply a shared inheritance from Proto-Hmong-Mien and does not imply any special relationship among these groups.

With regard to the merger of final glottal stop with tone C, Sagart (1984, 1985) has presented evidence that both tone B and tone C were glottalized in Chinese, Tai, and Hmong-Mien. If this is correct, then the merger of final glottal stop with tone C may turn out to be a very natural sound change which could easily have occurred independently in the different groups. Indeed even in Mienic, we find that in many dialects final glottal stop merges with tone C after *voiced initials, although remaining distinct after *voiceless initials. For example, in the Houei Sai dialect of Mun, [ɲu 21] 'chisel' (*-ʔ) differs only in initial from [du 21] 'hemp' (C2).

On the other hand, one area which I think does hold a lot of promise for elucidating subgroups of Hmong-Mien is the development of the finals. The old idea that "Miao" preserves initials and "Yao" preserves finals is too simple, as David Solnit has pointed out (1984:2). But a step by step investigation of each Proto-Hmong-Mien final may turn up cases of shared innovation which genuinely reflect common ancestry. We have already seen two possible examples. One was the merger of *iu and *i shared by Ho Nte and Hmongic which I discussed in connection with the word for 'fruit'. The other was the array of forms which I cited in connection with the word for 'sun'. Mien shows several different finals in these forms: -am, -ɔm; -ɔn, -un; -aŋ, -a:ŋ, -ɔŋ, -uŋ. Even in Biao Min, which is noted for its greatly reduced system of finals, we have different finals in [jɔ 44] 'young plant' versus [taŋ 24] 'to break'. But in Ho Nte, Hm Nai, Pa Hng, and Hmongic all these finals appear to have merged.

7. Conclusions

(1) Both Na-e and "Tahua Yao" (Xishanjie) appear to belong to the Pa Hng subgroup of Hmong-Mien.

(2) Pa Hng may have split off from the Hmongic stock prior to the period represented by Wang's reconstruction of Proto-Hmongic.

(3) Hm Nai, Yu Nuo, and Kiong Nai may also have split off from Hmongic at an early date. These languages deserve more research.

(4) There may be some special connection between Pa Hng and Hm Nai: see 'dog', 'finger', 'red', 'two', 'nine'.

(5) The earliest split in the Hmong-Mien family may have

been between Mienic and everything else but the evidence for this is not yet conclusive.

Benedict suggests, rather, that the earliest split might have been between Na-e (read "Pa Hng") and everything else. He bases this suggestion on the labial initial in the Na-e form for 'bird'. To this evidence we may now be able to add the labial initial in the Na-e form for 'hemp'. On the other hand, as Benedict admits, the change from *mr to [n] could have occurred independently in different languages rather than being a shared innovation implying any special connection among the non-Pa Hng languages. Indeed a similar change to a dental initial would seem to have occurred even within Pa Hng in the Xishanjie form for 'hemp'. I agree with Benedict, however, that more data on the Pa Hng languages are needed before questions such as these can be resolved.

ACKNOWLEDGEMENTS

I am both grateful and honored that Paul Benedict has asked me to write these comments on his paper. Ever since I began doing research on Hmong-Mien languages some years ago Paul has shown interest in my ideas and suggested lines of inquiry. This has meant a great deal to me.

I am also deeply grateful for the encouragement and guidance I have received from André Haudricourt, Kun Chang, Wang Fushi, Martha Ratliff, and Brenda Johns. I also owe a special debt to Martha for making me xerox copies of many of the sources used in preparing this paper.

Finally, I would like to take this opportunity to thank my Mun informant, Boonchan Chang, for his kindness and interest in my research.

APPENDIX: Sources

- Proto-Hmongic: Initials and tones from Wang (1979). Reconstructions of finals are my own. (Wang assigns the finals numbers but does not reconstruct the actual pronunciations.) See also Chang (1947, 1953, 1972, 1976) and Purnell (1970).
- Hmongic:
- Green Mong of Laos: Xiong, Xiong, and Xiong (1983).
 - Shiban: Chen (1984:15, dialect #7). Dialect of Shibanzhai, Kaili county, southeastern Guizhou.
 - Pu Nu: Mao, Meng, and Zheng (1982). Dialect of Meizhu, Du'an county, central Guangxi.
 - Other: Wang (1979).
 - Na-e: Bonifacy (1905). Dialect of Bao-lac, northern Vietnam.
 - Wenjie: Mao, Meng, and Zheng (1982:118, 121-123). Dialect of Wenjie, Sanjiang county, northeastern Guangxi.
 - Institute wordlist: Institute for the Study of Minority Languages (1959), cited in Moskalev (1978:15).
 - Xishanjie: Chang (1947, 1953, 1972, 1976). Dialect of Xishan-

- jie, near Yongcong, southeastern Guizhou.
- Gundong: Chen (1984:17, dialect #13). Dialect of Gundong, Liping county, southeastern Guizhou.
- Hm Nai: Institute for the Study of Minority Languages (1959); Mao, Meng, and Zheng (1982:119, 121-123); Mao and Meng (1982:77-80). Hm Nai is spoken in western Hunan.
- Kiong Nai: Institute for the Study of Minority Languages (1959); Mao, Meng, and Zheng (1982:119, 121-123); Mao and Meng (1982:77-80). Kiong Nai is spoken in Jinxiu county, eastern Guangxi.
- Yu Nuo:
- (1) Mao, Meng, and Zheng (1982:120-123): dialect of Liutian, Longsheng county, northeast Guangxi.
 - (2) Chen (1984:17, dialect #12): dialect of Huangluo, Longsheng county, northeast Guangxi.
 - (3) Institute for the Study of Minority Languages (1959): Longsheng county, northeast Guangxi, exact locality not specified.
- Ho Nte (Boluo): Mao and Meng (1982). Dialect of Boluo county, southeast Guangdong (near Hong Kong).
- Ho Nte (Zengcheng): Chen (1982; 1984:18, dialect #15). Dialect of Xiashui, Zengcheng county, southeast Guangdong (near Hong Kong).
- Mien (Chiang Rai): Funnell (1970); Lombard (1968). Dialect of Chiang Rai province, northern Thailand.
- Mien (Dapingjiang): Mao, Meng, and Zheng (1982). Dialect of Dapingjiang, Jiangdi commune, Longsheng county, northeastern Guangxi.
- Mun: author's fieldnotes. Dialect of Ban Nam Yong, Houei Sai district, Houa Khong province, northwest Laos.
- Biao Min: Solnit (1984); Mao, Meng, and Zheng (1982:63-69). Dialect of Quanzhou county, northeast Guangxi.
- Dzao Min: Wong (1939); Yu and Chao (1984). Dialect of Youling, Liannan county, northwest Guangdong.

REFERENCES

- Benedict, Paul K. 1975.
 Austro-Thai Language and Culture with a Glossary of Roots.
 HRAF Press.
- _____. 1985.
 "Early Sino-Tibetan/Miao-Yao Loan Relationships". Panel on
 Hmong Language and Linguistics, Southeast Asian Studies
 Summer Institute Conference, Ann Arbor.
- Bonifacy, [Auguste Louis]. 1905.
 "Etude sur les langues parlées par les populations de la
 haute Rivière Claire". Bulletin de l'Ecole Française
 d'Extrême Orient 5:306-327.
- Chang, Kun. 1947.
 "Miao-Yao yǔ shēngdiào wèntí" (On the Tone System of the
 Miao-Yao Languages). Bulletin of the Institute of History

- Chang, Kun. 1947.
and Philology, Academia Sinica, XVI:93-110.
- _____. 1953.
"On the Tone System of the Miao-Yao Languages". *Language* 29:374-378.
- _____. 1972.
"The Reconstruction of Proto-Miao-Yao Tones". *Bulletin of the Institute of History and Philology, Academia Sinica*, 44:541-628.
- _____. 1976.
"Proto-Miao-Initials". *The Bulletin of the Institute of History and Philology, Academia Sinica*, XLVII:2:155-218.
- Chen Qiguang. 1982.
"Shèyǔ zài Miào-Yáoyǔlǐ'de dìwèi" (The Position of She in Miao-Yao). 15th International Conference on Sino-Tibetan Languages and Linguistics.
- _____. 1984.
"Gǔ Miào-Yáo yǔ bī guān bīshè yīn shēngmǔ zài xiàndài fāngyán zhōng fǎnyǐng xíngshì'de lèixíng" (The Reflected Patterns in Modern Dialects of the Prenasalized Stop Initials of Proto-Miao-Yao). *Mǐnzǔ Yǔwén* 1984:5:11-22.
- _____. and Li Yongsui. 1981.
"Hányǔ Miào-Yáoyǔ Tóngyuán Lǐzhèng" (Some Examples of the Genetic Affinity Between Han and Miao-Yao). *Mǐnzǔ Yǔwén* 1981:2:13-26.
- GSR = *Grammata Serica Recensa* (Karlgren 1957).
- Haudricourt, André G. 1954.
"Introduction à la phonologie historique des langues miao-yao". *Bulletin de l'École Française d'Extrême Orient* 44: 555-576. Reprinted in André G. Haudricourt, 1972, *Problèmes de phonologie diachronique*, Société pour l'étude des langues africaines, Paris, pp. 183-208.
- _____. 1971.
"Les langues miao-yao". *Bulletin du Centre de Documentation et de Recherche, Equipe Atlas*, fascicule 1, 31-47, 51.
- Institute for the Study of Minority Languages. 1959.
Zhōngguó Shǎoshù Mǐnzǔ Yǔyán Jiǎnzhī: Miào-Yáo Yǔzǔ Bù'fēn (A Sketch of the Languages of China's Minorities: Section on the Miao-Yao Language Family). Chinese Academy of Sciences, Institute for the Study of Minority Languages. Cited in Moskalev (1978).
- Karlgren, Bernhard. 1957.
Grammata Serica Recensa. The Museum of Far Eastern Antiquities, Stockholm, Bulletin No 29. Reprinted 1964 by Elanders Boktryckeri Aktiebolag, Göteborg.

- Lombard, Sylvia. 1968.
Yao-English Dictionary. Compiled by Sylvia J. Lombard. Edited by Herbert C. Purnell, Jr. Linguistics Series II, Data Paper: Number 69, Southeast Asia Program, Department of Asian Studies, Cornell University, Ithaca, New York.
- Lombard-Salmon, Claudine. 1972.
Un exemple d'acculturation chinoise: la province du Gui zhou au XVIIIe siècle. Appendix VII: Connaissance linguistique des Han: les quatre lexiques du Miao fang bei lan, pp. 301-321. Publications de l'Ecole Française d'Extrême Orient, volume LXXXIV.
- Mao Zongwu and Meng Chaoji. 1982.
"Bólúo Shěyǔ gǎishù" (A Brief Description of the She Language in Boluo District). Mínzú Yǔwén 1982:1:64-80.
- Mao Zongwu, Meng Chaoji, and Zheng Zongze. 1982.
Yáozú Yǔyán Jiǎnzǐ (A Sketch of the Languages of the Yao People). Mínzú Chūbǎnshè (Nationalities Press), Peking.
- Moskalev, A. A. 1978.
Jazyk duan'skix jao (jazyk nu) (The Yao language of Du'an (The Nu language)). "Nauka" Press, Moscow.
- Nguyen, Dinh-Hoa. 1985.
"Vietnam's Language Policy towards Ethnic Minorities". 18th International Conference on Sino-Tibetan Languages and Linguistics.
- Purnell, Herbert C. 1970.
Toward a Reconstruction of Proto-Miao-Yao. Ph.D. dissertation, Cornell University. University Microfilms, #70-17,096.
- Sagart, Laurent. 1984.
"On the Departing Tone". 17th International Conference on Sino-Tibetan Languages and Linguistics.
- _____. 1985.
"Some Further Evidence on the Glottalization of the Chinese Departing Tone". 18th International Conference on Sino-Tibetan Languages and Linguistics.
- Solnit, David B. 1984.
"Introduction to the Biao Min Yao Language". 17th International Conference on Sino-Tibetan Languages and Linguistics.
- Wang Fushi. 1979.
Miáoyǔ Fāngyán Shēng Yǔn Mǔ Bǐjiǎo (The Comparison of Initials and Finals of Miao Dialects). Monograph distributed at the 12th International Conference on Sino-Tibetan Languages and Linguistics.

- Wang Fushi. 1983.
 "Miáoyǔ fāngyán huàfēn wèntí" (On the Dialect Divisions of Miao Language). *Mínzú Yǔwén* 1983:5:1-22.
- _____. 1985a.
 Miáoyǔ Jiǎnzǐ (A Sketch of the Miao Language). *Mínzú Chūbǎnshè* (Nationalities Press), Peking.
- _____. 1985b.
 "Miáo-Yáoyǔ de xī shǔ wèntí chū tǎn" (A Preliminary Investigation of the Genetic Affiliation of the Miao-Yao Languages). Symposium on the Languages, Cultures and Histories of the Minorities of China, University of California, Santa Barbara. English translation by Mark Hansell.
- Wong Sik-ling. 1939.
 "Phonetics and Phonology of the Yao language: description of the Yau-ling dialect". *Lingnan Science Journal* 18:425-455.
- Xiong, Lang; Joua Xiong; and Nao Leng Xiong. 1983.
 English-Mong-English Dictionary. Privately printed, Milwaukee.
- Yu Weiwen and Chao Zongqi. 1984.
 "Yóulǐng Yáo huà gǎishù" (A brief description of the Yao dialect in Youling). *Journal of Sun Yat-sen University, social sciences edition*, 1984:3:115-126, 62.

Some Comments on Benedict's "Miao-Yao Enigma": Addendum

David Strecker

I ended my paper with the words "I agree with Benedict, however, that more data on the Pa Hng languages are needed before questions such as these can be resolved." Some of these additional data have now become available. On the very day that I mailed my paper off to be considered for publication I received from Wang Fushi a comparative wordlist including 80 words in the dialect of Pa Hng spoken in Wenjie. 57 of these are new words, not given in either Mao, Meng, and Zheng or the Institute wordlist. The kindness and generosity of Professor Wang now make it possible for me to fill in some gaps and correct some mistakes in my treatment of Wenjie.

In addition, Dr. Benedict has sent me several pages of detailed comments and corrections on my paper. Discussing these will require a separate paper, but I will mention a few of Benedict's most important corrections below.

1. Preface

With regard to my schema of Hmong-Mien subgroups, Wang agrees with me that on purely linguistic grounds Pu Nu would be considered a subgroup of the Sichuan-Guizhou-Yunnan branch of Hmongic. He goes on to explain, however (in his letter accompanying the comparative wordlist), that Chinese writers base their classifications on non-linguistic criteria as well: "The policy of our country holds that the speech of each national minority is named according to the official name of the national minority." Thus Chinese writers divide Hmongic into Miào yǔ, namely those languages whose speakers are culturally as well as linguistically Hmong, and Bùnmǎ yǔ, namely those languages whose speakers are linguistically Hmong but culturally Yao. I think it is important to understand that we are not dealing here with any disagreement about the histories of the languages but merely with different types of classification, based on different criteria and fulfilling different purposes.

Wang disagrees with Benedict's and my idea that Pa Hng should be a separate branch: "the materials show the close relation between pa 31(3) ɣɔ 35(1) and pu 54(3') nu 24(2')." Thus Wang would retain Pa Hng within Hmongic.

Benedict rightly questions my raising Hm Nai, Kiong Nai, and Yu Nuo to the status of separate branches without giving any evidence for this. Let me therefore rephrase my suggestion thus:

Hm Nai, Kiong Nai, and Yu Nuo have been considered to be branches of Hmongic. I know of no evidence to the

contrary, but it is difficult to be certain because not very much has been written about these languages. Mao, Meng, and Zheng, while grouping them with Pu Nu, do point out that they are rather distinctive. Thus I hope very much that someone will do a systematic historical-comparative study of these languages along the lines of Wang (1979).

So let me give a more conservative schema of subgroupings:

I. Hmongic

- A. Eastern Guizhou (Purnell's "Eastern") [HM]
- B. West Hunan (Purnell's "Northern") [HM]
- C. Sichuan-Guizhou-Yunnan (Purnell's "Central" and "Western")
 - 1. Sichuan-Guizhou-Yunnan (Purnell's "West A") [HM]
 - 2. Northeast Yunnan [HM]
 - 3. Guiyang [HM]
 - 4. Huishui [HM]
 - 5. Mashan [HM]
 - 6. Luobo River [HM]
 - 7. Eastern or Chong'an River [HM, Gedou]
 - 8. Pu Nu (Tung Nu) [BN]
 - 9. Nu Nu [BN]
 - 10. Pu No [BN]
 - 11. Nao Klao [BN]
 - 12. Nu Mhou [BN]
- D. Probably Hmongic but exact classification uncertain:
 - 1. Hm Nai [BN]
 - 2. Kiong Nai [BN]
 - 3. Yu Nuo [BN]
 - 4 - 12. Nine additional major groups [HM]: see Wang (1983:1).

II. Ho Nte

III. Mienic

- A. Mien-Kim: Mien (Iu Mien), Mun (Kim Mun), Biao Mon
- B. Biao-Chao: Biao Min, Chao Kong Meng
- C. Dzao Min

IV. Classification uncertain: Pa Hng [BN], Na-e

Within the Hmongic group, languages marked HM are spoken by people who are culturally Hmong and are called "Miao" by Chinese writers, and languages marked BN are spoken by people who are culturally Yao and are called "Bunu" by Chinese writers. Speakers of the Chong'an River dialect group call themselves [phoŋ 33 (1)] in their own language, and are included under "Miao" by Chinese linguists, but they, themselves, when speaking Chinese, say that they are Gedou

people rather than Miao people (Wang 1979, pp. 27-28).

2. Transcription

Mao, Meng, and Zheng give only the etymological values of the Wenjie tones, following the usual convention: 1 = A1, 2 = A2, etc. Wang gives the pitch contours as well, using the usual five point scale with 1 the lowest pitch and 5 the highest: 55 = high level, 54 = falling from high to mid high, etc. This makes it possible to compare the pitch contours of Na-e, Wenjie, and Xishanjie:

	Na-e	Wenjie	Xishanjie
A1	interrogative	35	34
A2	descending	33	33
B1	descending	31	22
B2	descending	31	11
C1	acute	55	55
C2	descending	44	44
D1	acute	53	54
D2	descending	32	31

In Wenjie and probably in Na-e syllables with lower register tones, A2, B2, C2, D2, are pronounced with whispery voice. In Wenjie the whispery voice serves to distinguish B2 from B1, which has the same pitch contour. (See section 4.)

3. Evidence that Na-e belongs to the Pa Hng group

(1) *r and *ʀr become yod: Wenjie [jo 35(1)] 'stone'.

(2) Velar and uvular initials merge into a single series, realized phonetically as uvular in Wenjie. But in some words Wenjie has a velar initial. I have no explanation for this.

	Proto-Hm. Prepalatal	Proto-Hm. Velar	Proto-Hm. Uvular
Wenjie uvular	Nqo 35(1) 'mushroom' Nqo 3 'rice bread'	q ⁵ 35(1) 'horn' q ⁵ 1 'warm' q ^ʀ 35(1) 'insect' qo 3 'road'	q ⁵ 35(1) 'star' qa 31(3) 'excrement' qo 55(5) 'old' qh ⁵ 31(3) 'hole' Nqo 35(1) 'pigeon' Nqh ^ɛ 33(2) 'meat'
Wenjie velar	k ^h o 33(2) 'nine' ŋ ^h 33(2) 'cow'	k ⁵ 35(1) 'needle' k ^h u 32(8) 'ten' ŋ ^h k ^h r 32(8) 'pair'	

(4) Proto-Hmongic final -5: Wenjie [mpjo 31(4)] 'fish', [jo 35(1)] 'stone'.

Proto-Hmongic final -6: Wenjie [phe 53(?)] 'daughter' (compare Na-e piē 1 dèn 1 'younger sister'), [Nqhē 33(2)] 'meat'.

Proto-Hmongic final -23: Wenjie [t⁵ 35(1)] 'son', [t⁵ 55(5)] 'to snap, break', [l⁵ 35(1)] (classifier), [q⁵ 35(1)] 'horn', [q⁵ 35(1)] 'star'.

(11) 'two': Wang gives Wenjie [va 35(1)] with the historically expected tone, A1. The form in Mao, Meng, and Zheng with tone D1 is perhaps a misprint.

3a. Additional characteristics shared by the Pa Hng languages

(1) The ethnonym [p̄h̄], with tone A1: Wenjie [pa 31(3)] [p̄h̄ 35(1)] 'Pa Hng', Xishanjie [h̄h̄ 34(1)] 'people'. (It is not clear from Chang whether the Xishanjie form means people in general or specifically Pa Hng people.) Is [pa 31(3)] [p̄h̄ 35(1)] related to the Vietnamese name Pāthen? And what is the source of the name Nā-é?

(2) The word for 'thousand' has tone C1 instead of the historically expected tone A1: Wenjie [p̄ē 55(5)], Xishanjie [p̄o 55(5)]. For Na-e Bonifacy writes de 1 [ðe] with tone 1. Unfortunately Bonifacy's tone 1 seems to represent cases where he simply failed to hear the tone.

(3) Wang's Proto-Hmongic final -15 generally corresponds to Pa Hng [ū], [u], or [u]:

	Na-e	Wenjie	Xishanjie
bean	---	---	tu 31(8)
iron	au ³ [d̄u]	tu 55(5)	tu 55(5)
six	ku ³	təu 55(5)	təu 55(5)
mouth	---	ntəu 33(2)	nu 33(2)
ten	ku ¹	khu 32(8)	ku 31(8)
blind*	ku ¹	---	---

But 'to drink' has [ɔ]:

	Na-e	Wenjie	Xishanjie
drink	hɔ ³	hɔ 53(7)	hɔ 54(7)

This word may be a recent loan from Chinese. Compare Cantonese [hɔ:t 33(7)].

*Not in Wang (1979). See the Layiping and Yanghao forms in Wang (1985a:170) -- the vowel in the Dananshan form is irregular.

4. Some further remarks on initials

The pattern of development of prenasalized stops in Wenjie appears to be as follows:

*mph- > ph-: [phe 53(7)] 'daughter'. But [ntəe 31(3)] 'blood':

*mp- > mp-: [mpe 55(5)] 'pig', [mpjo 35(1)] 'green',
[nte 31(3)] 'to wash (hands)', [ntəe 31(3)] 'early',
[ntɪ 35(1)] 'cloth', [nta 53(7)] 'to weave', [Nqo 35(1)]
'mushroom', [Nqo 3] 'rice bread', [Nqo 35(1)] 'pigeon'.

*mb- > mph-: [-mphjo 2] 'ear' (sc. [mph-]; see below),
[mphjo 44(6)] 'nose', [mphɪ 32(8)] 'peppery',
[mphjo 33(2)] 'leaf', [ntəhu 33(2)] 'mouth',
[ŋkɸɪ 32(8)] 'pair', [Nqhɛ 33(2)] 'meat'.

The forms [mpjo 31(4)] 'fish' (Wang) and [-mpi 8] 'fan' (Mao, Meng, and Zheng) are perhaps typographic errors for forms with [-h-]: see below.

Wenjie does not share with Na-e the Northern-Pa-Hng-like development in mb 1 'leaf'. Wenjie has [mphjo 33(2)]. But Wenjie does show a Northern-like development in [nho 33(2)] 'boat' and [nhɛ 31(4)] 'lazy', both from *ŋg-. Note also Wenjie [lhja 31(4)] 'to flow', from *NGL-.

I also said that in Southern Pa Hng initials are regularly aspirated in syllables with lower register tones. I used the word "aspiration" loosely to refer to both true aspiration, that is, delay of voice onset, and "voiced aspiration", that is, whispery voice, because I was not certain which type of aspiration Mao, Meng, and Zheng meant to represent by the symbol [h]. Wang's list clearly shows whispery voice, [h̥], in all lower register tones, including D2:

A2: [pɸɪ 33(2)] 'flower', [mphjo 33(2)] 'leaf', [nhɔ 33(2)] 'eat', [tɸh̥ 33(2)] 'door', [lh̥ɪ 33(2)] 'farmland', [kɸo 33(2)] 'nine', [ŋh̥ 33(2)] 'cow', [ntəhu 33(2)] 'mouth', [nho 33(2)] 'boat, ship', [Nqhɛ 33(2)] 'flesh', [kwɸɪ 33(2)] 'yellow'.

B2: [mhɪ 31(4)] 'horse', [vh̥e 31(4)] 'boiler', [tɸɪ 31(4)] 'fire', [nh̥u 31(4)] 'he, she, it', [nh̥ɛ 31(4)] 'lazy', [lh̥ja 31(4)] 'flow'.

C2: [mhɪ 44(6)] 'sell', [mphjo 44(6)] 'nose', [tɸe 44(6)] 'die', [tɸh̥a 44(6)] 'seven'.

D2: [vh̥ɛ 32(8)] 'ten thousand' (D2 in Pa Hng and some Hmongic languages; generally C2), [mphɪ 32(8)] 'peppery', [jh̥ɪ 32(8)] 'eight', [kɸu 32(8)] 'ten', [ŋkɸɪ 32(8)] 'pair', [jh̥u 32(8)] 'one' (compare Mien [jet 12(8)]).

The Institute wordlist, as reproduced by Moskalev, likewise shows whispery voice:

A2: [-mpɦio 2] 'ear'

B2: [-pɦu 4] 'hand; arm', [-tɦɿ 4] 'fire', [-pɦi 4] 'head'

C2: [-mɦi 6] 'eye'

D2: no examples.

There are three exceptions, which may be typographic errors:

Wang [mpjo 31(4)] 'fish', with unaspirated initial.

Mao, Meng, and Zheng [-mpi 8] 'fan', with unaspirated initial.

Institute wordlist (in Moskalev) [-mphjo 6] 'nose', with [h] instead of [ɦ].

Thus I interpret Bonifacy's mhu 1 'bird' as [mɦu] and assert that the [ɦ] is a predictable correlate of the tone and does not imply any sort of prefix. This interpretation is better for Benedict's overall argument since it makes for a more straightforward development from Proto-Austro-Tai *mrok to Na-e mhu 1. I am happy to report that Benedict and I are now in agreement on this interpretation of Bonifacy's form.

Finally, Wenjie agrees with Xishanjie and disagrees with Na-e in having prepalatal initials in 'seven', 'blood', and 'six':

	Proto-Hmongic	Wenjie
seven	*dz-	tɕɦā 44(6)
blood	*ɾtɕh-	ɾtɕe 31(3)
door	*d̥l-	tɕɦĩ 33(2)
fir	*tɕ-	tɕĩ 35(1)
thrush	*tɕ-	tɕĩ 35(1)
wine	*tɕ-	tɕɿ 31(3)
mouth	*ɾd̥ɿ-	ɾtɕɦu 33(2)
six	*t̥l̥ _o -	tɕu 55(5)

5. Evidence for separating Pa Hng from Hmongic

(1) Proto-Hmongic dentals = Pa Hng dentals: Wenjie
 [tɿ 35(1)] 'earth', [tẽ 35(1)] 'skirt', [tĩ 35(1)] 'son',
 [tɦɿ 31(4)] 'fire', [tɦe 44(6)] 'die', [nẽ 35(1)] 'snake',
 [nĩ 55(5)] 'cold', [ŋe 35(1)] 'sun, day', [ŋe 35(1)] 'ear of

grain', [nɦu 31(4)] 'he, she, it', [nɦɔ 33(2)] 'eat', [ntɪ 35(1)] 'cloth', [nta 53(?)] 'weave', [lɿ 35(1)] (classifier), [ʔa 55(5)] 'moon', [ʔu 55(5)] 'iron'

Proto-Hmongic dentals = Pa Hng labials: no examples in Wang's list.

I may have been mistaken in including 'hemp' among the examples of Proto-Hmongic dentals corresponding to Pa Hng labials. Benedict has reminded me that Proto-Hmongic *ndo C 'hemp' may be related to Archaic Chinese 紵 *dfio B (a kind

of hemp) (GSR 84e; see also Chen and Li 1981, 1^o7). If we accept the view of Chinese linguists that Hmong-Mien is genetically related to Chinese, then 'hemp' may have had some sort of cluster in Proto-Sino-Tibetan which became labial in Na-e and which became dental in other Hmong-Mien languages and in Chinese. If we accept Benedict's view that Hmong-Mien is a branch of Austro-Tai, then the lexical resemblances between Hmong-Mien and Chinese are due to borrowing. Then we have two possibilities:

(1) Benedict's view is that the Proto-Hmong-Mien word was borrowed from some dialect of Archaic Chinese and therefore always had a dental initial in Hmong-Mien. Na-e bó } is completely unrelated to Proto-Hmongic *ndo C.

(2) Conversely, the Chinese word may have been borrowed from a Hmong-Mien language. In that case, it is still possible that the Proto-Hmong-Mien word had some sort of cluster which became labial in Na-e and dental everywhere else. The Chinese form would simply have been borrowed from a language other than Na-e after the change to a dental initial had already taken place.

I am not prepared at the present time to enter the controversy over whether Chinese borrowed words from Hmong-Mien (Wang 1985b:30-31); I am merely raising this as a theoretical possibility. I do not know whether this is relevant or not, but I learn from The Times Atlas of China (1974 edition, page 81) that hemp fiber is an important crop in the region of Guangxi where Pa Hng is spoken.

(2) Proto-Hmongic prepalatals = Pa Hng prepalatals:
Wenjie [tɕi 35(1)] 'fir', [tɕɔ 35(1)] 'thrush', [tɕɿ 31(3)] 'wine', [lɿ 35(1)] 'wife', [ntɕɦu 33(2)] 'mouth', [lɕɦi 32(8)] 'eight'.

Proto-Hmongic prepalatals = Pa Hng velars or uvulars:

Velar: Wenjie [xɦo 33(2)] 'nine', [ɣɦ 33(2)] 'cow'.

Uvular: Wenjie [Nqo 35(1)] 'mushroom', [Nqo 3] 'rice bread'.

Notice that 'mushroom' has a uvular initial in Wenjie versus a prepalatal initial in Xishanjie.

(3) Na-e stands alone as our sole witness for a velar initial in *ku 3* 'six'. Both Wenjie and Xishanjie have a prepalatal initial: Wenjie [təu 55(5)], Xishanjie [təu 55(5)].

(4) Wang's final -13:

	Na-e	Wenjie	Xishanjie	Gundong
to see	vã ¹	---	---	---
mouth	-lhã ³	---	la 55(5)	---
strength	---	---	jæ 44(6)	---
to weave	---	nta 53(7)	næ 54(7)	na ⁷
to laugh	t·ã ¹	---	təæ 54(7)	---
duck	---	---	tæ 54(7)	---
to see	-pã ²	---	pa 31(8)	---
thunder	bð ²	---	---	---
deep	---	---	to 34(1)	---
big	---	tjo 35(1)	---	---
early	---	təo 31(3)	---	no ³
to come	---	---	lo 11(4)	---
buffalo	-hð ¹	---	---	---

I made a mistake in including the word for 'old' as an example of final -13. I believe it should be final -12. Wang gives the Wenjie form of this word as [qo 55(5)], with [o] rather than [a]. Wang's form agrees with other examples of final -12: Wenjie [mpjo 35(1)] 'green', [so 1] 'on', [kfo 33(2)] 'nine', [Nqo 3] 'rice bread', [Nqo 35(1)] 'pigeon'. Therefore the form [qa 5] in Mao, Meng, and Zheng may be a misprint.

Wang's final -18:

	Na-e	Wenjie	Xishanjie
wine	---	tɕɿ 31(3)	tɕɿ 22(3)
fire	tɕ ¹	tɕɿ 31(4)	tɿ 11(4)
correct	---	---	ɲɿ 54(7)
pair	---	ŋkɕɿ 32(8)	---
white	kuã ⁴	---	kuo 34(1)
paper	dɕ ²	---	---
to open	---	---	pɕ 54(7)

(6) 'sun/day' does not rhyme with 'son', 'horn', etc.:
 Wenjie [tɕ¹ 35(1)] 'son', [lɕ¹ 35(1)] (classifier), [qɕ¹ 35(1)]
 'horn', [qɕ¹ 35(1)] 'star', [tɕ¹ 55(5)] 'break' versus
 [pe 35(1)] 'sun, day'.

6. Mienic versus non-Mienic

Final glottal stop merges with tone C:

	Wenjie	Mien (Chiang Rai)
sleep (C1)	py 55	pwei 25
six (*-?)	tɕu 55	cu? 55
hundred (*-?)	pe 55	pe? 55
weave (D1)	nta 53	dat 55

I said that if Benedict's Austro-Tai etymology for *mpai C 'pig' was correct, then this form already existed in Proto-Hmong-Mien and did not imply any special relationship among Hmongic, Pa Hng, Ho Nte, and so on. Unfortunately, I had forgotten that Benedict himself now rejects an Austro-Tai etymology for *mpai C. Both Benedict (1985:8) and Wang (1985b:34) have now suggested that the word may be related to Archaic Chinese 豕 *pa^{GA} 'sow, pig' (GSR 39d).

If this etymology is correct, and if we accept the view of Chinese linguists that Hmong-Mien is genetically related to Chinese, then the original argument holds: *mpai C goes back to Proto-Sino-Tibetan, therefore, a fortiori, it goes back to Proto-Hmong-Mien, therefore it is a shared inheritance in the non-Mienic branches, not a shared innovation.

On the other hand, if we accept Benedict's view that the Hmong-Mien words which have Chinese cognates are loans,

then the loan *mpai C is a shared innovation uniting the non-Mienic branches.

In this connection I should mention that Benedict has also suggested a possible external source for Mienic *duj B 'pig'. He says that there are similar looking forms in the Karenic languages (Benedict 1985:8).

7. Conclusions

In some respects the new Wenjie material provided by Wang corroborates the testimony of the old material, particularly in the development of the finals. It even adds some new evidence, such as the C1 tone in 'thousand', shared by Wenjie and Xishanjie. In other respects the new material introduces complications which may require changing my conclusions:

- (1) The split between velar and uvular initials in Wenjie has no parallel in Xishanjie or Na-e.
- (2) The word for the numeral 'one' is quite different in Wenjie and Na-e: Wenjie [j^hu 32(8)], Na-e i.
- (3) Wenjie has a prepalatal initial in [t^hu 55(5)] 'six'. It does not share the distinctive velar initial of Na-e ku.
- (4) 'Mushroom' has a uvular initial in Wenjie versus a prepalatal initial in Xishanjie. If the variation between prepalatal and uvular/velar initials occurs even within Pa Hng, I may have been wrong to use it as an argument for separating Pa Hng from Hmongic.

The discrepancies between Na-e and Wenjie (e.g. 'six') or between Na-e and Xishanjie (e.g. 'fruit') have at least three possible explanations:

- (1) They are errors in Bonifacy. We should forget about Na-e until more accurately recorded material becomes available.
- (2) Na-e is a member of the Pa Hng group that has undergone some aberrant local developments that are of no significance for the overall classification and subgrouping of Hmong-Mien. For example, the velar initial of 'six' might have an explanation purely internal to Na-e, such as analogy or avoidance of homophony.
- (3) Benedict is right that the first split in the Hmong-Mien family was between Na-e and everything else. Subsequently, however, Na-e came into intensive contact with Pa Hng, borrowed many words from it, and shared some areal phonological developments with it.

In other words, I am suggesting as a third possibility that Na-e may not really be a member of the Pa Hng group

after all. If Benedict is right that Na-e is something very distinctive within the Hmong-Mien family, then the position of Na-e in Hmong-Mien might be analogous to the position of Saek in Tai.

Saek is a Tai language spoken in a small area in Thailand and Laos. It is generally considered to be a geographically displaced member of the Northern branch of Tai (the rest of the Northern branch is found in China and a few places along the northern border of Vietnam), but it has certain peculiarities that suggest that Saek may originally have been a wholly separate subdivision of Tai which later came into intensive contact with Northern Tai.

The difference is that for Saek we have thousands of words recorded with meticulous accuracy by William J. Gedney, whereas for Na-e we have 150 words recorded with highly dubious accuracy by Bonifacy. So we are back where we started: "in urgent need of more data on the language".

A Note on the Phonemic Status of [ə] in Biao Min Yao

David B. Solnit
 Cornell University

In Solnit 1985 I presented the rhyme system of Biao Min as:

a	ai	au	an	aŋ
ə	ai	au	an	
ɛ			(y)ɛn	- ɛŋ
i			in	
ɔ				ɔŋ
u	ui		un	uŋ

There is also a sequence [-iu] which is analyzed as medial plus simple vowel: /-yu/.

It is evident that the phonetic rhymes [ai au] could be phonemic /ɛi ɔu/. Mao, Meng & Zheng (MMZ) make this choice for [ai]/ɛi/, as shown by the cited forms (p.67-9) *mɛi*¹ 'vine', *hɔjwɛi*⁷ 'child', *beɪ*⁵ 'hot', and *mɛi*² 'you'. The only instance in MMZ of the corresponding rhyme with back vowel element is *lɔjau*⁴ 'successful, complete action' (Mandarin *le*). Finally MMZ have a sequence -wei in *ko*⁴ *dwei*³ 'tail', but this may be a misprint: I have *ko*⁴ *ɔwai*³ for this word (the choice of *o* vs. *ɔ* is unimportant).

Of the remaining rhymes with *a*, -an can be assimilated to the vowel *ɔ*: /ɔn/=[an]. The unrounding before -n has a parallel with *u* after labials; /un/→[un] / [+labial] _

Thus all except the simple -a can be assimilated to rhymes with *ɛ* and *ɔ* (while we are at it, we could also write these /e o/, there being no contrast *ɛ/e* or *ɔ/o*). Examining the distribution of -a brings out several interesting points.

1. the rhyme -a occurs after

- w, initial and medial
- l, medial only
- p
- m

and not after simple (non-cluster) dentals, prepalatals, or velars.

2. the rhyme -ɔ occurs with all simple initials *except* p, b and m. So we seem to have complementary distribution, with *ɔ* following aspirates, *a* following 'plain' stop and nasal, and neither occurring with voiced stop:

	p	ph	b	m	hm
a	pə	-	-	mə	-
ɔ	-	phɔ	-	-	hmɔ

3. However, both -a and -ɔ occur with medial -l-, e.g.

<i>klə</i> ¹ rainbow	<i>blə</i> ² chest
<i>klɔ</i> ¹ horn	<i>blɔ</i> ³ thin (not thick)

4. /ou/ (formerly -au) does not occur with medial /-/. This makes it possible to eliminate a as a phoneme altogether:

[pə] = /pə/
 /phə/ = [phə]
 [mə] = /mə/
 /hmə/ = [mə]
 /plo/ = [plə]

/plou/ = [plə]

In other words, phonetic [-a] after labials is really /-ə/, and after medial /l/ it is really /-ou/.

This reanalysis has some historical backing. BM [-a] generally corresponds to Yu Mien (YM) back rounded vowels followed by ɲ:

	BM	YM	
fly (n.)	mə ⁴	mɯŋ ⁴	
rain	bɬə ⁴	byɯŋ ⁶	
rainbow	kɬə ¹	cɯŋ ¹	
pig	twa ⁴	tɯŋ ⁴	
deaf	dwa ¹	dɯŋ ¹	
room	pə ²	pɯŋ ²	(房)
towards	hwa ⁵	huŋ ⁵	(向)
thick (liquids)	nwa ²	noŋ ²	(滾)
forehead	plə-	pyɯŋ ²	

('forehead' is the first syllable of a compound, *plə-tau²*, and so occurs only in the sandhi tone, phonetically equal to tone 7)

Comparative data also shows that the factor conditioning the allophone [ɔ] was originally not aspiration, but final stops, a less unusual feature:

	BM	YM	
tree	hmə ⁴	-	Ch. 木 muk
hail	phə ⁴	po ⁸	
six	kɬə ⁷	cɯ ⁷	
finger	də ⁷	dɯ ⁷	
few	tshə ⁴	tsu ⁸	

Thus the general development after labial initials and medial is *uŋ > ə, *uk > ɔ. The final stop also causes aspiration of *voiced initials, then disappears, leaving the aspiration as the apparent factor conditioning the complementary variant [ɔ]. Actually the aspiration and the vowel distribution are both effects of a vanished cause.

References

- Mao Zongwu, Meng Chaoji, & Zheng Zongze. 1982. *Yaozu yuyan jianzhi* [Brief account of the languages of the Yao people]. Beijing: Minzu Chubanshe.
- Solnit, David B. 1985. Introduction to the Biao Min Yao language. *Cahiers de Linguistique Asie Orientale* XIV #2, 175-191.

AN INVESTIGATION OF TWO ALVEOLAR STOP CONSONANTS
IN WHITE HMONG¹

Nerida Jarkey

1. INTRODUCTION

White Hmong has a fairly large inventory of stop consonant phonemes all of which occur in initial position only in the basically CV syllable structure. With only one exception, at each place of articulation where stops (and affricates) occur, the following series is found --- voiceless, voiceless aspirated, pre-nasalized and pre-nasalized with aspiration. The exception is the alveolar set, which has only two members, neither of which belong to any of these four categories. The purpose of this paper is to discuss the auditory, acoustic and articulatory properties of these two alveolar stops in White Hmong.

In the White Hmong orthography, devised in the early 1950's by missionaries working in Laos and Thailand, these stops are symbolized as 'd' and 'dh' (see Smalley (1976)). These are the symbols which will be used to refer to them throughout this paper.

Previous attempts have, of course, been made to characterize the auditory and articulatory qualities of these two stops, although to my knowledge no work has been done on their acoustic properties. In his phonetic transcription of the letters of the White Hmong alphabet, Mottin (1978) symbolizes 'd' and 'dh' simply as [d] and [dh] respectively, and in a brief explanation of each symbol he describes 'd' as "d comme dans doigt" ('d' as in doigt) and 'dh' as "d aspiré" (aspirated 'd'). Smalley (1976) refers to these two stops as "glottalized", but does not explain exactly what he means by this term, which has been used to refer to various types of sounds, including implosives, ejectives and creaky voice, as well as ordinary stops accompanied by glottal stops (Ladefoged 1971: 28). Smalley symbolizes 'd' as [ʔd] and 'dh' as [ʔth]. Heimbach (1979), in his "Guide to Pronunciation", is somewhat more explicit:

"d like the d in dream but with a glottal stop before it. That is, complete stop of the air flow in the throat before the beginning of the sound.

dh the Hmong sound 'd' with a brief puff of air after it".

My own auditory impressions of these two consonants do not entirely concur with this description. I cannot detect any glottal stop before the beginning of the sound (presumably before the oral closure) as Heimbach suggests. However, what can be detected, particularly in the case of 'dh' but also occasionally in the case of 'd', is the impression of a slight 'catch' in the voice, a momentary hesitation in the flow of the sound, just before the release of the stop, and it is this that I assume both Heimbach and Smalley are trying to capture with their references to "a glottal stop" and "glottalization".

Apart from this very slight sense of a 'catch' in the voice in some tokens, the consonant 'd' sounds simply like a fully voiced alveolar stop. 'dh', on the other hand, is not quite so easily categorized. There is no doubt that, in some cases, 'dh' sounds quite distinctly aspirated. In others, however, no aspiration can be detected. The part of the stop before the release is even more puzzling. Careful listening reveals that there is a fleeting period of whispery voicing, often extremely brief, and then the distinct 'catch' in the voice before the burst.

The experiment, described and discussed below, was designed with a view to investigating these phenomena. As not all scholars with an interest in Hmong language can be expected to have training in acoustic phonetics, I have endeavoured to make the discussion as clear and uncomplicated as possible, with this audience in mind. I apologize in advance to those already familiar with the terminology and concepts explained, and who would perhaps prefer more attention to be devoted to other issues.

2. PURPOSE OF THE INVESTIGATION

- to try to ascertain the acoustic characteristics of the alveolar stops 'd' and 'dh' in White Hmong.
- to try to draw inferences about the articulatory characteristics of these stops from the acoustic data.

3. EXPERIMENTAL PROCEDURE

3.1 The first task in the investigation was to select morphemes exhibiting the relevant stops followed by the three maximally differentiated vowels in White Hmong, namely [i, a, u]² ('i,a,u'), or the laxer versions of these vowels which occur in diphthongs.

Only morphemes exhibiting the low level tone (orthographically 's'), the mid level tone (orthographically unmarked) and the mid rising tone ('v') were chosen. There were two important reasons for this limitation on the selection. Firstly, there is no great variation in the pitch or voice quality at onset. It was therefore considered unlikely that they would influence the acoustic properties of the syllable initial consonants in different ways. The second reason is the extreme rarity of the consonant 'dh' in the vocabulary. Heimbach (1979) has only twelve entries for it in his Hmong-English dictionary, and it seems that this consonant simply does not co-occur with a certain number of the tones. The need to select minimal or near-minimal pairs for the study meant that the least common of the two consonants dictated the range of choice.

3.2 Five tokens of the English translations of each of the morphemes chosen were written on cards, first ensuring that the informant was able to confidently translate the words into Hmong. The words were written in English rather than Hmong in order to avoid any possibility of the spelling of the word influencing the informant's pronunciation.

3.3 The cards were then presented to the informant one by one in random order, and high quality recordings³ made of each token, uttered first in isolation and then in a frame. The frame used was: [haɪL_____dɔ̃aɪ] (hais_____dua) ('say_____again'). The cards were presented in random order to decrease the likelihood of list intonation being used.

3.4 The informant was asked to check the recordings for acceptability as normal Hmong utterances.

3.5.1 Wide band bar spectrograms were produced with both high shaping (to examine the general spectral characteristics of the utterances recorded, particularly in the higher frequency range) and flat shaping (to examine voice onset time and other low frequency features of glottal activity).

3.5.2 Wide band spectra⁴ (with flat shaping) were taken just after the release of the stops in order to examine the spectral properties of the bursts.

3.5.3 Average amplitude tracings were made to check if the consonants were responsible for any difference in amplitude at vowel onset.

3.5.4 High speed oscillograms⁵ of the wave forms were photographed with a polaroid camera, to check the fundamental frequency at vowel onset.

(Steps 3.5.2 - 3.5.4 are not discussed in detail below. Some of the inferences drawn from them simply serve to confirm those which could also be drawn from the bar spectrograms (3.5.1), while others are somewhat beyond the scope of the present discussion.)

3.5.5 Finally, the results were checked by following the same procedure (steps 3.3 - 3.5) with another informant, but recording only one token of each *morpheme*.

4. THE INFORMANTS

The choice of informants available for the experiment was extremely restricted, simply due to the scarcity of adult Hmong in the region (Canberra, Australia). However, although further acoustic analysis was not possible, it seems from auditory impressions that other speakers of White Hmong, now living both in other parts of Australia and in America and originally from different provinces, do pronounce the stops in question in basically the same way.

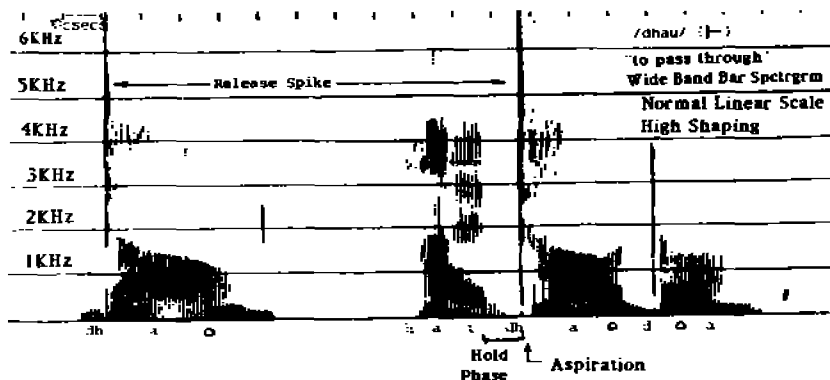
The main informant recorded for the experiment was Ntxawg Lis (known in English as Yeu Lee), a 24 year old male from Xieng Khouang province in Laos. The second informant was Yeu's mother Maiv Yaj (May Yang), also from Xieng Khouang.

5. EXAMINATION AND DISCUSSION OF BAR SPECTROGRAMS

5.1 High Shaped Wide Band Bar Spectrograms

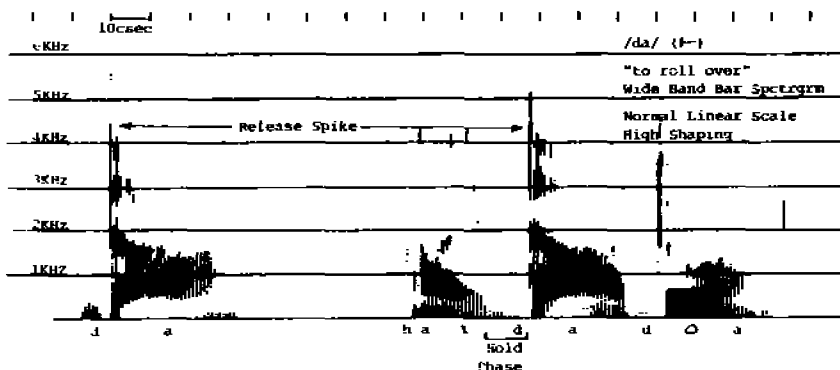
In order to observe the general spectrographic features of the utterances recorded we will begin by examining representative examples of the bar spectrograms with high shaping --- that is, with pre-emphasis designed to enhance the upper frequency range.

FIGURE ONE



High shaped wide band bar spectrogram illustrating the alveolar stop 'dh'.

FIGURE TWO

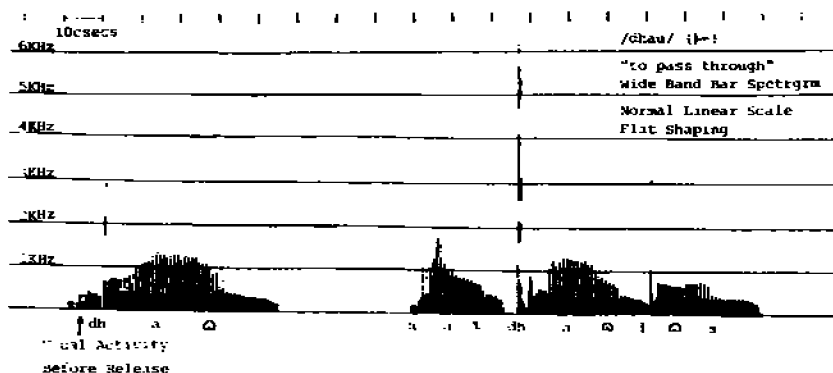


High shaped wide band bar spectrogram illustrating the alveolar stop 'd'.

5.2 Flat Shaped Wide Band Bar Spectrograms

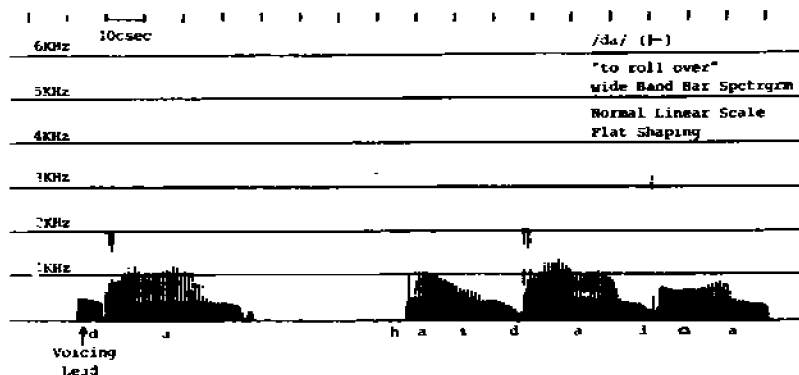
In order to examine voice onset time and other aspects of glottal activity in more detail, we will now turn to the bar spectrograms with flat shaping. Flat shaped spectrograms record a true impression of the relative amounts of energy throughout the frequency range, with no pre-emphasis of the upper frequencies. Thus they provide more detail in the lower frequency range, where vocal activity occurs, than do their high shaped counterparts.

FIGURE THREE



Flat shaped wide band bar spectrogram illustrating the alveolar stop 'dh'.

FIGURE FOUR



Flat shaped wide band bar spectrogram illustrating the alveolar stop 'd'.

Figure One shows two tokens of the word dhau (↳) 'to pass through'. As in all other examples below the left hand side of the spectrogram shows the word uttered in isolation, and the right hand side shows the word in the frame 'hais ___ dua' ('say ___ again'). Thus the word in question constitutes the central portion of the longer utterance on the right in each case. The similarity of this portion to the isolated example of the word on the left should be clear.

The horizontal axis of each spectrogram represents the time domain, the vertical strokes along the upper edge marking intervals of 10 centi-seconds. The fairly regular vertical striations seen particularly clearly below the 2KHz line indicate phonation --- each striation caused by one glottal pulse.

The vertical axis of the spectrograms represents the frequency domain, and shows the distribution of energy throughout the frequency range to 6000 Hz. The slightly darker bands running roughly horizontally throughout each word show the principal vowel formants. For example, the first four formants of the last part of the diphthong [aɪ] in the word hais [haɪL] in Figure One can be seen at approximately 650 Hz, 2000 Hz, 3000 Hz and 4000 Hz.

Our particular interest is in the region just before and after the release of the consonantal stricture (e.g. of 'dh' in Figure One) --- that is, the brief 'burst' of energy throughout the frequency range, shown on the spectrograms by a dark vertical line. This line is sometimes referred to as the release 'spike'.

The hold phase of the stop can be seen directly to the left of the release spike, beginning about 10 csecs before, from the abrupt cessation of the second and higher formants of the preceding vowel. Notice, however, that at the beginning of the hold phase some energy is still present in the very low frequency range, below 300 Hz. This shows that some phonation continues even after the closure for the stop has been made. In a voiceless or voiceless aspirated stop (like 't' or 'th' in Hmong) we would expect virtually no energy to be present during the hold phase.

If we now examine the period after the release of 'dh' in the frame in Figure One, we can see that the regular striations indicative of voicing do not begin again until after a lapse of approximately 3 csecs. This lapse, accompanied by the random stippling evident in the higher frequencies, can be attributed to a short period of aspiration after release, before the onset of the vowel, and is known as "Voicing Lag". Notice that there is almost no voicing lag in the accompanying example of the same morpheme uttered in isolation.

Turning to the representative example of a morpheme beginning with 'd' in Figure Two (da (↳) 'to roll over'), we can see that there is no voicing lag at all, the voicing resuming immediately after the burst. In fact, phonation continues virtually unbroken throughout the articulation of the stop, and can be clearly observed before the burst. This activity before the release of a stop is known as "Voicing Lead". (The decreasing amplitude and small break in the glottal pulse striations directly before the release of 'd' will be explained below (see 6.1).)

5.2.1 Voice Onset Time

Following Lisker and Abramson (1964), the Voice Onset Time (V.O.T.) of an occlusive is defined as the time between the release of the supra-glottal stricture and the onset of voice. Voicing lag, characteristic of aspirated stops, is seen on spectrograms as a period between the release spike and the (re-)commencement of the regular glottal pulse striations. Voicing lead, associated with voiced stops, is seen as a period before the release spike where the only energy present is in the lowest frequency range, where glottal activity is recorded.⁶

For each token of each morpheme recorded, both in isolation and in the frame, this period was measured from the flat shaped spectrograms, and the time it represented was calculated in centiseconds. (Voicing lag is recorded as a positive value, and voicing lead as a negative value.) An average was then taken of the voice onset time of the five tokens of each morpheme, both in the frame and in isolation, giving a mean V.O.T. value for each morpheme in both environments. These values, along with the standard deviation (S) from the mean in each case, are recorded in Table One.

TABLE ONE

Morpheme Recorded	Average V.O.T. in Isolation	S	Average V.O.T. in Frame	S
di	-9.0	1.6	-11.6	0.6
duas	-6.4	0.9	-10.4	0.8
dua	-6.1	0.8	-10.0	0.7
da	-7.3	0.7	-10.4	1.0
dhia	+4.6	1.9	+ 6.9	1.4
dhuas	+1.8	0.5	+ 3.3	1.0
dhuav	+3.7	1.8	+ 6.2	2.3
dhau	+2.0	2.2	+ 3.6	1.5

Table One. Mean Voice Onset Time values for five tokens of each morpheme, in isolation and in the frame *hais_____dua [hail_____dóat]* ('say_____again'). Values are in centiseconds, rounded off to one decimal place, reflecting the degree of accuracy of the measurements.

Returning briefly to the representative example of the consonant 'd' in the flat shaped spectrogram above, we can clearly see the voicing lead, confirming the auditory impression of 'd' as a voiced stop. The mean values for the voicing lead of 'd', shown in Table One, vary predictably between -11.6 and -6.1 csecs, depending on the vocalic environment and on whether the morpheme occurred in isolation or in the frame. The reasons for this variation will not be discussed here, but explanations can be found elsewhere in the literature.

Examining the flat shaped spectrogram illustrating 'dh' above, we can see a clear voicing lag when the morpheme occurs in the frame, but in isolation almost no lag is evident. In other spectrograms however, this consonant exhibits voicing lag sometimes in both environments and sometimes in neither. The range recorded for the mean values of voice onset time for 'dh', shown in Table One, was from +1.8 to +6.9 csecs. The overall mean V.O.T., disregarding environment, was +4 csecs with a standard deviation of 1.7 csecs.

To understand the significance of these figures it is helpful to compare the voicing lag of 'dh' with that of two other stops in Hmong, which have a similar place of articulation. The data for these stops were recorded using an identical experimental procedure. The voiceless aspirated dental stop 'th' [t^h] in Hmong, a stop which is invariably perceived as aspirated, was found to have an overall mean V.O.T. value of 6.2 csecs with a standard deviation of 1.4 csecs. The overall mean V.O.T. of the voiceless unaspirated counterpart of this stop, 't' [t], was 1.2 csecs with a standard deviation of 0.4 csecs. These figures are summarized in Table Two.

TABLE TWO

Stop	Overall Mean V.O.T. Value	Standard Deviation
d	-8.9	1.9
t	+1.2	0.4
dh	+4.0	1.7
th	+6.2	1.4

Table Two. Overall Mean Voice Onset Time Values for four dent-alveolar consonants. Values are in centiseconds, rounded off to one decimal place, reflecting the degree of accuracy of the measurements.

It should be clear from the information recorded in Table Two that the V.O.T. lag of the alveolar stop 'dh' falls roughly between that of the voiceless unaspirated and voiceless aspirated dental stops in Hmong. This concurs with the auditory impression, mentioned in the introduction, that some tokens of 'dh' seem aspirated while others do not. The t-Test indicates that the mean V.O.T. value is significantly different at the 5% level from that of both 't' and 'th'. As there are no obvious environmental factors which seem to determine the length of V.O.T. for 'dh', it is tentatively concluded that a relatively short period of aspiration is optional for this consonant.

5.2.2 Glottal Activity During the Hold Phase

Although the consonant 'dh' was observed to have an optional voicing lag, it is also clear from the spectrograms (Figures One and Three) that some glottal activity also takes place during the hold phase of this stop. Comparing this with the activity observed during the hold phase of 'd' (Figure Four), two important differences can be noted.

The first obvious difference is in the nature of this glottal activity. Whereas the striations recorded during the hold phase of 'd' are clear and regular, those of 'dh' are messy, uneven and irregular. This is particularly obvious in the examples of the morphemes uttered in isolation, where the vocal cords have not already been vibrating in a regular pattern before the closure of the stop. Secondly, while the voicing during the hold phase of 'd' is maintained almost to the point of release, that of 'dh' tends to die out much more quickly, often well before the burst. The significance of both of these aspects of the glottal activity during the hold phases of the two consonants will be discussed in the following section.

6. ARTICULATORY INFERENCES AND EXPLANATIONS

(Why a 'catch' in the voice?)

6.1 'd'

From the evidence of the spectrograms, discussed above, it is concluded that the consonant 'd' is simply a fully voiced alveolar stop. The auditory impression of a slight 'catch' in the voice in some tokens of this stop can be attributed to a very brief cessation of voicing before the release of the oral stricture. This was observed in the spectrograms as a small break in the glottal pulse striations directly before the release spike. Only in those tokens of 'd' where this break is long enough to be audible will a 'catch' in the voice be perceived.

The brief cessation of voicing described can be attributed to an interruption to the air flow caused by an equalization of air pressure above and below the larynx. This will occur sooner or later in any voiced stop if the oral stricture is not released within a certain length of time, and can be explained as follows. The process of voicing can begin when the vocal cords are sufficiently tense and slightly parted. The air, flowing from the sub-laryngeal vocal tract (an area of higher pressure) to the supra-laryngeal vocal tract (an area of lower pressure) is accelerated as it passes through the narrow opening between the two vocal cords (or 'folds'). This acceleration causes a slight reduction in pressure along the edges of the folds, which are consequently drawn together (the Bernoulli effect).

However, the build-up of pressure beneath the glottis, in conjunction with the natural elasticity of the vocal folds themselves, causes them to spring apart again a fraction of a second later, allowing the process to begin once more. This process can often go on repeating itself, as long as the speaker's breath lasts, each repetition corresponding to one of the glottal pulses we observed in the spectrograms. However, if it occurs in conjunction with an oral stop, it will not be long before the air pressure begins to build up behind this closure. The flow of air can be made to continue for a short while longer by strategies such as larynx lowering, which increases the volume and hence decreases the pressure in the supra-laryngeal vocal tract. Even so, unless the oral stricture is released it will not be long before the pressure above and below the larynx equalizes totally, the air flow ceases and voicing dies out. The hold phase of the stop may then continue long enough for this cessation of voicing to be perceived as a brief 'catch' in the voice, before the release of the stop allows the air flow, and hence the voicing, to begin again.

6.2 'dh'

The information obtained from the spectrograms concerning the acoustic properties of 'dh' is more difficult to interpret than in the case of 'd'. However, the data do seem to support the auditory impressions that 'dh' is a whispery voiced alveolar stop, with optional aspiration. The feature of optional aspiration has already been discussed at some length (5.2.1). However, the suggestion of a whispery voice quality for this consonant will require further explanation.

It will be helpful to begin by comparing the articulatory gestures involved in modal (normal) voice with those used to produce whispery voice. It should first be noted that scholars disagree somewhat both about the terminology that should be used (some using the terms "breathy voice" or "murmur" for what is here referred to as "whispery voice") and about the exact features relevant in the production of this phonation type. Such differences would not effect the general conclusions.

The vocal cords consist of two parts -- the ligamental vocal cords towards the front of the larynx, and the cartilaginous vocal cords (the arytenoid cartilages) towards the back. In modal voice both sections of the vocal cords vibrate as one unit. However, whispery voice is produced, according to one hypothesis, by drawing the arytenoid cartilages apart, so that vibration occurs only along the ligamental vocal cords. The irregular tension of the vocal cords results in irregular vibration and the gap between the arytenoid cartilages allows direct current flow through the larynx, unlike the regular alternating current flow of modal voice.

The suggestion that whispery voice is involved in the production of the consonant 'dh' can account for the two main acoustic differences noted on the spectrograms between 'd' and 'dh'. Firstly, it accounts for the 'messiness' or irregularity of the vocal activity evident during the hold phase of 'dh', unlike the neat, regular striations of 'd'. Secondly, it accounts for the far more rapid cessation of vocal activity during the hold phase of 'dh' than of 'd'. Recall that the cessation of voicing before the release of 'd' was attributed to the equalization of air pressure above and below the larynx. Naturally this equalization will take place much more quickly in the case of 'dh', due to the direct current flow between the cartilaginous vocal cords.

The reader may wonder how this whispery voiced stop in Hmong differs from the so-called "voiced aspirates" in languages like Hindi and Marathi. These stops are often referred to as "breathy" or "murmured" but, as mentioned above, these terms are sometimes used to describe what is here termed whispery voice. Lisker and Abramson (1963) give the following description of their auditory and acoustic characteristics:

"Auditory impressions suggest that the voiced aspirates are released with breathy voice or murmur. These impressions are supported by the spectrograms in which, upon release of the stop, the voicing is seen to take on a special character. There is a period of glottal periodicity, sometimes intermittent, mingled with random noise in the formant regions, all at relatively low amplitude." (Lisker and Abramson 1963:418 - 419)

Unfortunately it is not possible to say to what extent the "sometimes intermittent" glottal periodicity Lisker and Abramson describe differs from the vocal activity observed for 'dh' in Hmong. However, whereas it is during the hold phase of the stop that whispery voice apparently occurs in Hmong 'dh', it is after the release that it is evident in Hindi. Furthermore, while the effect of this phonation type is very brief in Hmong, so that 'dh' occurs with modally voiced vowels, the effect of the Hindi "breathy" consonants may last well into the following vowel.

7. COMMENTS ON THE LIMITATIONS OF THE STUDY AND SUGGESTIONS FOR FURTHER INVESTIGATION

The approach used in this study has been to examine information about the acoustic characteristics of the stops in question, principally that encoded on wide band bar spectrograms, and to draw inferences from this data as to the physical processes involved in their articulation. While this approach has enabled us to make several important statements, with a reasonable degree of confidence, about the articulation of 'd' and 'dh' in White Hmong, there are some significant limitations on this study which should be noted. These observations naturally also serve to point to several avenues for further research.

The first and most obvious of these limitations is the small number of informants. The acoustic data used in the analysis were based on material recorded from only one speaker, and the 'normality' of these measurements was checked with material from only one other speaker, closely related to the first. As noted above, however, fairly extensive checks were made to ensure that the auditory impressions of the relevant consonants in these two idiolects matched those of other speakers. It is relevant to note, parenthetically, that although all the speakers thus observed were presumably aiming at approximately the same auditory target, and achieving a similar output, it is not necessarily the case that we can therefore assume invariance at the production level. Furthermore, as Ladefoged and

Antoñanzas-Barroso (1985) clearly demonstrate, distinctions involving voice quality may be relative rather than absolute --- what is distinctly "breathy" for one speaker may fall within the range of modal voice for another, who normally has a slightly "breathy" articulation. A large number of informants would be needed in order to investigate phenomena such as these.

Another important limitation on the present study is the fact that some of the central conclusions, notably concerning the random noise element in the hold phase of 'dh', depend on visual impressions only of the spectrographic data. Ideally these impressions should be supported by accurately quantifiable forms of measurement, along the lines suggested for the "breathy"-modal distinction by Ladefoged and Antoñanzas-Barroso (1985).

Finally, it is extremely important to realize that the tentative conclusions drawn concerning the articulatory gestures involved in the production of the two consonants are inferences only. In order to ascertain the exact configuration of the vocal cords during the production of these or any other sounds, it would be necessary to undertake direct physiological observation. For example, fiberoptic study would be appropriate in this case.

8. CONCLUSION

The auditory impression of a slight 'catch' in the voice during the articulation of the alveolar consonants 'd' and 'dh' in White Hmong has previously been misinterpreted as a "glottal stop" or "glottalization". The experiment described above has shown that it can actually be correlated with a brief cessation of vocal cord activity immediately before the release of these stops, which was attributed to an equalization of air pressure above and below the larynx.

The auditory impression noted is more marked in the case of 'dh', because the vocal activity during the hold phase of this stop dies out more quickly than in the case of 'd'. It has been suggested that this extremely rapid cessation can be attributed to the whispery voiced quality of this vocal activity. The whispery voiced phonation type may involve a fairly large opening at the back of the larynx, allowing air to escape more rapidly than in the case of the modal (normal) phonation type used for 'd'.

On the basis of this study it is recommended that the White Hmong orthographic 'd' be described as a fully voiced alveolar stop (symbolized phonetically as [d]) and orthographic 'dh' as a whispery voiced alveolar stop with optional aspiration (symbolized phonetically as [d^h]).

NOTES

1. An earlier version of this paper was presented at the Southeast Asia Summer Studies Institute Conference (Hmong Linguistics Panel), Ann Arbor, Michigan, 2 August, 1985. My attendance at the Institute was made possible by a tuition fellowship from the University of Michigan and a grant from the Faculty of Arts, Australian National University.

The experimental work for this study was conducted at the Department of Linguistics, School of General Studies, Australian National University, under the guidance of Dr. Philip Rose. Sincere thanks are due to him for his invaluable help.

I am extremely grateful to Dr. Leo Papademetre, Dr. Philip Rose and Dr. Michael Walsh for reading and commenting on earlier drafts of this paper. I would also like to thank Ntxawg Lis and Maiv Yaj for acting as informants for the study, and Honorine Jarkey for typing the paper for me.

2. The symbol [a] is here used to refer to a low central vowel rather than a low front vowel, there being no more suitable symbol available in the International Phonetic Alphabet.
3. Recorder: Nagra 4.2; Microphone: Nakamichi CM300; Recording Speed: 7.5 ips; Recording level: 85db; Approximate distance of speakers mouth from microphone: 150cm.
4. 300 Hz $\frac{1}{2}$ power bandwidth.
5. Oscilloscope: Tektronix 5441 Storage Oscilloscope; Amplifier: 5A18N Dual Trace Amplifier; Time Base: 5B12M Time Base.
6. It is suggested in this paper (see 5.5.2) that some tokens of the consonant 'dh' exhibit both the feature of voicing lag and of voicing lead. However, it should be noted that this is not consistent with the conventional use of these terms, which regards voicing lag and lead to be mutually exclusive. Perhaps some new terminology, such as 'intermittent voice onset' is required for consonants such as 'dh', but for the purposes of this paper the conventional terminology is used, in a slightly unconventional way.

BIBLIOGRAPHY

- Catford, John C. 1977. *Fundamental Problems in Phonetics*. Edinburgh: Edinburgh University Press.
- Fant, Gunnar. 1960. *Acoustic Theory of Speech Production*. The Hague: Mouton.
- Fischer-Jørgensen, Eli. 1954. 'Acoustic Analysis of Stop Consonants' in: *Readings in Acoustic Phonetics*, ed. by Ilse Lehiste, Cambridge Mass.: MIT Press, 1967. Reprinted from: *Miscellanea Phonetica* Vol. II, 1954.
- Heimbach, Ernest E. 1979. *White Hmong-English Dictionary*, Rev. ed., Southeast Asia Program Data Paper No. 75, Ithaca: Cornell University Department of Asian Studies.
- Halle, Maurice, G. W. Hughes and J-P. A. Radley. 1957. 'Acoustic Properties of Stop Consonants' in: *Readings in Acoustic Phonetics*, ed. by Ilse Lehiste, Cambridge Mass.: MIT Press, 1967. Reprinted from: *The Journal of the Acoustic Society of America* Vol. 29, No. 1 (January, 1957).
- Ladefoged, Peter. 1971. *Preliminaries to Linguistic Phonetics*. Chicago: University of Chicago Press.
- Ladefoged, Peter and Norma Antoñanzas-Barroso. 1985. 'Computer Measures of Breathily Voice Quality' in: *UCLA Working Papers in Phonetics*, 1985, pp. 79-86.
- Laver, John. 1980. *The Phonetic Description of Voice Quality*. Cambridge: Cambridge University Press.
- Lisker, Leigh and Arthur S. Abramson. 1964. 'A Cross Linguistic Study of Voicing in Initial Stops: Acoustical Measurements' in: *Word* Vol. 20, No. 3 pp. 384-422.
- Mottin, Jean. 1978. *Éléments de Grammaire Hmong Blanc*. Bangkok, Thailand: Don Bosco Press.
- Smalley, William A. 1976. *Phonemes and Orthography: Language Planning in Ten Minority Languages in Thailand* Series C, No. 43. Canberra: Pacific Linguistics.
- Zue, Victor W. 1980. *Acoustic Characteristics of Stop Consonants: A Controlled Study*. Bloomington: Indiana University Linguistics Club.

TONE SANDHI COMPOUNDING IN WHITE HMONG

Martha Ratliff
The University of Chicago

1. Tone Sandhi and Compound Formation¹

The term "sandhi" refers to phonetic changes occurring in words which are caused by certain phonetic characteristics of contiguous words. Thus "tone sandhi" is the change of tone in one word caused by the tone of a neighboring word. According to Kenneth Pike (1948: 25), "regular tone sandhi" narrowly described is "forced meaningless substitutions of one toneme for another . . . in which one toneme is perturbed by another." Eugénie Henderson (1967: 174) refers to "tonal alternation and compounding" and differentiates "tonal alternation," which affects meaning, from "tone sandhi," which she takes in the narrow sense given above.

White Hmong tone sandhi does not conform to either of the above descriptions of tone sandhi or tonal alternation exactly. It is of Pike's "arbitrary type" (1948: 26), in which the proper phonetic environment is not sufficient to guarantee tone change. Grammatical category and the particular lexical items involved also play a role. The compounds which result from tone change are sometimes different in meaning from a phrase involving the two unaltered words, but usually they are not. I choose to call tonal alternation in White Hmong "tone sandhi" (1) to emphasize the syntagmatic nature of the alternation, and (2) to emphasize the historical connection between the White Hmong "system relic" and the "regular" [mechanical] tone sandhi system from which it came (see Section 3. below). For reasons concerning the semantic unity of the members of the resultant tonally defined pairs and the behavior of a few of them as syntactic units (see Section 2.5.), I choose to call these pairs "compounds" after Lyman², and indicate their compound status with a linking hyphen in the orthography. Inasmuch as tone sandhi serves to create new words in White Hmong, it constitutes one of the morphological functions of tone in the language.

2. A Synchronic Account of White Hmong Tone Sandhi

2.1. On the Nature of White Hmong Tone Sandhi

Many words with high falling, low level, low checked, mid rising, and mid level tones (those words marked with *-j*, *-s*, *-m*, *-v*, and *b*) respectively) have alternate tonal realizations when they enter into a particularly close relationship with a preceding word which has either a high level or high falling tone (*-b*, *-j*). Almost five hundred pairs of words which involve alternate tonal realizations, i.e., tone sandhi compounds, are listed in Appendix I of my dissertation (not

¹ This paper is simultaneously an expansion of a paper presented at the 1985 SEASSI Conference Panel on Hmong Linguistics and a reduction of Chapter II of my Ph.D. dissertation, "The Morphological Functions of Tone in White Hmong" (University of Chicago: 1986).

² Unpublished notes of Thomas A. Lyman, 1963, quoted in Heimbach, p. 454.

reproduced here in the interest of economy). A few examples here will serve to give an introduction to the nature of these tonally-defined compounds:

Noun-Verb Attribute (no change of style or meaning)	<i>dej stav</i> water cooked "boiled water" (H 298) ³	<i>dej-sia</i> "boiled water" (H 34)
Noun-Noun Attribute (with reported stylistic difference)	<i>hnoob tes</i> sun hand "sunray" (B <i>taw</i>)	<i>hnoob-teg</i> "sunray" (B <i>tes</i>)
Verb-Noun (with reported meaning difference)	<i>poob dej</i> fall water "to fall into water" (native speakers)	<i>poob-deg</i> "to drown" (B <i>dej</i> ; H 447)

There are both strict phonological (Section 2.2.) and syntactic (Section 2.3.) conditions on the occurrence of tone sandhi compounding; I have also observed that certain lexical items are more likely than others, given the same phonological and syntactic environments, to enter into such compounds (Section 2.4.). However, even after long familiarity with the compounds that do exist and the rules which limit their occurrence, it is impossible for me to predict which pairs of words must always compound, which (of those that meet the minimal criteria) must never compound, and which, as those examples cited above, can occur either way. Of those pairs of words which have been found both in a collocation of independent morphemes and in a tonally defined compound, some have concomitant meaning differences and some do not.

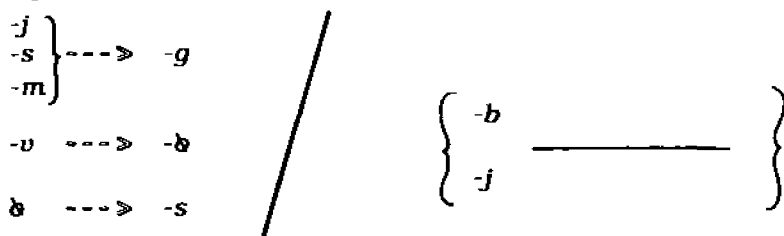
Apart from a number of frozen compounds which all sources and speakers seem to agree on (such as *teb-chaws* "country," from *teb* "land" and *chaw* "place"), the sporadic appearance of these compounds in the lexicon and the equal acceptability of both the compounded and the uncompounded forms in many cases indicates to me the truth of what Gordon Downer (1967) postulates: the tone sandhi system in White Hmong is an historical relic of a system that dates back no further, probably, than Proto-West-Hmongic (the Sichuan-Guizhou-Yunnan *fangyan* protolanguage), and which is in the process of dying out. The existence of a Proto-West-Hmongic sandhi system is inferred by Downer from a comparative study of five dialects of that branch including White Hmong, in which the system is eroded the furthest (see Section 3.1. for an expanded comparative analysis). Downer believes the sandhi system may someday disappear completely: it is possible for the original forms to replace the tonally derived forms in most cases, since speakers rarely cease to identify the base form and the sandhi form as the same "word" (1967: 596). Young speakers, as represented by the two young men I have had the greatest contact with, are beginning to use words with the base tone even in the syntactic collocation which most predictably gives rise to tonal compounding: the numeral-classifier collocation. They often meet questions about which tone is acceptable or better in certain collocations with a shrug of the shoulders, meaning that either is all right, and that neither will obscure communication.

³ "H" and "B" refer to the Heimbach and Bertrais White Hmong dictionaries.

For reasons that will be discussed fully in Section 3. below, I believe that compound formation by tone change is a new use being made of an old, once more regular and thoroughgoing, phonetically motivated tone sandhi system. The points of interest in a synchronic analysis of White Hmong tone sandhi are why it occurs when it does, what kinds of words are likely to be involved within the recognized phonological and syntactic constraints, and how these tonally defined compounds may differ in meaning and behavior from collocations made of the base forms of their component parts.

2.2. Phonological Conditions

White Hmong tone sandhi is of the progressive type: that is, the trigger word precedes the word that undergoes the tone change⁴. Although the changes were most likely due to neutralization under loss of stress in non-prepausal environments originally (see Section 3.3.), the tones that these changes gave rise to have become associated with certain tones in the basic inventory of lexical tones, so that the system in White Hmong today is one of paradigmatic replacement of one tone by another. The "neutralization effect" is still evident in the fact that five different tones collapse into a system of three. A representation of the specific changes which take place appears below:



It is important to remember that the above description of the phonological environment and changes captures the necessary phonological facts about tone sandhi, but that the proper phonological conditions alone are no longer sufficient criteria for change. As mentioned above, sandhi takes place optionally and, although its likelihood can be assessed, it takes place unpredictably. The focus is properly on its occurrence, therefore, rather than on its nonoccurrence. Nonetheless, there are three important exceptions to the scheme above, which are fairly easily explained.

2.2.1. Low Level (-s) Tone Exceptions

In an account of the White Hmong tone system from a diachronic perspective, we need to differentiate two low level (-s) tones. In White Hmong, the reflex of category B2 (the second of the three Proto-Hmong-Mien tones A, B, and C; the "2" indicates an ancient voiced initial consonant) merged with the reflex of category D1 (the D tone belongs to words that had final *-p or *-t in Proto-Hmong-Mien and arose later than the first three; the "1" indicates an ancient voiceless initial consonant). Only the words with low level tones which

⁴ Tone sandhi in the Micnic branch, on the other hand, is regressive. See Lu Yichang, [The Tonal Change in Morphology in Biaomin Dialect of Mien Language] MZYW 6 (December 1985), pp. 16-20, and Solnit (1985:175-91). Interestingly, tone sandhi in North Hmongic is also regressive (Ying 1962:73; Institute 1962.5).

can be assigned to tone category B2 on the basis of comparative evidence undergo tone sandhi. This indicates that the period when the tone sandhi process was productively generating new compounds predated the White Hmong tonal merger, since no low level (-s) tone word which can be assigned to tone category D1 undergoes tone sandhi. To illustrate the different behavior of the -s¹ tone (< B2) from the -s² tone (< D2), the following comparison between White Hmong and Green Hmong, which did not undergo the same tonal merger, is given below:

tone sandhi with

White Hmong -s¹ (< B2) corresponding to Green Hmong -g (< B2)

<i>tus</i> :	<i>ib-tug</i>	"1-animate- "	<i>tug</i>
<i>tais</i> :	<i>ib-taig</i>	"1-bowlful- "	<i>taig</i>
<i>tes</i> :	<i>pob-teg</i>	"round part-hand"	<i>teg</i>
		(=wrist bone)	

no tone sandhi with

White Hmong -s² (< D1) corresponding to Green Hmong -s (< D1)

<i>caws</i> :	<i>ib caws</i>	"1 leap- "	<i>caws</i>
<i>qais</i> :	<i>ib qais</i>	"1 skein- "	<i>qais</i>
<i>kws</i> :	<i>pob kws</i>	"corn"	<i>kws</i>

Similarly, tone sandhi does not affect the following D1 -s tone words:

<i>nqajj dais</i>	"meat bear" (=bear meat) (cf. <i>nqajj-npuas</i> "pork" and <i>nqajj-nyug</i> "beef") (H 450)
<i>paj kws</i>	"flower corn" (=popcorn) (cf. <i>paj-npleg</i> "puffed rice" (H 449)
<i>sib ntaus</i>	"(recip)-strike" (=to strike each other) (cf. <i>sib-nraus</i> "to butt each other") (H 185)
<i>sib zas</i>	"(recip)-face off" (=to face off against each other) (cf. <i>sib-zeg</i> "to tease each other") (H 450)

2.2.2. The Voiceless Aspirated Stop, Voiceless Fricative/Breathy (-g) Tone Cooccurrence Restriction

The voiceless aspirated stops and the voiceless fricatives constitute a natural class in White Hmong, in that they share a feature which might be called "heightened voiceless noise": impressionistically, "heightened air flow."⁵ Since aspiration is distinctive in White Hmong, the voiceless aspirated stops have a pronounced period of aspiration before the onset of voicing in the vowel. The voiceless fricatives *f*-, *h*-, *hl*-, *s*-, *x*-, *xy*-, are characterized, of course, by voiceless airflow as well as by friction. The phonological distinctions based on "heightened voiceless noise" would be difficult to maintain if the syllable in question were to be pronounced with the breathy (-g) tone, the main feature of which is also heightened voiceless noise.⁶ Therefore, no Hmong word which begins with one of these initials will carry the breathy (-g) tone.⁷ As a corollary

⁵ These consonants constitute a class in the Tai languages as well, and are referred to as the "high consonants." Certain dialects, including Siamese, underwent a tone split on the basis of whether or not the initial was of this class (Li Fang-Kuei, *A Handbook of Comparative Tai* [Hawaii: University Press, 1977], p.29).

⁶ This tone is a more "whispered" than "murmured" breathy tone.

⁷ Bertrais, introduction: "Les mots qui ont une aspiration ne portent jamais le ton 'nceg'." Motin, p. 16. "Notons qu'aucun mot aspiré ne se prononce sur ce ton [-g]."

to this fact, no word beginning with either a voiceless aspirated stop or a voiceless fricative will undergo these sandhi changes:

$$\begin{array}{l} -j \\ -s \text{ ---} \rightarrow -g \\ -m \end{array} \quad / \quad \left\{ \begin{array}{l} -b \\ \text{---} \\ -j \end{array} \right\}$$

Examples of this cooccurrence restriction appear below:

<i>ib phaj mov</i>	"1 plate of rice" (H 448)	(cf. <i>ib-roog mov</i> ,
1 plate rice		"1 tableful of rice")
<i>ib phom</i>	"1 shot" (H 244)	(cf. <i>ib-nplawg</i> ,
1 gun/shot		"1 blow")

Similarly:

<i>ib chim</i>	"a little while" (H 24)
<i>ib faj</i>	"1 of a pair" (H 44)
<i>ib feem</i>	"1 part; 1 portion" (H 44)
<i>ib fij</i>	"a shift of work" (Vwj 1983: 44-45)
<i>ib hom</i>	"1 kind" (H 54)
<i>ib hwm</i>	"an ensemble of pieces" (B <i>hwm</i>)
<i>ib phaum mob</i>	"1 sickness" (H 450)
<i>ib phiaj</i>	"1 row; 1 set" (H 243)
<i>ib sas</i>	"1 spurt" (H 287)
<i>ib sij</i>	"continually; repeatedly" (H 293)
<i>ib sim neej</i>	"1 lifetime" (H 294)
<i>ib suam nag</i>	"1 shower" (B <i>suam</i>)
<i>ib thaj</i>	"1 stretch of land" (B <i>thaj</i>)
<i>ib thaj neeb</i>	"1 session of spirit worship" (B <i>thaj</i>)
<i>ib thooj</i>	"1 lobe; 1 clump; 1 packet" (B <i>thooj</i>)
<i>ib tshaj</i>	"1 case at law" (H 448)
<i>ib tshooj</i>	"1 level; 1 story" (H 369)
<i>ib txhij</i>	"at the same time" (H 392)
<i>sib fim</i>	"to be acquainted with each other" (H 45)
<i>sib foom</i>	"to set a curse on each other" (H 46)

Although the cooccurrence restriction between the native initials *hl-*, *s-*, *x-*, and *xy-* and the breathy (-*g*) tone seems to be a restriction of a solely phonological nature, the restriction on the cooccurrence of many *f-* and *h-* words and the breathy tone is apparently due both to phonology and to the fact that words with these initials are often loans from Chinese. Recent loans came into the language when the tone sandhi process was no longer productive, and are hence less likely to undergo change (see below).

2.2.3. The Non-Involvement of More Recent Chinese Loans

Contact between the Chinese and the Hmong has existed over centuries and the traffic in loanwords has been extensive. This is a matter of general knowledge, shared both by those who claim a genetic relationship between the two families and those who do not. The difficulty in analyzing these loans lies in identifying the dialect(s) of Chinese from which the loans came, and the times at which they entered the language. The Southwest variety of Mandarin as spoken in Yunnan province is the form of Chinese used as a probable source of recent

loans in Lyman's Green Hmong dictionary. Studies of Chinese loans in Hmong-Mien languages include Institute 1962 (loans in Mien, Bunu, and Lakkja) Ying 1962 (loans in several Hmongic dialects), Downer 1973 (strata of loanwords in Mien), and Benedict 1985 (early Sino-Tibetan/Hmong-Mien loan relationships).

Ying (1962:74) and Ruy and Kuan (1962:525) both mention that Chinese loans do not undergo tone change in XIANJIN and XUYONG, respectively. The same is true in White Hmong. Such loanword exceptions in the most exceptionless environment for tone sandhi, the collocation numeral-classifier, are listed below (the Southwest Mandarin forms are from Lyman's dictionary):

<i>lum</i> (< SWMan <i>er</i> "two")	"another; times, occasions"
- <i>nyeem ib lum</i>	"read one time" (H 123)
- <i>daw hlau yuj ob peb lum</i>	"the plane makes 2-3 turns" (B <i>yuj</i>)
hawk iron hover 2 3	
<i>tiam</i>	"generation"
- <i>ib tiam neej</i>	"1 generation" (Downer 1967: 594;
1 gen. person	B <i>tiam</i>)
<i>vam</i> (< SWMan <i>wán</i> "10,000")	"tens of thousands" ⁸
- <i>ib vam</i> ; <i>ob vam</i>	"10,000"; "20,000" (H 398)
<i>yam</i> (<SWMan <i>yáng</i> "thing")	"kind, sort, type"
- <i>ib yam</i>	"1 thing; the same thing" (H C19)

The loanword exception also explains the absence of sandhi forms in compounds such as the following:

<i>cuw mem</i> (< SWMan <i>mě</i> "ink")	"pencil; pen" (H 21)
pointed ink stick	
<i>sib cam</i>	
(recip) argue	"argue with each other" (H 7)

Although I will not attempt to read too much into the fact that the six examples cited above bear the checked (-*m*) tone, I think that it suggests that a proper study of "loan-tones" in Hmong would be fruitful.

2.3. Syntactic Conditions

The following word class conditions on tone sandhi were recognized by Lyman (of Green Hmong) in 1963, and were reproduced by Heimbach in his dictionary (p. 454):

TONE SANDHI

At the present stage of research, it would seem that sandhi changes occur when members of certain word classes are joined syntactically, in what may be called "compounds." The following have so far been recorded:

1. Numerals joined to classifiers: *ob-leeg* 'two persons'

⁸ As David Strecker points out (p.c.), the word for "10,000" has tone C2 in seven of the nine dialects represented in Wang 1979, and tone D2 only in White Hmong, Green Hmong, and SHIMEN. It may be that it was an early loan in the majority of the dialects, and a later loan in the latter three.

2. Numerals joined to numerals: *ib-puas* 'one hundred'
3. Nouns joined to modifying words (these latter being either Nouns or Verbs): *teb-npleg* 'rice field'
4. Autotelic Verbs joined to Nouns (as locative objects):
poob-deg 'fall in the water'

(The label "Noun" is here taken to include localizers, a subgroup of nouns which correspond to English prepositions, e.g., *qab* 'bottom' or 'under'. The label "Verb" is here taken to include words corresponding to English adjectives, e.g., *liab* 'to be red'. An "Autotelic Verb" is a verb which may or may not take a noun object. In the latter case, the object modifies the verb.)

Heimbach's only objection to the above account of Hmong tone sandhi is that it does not explain the countless times tone sandhi fails to occur in such syntactically defined compounds. In his introduction to the above quotation, Heimbach states that the two factors of word class and juncture are important in determining whether a change of tone is "required" or not. I would go further than Heimbach: to assert that word class and juncture are sufficient to insure that tone sandhi will take place ignores the frozen collocations without tone change, certainly compounds in terms of meaning and behavior, such as *pojniam* "woman; wife" (literally "woman-woman"). I think it is better to say that tone change indicates close juncture rather than that close juncture is a prerequisite of tone change. (See Section 3. on why it is impossible to fully specify the conditions under which tone sandhi will occur.)

Gordon Downer (1967: 592-93) lists the following:

. . . constructions in which the modified tones are commonly found. As the grammatical analysis of WM [White Miao] is far from complete, all grammatical labels must be taken as purely provisional.

- (1) Num. + Quant. (when the Num is *i*¹ 'one'
*ɔ*¹ 'two,' *pe*¹ 'three,' *plɔ*¹ 'four,' *tʃi*¹
'five,' *kyua*² 'nine,' and *ti*¹ 'only one')
- (2) Nominals (the first syllable is Noun, the second may be Noun, Verb, or Adj.; both subordinative and coordinative constructions occur)
- (3) Quant. + Noun (only three examples found)
- (4) VPref. + Verb (the only VPref. is *fi*¹ or
*fi*⁴ 'each other')
- (5) Verb + Noun

Downer recognized two important constructions which Lyman omitted: the reciprocal-verb construction, and the rare, though very interesting, classifier (Downer uses the term "quantifier")-noun construction (see Sec. 4.2.).

Mottin (p. 19) recognized roughly the same constructions as Downer:

Cela [changement de ton] semble être la plupart du temps le cas quand les mots forment une entité, un syntagme:

- avec un numéral: *ib-qho* = un lieu, une chose
- avec un verbe ou nom pris comme adjectif:
dej sov = l'eau est chaude

dej-so = de l'eau chaude
 kauf-ntseg = des pendants d'oreille

-avec un verbe normalement intransitif, mais qui peut être
 directement suivi d'un nom avec lequel il fait comme expression:
 poob-deg = tomber (dans) l'eau

-avec des mots comme "sib":

sib-tog = se mordre l'un l'autre⁹

2.3.1. Constructions Involving Numerals

As Heimbach observed (p. 446), tone sandhi is "particularly noticeable in words preceded by one of the first five numerals."¹⁰ A large number of examples involving each one of the possible changes is easy to find:

<i>ib-rag</i>	(< <i>raj</i>)	<i>dej</i>	"1 tubeful of water"
<i>ib-pluag</i>	(< <i>pluas</i>)	<i>mov</i>	"1 meal (of rice)"
<i>ob-daig</i>	(< <i>daim</i>)	<i>ntawv</i>	"2 sheets of paper"
<i>peb-kw</i>	(< <i>kwv</i>)	<i>taws</i>	"3 shoulderloads of firewood"
<i>ib-los</i>	(< <i>lo</i>)	<i>lus</i>	"1 mouthful of language" (=1 word)

The above collocations are all numeral-classifier.

The collocation numeral-numeral also gives rise to tone sandhi when the second numeral is *caum* ("10's" - 30 and above) or *pua* ("100") and the first numeral is 1-5 or 9 (with the interesting exception *cuaj caum* "90"¹¹):

<i>plaub-caug</i>						"40"
4	10					
<i>ib txhiab cuaj-puas yim caum rau</i>						"1986"
1	1,000	9	100	8	10	6

Notice that *vam* "10,000," perhaps in White Hmong a more recent Chinese loan, does not undergo tone sandhi, as mentioned above (Section 2.2.3.):

<i>ib vam</i>	"10,000"
<i>ob vam</i>	"20,000"

A third category of words is affected by the numeral *ib* "one": nouns used as measure words, or nouns the entirety of which is being expressed:

<i>ib-hmos</i>	(< <i>hmo</i>)	"1 night" (H 452)
<i>ib-vog</i>	(< <i>vof</i>)	"1 circle" (H 401)
<i>ib-ce</i>	(< <i>cev</i>)	"the entire body" (B <i>cev</i>)
<i>ib-zog</i>	(< <i>zos</i>)	"the whole village" (H 449)

Finally, there are two words which are "numeral-like" in meaning, bear the right tones, and can effect tone sandhi: *tib* "single, sole; 1 blow of . . ." and *thawj* "first; head":

<i>tib-leeg</i>	(< <i>leej</i>)	"sole person" (B <i>leej</i>)
<i>tib-qho</i>	(< <i>qhov</i>)	"sole thing" (H 315)
<i>tib-tug</i>	(< <i>tus</i>)	"sole person" (B <i>tib</i> ; H 315, 325)
<i>tib-plhaws</i>	(< <i>plhaw</i>)	"1 jump" (B <i>plhaw</i>)

⁹ I have added hyphens to the quotation to indicate when tone sandhi has taken place, since the base forms are not given.

¹⁰ In at least one instance, the numerals themselves, in ordinal postposition, are changed: in the compounds referring to the days of the waxing of the moon, *xiab-raus* "the sixth day of the moon's increase," *xiab-xyas* "the seventh . . .," *xiab-yig* "the eighth . . .," *xiab-cuag* "the ninth . . .," *xiab-kaug* "the tenth . . ."

¹¹ Heimbach, p. 457; Downer 1967: 592; Vwj 1983: 37; Xab Xyooj (p.c.)

<i>tib-riag</i>	(< <i>riam</i>)	"1 stroke of the knife" (H 450)
<i>tib-teg</i>	(< <i>tes</i>)	"1 blow of the hand" (H 316)
<i>thawj-caw</i>	(< <i>cawv</i>)	"'head' of the whiskey" (B <i>cawv</i>)
<i>thawj-zaug</i>	(< <i>zaum</i>)	"first time" (B <i>thawj</i> ; H 335)

When I first started looking at tone sandhi in White Hmong, it seemed as though these numeral collocations came fairly close to constituting an exceptionless environment for tone change, probably because the most common classifiers (for example, *tus* "animate, long and slender," *leej* "human," *nkawm* "pair," *daim* "flat object") seem to change exceptionlessly. On a careful reading of both Bertrai's and Helmbach's dictionaries, though, I discovered a great number of exceptions in this syntactic collocation, too, bringing it more in line with the other syntactic environments in which tone sandhi compounding may occur. A list of these exceptions follows:

<i>ib cim</i>	"1 turn of work, 1 season" (H 13, 90; B <i>cim</i>)
<i>ib com</i>	"1 measure of 6 kilos" (B <i>com</i>)
<i>ob cum</i>	"2 sides, groups, clans" (H 17)
<i>ib cuam teeb</i>	"opium smoking utensils" (B <i>cuam</i>)
<i>ib cuam nqajj</i>	"1 'stretching' of meat" (B <i>cuam</i>)
<i>ib cuam nplooj</i>	"1 length of roofing leaves" (H 20)
<i>ib cuam</i>	"1 time" (B <i>cuam, rrauj</i>)
<i>ib cham neeg</i>	"1 group of people" (B <i>cuam</i>)
<i>ib chaw</i>	"1 place" (B <i>chaw</i>)
<i>ib daw hlau</i>	"1 plane/ful (of goods)" (B <i>thawj</i>)
<i>ib kauj hlau</i>	"1 coil of rope" (H 77, 448)
<i>ib kuam</i>	"1 hand (of bananas)" (H 90; B <i>kuam, thijj</i>)
<i>ib kheev</i>	"1 bundle" (B <i>kheev</i>)
<i>ib leej</i>	"1 line (of houses)" (H 110, 448; B <i>leej</i>)
<i>ib lo lus</i>	"1 mouthful of language" (B <i>lo</i>)
<i>ib naj zia</i>	"1 gummy stick" (B <i>naj</i>)
<i>ib ncauj hntav</i>	"1 mouthful of teeth" (H 145, 448)
<i>ib ncauv plas</i>	"1 slap" (H 149)
<i>ib ntsawv</i>	"1 clump" (B <i>ntsawv</i>)
<i>ib ntsis</i>	"1 moment" (H 201)
<i>ib ntsuj teb</i>	"1 portion of field" (H 204)
<i>ib ntshua</i>	"1 cluster" (B <i>ntshua</i>)
<i>ib ntxiuj</i>	"1 slab" (B <i>ntxiuj</i>)
<i>ib nbawj tsua</i>	"1 out-jutting of rock" (H 448; B <i>nixawj</i>)
<i>ib nyuam qhuav</i>	"1 moment" (H 222, 272)
<i>ib pam quav</i>	"1 passage of defecation" (H 224)
<i>ib phau ntawv</i>	"1 book" (B <i>phau</i>)
<i>ib phev teb</i>	"1 slope of field" (B <i>phev</i>)
<i>ib plhaw</i>	"1 jump" (H 254)
<i>ib qhov tshtab</i>	"1 new thing" (B <i>qhov</i>)
<i>ib qhoo</i>	"1 bundle" (B <i>qhoo</i>)
<i>ib rauj rau</i>	"1 hammer blow on" (B <i>rauj</i>)
<i>ib rawv</i>	"1 cluster (of fruit)" (B <i>reuj</i>)
<i>ib raav</i>	"1 square (of embroidery)" (B <i>rauv</i>)
<i>ib rooj</i>	"1 tableful" (B <i>rooj</i>)
<i>ib ruam</i>	"1 step" (B <i>ruam</i>)
<i>ib tauw nroj</i>	"1 clump of weeds" (B <i>tauw</i>)
<i>ib tauwm tsheb</i>	"1 train of cars" (H 311)
<i>ib tom</i>	"1 time" (H 322)
<i>ib twv huab</i>	"1 cloud" (B <i>twv</i>)
<i>ib tsam</i>	"1 period of several hours" (B <i>tsam</i>)
<i>ib tsum mov</i>	"1 tableful of food" (Yaj. forthcoming)
<i>ib txwm</i>	"a long time" (H 386)

<i>ib txwm ntawv</i>	"1 letter of the alphabet" (H 386)
<i>ib txhia</i>	"some; a portion" (H 393)
<i>ib voj-teg</i>	"a circle made with the 2 hands" (B tes)
<i>ib yoj thee</i>	"1 ovenful of charcoal" (B yoj)

Many of these unmodified words have been found with the expected sandhi forms. The places where the exceptional unmodified forms were found is indicated. As of now, I have no explanation for these exceptions, except to say that a number of these words will doubtlessly prove to be Chinese loanwords (those words with initials *h-*, *f-*, *y-*, and/or tone *-m* are suspect, for example). But if tone sandhi compounding is no longer a live process in the language, it stands to reason that there would be exceptions in every eligible syntactic collocation. The proper focus here, too, then, is on the occurrences rather than the non-occurrences of the tone change.

An interesting re-analysis of tone sandhi involving the numerals was offered by Lauj Pov Vaj, who was a 21-year old instructor in Hmong at the Southeast Asian Studies Summer Institute at the University of Michigan, summer 1985. He used the sandhi tone with the common classifier *tus* (> *tug*) after the numerals, but said it was not necessary to do so, and that both tonal realizations sounded equally good to him. Then, seemingly as an afterthought, he added that perhaps *tus*, the base form, was the singular form (*ib tus dev* "1 dog"), and that *tug*, the sandhi form, was the plural form (*ob tug dev* "2 dogs"). His own speech did not bear out his analysis, but it was a familiar attempt to make sense out of the remains of an ancient process.

2.3.2. Reciprocal-Verb

Roughly as often as not, the reciprocal *sib* induces tone sandhi in the following verb, if it bears one of the proper tones.

It is not clear yet whether or not for some speakers there is a meaning differentiation correlated with the base form as opposed to the sandhi form of the verb. For example, with the verb *tom* "to bite," does *sib tom* mean "to bite each other" (the sum of its component parts) whereas *sib-tog* means "to fight" (with specific acts of biting raised to the main event of which they are a part)?

With regard to this collocation, it is very interesting to note that *sis*, an alternate form of the reciprocal, often appears with the sandhi form of the verb in Bertrais' examples. It is not a feature of the White Hmong of Laos to the exclusion of the White Hmong of Thailand, apparently, since Heimbach (p. 294) mentions that *sis* is a variant form of *sib* which is "often used." Heimbach gives fewer examples with *sis*, though, having regularized to *sib* for the most part, so the following examples all come from Bertrais:

<i>sis-ceg</i>	"to argue with each other"
<i>sis-ncag</i>	"to make a line with each other"
<i>sis-nraus</i>	"to butt each other"
<i>sis-qawg</i>	"to embrace each other"
<i>sis-tog</i>	"to bite each other"
<i>sis-tuas</i>	"to kill each other"
<i>sis-tuag</i>	"to kick each other"
<i>sis-txig</i>	"the same height"
<i>sis-xyaws</i>	"mixed up together"
<i>sis-zeg</i>	"to tease each other"

Note that *sis* bears the wrong tone for a tone sandhi trigger: it is extremely unusual to have the sandhi tones follow anything other than a word with a *-b* or a *-j* tone, and it always suggests something interesting about the history of the language. The other situation where this occurs is "sandhi form promotion," where the sandhi form of a word, generated in the usual fashion, has, over time, become detached from the trigger which gave rise to it and has gained base form status, either co-existing with the old base (often with an accompanying semantic split), or supplanting the old base entirely. To understand the appearance of sandhi forms following *sis*, it is necessary to look at Wang's comparative data on this root (1979:89):

YANGHAO	XIANJIN	SHIMEN	GAOPO	FUYUAN	FENGXIANG
çi ⁴ 4	çi ³³	hi ¹¹ /çi ¹¹	shoŋ ¹³	si ³¹	çou ³³
(C1)	(D1)	(D1)	(B1)	(A/D)	(A1)

The tonal reflexes of this root all indicate an ancient voiceless initial, yet all four Proto-Hmongic tones are represented. White Hmong *sib* is a reflex of category A1 and *sis* is a reflex of category D1. Although Wang writes that it is difficult to account for the vast discrepancy among the tones in the different localities, or to reconstruct the original tone, his discussion of the final-ŋ of the GAOPO form (1979: 124) may provide a clue as to the cause. He reports that in GAOPO, the final of the reciprocal will harmonize with the final of the following verb.

Accordingly,

shoŋ ¹³	zoŋ ⁴³	"to be good to each other"
shu ¹³	tçu ²²	"to meet each other"
shə ¹³	pə ⁵⁵	"to see each other"
shi ¹³	zi ²²	"to bind each other"

The intimate nature of the relationship between the reciprocal and its verb (both Wang [1979: 124] and Downer [1967:593] refer to the reciprocal as a "prefix") could explain the different tonal reflexes as well: in GAOPO, a live process of vowel harmony marks the relationship¹²; perhaps in ancient Hmong, a process of tone harmony marked the relationship. As this process faded out, one or another of the shifting forms became the sole form, or, as in White Hmong, two forms persisted. The connection thereafter was marked, for those languages of the West branch of Hmongic which preserved the A1-reflex reciprocal, by tonal modification of the right-hand member, the verb, rather than by modification of the left-hand member, the reciprocal.

The explanation for the tone sandhi forms following this variant could be, therefore, (1) the identification of *sib* and *sis* as trivially different manifestations of the same root, with no differentiation in meaning, and (2) the need, existing through reconstructible history, to signal this semantic relationship as being an unusually close one through the deformation of one or the other of the two words.

¹² We can only speculate as to how the vowel harmony in GAOPO developed. It is reported that some GAOPO prefixes also change to harmonize with the vowel of the root (nouns and ordinal numerals) to form disyllabic words (Institute 1962:5-6).

2.3.3. Noun-Modifier

In normal White Hmong word order, modifiers follow the words they modify. When the relationship between modifier and noun is a close, common, and conventionalized one (see Section 2.4. below) and the right phonological conditions obtain (Section 2.2.), a tone sandhi compound may result. Of the two kinds of modifiers, noun and verb, compounds with noun modifiers are four times as numerous as compounds with verb modifiers in the examples presented in Appendix I of my dissertation. Some examples follow:

<i>ciab-mu</i> (< <i>muv</i>)	"beeswax"
wax bee	
<i>dab-npuas</i> (< <i>npua</i>)	"pig trough"
trough pig	
<i>kab-teg</i> (< <i>tes</i>)	"lines of the hand"
line hand	
<i>vaubkib-deg</i> (< <i>dej</i>)	"turtle"
turtle water	
<i>nab-qa</i> (< <i>qav</i>)	"lizard"
snake frog	
<i>nqob-npleg</i> (< <i>nplej</i> -)	"the upper part of the rice stalk"
upper stalk rice	

Sandhi compounds with verb modifiers, however, are not unusual (there are approximately fifty in the examples presented in Appendix I). It is well known that a separate class of "adjective" does not exist in most Asian languages, including Hmong, since adjective-like verbs are predicated of nouns with no need for the support of a copula. Many of the noun-verb modifier compounds involve one of these "adjective-like" verbs. A few examples follow:

<i>dib-caug</i> (< <i>cauj</i>)	"early-bearing cucumber"
cucumber early	
<i>nkauj-mog</i> (< <i>mos</i>)	"young girl"
girl soft	
<i>nplooj-qhua</i> (< <i>qhuav</i>)	"dry leaves"
leaf dry	
<i>tiab-nres</i> (< <i>nre</i>)	"pleated skirt"
skirt to pleat	
<i>vab-tshaus</i> (< <i>tshau</i>)	"sieve"
tray to sift	

2.3.4. Noun-Noun

Repetition with a slight variation in four word coordinative constructions is typical of Hmong figurative language (P'an and Ts'ao 1958; Johns and Strecker 1982). It is found in miniature in a number of two-word sandhi compounds. Here neither word modifies the other, but together they form a coordinative construction; either a repetitive compound involving words with only slight meaning differences or a compound pair involving objects or people which belong together:

<i>hlab-hluas</i>	"cord-rope" (=viscera)
<i>hlab-kag</i>	"cord-band" (=tatters, rags)

<i>kab-ke</i>	"custom-way" (=custom, ceremony)
<i>kwj-ha</i>	"gully-valley" (=valley)
<i>liaj-ia</i>	"paddy field-mud" (=land)
<i>mab-sua</i>	"foreigner-foreigner"
<i>nyiaj-txiag</i>	"silver-money" (=money)
<i>plab-plaw</i>	"stomach-heart" (=character, intelligence)
<i>qeej-nruag</i>	"pipes-drums"
<i>teb-chaws</i>	"land-place" (=country)
<i>teb-nrag</i>	"land-plain" (=earth)
<i>tub-se</i>	"son-wife" (=wife and children)
<i>twj-taig</i>	"utensils-dishes" (=dishes)
<i>vaj-tse</i>	"garden-house" (=house and grounds)
<i>zeb-tsuas</i>	"stone-rock" (=rock[y])

Occasionally the four word coordinative constructions can be made up of two compounds, one or both of which involve tone sandhi. The same relationships that hold between the single words in the compounds listed above obtain here (repetition with slight variation; pair membership), but in this case, the relationships hold between compounds:

<i>hnoob-teg hnoob taw</i> sun hand sun foot	"hands and feet of the sun" (=sunrays)
<i>kauj vab kauj-le</i> tray mat	"to perform <i>tso plig</i> (spirit release) with basket and mat"
<i>khaub hlab khaub-hluas</i> cloth cord cloth rope	"ragged"
<i>liaj-ia teb-chaws</i> field mud land place	"lands (for subsistence)"
<i>lwj siab lwj-plaw</i> soft liver soft heart	"to be in turmoil"
<i>mob plaub mob-nqaig</i> hurt hair hurt flesh	"to speak harshly"
<i>muaj-kw muaj-tig</i> have brother have brother	"to have family"
<i>muaj mob muaj-nkeeg</i> have hurt have lethargy	"to be unwell"
<i>neej saub neej-see</i> life sour life lonely	"to have an unhappy life"
<i>noj-nqaig noj-hnos</i> eat meat eat rice	"a feast"
<i>nkim mov xob-hnos</i> waste rice ? rice	"to eat without working for it"
<i>npauj kub npauj-nyiaj</i> jewelry gold jewelry silver	"gold and silver jewelry"
<i>nplooj xyoob nplooj-ntoos</i> leaf bamboo leaf tree	"the leaves of bamboo and trees"

<i>qub-teg qub taw</i> ¹³ old hand old foot	"inherited things; old things"
<i>suab-qeeg suab-nruag</i> sound pipes sound drum	"funeral music"
<i>suab-quag suab-nyia</i> sound cry sound lament	"the sound of crying and lamentation"

In this discussion of sandhi compounds in figurative language, it is significant to report that my primary tutor considers the only difference between many pairs of words which can equally well exhibit sandhi or not, for example *lwj-plaw* as opposed to *lwj plawv* "soft heart" is that the former, in which *plawv* undergoes tone change, sounds more "poetic."

2.3.5. Verb-Noun

In earlier descriptions of White Hmong tone sandhi (as reproduced in Section 2.3. above) the verb-noun sandhi compounds were described as "autotelic verb [one which may or may not take a noun object]-(modifying) locative object" (Heimbach, p. 454) and "avec un verbe normalement intransitif, mais qui peut être directement suivi d'un nom avec lequel il fait comme expression" (Mottin, p. 19). Although it is true that many of the verb-noun compounds involve intransitive verbs, a number of the verbs in the verb-noun compounds listed in Appendix I are transitive. The most heavily involved transitive verb is *noj* "to eat," with five compounds. Transitive verbs that enter into one or two compounds include *faib* "to divide," *lajj* "to plow," *muab* "to grasp," *ncab* "to stretch," *npuaj* "to clap," *ntxuj* "to wave," *tshab/tshaj* "to spread abroad," *txhib* "to split," and *yaj* "to wave." Even *poob* "to fall" in its metaphorical uses can be transitive, in which case it means "to lose":

poob-nyiaq "lose money"
silver

poob-plhus "lose face"
cheek

Common to all verb-noun compounds, though, seems to be the fact that the noun involved is independent; that is, it is not a constituent of a noun phrase.¹⁴

A few examples of verb-noun compounds follow:

dai-ntseg "yellow face" (indicating physical problem)
yellow ear

kaj-ntug "morning"
bright sky

muaj-cag "to have a well-shaped (long) nose"
have nosebridge

(*dab*) *noj-hlts* "an eclipse of the moon"
spirit eat moon

qajj-ke "to lean to one side of the path (so that others may pass)"
lean trail

¹³ Note the interesting asymmetry in this construction and in *hnoob-teg hnoob taw* above: one would expect the sandhi form of *taw* "foot" as well (see Section 2.4, on the heavier involvement of hands as opposed to feet).

¹⁴ I am grateful to James McCawley for suggesting that I consider this as the significant generalization concerning the verb-noun compounds.

tshaj-xo "to spread news"
spread message

2.3.6. Minor Types

Although head-modifier is the dominant word order in Hmong, there are a few examples of modifier-head compounds which exhibit tone sandhi. Quite a few of these are with *qub* "of old," which can precede the noun it modifies (Heimbach, p. 265; Mottin p 48):¹⁵

qub-chaws "the old (original) place"
place

qub-ke "the old way"
way

qub-teg qub taw "old (inherited) things"
hand foot

qub-zog "the old village"
village

A few other compounds were found which seem to exemplify the atypical modifier-head order:

dej-cog "water-driven treadmill"
water treadmill

hnoob-teg "hands of the sun" (=sunrays)
sun hand

noob-qes "testicles"
seed egg

tej-zaug "other times, occasions"
other time

A second minor syntactic type for which there are few examples of tone sandhi compounding is verb-verb, either in a coordinative relationship:

stab-qig "(to examine) high and low"
high low

or in a subordinative relationship:

leej-tuag "willing to come"
willing come

quaj-taug "crybaby"
cry able

qheb-tsha "cleared up (of the sky)"
open bright

2.3.7. Lack of Involvement of "Weak" Word Classes

Pronouns never trigger tone sandhi in a following word, despite the fact that many of them bear a sandhi trigger tone: *koj* "you (singular)," *neb* "you (dual)," *nej* "you (plural)," *wb* "we/us (dual)," *peb* "we/us (plural)." This serves conveniently to disambiguate the homophones *peb* "three" (which almost always

¹⁵ Other verbs which, as modifiers, can precede are *tuam* "grand, great" (a Chinese loan), *me* "small," *zoo* "good" (Mottin, p. 48), and *niag* "great" (Heimbach, p. 140).

triggers tone change in a following classifier, see Section 2.3.1.) and *peb* "we/us" in phrases such as:

peb-tug npua "three pigs"
3 clf pig

peb tus npua "our pig"¹⁶
we clf pig

Classifiers, also a "weak" word class, have been said to be incapable of influencing a following word (Heimbach, p. 446; Mottin, p. 19). Although this is generally true, a number of exceptions, such as

ib lub-hlis "1 month"
1 clf moon

have been found.¹⁷

2.4. Lexical Selectivity

An examination of tone sandhi compounds arranged by their first words typically reveals one to three examples of compounds beginning with a particular word. This arrangement of the data would seem to indicate that tone sandhi compounding operates fairly shallowly through the lexicon, given the phonological and syntactic restrictions noted above, and involves a large number of words. Through working with the data, however, I developed a feel for those words most likely to occur in a new compound. To try to substantiate this sense I had of the inequality among words with regard to their ability to enter into tone sandhi compounds, I arranged the compounds by the second word of each as well. It became clear that a rather small number of words were involved as the second member in a large number of compounds. A combination of the information on high-frequency words in tone sandhi compounds drawn from both organizational formats (Appendix I and Appendix II in my dissertation) appears in Table 1. This table contains sixty words which were found in at least four compounds. Those which are especially well-represented (fifteen or more compounds) are highlighted in capital letters.¹⁸ The hyphens either before or after each word indicate whether the word appears as the left- or right-hand member of a compound (in most cases, this is predictable from the tone of the word). A number of nouns which bear the high falling (-j) tone can occur in either position, and are marked with a hyphen on either side. For these nouns, the number of compounds attested has been broken into two parts: attestations as the left-hand member and ("+") attestations as the right-hand member. The last column indicates whether or not the word was included in Wang's 1979 comparative study, and, for those included, the number of the reconstructed initial. It is significant that at least forty-six of these sixty words can be reconstructed for Proto-Hmong (see Section 3., where the historical development of the tone sandhi system is discussed).

¹⁶ Downer (1967: 594) also makes this observation.

¹⁷ See Rauliff 1986: 116-18 for further examples and discussion.

¹⁸ Since tone change after the numerals *one* through *five* and *nine* is so widespread (Section 2.3.1.), examples are not included in the data here.

TABLE 1
WORDS APPEARING WITH HIGH FREQUENCY IN WHITE HMONG TONE
SANDHI COMPOUNDS

dab-	spirit	5	W113
-daj	yellow	8	W116
-DEJ-	water	10 + 12	W110
hlab-	cord	5	W54
-hneev	crossbow	4	W48
-KEV	road	15	W94
kwj-	ditch	4	
-laus	old	11	W55
-l1aj-	paddy field	2 + 6	W74
-mos	soft	8	W6
muaj-	to have	10	W6
-noj-	to eat	5 + 1	W49
noob-	seed	4	W59
-NCAUJ-	mouth	6 + 11	W90
ncej-	post	7	W90
nkauj-	girl	6	
-NPUA	pig	18	W7
-NPLEJ-	rice	5 + 10	W34
nplooj-	leaf	4	W34
-nqaij-	meat	4 + 3	W108
-nras	plain	4	W64
nruab-	middle	10	
ntaub-	cloth	4	W50
-ntoo	tree	12	W50
-NTUJ-	sky	7 + 10	W52
-ntsej-	ear	3 + 8	W27
-nyiaj-	silver	1 + 4	W87
-NYUJ-	cow	4 + 13	W87
-paj-	flower	9 + 2	W3
pob-	bump/blob	8	W1
poob-	to fall	7	
plab-	stomach	6	W28
QAB-	behind/under	18	W103
-qeej-	keng	1 + 8	W110
qub-	old	4	
-qhuav	dry	12	W104
-raj-	tube	1 + 3	W58
roob-	mountain	5	
rooj-	door/gate	7	W1
-ruv	roof ridge	4	
SIB-	(reciprocal)	22	W81
-siav	life	4	
-sov	warm	4	W81
suab-	sound	4	
taub-	gourd-like	9	W44
-taw	foot	6	W44
-taws	firewood	4	W46

Table I con't.

teb-	land	4	W44
-TES	hand	20	W15
tib-	single	6	
tub-	son/boy	5	W44
-tsev	house	8	W22
-tsua	rock	8	
-tshauv	ash	4	W76
-tsho	jacket	4	
-txaj-	bed	1 + 7	
-vaj-	garden	2 + 4	W12
xyoob-	bamboo		
-zaj-	dragon	1 + 3	W67
-zoov	forest	8	W65

These sixty words appear in well over three-quarters¹⁹ of the tone sandhi compounds. They are "core vocabulary" in the sense that they represent important natural items ("water," "road," "ditch," "seed," "rice," "leaf," "plain," "tree," "sky," "flower," "mountain," "firewood," "land," "rock," "ash," "bamboo," "forest") or cultural items ("spirit," "crossbow," "paddy field," "post," "pig," "cloth," "silver," "cow," "keng," "roof ridge," "house," "jacket," "bed," "garden," "dragon"). Common verbs ("to have," "to eat," "to fall"; "yellow," "old," "soft," "dry," "warm") are also included. Body parts included are "mouth," "ear," "stomach," "foot," and "hand." Certain prefix-like words which specify part, shape, or sex include "cord," "mouth," "girl," "leaf," "middle," "expanse (sky)," "blob," "behind," "tube," "gate," "gourd," and "boy." The reciprocal comprises a class of its own (see Section 2.3.2. above). Assuming that Wang's five hundred and ninety cognate sets represent important core vocabulary as well as simply the available vocabulary, it is significant to note that only fourteen of the above words are not included in his study.

There are, however, interesting asymmetries in the list. *Tes* "hand" enters into at least twenty compounds, whereas *taw* "foot" enters into only six, as evidenced by the dictionaries. Of the six "foot" compounds, my primary tutor accepts none. In a number of cases, this asymmetry does not seem to relate to the versatility and prominence of the hand over the foot, since the same trigger words yield sandhi compounds with *tes* (*pob-teg* "wrist bone"; *taub-teg* "fingertip"; *xib-teg* "palm of the hand") but, for my tutor at least, not with *taw* (*pob taw* "ankle bone"; *taub taw* "toetip"; *xib taw* "sole of the foot"). The particular tones involved do not seem to play a role either, since the change of mid level (b) to low level (-s) occurs quite commonly for my tutor with *npua* "pig," *ntoo* "tree," and *tsua* "rock." Another interesting asymmetry is seen in the involvement of words referring to animals. *Npua* "pig" with eighteen compounds and *nyuj* "cow" with seventeen are heavily involved (see Appendix II). On the other hand, no compounds found thus far involve *dev* "dog" or *nees* "horse," only two have been found for *twm* "water buffalo," and only three for *ntses* "fish." There is only one tone sandhi compound with *qaitb* "chicken" due to its tone (*npua* always, and *nyuj* most often are involved as the right-hand member of a compound, and high level (-b) tone words do not undergo change in

¹⁹ Three hundred and ninety-two out of four hundred and ninety-four, or 79.35 per cent of the compounds listed in Appendix I of my dissertation.

this position). *Tshis* "goat" does not participate in these compounds since -s changes to -g in the right-hand member, and there is a cooccurrence block between aspirated initials and the breathy (-g) tone (see Section 2.2.2.). Thus we have *nqajj-npuas* "pig meat" and *nqajj-nyug* "cow meat," but *nqajj twm* "buffalo meat" and *nqajj ntse* "fish meat." My tutor explained that the relative prominence of "pig" and "cow" in compounds had to do with their central role in everyday life, and the time and effort expended in their care. Poultry, goats, pigs, buffalo, cattle, and horses constitute the chief domestic animals of the Hmong of highland Laos (Barney 1967: 284-85). The explanation for the involvement of "pig" and "cow" seems good, but still does not explain the lack of involvement of *nees* "horse" and *twm* "water buffalo." Similarly, *zaj* "dragon" is involved in four tone sandhi compounds, but the culturally important *tsov* "tiger," even more central a figure in Hmong folk tales, has not yet been found in one compound.

The only way to account for this selective involvement of certain lexical items in tone sandhi compounds over their semantically related cousins is to mark in the lexicon the fact that these words, namely the sixty reproduced above, or some subset thereof, are particularly powerful when it comes to inducing tone change and thereby creating a compound (if it is a -b or a -j word), or are particularly susceptible when it comes to being influenced by a preceding -b or -j word (if they bear the tones -j, -s, -v, or -ð). I think it possibly significant that no word with a low checked (-m) tone appears in the table of high frequency words (which may be related to the limited involvement of *twm* "water buffalo" mentioned above). This tone is the reflex of historical category D2, the words of which can be shown to have once had a final *-p or *-t. It may be that the -m tone does not play as great a role in tone sandhi compounding due to its special development from checked syllables, or due to its possible role as a "loan tone" for borrowed words (see Section 2.2.3. above), or both.

Although a number of semantic relationships hold between the members of a tone sandhi compound, it is important to describe some of the typical relationships in order to come to an understanding of the above high-frequency words in a different way. These are specialized relationships which can be considered subordinate to the general syntactic relationship of noun-modifier described in Section 2.3.3. Although it still may be desirable to analyze the syntactic relationship between the two members of the compound as noun-modifier in the following categories of specialized relationships, semantically the first member of these compounds specifies and delimits the broader category represented by the second member.

2.4.1. Part-Whole

In the following compounds, if the left-hand member is "A" and the right-hand member is "B," each can be glossed as "the A of the B." Thus *nplooj-ntoos* can be understood as "the leaf of the tree" (that part of the tree), or as "tree leaves" (as opposed to rice leaves). This group includes spatially delimiting words also: *nrab/nruab* "in the middle of" and *qab* "on the underside of; at the base of; on the downhill side of." The following examples are representative:

<i>ncej-cog</i>	"treadmill post"
post treadmill	

<i>ncauj-ke</i> mouth road	"entrance to the road"
<i>nruab-ntug</i> middle sky	"in the heavens"
<i>qab-pag</i> bottom lake	"the bottom of the lake"
<i>qib-hnee</i> trigger crossbow	"the trigger of the crossbow"
<i>rooj-ntxas</i> gate grave	"the opening of the grave"
<i>taub-qeeg</i> gourd keng	"the body of the keng"

2.4.2. Object-Material

Another type of tone sandhi compound in which the left-hand member delimits the right-hand member is the object-material type. As with the part-whole compounds, the relationship between the two members can be understood in one of two ways. *Khawb-hlaus* can be understood as "a link of iron," where "link" serves as a unit of measure of an infinite quantity of iron, or as "an iron link" (as opposed to a silver link).

<i>hleab-ntoos</i> coffin tree	"wood coffin"
<i>khawb-hlaus</i> link iron	"iron link"
<i>khawb-nyiaq</i> link silver	"silver chain"
<i>npauj-nyiaq</i> jewelry silver	"silver jewelry"
<i>nqaj-hlaus</i> bar iron	"iron bar"
<i>ntaub-pag</i> cloth cotton	"cotton material"
<i>roj-a</i> fat earth	"fat of the earth" (=oil)
<i>roj-npuas</i> fat pig	"pig fat"
<i>roj-nyug</i> fat cow	"cow fat"
<i>voj-hluas</i> circle rope	"circle made of rope" (=lasso)
<i>voj-teg</i> circle hand	"circle made with the two hands"

2.4.3. Shape-Object

What I consider "shape prefixes" in these compounds are actually nouns, but semantically they seem to fall somewhere between classifiers such as *lub* "round, bulky object," *txoj* "long object (often abstract)," and *tus* "long, slender object (shorter than *txoj*)" on the one hand and the semantically meatier nouns with which they are paired on the other. They are not classifiers because they occur with classifiers, and, secondarily, because classifiers generally do not trigger tone sandhi (but see Section 2.3.7.). They serve to describe what facet of a multifaceted object is under consideration. For example, *tes* "hand" is probably better understood as "the protuberance at the end of the forearm" because it includes a body part which we do not include in our notion of "hand," namely the wrist. Thus we have *dab-teg* "wrist" and *pob-teg* "wrist bone" as well as *taub-teg* "fingertip" (from *dab* "narrowing," *pob* "round object," and *taub* "gourd-shaped object"). Similarly, for *ntsej* "ear" we have *nplooj-ntseg* "the outer ear" (leaf-shaped) and *taub-ntseg* "the earlobe" (gourd-shaped). A few more examples, arranged by shape prefix, appear below:

<i>pob</i> "round object"	
<i>pob-a</i>	"clod of earth"
earth	
<i>pob-ntoos</i>	"tree stump"
tree	
<i>pob-ntseg</i>	"ear" (the whole thing)
ear	
<i>pob-tsuas</i>	"rock mass"
rock	
<i>taub</i> "gourd-shaped"	
<i>txiv taub-ntoos</i>	"papaya"
fruit tree	
<i>taub-nkawang</i>	"mass of hornets" (either swarm
hornet	or nest)
<i>tswb</i> "bell-shaped"	
<i>tswb-tsaig</i>	"bell jaw" (=jowls)
jaw	
<i>txoj</i> "long" ²⁰	
<i>txoj-hmoo</i>	"fortune" (regarded as a length)
luck	
<i>txoj-ke</i>	"road" (regarded as a length)
road	
<i>txoj-sia</i>	"life" (regarded as a length)
life	

2.4.4. Sex/Agent Designators

Tub "son, boy, male," *nkauj* "girl" and *poj* "woman" are involved in tone sandhi compounds as agentive prefixes or simply as sex designators. *Tub* "son" is normally matched with *ntxhais* "daughter," which cannot effect tone change. *Poj* "woman" is matched with *txiv* "man, father," which also is powerless to

²⁰ *Txoj* has been labelled a classifier, but its behavior in these compounds is much more noun-like.

effect tone change. Thus we do not have parallel tone sandhi compounds according to sex, with the exception of *tub-qhe* "male servant"/*nkauj-qhe* "female servant":

<i>tub-nkeeg</i>	"a lazy person"	<i>nkauj-npuas</i>	"female pig"
boy lethargic		girl pig	
<i>tub-ntsog</i>	"an orphan"	<i>nkauj-qhe</i>	"female servant"
boy ?		girl servant	
<i>tub-qhe</i>	"male servant"	<i>nkauj-zag</i>	"female dragon"
boy servant		girl dragon	
<i>tub-txawg</i>	"an able person"		
able		<i>poj-cuag</i>	"mother of child's spouse"
		woman (kin)	
		<i>poi-sua</i>	"foreign woman"
		woman foreign	

2.4.5. Body Part Designator

Although it has been found in only one tone sandhi compound so far, the prefix *caj*, with no independent meaning (*caj dab*) "the neck," *caj tw* "buttocks," *caj npab* "the upper arm"), is the clearest example of a prefix triggering tone sandhi. That one compound is *caj-pas* "wind pipe" from *pa* "breath."

2.5. Compounding and Semantic or Syntactic Shift

In the majority of cases, if tone sandhi compounding is optional, it either results in no change in meaning, or only a stylistic change, the compound being perceived as "smoother," "gentler," or "more poetic" than the uncompounded collocation (see Section 2.3.4.). In a few interesting cases, though, either meaning or meaning and structure seem to be changed when the tone is changed. These five cases are discussed in some detail below:

(1) *Zaub* "vegetable + *ntsim* "peppery."

Both Bertrais and Heimbach record *zaub-ntsig* as the name for a particular plant: Bertrais (*zaub*) simply writes "espèce de légume" and Heimbach (p. 200) describes it as "a kind of pickled peppery vegetable prepared from the tops of mustard greens." This is to be contrasted with the following sentence from the same entry in Heimbach:

Zaub ntsim ntsim li kuv tsis noj
vegetable peppery peppery this I not eat
way

"I don't eat peppery vegetables."

Here the subject is not a particular peppery dish, but rather those vegetables of which a peppery flavor is being predicated. The semantic contrast is particular versus general, the syntactic contrast is noun-modifier versus subject-predicate.

(2) *Kub* "horn" + *twm* "water buffalo."

Both Bertrais (*twm*) and Downer (1967: 594) record a contrast in meaning between the compounded and uncompounded collocations of the above two words. *Kub twm* is a water buffalo's horn still attached to the buffalo,

whereas *kub-twg* is the horn used as "utensil." or "coupée, devenue matière première." notably a musical horn used in funeral services or in the hunt [Bertrais *kub*].²¹ This semantic shift from the neutral designation of one of the body parts of the buffalo to the particularized, independent role it plays once separated from the body is analogous to the semantic shift of *zaub ntsim* (category of plant) to *zaub-ntsig* (a specific plant).

There are three sources which do not support the analysis of "buffalo horn" presented above, however. My primary tutor accepts only *kub twm*. In Vwj et al., 1983 (p. 31), there are two sentences which refer to detached animal horns, neither of which shows compounding:

Kuam tau ib tug kub kauw.
 Kuam get 1 clf horn deer
 "Kuam got a deer horn."

peb tus kuam yog kub twm.
 we clf divining be horn buffalo
 horn

"Our divining horns are buffalo horns."

Heimbach (p. 450) cites *kub-twg*, but notes that "this may just be poetic."

(3) *poob* "to fall" + *dej* "water."

Poob dej "to fall into water" has been altered through compounding and semantic extension to *poob-deg* "to drown" according to two of my tutors, although *poob-deg* can also simply mean "to fall into water" (Heimbach, p. 447; Bertrais *poob*). To indicate that "to drown" is indeed meant, *tuag* "to die; dead" may be added: *poob-deg tuag* "to fall into water to death," i.e., "to drown" (Bertrais *dej*). That *poob-deg* alone can mean "to drown" is evident from the following sentence taken from Vwj et al. 1983 (p. 44):

Tus neeg poob-deg laww muab faus lawm.
 clf person drown they take bury perf.

"As for the drowned person, they took (him) and buried (him)."

This is an action they would presumably not have taken unless the person were also *tuag* "dead."

(4) *Muaj* "to have" + *nyiaj* "silver"/*txiaj* "money."

Downer (1967: 594) reports that tone sandhi signals a meaning difference in the following two sentences:

Nws muaj-txiaj heev. "S/he is very rich."
 3ps rich very

Kuw tsis muaj txiaj li. "I have no money at all."
 I not have money at all

²¹ Compare the same distinction marked in the opposite way with regard to animal legs: *ceg* (the sandhi form historically, now promoted to base form in White Hmong) "leg; branch" versus *ces* (with the historical tone) "leg separated from the body".

The same difference between "having money" (no tone sandhi compounding) and "being rich" (tone sandhi compounding with semantic extension) is recorded for the collocation with *nyiaj* by Bertrais:

<i>Koj puas muaj nyiaj</i>	"As-tu de l'argent?" (B <i>muaj</i>)
2ps Q have money	
<i>muaj-nyiaj</i>	"être riche" (B <i>muaj</i>)
<i>neeg muaj-txiaj</i>	"personne riche" (B <i>txiaj</i>)

The distinction is a fine one, however, and Heimbach cites:

<i>muaj muaj nyiaj</i>	"wealthy" (H 219)
<i>muaj muaj txiaj</i>	"wealthy" (H 380)

The extension of meaning here is due to the reduplication of the verb.

(5) *Dej* "water" + *sov* "warm"/*txias* "cold."

Although Downer (1967: 593) cites both the compounded and uncompounded forms of "warm water" and "cold water," saying that there is "no perceptible difference in meaning," Mottin (p. 19) makes a distinction between the two:

<i>dej sov</i>	"l'eau est chaude"
<i>dej-so</i>	"de l'eau chaude"

This distinction is the same as the one noted above for *zaub ntsim* "the vegetables are peppery" and *zaub-ntsig* "(a particular) peppery vegetable," i.e., without tone sandhi compounding, warmth is predicated of the water; with tone sandhi compounding, warmth is attributed to the water. Bertrais (*sov*), however, did not find this distinction, and glosses both *dej sov* and *dej-so* as "l'eau chaude." To put this distinction on an even more tenuous footing, Heimbach (p. 452 and pp. 381, 449) cites *dej sov* "warm water" (no compounding), but *dej-txiaj* "cold water" (compounding).

As seen in the above five examples, there is plentiful evidence that these semantic and syntactic distinctions are neither widely nor deeply felt. We may draw the conclusion that tone sandhi has the potential to encode distinctions of this type, but that it is not now being used to do so in a systematic way. That it does so in a few intriguing instances, though, may suggest a possible line of development for this system, which was probably phonetically motivated to begin with (see Section 3.3.).

3. Explanations for the White Hmong Tone Sandhi System

3.1. Downer 1967

In his paper "Tone-Change and Tone-Shift in White Miao," Gordon Downer outlines a number of important tonal phenomena in White Hmong, and proposes historical explanations for them. About White Hmong tone sandhi (one variety of "tone-shift," his term for syntagmatic tone change, as opposed to "tone-change," paradigmatic change which affects the basic inventory of tones), Downer emphasizes the numerous exceptions to the phonological and syntactic "rules" described in Section 2 of this chapter (pp. 593-94). He also shows how the same historical tone categories are involved in three of the five dialects he examines (p. 595):

Magpie [XUYONG]	XIANJIN	White Hmong
A2 > C2	A2 > C2	A2 > C2
B2 > C2	B2 > C2	B2 > C2
---	D2 > C2	D2 > C2
B1 > C1	B1 > C1	B1 > C1
---	C1 > D1	C1 > D1

On this striking correspondence of tone categories involved, despite the differences in the phonetic values of the reflexes in each dialect, and its significance for White Hmong, Downer writes as follows (pp. 594, 595):

Other West Miao dialects . . . possess much more thoroughgoing tone-sandhi systems than White Miao, which operate throughout the sentence. Such systems are suggested by the examples given of tone-sandhi in the Xianjin dialect, and are well documented for Magpie Miao. In the latter dialect, it appears that following high falling and low falling tones (the Magpie Miao reflexes of Proto-West-Miao tones [A1] and [A2]) certain tone-shifts occur regardless of the constructions involved.

. . . taking into account the fact that in nearly every way Magpie Miao and the Xianjin dialect are closer to each other than to WM, it may eventually be possible to demonstrate that the Xianjin sandhi system may be projected in its entirety back to Proto-West-Miao, although the complications of the Weining sandhi system, when it is fully described, may necessitate some modifications.²²

A plausible historical hypothesis that postulates some such sandhi system for Proto-West-Miao and explains both the thoroughgoing sandhi found in Magpie and Xianjin dialects as well as its limited occurrence in WM might be that WM has gone through three stages: first, a stage in which sandhi proceeded right through the sentence; second, a stage in which tone-sandhi is largely restricted to certain constructions; and lastly, the present stage of WM, in which tone-sandhi is further restricted to certain collocations within these constructions, and in which many of the hitherto modified items are apparently being gradually replaced by the basic tones of those items.

Downer's belief that the Xianjin system "may be projected in its entirety back to Proto-West-Miao" is supported by data drawn from Wang 1979 concerning the tone sandhi systems of other West Hmongic (Sichuan-Guizhou-Yunnan) dialects.²³ The data from both sources is presented in Table 2.

Downer's discovery that the historical category of the tone determines the Proto-West-Hmongic tone sandhi system, and his contention that the system is dying out in White Hmong explain two important things: (1) the difficulty of describing the system in purely phonetic terms (see Section 3.2. below), and (2) the sporadic occurrence of tone sandhi compounding in the language. The variation encountered in the source books, among speakers and across styles and

²² See Wang and Wang 1984 on the Weining (SHIMEN) tone sandhi system..

²³ In Institute 1962 (p. 5) it is mentioned that most types of Hmong have tone sandhi, with the major exception of East Hmongic and the minor exception of the Eastern sub-fangyan of West Hmongic. As was noted in a footnote in Section 2.2., tone sandhi in North Hmongic is of an entirely different type.

Table 2
West Hmongic Tone Sandhi Systems

Tone changes after an A-tone reflex	West Hmongic dialects				SHIMEN ^a	QINGYANG	GAOPO	ZONGDI
	WHITE	GREEN	XUYONG (Ruey & Kuan)	XIANJIN (Wang #3)	(Wang #4)	(Wang #5)	(Wang #6)	(Wang #7)
A2 ----> C2	X	X	X	X		X	X	X
B2 ----> C2	X		X	X	X	X	X	X
D2 ----> C2	X	X		X				
B1 ----> C1	X	X	X	X	X			X
C1 ----> D1	X	X		X	X			X
	environment: A1- or A2- ____				environment: A1- ____			

^a SHIMEN also has A2 ----> A1 in the environment A1- ____.

Future work with more data, and at a consequently greater time depth, will reveal what significance these observations may have on our ability to explain the evolution of the Proto-Hmong-Mien C tone and the proper dating of this particular tone sandhi system.

3.2. Sprigg 1975

R. K. Sprigg discusses the data on White Hmong tone sandhi presented in Downer in an article entitled "The inefficiency of 'tone change' in Sino-Tibetan descriptive linguistics." He is interested in presenting a purely synchronic account of White Hmong tone sandhi, and of all the information presented by Downer, uses only the information about the phonetic values of the White Hmong tones and the phonetic conditioning of the sandhi changes in White Hmong. Rather than say that one tone "changes into" another, Sprigg chooses to say that each (abstract) tone may have more than one phonetic realization, and that the occurrence of a particular realization is determined by the phonetic context, and is predictable by rule.

I would claim for my "phonetic-overlapping" type of analysis that it has the advantage of making a single tone classification possible for each lexical item, thereby making the distinction between "basic tones" and "modified tones" or "sandhi tones" unnecessary. The price to be paid for it I believe to be modest: familiarizing oneself with two phonetic exponents instead of one for certain tones [-j, -v, -s, -ɔ̃, -m] and accepting some degree of overlapping in the phonetic exponency of, in this White Miao example at least, tones [-j, -s, -m] with each other and with tone [-g], and, though not under comparable conditions, of tone [-v] with tone [-ɔ̃] and of tone [-s] with tone [-ɔ̃]. Consequently, a single tone classification will stand, by my analysis, for each lexical item (p. 177).

I see the variation . . . as a problem of pitch harmony and exponent (or realization) harmony, operating at the phonetic level, without any need to go to such drastic lengths as stating the variation as change of tone, which is, of course, a phonological change, reflecting a significant structural difference (p. 175).

. . . it is tones [-j, -v, -s, -ɔ̃, and -m] that I wish to state as having alternative pitch exponents according as these five tones harmonize, in junction, with the preceding high-register pitch (level or falling, 55 or 51) of tones [-b] and [-j] . . . (p. 175).

I believe that Sprigg's preference for this kind of analysis stems from his familiarity with languages with much more restricted tone inventories, where a low tone in context may become "non-low" and therefore much closer in phonetic value to a lexically distinctive high tone also occurring in the system. In a system with few tones, each tone has accordingly more "tone space" and a correspondingly greater amount of variation can be tolerated. If the neutralized low and high tone are similar (or the same), it may be said to be an accidental case of "phonetic overlapping." But to describe the complex system of paradigmatic replacement of the reflex of one tone category by another, as so clearly elucidated by Downer, in a language of a family characterized by languages

and dialects with large tone inventories as a matter of "pitch harmony," simply because the trigger tones both happen to be high, seems to deliberately ignore certain facts in order to advance a particular theoretical position.

The important question is how well his idea of "pitch harmony" works for White Hmong tone sandhi, and whether it is a simpler analysis, as he claims. I reproduce his chart of the phonetic exponents of the five tones which are borne by the right-hand members of tone sandhi compounds below (p. 176). "High word" refers to a disyllabic word (tone sandhi compound) which begins, as we have seen, with a *-b* (high level) or *-j* (high falling) tone (and corresponds to the sandhi tone); "low word" refers to two syllables adjacent to each other, the first of which begins with one of the other tones of the inventory (and corresponds to the base tone).

Tone	Phonetic exponents		Pitch figure
	'high word'	'low word'	
2 [-j] ²⁵	[mid	high	falling { 31 [-g] 51 [-j]
3 [v]	[level	rising	mid { 33 [-b] 35 [-v]
4 [-s]	[mid, falling	low, level	--- { 31 [-g] 11 [-s]
5 [-b]	[low	mid	level { 11 [-s] 33 [-b]
7 [-m]	[mid	lower-mid	falling { 31 [-g] 21 [-m]

Is it somehow harmonious for the pitches of tones to change in the following ways?

- Tone 2) High + High --> High - Mid (1st syllable depresses 2nd)
 Tone 3) High + Rising --> High - Level (1st syllable levels 2nd)
 Tone 4) High + Low --> High - Mid Falling (1st syllable raises and contorts 2nd)
 Tone 5) High + Mid --> High - Low (1st syllable depresses 2nd)
 Tone 7) High + Lower --> High - Mid (1st syllable raises 2nd)
 Mid

²⁵ The orthographic representations of the tones have been added in brackets; the values of the tones are unchanged from Sprigg's article.

In some cases a "high" tone raises and in some cases a "high" tone lowers the tone of a following word. Such an analysis has nothing to do with our understanding of "harmony" as exemplified by tone harmony phenomena such as downstep in African languages or by vowel harmony phenomena such as that of reciprocal and verb in GAPO Hmong (Section 2.3.2.). There are many more problems with Sprigg's analysis as well:

(1) The fact that the reflexes of tone category A1 and A2 are high in pitch (at the outset) in White Hmong may be accidental. A glance at the chart of the phonetic values for West Hmongic tone categories given in Section 3.3. below shows, for category A1, values of 53, 43, 55, 24 and 32. These A1 reflexes trigger tone sandhi in their respective dialects as well. An analysis that fails to account for the obviously related systems of near neighbors is not as good as one that does.

(2) The -b tone (55) and the -j tone (52) do not constitute a "high" class when we consider the endpoints of the two tones. It is this endpoint pitch, decidedly low in the case of -j, which should be expected to influence the pitch of the following syllable in a purely phonetic perturbation.

(3) Downer (p. 593) writes ". . . any attempt to treat this process of tone-shift as an automatically occurrent feature of present-day WM will be defeated by the overwhelming number of exceptions turning up, far outnumbering the cases of tone-shift." Sprigg simply ignores the chief characteristic of tone sandhi in White Hmong: it is sporadic, inconsistent, and idiosyncratic. This fact is elaborated at length in Downer's article and is at the heart of his argument about the nature of such cyclically rising and falling systems in the histories of Asian languages. Sprigg makes no mention of these exceptions to his system of "pitch harmony," despite the fact that Downer's article was the only source he used for his White Hmong data.

(4) It makes sense to retain the distinction between "basic tones" and "sandhi tones" since the sandhi tones represent a neutralization of the number of contrasts embodied in the basic tone inventory. In White Hmong there are seven basic tones (two of which never undergo tone sandhi), but only three sandhi tones.

(5) It also makes sense to say that one tone "changes into" another when it can be proved through comparative study, as Downer has done, that discrete tonal categories and their discrete reflexes are involved as counters in the same way in a number of different dialects, sometimes as basic tones and sometimes as sandhi tones. Although probably originally due to phonetic perturbation (see Section 3.3. below), there is no denying the role of these discrete categories at some intermediate stage of West Hmongic.

This is not a system that can be reconstructed all the way back to Proto-Hmong, however. Meng Chaoji has shown that in MEIZHU and LONGMO Bunu sandhi tones developed out of the phonetic matter of tones A1, A2, B1, and, in LONGMO, B2 (Mao, Meng, and Zheng 1982; Meng 1983). The sandhi tones are in contour identical to the basic tones from which they arose, differing only in being one step higher than the corresponding basic tones. The sandhi tones do not correspond to any of the basic tones of the language. The triggering environment is the presence of a preceding word with a category A1 or A2 tone, as in the other West Hmongic dialects. "Pitch harmony" can be used to explain tone sandhi in these Bunu dialects, but it cannot be used for White Hmong or for the dialects closely related to White Hmong.

3.3. Benedict 1985

Following a presentation I made on the topic of White Hmong tone sandhi at the Southeast Asian Studies Summer Institute Conference in Ann Arbor, Michigan (August 1985), Paul Benedict, in discussion, gave me an alternate explanation for the development of the system. I reproduce his explanation here exactly as it was given to me, but add the following caveat: this theory concerns the distant past, and many difficult questions, such as the role of stress in a tone language, are not addressed.

Benedict believes that the Proto-West-Hmongic tone sandhi system developed out of a "transphonologization of tonal close-juncture phrase" wherein initial stress on the two word phrase was realized as high initial pitch which fell throughout the phrase. The triggering mechanism was not only a high pitch on the initial word, but a lower pitch on the second word. Coupled with this analysis of the prosodic structure of Proto-West-Hmongic compounds is Benedict's reconstruction of the Proto-Hmong-Mien tonal shapes:²⁶

- **A: falling²⁷
- **B: rising
- **C: mid level

Benedict does not reconstruct the D tone at this stage since it corresponds to syllables that ended in a stop consonant in Proto-Hmong-Mien. He considers the early checked syllables "tonally undifferentiated" (personal communication). When the general bipartition of these tones took place upon loss of contrastive voicing in the initial consonants, high and low allotones of the Proto-Hmong-Mien tones became distinctive:

- **A1: high falling
- **A2: low falling
- **B1: high rising
- **B2: low rising
- **C1: high level
- **C2: low level

Upon loss of the final consonants in proto-Hmong, the old checked syllables developed distinctive tones, also with a high-low bipartition based on the feature of voicing in the initial:

- *D1: high level (+ ?)
- *D2: low level (+ ?)

Benedict feels that White Hmong is remarkable in its conservation, in large part, of these original tonal values:

- A1: high level (with a raised ending, as in the Cantonese A1 reflex)
- A2: *high falling*
- B1: *mid rising*
- B2: low level (with a lowered ending)

²⁶ These are the same as Benedict's reconstructed tonal shapes for Old Chinese (c. 500 B.C.).

²⁷ Two stars are used to represent proto-Hmong-Mien in this section; one star is used to indicate Proto-Hmongic.

- C1: *mid level*
 C2: mid falling (with whisper, a relic of the original
 voiced initial, < *11)
 D1: low level (< *55?; low level due to merger with B2)
 D2: low falling, checked

The White Hmong tone sandhi process (compound formation process) originated as follows (bear in mind that tone change occurred only if the second word was lower in pitch in this hypothesis):

- A1: High + 55 (-b) --> no change (2nd word high)
 A2: High + 52 (-j) --> *11 (C2) > 42 (-g)
 B1: High + 24 (-v) --> 33(C1) (b)
 B2: High + 22 (-s) --> *11 (C2) > 42 (g)
 C1: High + 33 (-b) --> 22 (D1) (-s)
 C2: High + *11 (42 -g) --> no change (2nd word
 maximally low)
 D1: High + *55? (> 22 -s) --> no change (2nd word high)
 D2: High + 21? (-m) --> *11 (C2) > 42 (-g)

Benedict's analysis explains a number of things; most importantly, it provides a plausible original phonetic motivation for a system that seems highly unmotivated in the present day. Furthermore,

(1) It explains why tones -b, -g, and -s (< D1) do not change.

(2) If the reconstructed shape of C is accepted and the theory of an original falling contour across the phrase is accepted, it explains the fact that four out of the five changes are changes to the C tone.

(3) It supports my hypothesis that the one sandhi change that does not move toward C, C1 > D1, came later; according to Benedict's scheme above, tone C1 moves to the D1 reflex after D1 has merged with B2. That is, C1 (33) does not change to *55?, Benedict's value for the D1 tone originally, but to 22, the value of D1 following the B2-D1 merger. This cannot be determined by examination of the White Hmong data alone, of course, but in neighboring dialects where B2 and D1 are distinct, it is clear that C1 > D1.

There are, however, a number of problems, too:

(1) Although an inspection of the phonetic values of the tones of the dialects of West Hmongic (given below) does not invalidate Benedict's reconstructed tone shapes (**A falling, **B rising, **C mid level), in what way can it be said to support it, as opposed to another plausible reconstruction?²⁸

²⁸ Benedict has recently explained to me that he started with good evidence (from Chinese) that B was originally rising, and projected a falling/rising two tone system from it on the basis of simplicity and maximal differentiation (p.c.).

	**A Falling		**B Rising		**C Mid Level	
	1	2	1	2	1	2
WHITE	55	52	24	22	33	42
GREEN	55	52	24	42	33	42
XUYONG	53	21	51	11	55	33
XIANJIN	43	31	55	21	44	13
SHIMEN	55	35	55	33/11	33	53/31
QINGYAN	55	54	13	32	43	21/21
GAOPO	24	55	13	31	43	22
ZONGDI	32/22	53	42/232	11	55/35	33/13

(2) Assuming that this system developed after the loss of the initial voicing contrast in most dialects and the consequent bipartition of tones, why could C1 not have changed to C2 (33 > *11), too? Why is there a strict division of possible changes according to the feature of voicing in the Proto-Hmong initial (type 1 > type 1; type 2 > type 2)?

(3) According to this theory, why did D2 change? It was originally *11?, maximally low. This was what was supposed to have kept C2 (*11) from changing.

(4) On the basis of the data presented in Li, Ch'en, and Ch'en 1959 and Wang 1979, I had supposed that the merger of tones B2 and D1 went the opposite way, i.e., that 22 reflects the original value of D1 rather than B2. This is because in many Hmongic dialects both B2 and C2 (and, to a lesser extent, A2 and D2) are characterized by what Chinese linguists analyze as "voiced aspiration" or by what western linguists analyze as "breathy voice" (which both would derive, in this case, from the original voiced character of the initial). Since the White Hmong B2/D1 tone is low level, clear voice, I had thought that it was more likely to have reflected an original voiceless initial, and that the original breathy voiced B2 value had been subsumed.

3.4. Conclusions

From the many problems a purely synchronic account of White Hmong tone sandhi presents, such as the ones mentioned in Section 3.2. above, I would conclude that a description of tone sandhi as a system-relic is the only sensible and revealing kind of description possible. Even in a synchronic grammar, historical relics can best be described in terms of history. Both Downer and Benedict do that, and I find their explanations nicely complementary. Downer is describing the system at a more recent time (at a stage when the Sichuan-Guizhou-Yunnan sub-*fangyan* and the Northeast Yunnan sub-*fangyan* had not yet divided). He can prove his assertions about the tone categories involved, and can

support his ideas about the extent of the erosion of the system in White Hmong on the basis of comparative evidence. Benedict, typically, is comfortable speculating about an earlier state of affairs (Proto-West-Hmongic, Proto-Hmongic, and Proto-Hmong-Mien). Although his ideas cannot be proven directly, he has a plausible theory about the phonetic basis of the system.

REFERENCES

- Barney, G. Linwood. 1967. "The Meo of Xieng Khouang Province, Laos." *Southeast Asian Tribes, Minorities, and Nations*, Vol. 1, ed. by Peter Kunstadter, 271-94. Princeton: Princeton University Press.
- Benedict, Paul K. 1985. "Early Sino-Tibetan/Hmong-Mien loan relationships." Paper presented at the SEASSI Conference, Ann Arbor, 1-3 August 1985. (LTBA, this issue).
- Bertrais-Charrier, Yves, R. P. 1964. *Dictionnaire Hmong (Mèo Blanc)-Français*. Vientiane: Mission Catholique.
- Downer, G. B. 1967. "Tone-change and tone-shift in White Miao." BSOAS, The University of London 30: 589-99.
- , 1973. "Strata of Chinese loanwords in the Mien dialect of Yao." *Asia Major* 18: 1-33.
- Heimbach, Ernest E. 1979. *White Hmong-English Dictionary*. Revised edition. (Linguistic Series IV, Data Paper No. 75, Southeast Asia Program, Cornell University). Ithaca: Cornell Southeast Asia Program.
- Henderson, Eugénie J. A. 1967. "Grammar and tone in South East Asian languages." *Wissenschaftliche Zeitschrift der Karl-Marx-Universität (Leipzig)* 16: 171-78.
- Institute of Nationality Languages. 1962. "A brief description of the Miao language." *Miao and Yao Linguistic Studies. Selected Articles in Chinese* translated by Chang Yü-hung and Chu Kwo-ray (Linguistic Series V, Data Paper No. 88, Southeast Asia Program, Cornell University), ed. by Herbert C. Purnell, Jr., 1-25. Ithaca: Cornell Southeast Asia Program, 1972. (Originally published *Zhongguo Yuwen* 111: 28-37 [In Chinese].)
- Johns, Brenda and David Strecker. 1982. "Aesthetic Language in White Hmong." *The Hmong in the West: Observations and Reports*, ed. by Bruce T. Downing and Douglas P. Olney, 160-69. Minneapolis: Center for Urban and Regional Affairs.
- Li Yung-sui, Ch'en K'o-chung and Ch'en Ch'i-kuang. 1959. "Some problems concerning initials and tones in the Miao language." *Selected Articles*, Purnell, 83-110 (see citation at "Institute" above). (Originally published in *Yuyan Yanjiu* 4: 65-80. [In Chinese].)
- Lyman, Thomas Amis. 1974. *Dictionary of Mong Njua, A Miao (Meo) Language of Southeast Asia*. The Hague: Mouton.
- Mao Zongwu, Meng Chaoji, and Zheng Zongze. 1982. *Yaoyu Yuyan Jianzhi* [A Sketch of the Languages of the Yao People]. Beijing: Minzu Chubanshe [Nationalities Press]. [Translation of section on Bunu by Jiang Zixin.]
- Meng Chaoji. 1983. "Yaoyu Bunu-yu 1' zhi 4' diao de xingcheng he fazhan" [The origin and development of tones 1'-4' in the Bunu language of the Yao people]. *Minzu Yuwen* 1983 No. 2: 56-59. [Translation by Jiang Zixin.]
- Mottin, Jean. 1978. *Eléments de Grammaire Hmong Blanc*. Bangkok: Don Bosco Press.

- P'an Yuan-en and Ts'ao Ts'ui-yun. 1958. "Four syllable coordinative constructions in the Miao language of Eastern Kweichow." *Selected Articles*, Purnell, 211-34 (see citation at "Institute" above). (Originally published in *Shao Shu Minzu Yuwen Lunji* 1: 91-109. [In Chinese.]
- Pike, Kenneth L. 1948. *Tone Languages*. Ann Arbor: University of Michigan Press.
- Ratliff, Martha. 1986. *The Morphological Functions of Tone in White Hmong*. Ph.D. dissertation, University of Chicago.
- Ruey Yih-fu and Kuan Tung-kuei. 1962. *Chuannan Yaqiao Miao de hun sang lisu* [Marriage and Mortuary Customs of the Magpie Miao, Southern Szechuan, China]. Academia Sinica, the Institute of History and Philology Monograph Series A, No. 23. Taipei: Academia Sinica. [Translation of introduction to glossary by Jiang Zixin.]
- Solnit, David B. 1985. "Introduction to the Biao Min Yao language." *Cahiers de Linguistique Asie Orientale* 14: 2, 175-91.
- Sprigg, R. K. 1975. "The inefficiency of 'tone-change' in Sino-Tibetan descriptive linguistics." *LTBA* 2: 2, 173-81.
- Vwj Tsawb, Vaj P., Yaj S., T. Vwj, K. Yaj, N. Muas, Z. Xyooj, and Xyooj P. V. 1983. *Phau Qhia Nyeem Ntawv Hmoob Dawb* [White Hmong Literacy Primer]. Washington D.C.: Center for Applied Linguistics.
- Wang Fushi. 1979. *Miaoyu Fangyan Sheng Yun Mu Bijiao* [The Comparison of Initials and Finals of Miao Dialects]. Monograph presented at the 12th ICSTLL, Paris, October 1979. [Translation by Jiang Zixin.]
- Wang Fushi and Wang Deguang. 1984. "Gutzhou Weining Miaoyu de shengdiao" [The Tones of the Miao spoken in Weining, Guizhou Province]. Beijing: Institute of Nationalities (mimeographed).
- Yaj Txooj Tsawb. 1986. "Piav tus txheej txheem kab tshoob kev kos" [Outline of Marriage Rites]. Excerpt translated and annotated by David Strecker. *The Hmong World I*, ed. by Brenda Johns and David Strecker. New Haven: Yale Southeast Asia Studies.
- Ying Lin. 1962. "Chinese Loanwords in Miao." *Selected Articles*, Purnell, 55-81 (see citation at "Institute" above). (Originally published in *Zhongguo Yuwen* 115: 218-29. [In Chinese.]

Lexical and Phonological Sources
of Hmong Elaborate Expressions

Brenda Johns and David Strecker

In this paper, we would like to detail the ways in which White Hmong elaborate expressions are built. These expressions are sometimes in the form of a four-morpheme array with particular semantic and phonological constraints for aesthetic heightening of language (that is, they are found to be prominent in poetry, but may certainly occur as a rhetorical device in ordinary conversation, as well). Typical examples are the following:

khv*v* iab khv*v* daw 'arduous toil'
toil bitter toil salty

kav teb kav chaw 'to rule a country'
rule land rule place

nab nkhaus niv nkhaus nom
snake bent (intensifier) bent (intensifier)
'le serpent fait des sinuosités' (Bertrais)

There are two lexical sources in Hmong for quadrisyllabics. Hmong contains a large number of intensifiers, words describing the qualities of verbs. That is, these intensifiers function adverbially. Ratliff (1986a;1986c:Chapter IV and Appendix III) has examined the properties of these intensifiers in detail. The first noticeable feature is the alliteration, as in the following examples:

liab n*nc*av nu*ij* nuas 'the monkey extends his
monkey extend (intensifier) long arms this way and
that'

n*ws* n*kees* n*kees* q*aug* z*og* lu*ij* lu*as*
s/he tired tired weak (intensifier)
's/he is lethargic and unsteady'

Foom ua tib tug fee du*j* das x*wb*
Foom make single CLF. turn (intensifier) only
'Foom merely turned his head, shaking his head and
shoulder "I don't know."'

These intensifiers also involve, to some extent, patterned tonal changes in the two parts of the intensifier. Ratliff found, for example, that it is possible to link meaning and

form with some of these pairings.

The elements of the intensifier can be made into an elaborate expression or quadrisyllabic by interpolation of other grammatical material, typically a verb like ua 'to do'. Thus, the expression used as an intensifier, dog dig 'badly, haphazardly', may be broken up so that the elements appear in the second and fourth slots of a four part expression:

ua dog dig 'to do badly, haphazardly'

ua dog ua dig idem

Quadrisyllabic expressions may also be formed from a combination of a lexical item and an intensifier. In such cases, it appears that the intensifier has the property of alliterating with the lexical item, and it seems to be thus formed to be paired with a particular word. Ratliff calls this "prosaic word incorporation" (Ratliff 1986c; section 2.1.1, pages 183-185). Examples include:

dig 'to be blind' ==> duj dig 'gropingly, feeling one's way like a blind person'

teev 'a drop' ==> tuj teev 'drop by drop'

e.g. dej nrog tuj teev 'the water drips drop by drop'
water drip drop'

A slightly more complex example is the expression tej chwb chim tej ntwb ntu meaning 'sporadically', 'from time to time', e.g.

"Sib Fi Xov" lub neej ploj qab lawm
CLF. life, disappear has ...-ed
existence

tej	chwb	chim	tej	ntwb	ntu
PLURAL		period	PLURAL		period
		of time			of time

'Sib Fi Xov [a Hmong newsletter] has ceased publication from time to time', 'Sib Fi Xov has appeared only sporadically' (Sib Fi Xov 28:1:8)

Here chwb and ntwb seem to have no independent meaning but serve rather as alliterative intensification of the nouns chim and ntu.

We have not found so many of these examples; it would be good to have a larger corpus in order to find out whether these intensifiers are further constrained phonologically.

For example, is there a tendency to any vowel or tone patterning? Indeed, Ratliff reports six tonal patterns which seem to convey particular types of sound or movement (Ratliff 1986c:188ff and Appendix III):

- b/-b: high pitched, short sounds
- g/-g: low-pitched, echoic, hollow, airy sounds
- j/-j: energetic, fast, short sounds; surface contact as opposed to contact and penetration
- s/-s: flat, continuous, unending sights and sounds
- j/-s: suggests both aspects of -j/-v (back and forth) and -s/-s (level and steady); used for sounds, movements, and attitudes
- j/-v: a double orientation (back and forth, up and down, in and out); used for sounds, movements, and attitudes.

A great many quadrisyllabics involve lexically paired words. They do not alliterate or show any special phonological relationship but rather are from the same semantic field. Typical examples include those we have cited earlier:

- ... iab ... daw 'bitter and salty' = 'arduous'
 e.g. khwv iab khwv daw 'arduous toil'
 toil bitter toil salty
- ... teb ... chaw 'land and place' = 'country'
 e.g. kav teb kav chaw 'to rule a country'
 rule land rule place

Jean Mottin points out that such pairings are of different semantic types:

Certain paired words add to one another, in the sense that they amplify or slightly change the meaning, e.g. "noj" (to eat) + "haus" (to drink) = "to subsist". But others simply repeat without adding anything to the meaning -- e.g. "txhij" (complete) + "txhua" (complete) = "complete" -- so that in many cases one of the two words seems to have been created simply to form a pair and can not be used by itself: thus "ki" is used only in the expression "tub ki" = children.

(Mottin 1978:198; our translation from the French.)

Mottin is making two points here. First, he makes the

distinction between paired synonyms like txhij and txhua, where the meaning of the whole is the same as the meaning of either part, as opposed to pairings of semantically related but not synonymous words like noj and haus, where the meaning of the whole is different from the meaning of either part. Second, he calls attention to bound forms like ki, that occur only as part of paired words. We will return to Mottin's second point at the end of this paper.

Some paired words both alliterate and show a semantic relationship, as in Mottin's example of txhij and txhua, or in

... daj ... dub 'yellow and black' = 'severe'
 e.g. cua daj cua dub 'tempest'
 wind yellow wind black

Another example is

ua yoj ua yees 'to quiver'
 do swing from do quiver,
 side to side, move back
 wave, quiver and forth
 e.g. ntuj ua yoj teb ua yees
 sky do swing earth do quiver
 'the earth trembles'

Are such examples essentially semantic pairings which only coincidentally alliterate, or are the semantics and the phonology reinforcing one another?

Of particular interest are non-alliterating pairs involving what we may call a bound form. Consider, for example, ... qoob ... loo 'grains, crops' as in

ua qoob ua loo 'to raise crops' (Heimbach, Bertrais)
 sau qoob sau loo 'to harvest crops' (Johnson, pp. 122,
 123)

Qoob occurs independently in the meaning 'crops, grains', e.g. sau qoob 'to harvest crops' (Bertrais), but loo seems to occur only in collocation with qoob. In some other Hmongic languages, however, cognates of loo occur as independent morphemes meaning 'earth' or 'field', for example

QoXiong lut [lu⁵³] 'dry field' (as opposed to paddy field)

(Wang 1985:160; cf. Chang 1947:101, 1972: Chart 13; Funnell 1970, Appendix #314). (In terms of the reconstruction of

Wang 1979 this word would be Proto-Hmongic *ʔ1-(32)-C.)

It seems likely, therefore, that an expression like sau goob sau loo originally meant literally 'to harvest the grains and harvest the fields'. Later, loo meaning 'dry field' was replaced by teb so that loo survived only in collocation with goob and its original meaning was forgotten.

Another example is ... tub ... kiv (or ki) 'child, offspring', which we met earlier in the quote from Mottin. (The variant form ki is an example of sandhi form promotion: see Ratliff 1986b:6-8, 1986c: Chapter II, section 4.1.) This expression occurs in such quadrisyllabics as

muaj tub muaj ki 'to have children'

tseg tub tseg ki 'bereft of children'

luag tub luag kiv 'other people's children'

Tub occurs independently meaning 'son' as in this example from Yaj Txooj Tsawb's "Outline of Marriage":

Yus	muaj	ib	tug	tub.	Luag	tom	ub	muaj
one	have	one	CLF.	son	others	at	yonder	have
	ib	tug	ntxhais.					
	one	CLF.	daughter					

'Suppose that you have a son, and say that those other people over there have a daughter.'

(Yaj 1974:1, 1977:3:114, 1978:3:114, 1986:102-103.) In White Hmong and Green Mong and closely related languages of China (that is, those belonging to the Sichuan-Guizhou-Yunnan subgroup of the Sichuan-Guizhou-Yunnan group of Hmongic), kiv seems to occur only in collocation with tub, but in the more distantly related Mhu language of southeastern Guizhou there is a form

jid [tɕi³⁵]

which appears to be cognate and which means 'son-in-law'. * (Lyman 1974, s.v. /tɕi/; Xiong, Xiong, and Xiong 1983, s.vv. ki, tub; Hmongb-Shuad Jianming Cidian 1958, s.v. dob gid; Hmub-Diel Jianming Cidian 1958, s.v. jid; Wang 1985:173.) It is possible, therefore, that ... tub ... kiv originally meant 'sons and sons-in-law'. If this is true, then the shift in meaning to 'sons and daughters' may tell us something about the history of Hmong gender roles and social structure.

*In terms of the reconstruction of Wang 1979 this word would be Proto-Hmongic *k-(1)-B. For the palatalization in Mhu, compare 'medicine', 'gold', 'needle'.

REFERENCES CITED

- Bertrais, Yves. 1964.
Dictionnaire Hmong-Français. Mission Catholique,
Vientiane. Reprinted 1979 by Assumption Press, Bangkok.
- Chang, Kun. 1947.
Miao-Yao yu shengdiao wenti (On the Tone System of the
Miao-Yao Languages). Bulletin of the Institute of
History and Philology, Academia Sinica 16:93-110.
- _____. 1972.
The Reconstruction of Proto-Miao-Yao Tones. Bulletin of
the Institute of History and Philology, Academia Sinica
44:541-628.
- Heimbach, Ernest E. 1979.
White Hmong-English Dictionary. Revised edition.
Data Paper 75, Southeast Asia Program, Department
of Asian Studies, Cornell University.
- Hmongb-Shuad Jianming Cidian (Hmong-Chinese Pocket Dictionary).
Guizhou Nationalities Press, 1958.
- Hmub-Diel Jianming Cidian (Mhu-Chinese Pocket Dictionary).
Guizhou Nationalities Press, 1958.
- Johnson, Charles, editor. 1985.
Myths, Legends and Folk Tales from the Hmong of Laos.
Privately published.
- Lyman, Thomas Amis. 1974.
Dictionary of Mong Njua. Mouton, The Hague and Paris.
- Mottin, Jean. 1978.
Éléments de Grammaire Hmong Blanc. Don Bosco Press,
Bangkok.
- Purnell, Herbert C., Jr. 1970.
Toward a Reconstruction of Proto-Miao-Yao. PhD disser-
tation, Cornell University. University microfilms
70-17,096.
- Ratliff, Martha. 1986a.
Two-Word Expressives in White Hmong. In Glenn L. Hen-
dricks, Bruce T. Downing and Amos S. Deinard, editors,
The Hmong in Transition. Center for Migration Studies
of New York, Inc. The Southeast Asian Refugee Studies
of the University of Minnesota. Pages 219-236.
- _____. 1986b.
An Analysis of Some Tonally-Differentiated Doublets in
White Hmong (Miao). Linguistics of the Tibeto-Burman
Area 9.2:1-35.

- Ratliff, Martha. 1986c.
The Morphological Functions of Tone in White Hmong.
PhD dissertation, The University of Chicago.
- Wang Fushi. 1979.
Miaoyu fangyan shengyunmu bijiao (The Comparison of Initials and Finals of Miao Dialects). 12th International Conference on Sino-Tibetan Languages and Linguistics.
- _____. 1985.
Miaoyu jianzhi (A Sketch of the Miao Language).
Nationalities Press, Peking.
- Xiong, Lang, Joua Xiong and Nao Leng Xiong. 1983.
English-Mong-English Dictionary. Mong Volunteer Literacy, Winfield, Illinois.
- Yaj Txoob Tsawb. 1974.
Piav tus txheej txheem kab tshoob kev kos (Outline of Marriage Rites). In Kab Tshoob Kev Kos (Wedding Rites). Cyclostyled. Vientiane. Pages 1-19.
- _____. 1977.
With French translation by Yves Bertrais. In Yves Bertrais, Le mariage traditionnel chez les Hmong blancs du Laos et de la Thaïlande. Hmong Center, Chiang Mai, Thailand. Pages 3:114-3:133.
- _____. 1978.
With English translation of Bertrais's French translation. In Yves Bertrais, The Traditional Marriage among the White Hmong of Thailand and Laos. Hmong Center, Chiang Mai, Thailand. Pages 3:114-3:133.
- _____. 1986.
With new English translation by David Strecker. The Hmong World 1:102-111 (excerpt of first part of the Outline only). Yale Southeast Asia Studies, New Haven.

TOPIC MARKERS IN HMONG

Judith Wheaton Fuller
University of Minnesota

0. Introduction

The data that I will present in this paper arise from my investigations of Hmong syntax in light of the claims made by Li and Thompson (1976) about topic-prominence vs. subject-prominence in languages.¹ Li and Thompson state that the sentences of some languages can be more insightfully described in terms of subject-predicate organization, and they call these subject-prominent languages. Indo-European languages, including especially languages like English and French, fall into this category. In other languages, by contrast, sentences are more insightfully described in terms of topic-comment organization, according to Li and Thompson, and they call these topic-prominent languages. Lisu, described by Hope (1974), is a prototypical topic-prominent language, and Lahu (Matisoff 1973) and Chinese (Tsao 1979) are also quite topic-prominent.

Topic-prominent languages are characterized by Li and Thompson as giving a less prominent role in sentence organization to the grammatical subject, to the extent that the subject may be omitted in many instances, and a more prominent role to the topic. One of the ways in which the sentence topic is made more prominent is that it is marked overtly, either by position or morphological marker or both. While every language presumably has ways of marking topics, in topic-prominent languages a sentence structure consisting of a topic followed by a comment (which may or may not include a subject) is the basic (or a basic) sentence pattern of the language.

I have defined the notion of "topic" pragmatically, following Gundel (1985:4), in terms of the speaker's intentions. In particular, Gundel states that "an entity, E, is the pragmatic topic of a sentence, S, iff S is intended to increase the addressee's knowledge about, request information about or otherwise get the addressee to act with respect to E." Constituents presented as topics in this paper are evaluated by this definition. Since *topic* is defined pragmatically, the association of topics with syntactic or morphological characteristics such as position and special markers is an empirical question.

Since Hmong is a language with some of the characteristics associated with topic-prominent languages (see Fuller 1985a, 1985b, 1985c), it is reasonable to ask whether

Hmong marks topics overtly in the manner of topic-prominent languages. In Li and Thompson's sample, if a language uses position to mark topics, it always uses initial position. However, many languages, including Hmong, also use position to mark grammatical relations. Hmong has a fairly rigid Subject-Verb-Object order of grammatical relations, with the result that the subject frequently appears in sentence-initial position. If the subject is the topic, as often occurs, then both roles are embodied in one constituent, and no position conflict occurs. If topic and subject are distinct, these two roles are in competition for sentence-initial position. When both of these roles are overtly expressed, topic precedes subject. Topics can thus be identified by position; the question remains as to the possibility of marking topics morphologically. That question is addressed in the present paper.

Hmong makes use of a number of particles with a variety of functions. Two of these particles in particular are candidates for topic particles, mas and ces. I will first present information in Hmong dictionaries and grammars available to me about these particles. I will then provide evidence that they are not pause markers or subject markers. By looking at texts, I will show that these particles have a topic function.

1. Previous characterizations of mas and ces.

Heimbach's (1969) dictionary lists mas as a pause particle, functioning like the English comma; when it occurs at the beginning of a sentence, according to Heimbach, it functions like 'well' or 'and' in English. He lists ces as an initial particle which carries the action forward. Xiong (1983), Bertrais (1964), and Lyman (1974) translate both words as 'then' or 'so', with Lyman adding some other (conjunctive and relative pronoun) meanings. Mottin (1978:141) lists both words under conjunctions which express consequence. He gives the meaning as 'then', with the examples listed in (1-3).

- (1) Nws tuaj, mas kuv zoo siab. (Mottin)
 s/he come PRT I happy
 'Il est venu, et je suis très heureux.'
 (He came, and I am happy.)
- (2) Yog nws tuaj, mas koj hais rau kuv. (Mottin)
 if s/he come PRT you speak to me
 'S'il vient, dis-le-moi.'
 (If he comes, you tell me.)
- (3) Thaum nws los, mas kuv yuav tuaj. (Mottin)
 when s/he come PRT I will come
 'Quand il viendra, je viendrai.'

(When he comes back, I will come.)

(It should be noted here that according to my principal informant, the pause in Mottin's sentences comes after mas rather than before it, and that while it is sometimes possible to pause before mas, it is more typical to pause after it.)

Mottin also lists a quant à 'as for...' meaning for mas, which takes the form of a left dislocation, with the noun phrase, followed by the particle, then a sentence with a coreferential pronoun as subject or object. An example from Mottin is shown in (4).

- (4) Kuv, mas kuv nyob zoo thiab. (Mottin)
 I PRT I stay good and
 'Quant à moi, eh bien, je vais bien.'
 (Me, I am happy.)

Yang Dao (1980) also lists mas under quant à 'as for'. An example from Yang is (5).

- (5) Ntawm nej lub tsev mas kuv mam saib xyuas. (Yang)
 LOC your CLF house PRT I PRT look visit
 'Quant à votre maison, je m'en occuperai.'
 (As for your house, I'll take care of it.)

2. Rejection of Mas as a Pause Marker

Since it has been suggested that mas is a pause particle, I would like to address the question of the relationship between the particle mas and possible pauses in the discourse. If mas is simply the Hmong 'comma', then possible pause locations should be all and only those locations where mas can occur. If this is the case then the relationship of the particle mas to the topic would be only that one can typically pause after a topic. However, the set of pause locations and the set of mas locations do not correspond exactly. I first present examples where a pause is possible but mas is not. These are sentences (6-8).²

- (6) "Tub <pause/*mas> sab Luang Prabang mas av zoo tsis zoo."
 son side land good not good
 'Son, (on the) Luang Prabang side, the land is very good.' (HT)

- (7) Hmoob thiaj qm zom zaws hais tias: <pause/*mas>
 then yell all saying if egg

"Yog tuav qe..."
 if break egg

'The Hmong then yelled all at once saying, "If (you)

break
the egg...? (HT)

(8) Nws cov me nyuam mas, ib tug yog xib fwb <pause/*mas>
s/he grp child PRT 1 CLF BE teacher

ib tug yog tus kws tshuaj <pause/*mas> ib tug yog
1 CLF BE CLF doctor 1 CLF BE

tus coj zos.
CLF leader village

'His children, one is a teacher, one is a doctor, (and)
one is a village leader.'

In (6), Tub is an appositive, which may be followed by a pause, but not by mas. In (7), a pause but not mas may follow the expression introducing a direct or indirect quote. In (8), a pause but not mas may follow each comment about the children. In none of these cases could the material preceding the pause where mas is unacceptable be interpreted as topical. The unacceptability of mas in these pause contexts shows that although mas is often used where there is a pause, it is not a pause particle. The fact that the places where mas cannot occur are precisely those pauses which do not follow topics strengthens my argument that mas is used to mark topics.

The opposite situation, where mas is possible but a pause is not, also occurs. If mas marks topics, after which pauses are possible at least in prototypical cases, then it could be the case that wherever mas occurs, a pause is possible. This does not seem to be quite true, but the reason for the unacceptability of a pause in certain contexts with mas appears to be related to the stylistic factor of constituent length rather than to the syntactic/pragmatic factor of the relationship of the topic to sentence containing it. This is shown in (9-11). In these examples, (*pause) means a pause cannot occur either before or after mas.

(9) ...ua qoob mas (*pause) zoo tsis zoo; (HT)
do crops good not good
'...raising crops (was) very good;'

Taub dag mas (*pause) loj tsis loj...
pumpkin big not big
'Pumpkins (were) very big...'

(10) Lawv los mas (*pause) muaj thaj txoob ... (HT)
they come have CLF palm

'(When) they came there was one grove of palms'

- (11) Nti^m su mas, po_j niam mas (*pause) noj loj nawb. (HT)
 fix lunch woman eat big EMPH
 '(If you) carry a lunch, the women (want to) eat a lot.'

If the two sentences in (9) occur together, a pause (and mas) may occur between them, but no pauses occur in the individual sentences. In (10), no pause occurs until the end of the sentence. In (11), no pause can occur at the second mas (presumably since a pause occurs after the first mas), but if the sentence begins with poj niam, a pause can occur with mas. These facts suggest that while pauses occur only at constituent breaks, an additional constraint is maintenance of a space between pauses, and that with two constituent breaks where pauses are possible, the pause will occur at the larger break. Although further examination of this point is beyond the scope of this paper, it is clear from the examples above that the relationship between mas and the pause is an indirect one. The places where mas occurs are typically places where pauses also occur, but since a pause can occur where mas cannot, and (less clearly) mas can occur where a pause cannot, mas is not a true pause marker.

3. Rejection of mas as a Subject Marker

Since subjects and topics frequently coincide in a sentence, the occurrence of mas after the subject/topic NP might be interpreted as a subject marker. However, mas does not consistently mark subjects, as I show here. First are examples where the subject cannot be followed by mas, as shown in (12-14).

- (12) Nplias (mas) kuv (*mas) twb muab paj tshab rau nws lawm.
 Blia PRT I PRT already give present to her COMPL
 'Blia, I already gave a present to her.'

- (13) Phau ntawv no thiab phau ntawv ntawd (mas) kuv (*mas)
 CLF book this and CLF book that PRT I

yuav yuav tib si.
 will buy both

'This book and that book, I will buy both.'

- (14) Nag hmo (mas) Xia (*mas) tuaj xyuas kuv.
 yesterday PRT come visit me
 'Yesterday Xia came to visit me.'

In these sentences there is a sentence-initial topic followed by an optional particle and then the subject.

It is also the case that constituents other than subjects may be marked with mas. This is true in (12) and (13), where the constituent marked with mas is co-referential with an object, and in (14), where the constituent marked with mas is an adverbial.

Since subjects cannot always be marked with mas, and constituents other than subjects can be marked with mas, it is clear that mas is not a subject marker.

4. Mas and Ces in Two Texts

In order to examine how these particles functioned in connected discourse, I examined all of the occurrences of mas and ces in two of my collected texts, HT1 and HT2, both eight page narratives. I tabulated each occurrence of mas and ces according to the constituent category of its preceding context. The categories where mas and ces occurred in these texts were noun phrases, prepositional phrases (of place), time clauses, 'if'-clauses, and sentences with a 'then' relationship, a 'so' relationship, or an unspecified relationship to the sentence following the particle. In HT1, the contexts in which mas and ces appeared were, except for one instance, in completely complementary distribution: ces appeared in the 'then' and 'so' contexts, and mas in the others. The one exception was an instance of mas in a 'so' context. In HT2 the results were not quite so clear-cut; there were five instances of ces in NP contexts, and five instances of ces in 'if'-clauses. The trends in HT2, however, were the same as in HT1. Charts showing these tabulations are given in (15) and (16).

(15) Tabulation of instances of Mas and Ces in HT1.

<u>Preceding Context</u>	<u>Mas</u>	<u>Ces</u>
Noun Phrase	30	0
Place Phrase	3	0
Time Clause	4	0
If-Clause	7	0
Sentence (unspecified)	13	0
Sentence ('then')	0	15
Sentence ('so')	1	14

(16) Tabulation of instances of Mas and Ces in HT2

<u>Preceding Context</u>	<u>Mas</u>	<u>Ces</u>
Noun Phrase	22	5
Place Phrase	0	0
Time Clause	11	0
If-Clause	7	5
Sentence (unspecified)	3	1
Sentence ('then')	0	16
Sentence ('so')	0	3

Of these contexts, I will discuss in this paper only the use of the particles in noun phrases.

The discussion of particles with noun phrases is largely limited to mas, since ces appears so infrequently with noun phrases in the texts I have examined. (Because of this distribution of mas and ces, and because of the dictionary definitions I have cited, and my examination of other texts, I have come to believe that mas is more closely associated with topics than ces. Ces often has a ready translation into English as 'then' or 'so', as the tabulations of the texts would indicate, whereas the translation of mas is problematical. However, native speaker judgments about these two particles, which I will not present in this paper, do not fully agree with these intuitions.)

All the noun phrase contexts preceding mas could be identified as topics according to the definition of topic which I have presented. I sub-divided the noun phrases into four categories by type of topic according to Keenan and Schieffelin's (1976) categorization of discourse topics. Their categorization makes a primary division of continuous topics, which occurred in the previous sentence, and discontinuous topics, which do not. They divide continuous topics into COLLABORATING TOPICS, in which the topic is the same as the topic of the previous sentence, and INCORPORATING

TOPICS, in which the topic is part of the comment of the previous sentence. They divide discontinuous topics into RE-INTRODUCING TOPICS, in which the topic has appeared in the previous discourse, and INTRODUCING TOPICS, in which the topic is new in the discourse. The resulting categorization gives an indication of the functions that topics serve in discourse. In this case, since I have only tabulated noun phrase topics marked by mas and ces, the result does not give a total picture of the discourse pattern of topic function. In particular, it does not account for topic chains, which would produce a larger proportion of collaborating topics. What it does show is how noun phrases marked by particles function as discourse topics. This information for HT1 and HT2 is charted in (17) and (18).

(17) Categorization of NP Topics in HT1

	Mas	Ces
Collaborating	3	0
Incorporating	18	0
Re-introducing	4	0
Introducing	2	0

(18) Categorization of NP Topics in HT2

	Mas	Ces
Collaborating	2	0
Incorporating	7	2
Re-introducing	2	0
Introducing	6	0

From these charts it can be seen that the most frequent use of mas with noun phrases is to encode entities which have been introduced in the previous sentence and have been incorporated into the present sentence as a topic. Examples of these INCORPORATING TOPICS from the two texts are given in (19-22), where preceding sentences are given to show the context. In these examples, mas can appear only where it is indicated (that is, following the topic) and in no other place in the sentence.

(19) Ua tau teb pob kws thiab teb yeeb nkaus xwb.
 make able field corn and field opium single only
 '(We) could grow corn and opium only.'

Yeeb thiab pob kws mas zoo tas nrho.
 opium and corn PRT good completely
 'Opium and corn both grew well.' (HT1)

(20) ...Fab kis thiaj npaj tub rog tos.
 French then prepare soldier meet
 'The French then got the soldiers ready.'

Cov tub rog mas yog Nyab Laj coob xwb.
 grp soldier PRT BE VN many only

'The soldiers were mostly Vietnamese.' (HT2)

- (21) kuv txiv hnub nyooq muaj 5 xyooos los nyob Houei Kasao
 mv father age have year come live
 'My father was 5 years old (when he) came to live
 at H. K.'

mus txog hnub nyooq tiav hluas lawm,
 go arrive age mature young ASF
 'and became a young man'

...twb yeej ua kam lawm.
 already do business ASF
 '... already run a business.'

Thaum ntawd mas lawv ua kam rau Suav xwb nawb.
 time that PRT they do business to Chinese only EMPH
 '(At) that time, they did business with the Chinese
 only.' (HT1)

- (22) ...ua qoob mas zoo tsis zoo.
 do crops PRT good not good
 '...raising crops (was) very good.'

Taub dag mas loj tsis loj...
 pumpkin PRT big not big
 'Pumpkins (were) very big...' (HT1)

In (19), yeeb 'opium' and pob kws 'corn' were mentioned in the first sentence and became the topic in the second sentence. In (20), the same is true, with tub rog 'soldiers' mentioned in the first sentence and topic in the second. In (21), where the time is the topic, this time period is mentioned in the previous sentences and becomes the topic in the last sentence. In (22), the topic of the second sentence taub dag 'pumpkins' is a sub-set of qoob 'crops', which is the topic of the previous sentence.

COLLABORATING TOPICS marked by mas or ces are few, because a collaborating topic is not always overtly mentioned. When the topic does stay the same from one sentence to the next and is expressed by a full noun phrase, a particle may mark the topic, as in the examples (23) and (24).

- (23) Lawv los mas muaj ib thaj txoob, cov txoob uas rog rog
 they comePRT have 1 CLF palm grp palm REL fat fat

uas noj tau tam mov, thaum tsis muaj dab tsi noj.
 REL eat can rice when not have what eat

Cov txoob ntawd mas cuag hav tsawb nyob hauv tib
 grp palm those PRT reach valley banana LOC in 1

lub vox hav Cov txoob mas tej niag tug rog rog ...
 CLF valley grp palm PRT grp large CLF fat fat...
 (HT1)

'(When) they came, there was one grove of palms; those palms which were so big that (they could be) eaten as rice, when (the people) didn't have anything to eat. Those palms, (they were as many as) banana trees located in a broad level valley. Those palms, the group was large (and) each one was fat...'

In (23), txoob 'palms' is repeated three times after its introduction, the first time modified by a relative clause, and the second and third times with the particle mas. In all these instances it is clear that the speaker is saying something specifically about the palm trees, and marking this topic overtly by using both full noun phrases and particles. In the final sentence of this discourse, mas can only appear where it is shown.

In (24) the topic is initially expressed by a full noun phrase, then repeated with the particle mas.

(24) Xyoo twg, sau qoob tas thiab hlais yeeb ces Nplog
 year WH harvest crops finish and cut opium PRT Lao

tuaj sau se ntawm Hmoob txhia leej txhia tus.
 come collect tax LOC each person each one

Ib xyoos twg mas cov nom tswv Nplog tuaj mus
 1 year WH PRT grp leaders Lao come go

"sarave" ...
 inspection (HT2)

'Each year, after the crops are harvested and the opium cut, the Laotians come to collect taxes from every Hmong person. Each year, the Laotian leaders come (for) inspection...'

In this case it is also clear that it is the regularity of these events, expressed by a time phrase, that is the topic of both sentences. In these collaborating topics, the particle is used not for the introduction of a topic but for its maintenance.

The particle mas is also used for both the introduction and re-introduction of noun-phrase topics. Examples of INTRODUCING TOPICS are shown in (25-26).

(25) Cov Phu Phaib lub caij ntawd mas muaj tsawg tsawg...
 grp CLF time that PRT have few few

Lub Moos xaj mas nyob 7 yim Phu Phaib.
 CLF PRT live family (HT1)

'The Phu Phaib (at) that time were very few. (In)
 Moos xaj lived seven families of Phu Phaib.'

(26) Thaum ub, Tiao In thiab Chaomuong laus lawv mas xuas
 long ago and leader old they PRT take

nees mus thauj Hmoob cov nqaij npuas los dai tau
 horses go carry grp meat pig come hang

tsheej niag nqai nthab.
 (area full) beam storage

'Long ago, Tiao In and the old leader took horses to
 go and carry the Hmoob's seasoned meat back and hang
 it up in the ceiling storage area until the area is
 full.' (HT2)

Examples of RE-INTRODUCING TOPICS are shown in (27-28),
 where I present only the relevant sentence, omitting the
 discontinuous context.

(27) ...peb lub zos mas yog peb tib pang kwv tij xwb.
 our CLF village PRT BE our one grp. brothers only
 'Our village was only our group of cousins.' (HT1)

(28) Cov Suav mas muaj nyiaj heev.
 grp Chinese PRT have money very
 'The Chinese had a lot of money.' (HT1)

Summarizing the uses of mas and ces which I
 have discussed in the two texts, I find that mas in
 particular marks noun phrase topics. Among topic types,
mas is used most frequently to mark incorporating
 topics, that is, those noun phrases that are introduced as
 part of a comment and become the topic of the following
 sentence. Although time does not permit discussion of other
 constituents marked by mas, analysis of the
 relationship of these constituents to their following
 sentences reveals that the constituent preceding mas
 can be considered a topic of the following sentence. Thus
 the primary function of the particle mas appears to be
 that of marking the preceding constituent as the topic of the
 following sentence.

5. The Relationship between Mas and Ne

Additional support for the view that mas is a
 topic particle comes from the relationship between mas
 and ne. Ne is described by Heimbach (1969:136)

as a "final interrogative and exclamatory particle", with (29-30) as examples.

- (29) Koj tsis paub ne? (Heimbach)
 you not know
 'Don't you know?'
- (30) Kuv twb hais rau nej no ne! (Heimbach)
 I already say to you this
 'I did tell you!'

This particle also occurs in expressions like (31).

- (31) Koj ne?
 you
 'And you?'

The particle ne does not always occur sentence-finally, as shown in (32-33), from one of my interview texts.

- (32) Cov Nais Khus ne lawv ua lawv zaub mov noj ios? (HT4)
 grp teacher they do their food eat
 'What about the teachers? Did they cook their own food?'
- (33) Nyob li cov hluas ne, tsis yog cov me nyuam kawm ntawv,
 be like grp young not BE grp child study book

 lawv nrog cov Nais Khus tham ne lawv hais lub Hmoob
 they with grp teacher talk they speak CLF Hmong

 xwb los lawv hais lus nplog? (HT4)
 only or they speak word Lao

 'What about the younger persons, not the student, do they talk to the teachers in Hmong or in Lao?'

The left-dislocated phrases preceding ne in (32-33), as well as the translation (which was provided by a Hmong translator independently of the topic-particle hypothesis), strongly suggest a topic function for ne. This relationship is made explicit in sentences (34-35).

- (34a) *Chicken mas, lus Hmoob hu li cas?
 (34b) Chicken ne, lus Hmoob hu li cas?
 'How do you say 'chicken' in Hmong?'
- (35a) *Chicken ne, lus Hmoob hu ua qaib.
 (35b) Chicken mas, lus Hmoob hu ua qaib.
 'The Hmong word for chicken is 'qaib'.'

Mas and ne are in complementary distribution, mas being used in statements and ne in questions.

The (a) sentences are ungrammatical because mas has been used in a question and ne in a statement. Additional examples are shown in (36-37).

(36a) Koj lub tsev ne (*mas) koj puas muaj kiv cua?
 your CLF house you Q have spin wind
 'Your house, do you have a fan?'

(36b) Kuv lub tsev mas (*ne), muaj ntau ntau lub qhov rai.
 my CLF house have many CLF window
 'My house has lots of windows.'

(37a) Koj lub kev sib tham ne (*mas), puas mauj tej yam
 your CLF way RECIP talk Q have GRP kind
 txaus siab?
 (interest)

'Your meeting, did anything interesting happen?'

(37b) Lub kev sib tham mas (*ne) tsis muaj dab tsi
 not have anything
 (txaus siab).

'The meeting, nothing much happened.'

In these sentences, topics of statements are marked with mas, and topics of questions are marked with ne. The function of ne as an interrogative topic particle parallel to the function of mas in statements strengthens the argument that mas is a topic marker.

6. Gradations of Acceptability of Mas

If mas is a topic marker, the acceptability of mas following a given NP should indicate whether or not that NP can be interpreted as a topic. In (38), the NPs are marked for the acceptability of mas.

(38) (Hais txog) Maiv (mas), kuv (?mas) nco hais tias Xia
 speak about Mai I remember that Xia
 (mas) twb tau sau ib tsab ntawv (*mas) rau (nws) lawm.
 PRT already write 1 CLF letter to her ASP

'Speaking of Mai, I remember that Xia already wrote a letter to her.'

Mas is only acceptable with the left-dislocated NP Maiv and the subordinate subject Xia. Mas with the main clause subject kuv is questionable, and mas with the indefinite direct object is unacceptable.

These acceptability judgments for mas correspond at least roughly to intuitive judgments about what NPs in this sentence are candidates for topichood.

7. Conclusion

Evidence has been presented in this paper that the particles mas and ces in Hmong mark sentence topics. In particular, noun phrase topics marked with mas are shown to both introduce new topics and maintain old topics in discourse. The use of mas as a topic marker in statements is paralleled by the use of ne as a topic marker in questions. Degrees of acceptability of mas appear to correspond to degrees of acceptability of the preceding NP as a topic. These facts indicate that Hmong is characteristic of topic-prominent languages in marking topics overtly both by position and by a morphological marker.

NOTES

1. I would like to thank my informant, Kee Thao, for his significant contribution to this work, and Bruce Downing for his helpful suggestions on earlier versions. This work was partially supported by a grant from the University of Minnesota Graduate School.
2. The sentences marked (HT) are taken from spoken Hmong texts which were tape recorded and transcribed. Unmarked sentences were elicited from a native speaker.

REFERENCES

- Bertrais, Yves. 1964. *Dictionnaire Hmong (Méo-Blanc) - Français*. Vientiane: Mission Catholique.
- Fuller, Judith Wheaton. 1985a. On the passive construction in Hmong. *Minnesota Papers in Language and Linguistics*, ed. by Nancy Stenson.
- Fuller, Judith Wheaton. 1985b. Zero anaphora and topic prominence in Hmong. In *The Hmong in Transition*, ed. by Glenn Hendricks et al. Center for Migration Studies.
- Fuller, Judith Wheaton. 1985c. *Topic and Comment in Hmong*. Ph.D. dissertation, University of Minnesota.
- Gundel, Jeanette Kohn. 1985. 'Shared Knowledge' and Topicality. *Journal of Pragmatics*.
- Heimbach, Ernest E. 1979. *White Hmong - English Dictionary*. Revised edition. Linguistics Series IV. Data Paper: Number 75, Southeast Asia Program. Ithaca, New York: Cornell University.
- Hope, Edward. 1974. *The Deep Syntax of Lisu Sentences*. Ph.D. dissertation, Australian National University. Pacific Linguistics, Series B, No. 34.
- Keenan, Elinor Ochs, and Bambi B. Schieffelin. 1976. Topic as a discourse notion: a study in the conversation of children and adults. *Subject and Topic*, ed. by Charles N. Li, 335-384. New York: Academic Press, Inc.
- Li, Charles N., and Sandra A. Thompson. 1976. Subject and Topic: A New Typology of Language. *Subject and Topic*, ed. by Charles N. Li, 457-489. New York: Academic Press, Inc.
- Lyman, Thomas Amis. 1979. *Grammar of Mong Njua (Green Miao): A Descriptive Linguistics Study*. Published by the Author.
- Matisoff, James A. 1973. *The Grammar of Lahu*. California University Publications in Linguistics #75. Berkeley: University of California Press.
- Mottin, Jean. 1979. *Eléments de Grammaire Hmong Blanc*. Bangkok: Don Bosco Press.
- Tsao Feng-Fu. 1979. *A Functional Study of Topic in Chinese: The First Step Towards Discourse Analysis*. Taipei, Taiwan: Student Book Co., Ltd.
- Xiong Lang, Xiong Joua, and Xiong Nao Leng. 1983. *English-Mong-English Dictionary*. Milwaukee, WI: Hetrick Printing Inc.
- Yang Dao. 1980. *Dictionnaire Français - Hmong Blanc*. Paris: Comité National d'Entraide.

DEVELOPING PRACTICAL ORTHOGRAPHIES
FOR THE LU MIEN (YAO), 1932-1986: A CASE STUDY¹

Herbert C. Purnell
William Carey International University

I. INTRODUCTION

On the fringe of that field known as applied linguistics, there is a rather small area called "orthography." It is so small, in fact, that it has been overlooked in most descriptions of the field (e.g., Kaplan, 1980:58). Nevertheless, there are people working there. Many of them are linguists concerned with Bible translation (Nida, 1947, 1954; Pike, 1947; Smalley, 1963, 1976) and missionaries, but there are also a few sociolinguists (Berry, 1958; Sjoberg, 1966; Stubbs, 1980), and researchers (such as those from the Central Institute for Nationalities in Beijing), plus an occasional visitor or two examining writing systems in general or attempting to reform spellings.

Although orthography is not a large area, it is very much a part of applied linguistics since it attempts to solve, in Kaplan's words, "human problems stemming from various uses of language" (1980:63). The problem addressed is a society's lack of a writing system or its use of an inadequate one. The fact that a writing system should ideally be based on a thorough linguistic analysis can sometimes lead one to imagine that devising an orthography is *primarily* a linguistic problem. It is thus useful to have periodic reminders that a variety of *human* factors are the prime determinants in the acceptance or failure of practical orthographies (Berry, 1958; Sjoberg, 1966; Stubbs, 1980). In Stubbs' blunt words, "it does not matter how elegant, rigorous, or systematic your linguistic analysis is, if the native speakers do not like it, then it is a waste of time" (1980:71). Orthography, therefore, must look to several disciplines for insights, not only to linguistics but also to sociolinguistics, psycholinguistics, anthropology, history, and others.

To illustrate an applied linguistics approach to orthography design and modification, this paper will first review some of the main factors involved in such an undertaking. It will then briefly look at seven orthographies designed for the Lu Mien over the past 54 years and discuss some of the reasons why, until recently, none of them had received broad enough support to be fully accepted as the "official" Lu Mien script.

II. CONSIDERATIONS IN PRODUCING OR MODIFYING PRACTICAL ORTHOGRAPHIES

The following list of factors to be taken into consideration during orthography design draws heavily on Berry (1958), Smalley (1976), and Stubbs (1980). This list is not complete by any means, and the factors noted will not be further elaborated. The ordering of factors in the list is not significant for the purposes of this paper, since all have an important role to play. Nevertheless, those called "sociolinguistic" and "cultural" will receive greater emphasis in the case study which follows.

¹ This is a revised and updated version of a paper read at the 18th Conference on Sino-Tibetan Languages and Linguistics, Bangkok, Thailand. I would like to thank Christopher Court and David Solnit for their comments on it.

A. LINGUISTIC FACTORS

1. Is the phonological representation optimal with respect to economy, consistency, and unambiguous differentiation?
2. Is the phonological analysis supplemented by grammatical information on morpheme structure, grammatically conditioned tone change, etc.?

B. SOCIOLINGUISTIC FACTORS

1. Is the orthography maximally adaptive with respect to internal social and regional dialects?
2. Does the orthography accurately reflect the attitude of the people toward
 - a. their own language in its oral and written forms?
 - b. other writing systems for their language already in use?
 - c. the trade and/or national language (i.e., how similar to or different from these languages do they want their script to be)?

C. CULTURAL FACTORS

1. What is the attitude of the leaders toward the orthography?
2. Are the people motivated to read it?
3. Are there tensions in the culture which are being reduced or aggravated by the orthography?

D. PSYCHOLINGUISTIC FACTORS

1. Do the symbols fit the way that people feel about their language and the way it should be represented?
2. Does the orthography respect the psychological processes involved in reading and writing?

E. PEDAGOGICAL FACTORS

1. Are the symbols patterned in a culturally appropriate way?
2. How easy is it for adults to learn to read?

F. PRACTICAL FACTORS

1. Are all the special symbols or diacritics used really necessary?
2. Can the orthography be easily printed locally?

III. THE IU MIEN

The Iu Mien (also known simply as Mien) are one of the major branches of the peoples known as "Yao." They are found primarily in the upland areas of southern China and northern Southeast Asia. The total Iu Mien population is estimated to be 350,000 or more (Cushman, 1975). Beginning in the late 1970s, Iu Mien refugees began to arrive in Western countries (Canada, France, and especially the United States) where they now number approximately 12,000.²

The ancestral homeland of the Iu Mien is thought to have been either near Nanjing in Jiangsu (Lemoine, 1982) or further south around Dongting Lake in Hunan, largely on the basis of their legend about crossing the sea. Contacts with Chinese have had considerable effect on their language and culture (Cushman, 1975; Lemoine, 1982). For example, the Iu Mien have, in addition to their vernacular, a literary language and a ritual language, both of which were

² There are some 10,000 Mien in the United States, over 50% of whom live in California.

borrowed as separate systems written in characters, from two types of Cantonese but are now fully integrated into their complex language system (Purnell, 1975, 1986). Characters, on the other hand, were not used to write the vernacular which was left unwritten. A list of Iu Mien phonemes is given in Figure 1 found on page 138.

IV. THE THAI SYSTEM OF WRITING

Because three of the orthographies to be discussed below have used the Thai script, some general and highly simplified remarks on the major features of that script are necessary. Figure 2 lists the Thai phonemes in Roman script for typographical convenience. Note that in Thai (a) there are three classes of consonant symbols, two of which (High and Low) are used to symbolize the same set of phonemes; (b) tone marks are linked to consonant symbols, not to vowels as in many other systems; and (c) vowel symbols may occur in various satellite positions around the initial consonant(s).

V. EIGHT IU MIEN ORTHOGRAPHIES

A. ORTHOGRAPHY 1: THAI SCRIPT (1932)

To my knowledge, the first practical orthography for the vernacular language of the Iu Mien of Thailand/Laos³ was devised by a missionary couple, Mr. and Mrs. Trung, some time prior to 1932.⁴ In February of that year, the Gospel of Mark using this orthography was printed by the American Bible Society (Trung, 1932). This slim volume is the only extant specimen of the orthography.⁵ A one-page key to some of the nonstandard symbols used is found at the end of the book.

Much of the Trungs' orthography is a radical departure from the standard Thai writing system in a number of respects, affecting not only individual symbols but also general constitutive principles. For example, High Class consonants were eliminated, and the tone marks used with the Mid Class were used with fixed values for all consonants, though not always with the standard values; vowel symbols normally written to the left of initial consonants were put to the right of them, etc.

Orthography 1 is difficult to read even for someone who reads both Thai and Iu Mien, not only because of its odd choice of symbolization and the way it altered several fundamental rules of the system, but also because it failed to meet the criteria for a successful script: its linguistic base was inadequate, it had very little transfer value to the national language, it was not aesthetically pleasing, and it must have been very difficult for a typesetter to produce.

³ An earlier, Vietnamese-based orthography used to transcribe Yao is found in Savina's dictionary of the Kim Mun language (Savina, 1926) which is related to Iu Mien. In his introduction to the dictionary proper, Savina included an extensive wordlist comparing the two languages (pp. 14-25). Whether this orthography was ever used by Catholic missionaries in Vietnam to produce materials for the Kim Mun or Iu Mien is not known. In any case, Savina's work has had no effect on the development of Iu Mien practical orthographies and thus will not be considered further here.

⁴ Smalley (1976:8) gives the name as C. K. Trang. I follow the insert to the Gospel of Mark (Trung, 1932) which reads, in part, as follows: Friends of the Bible Society, We take pleasure in presenting you with this copy of the Gospel of Mark in Yao, just off the press. This is the first tangible result of a decade of prayer and labor on the part of the Agency Secretary and of nearly two years of agonizing residence among the Miao of Mr. and Mrs. Trung.

⁵ It also appears to be the first portion of the Bible translated into Mien.

ORTHOGRAPHIES 2 AND 3

The following two orthographies were developed by missionaries in northern Thailand at about the same time, one by the Overseas Missionary Fellowship using Romanized script, the other by the American Churches of Christ using Thai script. They were both based on essentially the same phonological analysis and are linguistically adequate.

To date, much of the literature produced in vernacular Iu Mien has been for the relatively small Christian population (numbering probably not much more than 5% of the roughly 35,000 Iu Mien in Thailand and, as refugees, in Western countries). Much of it has been printed in both orthographies to reduce the divisive effect which the presence of two scripts has had on the Christian community. The reasons why two orthographies were developed in the first place will be discussed next.

B. ORTHOGRAPHY 2: ROMANIZED SCRIPT (1954-PRESENT)

The Overseas Missionary Fellowship (OMF) and the American Churches of Christ (ACC) both began their work among the Iu Mien in Chiangrai province early in the 1950s, the OMF in the western part around Maechan and the ACC in the eastern section near Chiangkham (now part of Phayao province). Those in the OMF area were recent immigrants from Laos and, given the loose control exercised in the mountainous border regions by the Thai government at that time, they owed little allegiance to Thailand. Before long, the influential headman of the village of Maesalong and many of the villagers became Christians. The OMF missionaries began to learn the language and analyze it in preparation for developing a practical orthography.

C. ORTHOGRAPHY 3: THAI SCRIPT (1956-PRESENT)

As was mentioned above, the American Churches of Christ opened their work among the Iu Mien in what was then eastern Chiangrai. The ACC started off using the OMF Romanized script but soon switched to a Thai script when it became apparent that the Iu Mien themselves preferred it. Unlike those in western Chiangrai, the Iu Mien in the east had lived in Thailand for about a hundred years and, by the 1950s, saw their future as being there. Thus, to them, a Romanized script was not in their best interests. Instead, they preferred a transfer value Thai script. Accordingly, in 1956, the ACC missionaries began to develop a Thai orthography, staying as close to the standard Thai rules and symbol patterns as possible.

Orthography 3 is certainly more complex than Orthography 2, both because it is based on a more complex writing system and also because it aims for a high transfer product.

There is no question that the presence of two scripts for the Iu Mien in Thailand has been a source of tension within the Christian community. At least two factors contributed to this. First, a Christian community emerged sooner and grew much more rapidly in the OMF area than in the ACC area. This put pressure on the OMF workers to develop an orthography for Bible translation, etc. while they were still learning the language. Within a few years the amount of literature produced was sufficient to legitimize the script. A second factor was that although the script was never restricted along religious lines, only Christians were motivated to read it since Mien traditional ritual manuals are written with Chinese characters and used by specialists literate in that writing system. These factors, together with the western Chiangrai group's lack of

identification with Thailand, led to Orthography 2 being called the "Christian" script. The gradual emergence of a Christian community in the east who had a completely different preference for an orthography because of its different orientation to Thailand was therefore viewed by the western group as a challenge to group unity in general and to their leadership in particular.

From 1960 on, the Iu Mien in both areas began to have increasing contact with the Thai language and culture. Many villages moved out of the mountains and settled in the foothills or on the plains for various political and economic reasons. Most of the Christians in the western area joined this exodus from the hills, and their children began to attend Thai schools. Nevertheless, their opposition to using the Thai script for writing their language did not diminish. Even a move far south to a new homestead area in central Thailand by many of them in 1978 has done little to change their attitude.

In the late 1970s, however, time began to run out for the Romanized script. The Christian community in the eastern area was growing, Orthography 3 was being promoted among refugees in the camps located there, and the government saw the value of using the Thai script for writing minority languages as a bridge to literacy in standard Thai as well as a means to promote greater national unity among disparate ethnic groups.

By 1980, the Thai government had decided to forbid the printing of materials in the Romanized script among most minorities, including the Iu Mien. This obviously tipped the balance firmly in the direction of Orthography 3. A combination of more materials available in the Thai script and an aggressive literacy program has produced a number of new literates in the OMF area, primarily children and young people who have had contact with the Thai letters in school. A few adults are also becoming literate, but most of the church leaders firmly refuse either to give up the Romanized script or even to attempt to become literate in the Thai script. Nevertheless, it appears to be only a matter of time until the Romanized script ceases to be used entirely.

The Iu Mien at first wanted to use Chinese characters for their vernacular since China was their ancestral homeland and their ritual and literary languages were already written in characters. This was neither practical nor politically feasible, however. They therefore opted to have their language written in a Romanized script. The Thai script was unacceptable for two reasons. First, it was too highly localized; second, it carried few positive cultural attachments for the recent arrivals who had no citizenship and limited civil rights.

Once the decision was made to develop a Romanized orthography, the OMF missionaries attempted to produce one that was uncomplicated, without diacritics or special symbols, and restricted to a normal English typewriter keyboard so that the Iu Mien could eventually produce their own literature. They thus used letters at the end of syllables to represent tones, employed upper-case letters with different values from their lower-case counterparts in order to avoid consonant di- or trigraphs, and used three consonant symbols to write the vowels /æ/, /ə/, and /ɔ/ (see Figure 3).

Although there were a few indeterminacies, as well as some odd features such as the use of capital letters, the orthography worked well and was reasonably easy to learn. In fact, the apparent strangeness of the script when compared to English was not a factor in its development. After all, in the 1950s the Iu Mien were a rather remote highland ethnic group that had no foreseeable use for English. Hence, transfer value to English was not even considered.

At present, Orthography 2 is in rather limited use. Some of the older Christians, both in Thailand and among the refugees, continue to read in it, but

even some of these are beginning to learn another script (see discussions of other orthographies below).

D. ORTHOGRAPHY 4: THAI SCRIPT (1981)

The impending demise of the Romanized script has not automatically meant that the ACC transfer value Thai script has become firmly established. In actual fact, its ultimate future is still somewhat uncertain at this point because of a simplified Thai script developed as an alternative to it.

Since 1980, simplified Thai scripts have been developed and implemented by some missionaries for several minority languages in the north. Although the script prepared for Iu Mien has not yet been implemented, it will be described here because of its potential use.

The rationale given for developing simplified Thai scripts focuses on the actual and potential readers in the OMF area and stresses several points:

- a. The standard Thai orthography is very complex.
- b. Adult rural literacy programs, even for ethnic Thais, have not been particularly successful.
- c. Although Iu Mien children and young people will be able to learn Thai in school, church leaders, particularly those 30 and older, will find it difficult to learn a standard Thai orthography because of lack of time, ability, or motivation.
- d. Nevertheless, since these leaders will no longer have access to Romanized materials, they must learn some type of Thai orthography.
- e. Therefore, the most useful type of Thai script for church leaders and the adult Christian community is a simplified script.

Simplification has focused on the consonants and tones. Most of the High Class consonants have been dropped since they duplicate sounds in the Low Class, and the five tone marks previously used only with Mid Class consonants have been given fixed values for all consonants.⁶

There are two reasons why this simplified script has not yet been implemented among the Iu Mien. First, the transfer value script has been in existence for some 25 years and has been learned by perhaps as many as 200 people. Furthermore, it is known to Thai officials in several areas of the government. A second reason is that some in the missionary community objected that a simplified script would not be in the best interests of the Iu Mien over the long run. They favor the transfer value script and stress the deep nationalistic feeling Thais have toward their distinctive script in its standard form. They also question whether the low success rate in Thai adult literacy programs might actually be due more to ineffective teaching methods or a failure to sufficiently motivate students than to excessive complexity in the writing system itself.

The simplified Thai script has been shelved for the time being. The proponents of the transfer value script have been given a few years by their colleagues to show whether or not better teaching methods can, in fact, enable adults as well as young people to become literate. It is thus not at all certain that the regular Thai script orthography will, in the end, be adopted for use in

⁶ Interestingly, these were some of the changes made 50 years previously by the Trungs, whose work these later developers were unaware of.

all the Iu Mien areas.

E. ORTHOGRAPHY 5: ROMANIZED SCRIPT (1982-1984)

In June 1982, the Iu Mien Association of Oregon sponsored a conference to discuss matters relating to their written language as it affected their cultural preservation and adaptation. Several factors led up to the conference. First, the leaders of the refugee communities on the West Coast of the United States realized that their cultural identity was in danger. They were not particularly numerous, were not (at that time) recognized as a separate ethnic group, and unlike the Vietnamese, Cambodians, and even the Hmong (Meo, Miao), lacked the unifying force of a written language.

A second factor was a 1981 writer's conference for Iu Mien held in California for the purpose of producing a periodic newsletter that would enable refugees in Canada, France, and the United States to keep in touch with each other. It soon became apparent to the participants that they did not have an acceptable common orthography to use.

Third, there was a growing dissatisfaction with the suitability of their previous literacy experiences for their present situation. The Thai script which some of them had learned in the refugee camps was no longer relevant to their needs in America. The old Romanized script, on the other hand, looked very odd next to English and certainly had very little transfer value to their new language. Furthermore, it was still viewed as the "Christian" script and was thus unacceptable in principle to many non-Christians.

The Portland conference was attended by some 30 representatives from Oregon, Washington, and California along with three Iu Mien-speaking missionaries as non-participants. The Association asked me to present a variety of orthographic options for their consideration. They wanted, among other things, a Romanized script which would be acceptable to all the Iu Mien whatever their religious preference and which would have optimal transfer value for learning to read English.

F. PARENTHESIS-ORTHOGRAPHY 6: CHINESE PHONETIC SCRIPT (1982)

Toward the beginning of the Portland conference there was a serious attempt to move in a completely different direction. Mr. Yao Heng Saeteurn, a representative from Seattle, had spent a considerable amount of time and effort developing an orthography based on the old Chinese phonetic alphabet. Yao Heng's script was attractive to many at the conference for several reasons. First, it was devised by an Iu Mien. Second, it was so different from any of the previous scripts that they thought it might provide a new beginning for all factions. And third, it provided a visible link with things Chinese. Furthermore, Yao Heng's script appeared to be linguistically adequate. Nevertheless, after considering the practical and pedagogical implications of such a script for the refugee communities, the representatives voted to proceed with a Romanized script.

G. DISCUSSION OF ORTHOGRAPHY 5 RESUMED

The script that the conference finally adopted eliminated the use of upper-case letters to indicate separate phonemes,⁷ used appropriate symbols for vowels instead of the three consonant symbols which had been used for that

⁷ This bears out Smalley's prophetic criticism of such scripts: "I believe . . . that attempts to use capital letters as having distinct values from their lower-case counterparts will often be abortive, and will be rejected in time by members of the group who gain wider education." (1963:15).

purpose, and included symbols to write Lao and Thai words (see Figure 4). Soon after the conference nearly 80 people began attending literacy classes in several cities on the West coast. The script received the blessing of two influential Mien in Thailand, one the most well-known Christian, the other the titular political leader of the refugees from Laos. And yet its future was clouded because of personal misunderstandings.

One of the Mien invited to the Portland conference but who did not attend was Mr. A⁸, a prominent refugee leader in California whose father and grandfather had been powerful leaders in Laos. During 1979-80, one of his relatives had, in Mr. A's stead, worked as a language informant for a linguistics class at a nearby university, and the next spring another relative was similarly engaged. One of the Americans who had worked on Mien with these two relatives of Mr. A then went to China for field research on Yao. Although only a phonemic transcription had been used during the informant sessions, this linguistic activity, together with the later China research by one of the linguists involved, led to suspicion about Mr. A's non-involvement with the 1982 orthography.

Suspicion of Mr. A increased when, not long after the Portland conference, he approached another linguist at the same university and asked him to evaluate the newly-adopted script to see whether it did, in fact, have sufficient transfer value with English. He also expressed his concern that the new script was perhaps another Christian script and that those who followed Mien traditional religion might come under some pressure to become Christians if they used it.

Mr. A's absence at the 1982 Portland conference, together with these two matters, was taken personally by several of the representatives, especially one of the conference organizers, Mr. L. The two men had known each other well in Laos and had worked together on refugee matters, but some disagreements had arisen between them. Not long after the conference, when word came from China that work on a Romanized script was underway, Mr. L and several of the delegates were quite upset. First, they assumed that Mr. A did not attend the conference because he had produced a rival script through his contacts with linguists at the university. They were offended that Mr. A had apparently gotten this script to China, the ancestral homeland, and that it was this script that the scholars there had referred to. It appeared to them that in doing this Mr. A was trying to assert his personal superiority. Why, they asked, should the activity of one person working with a couple of linguists be able to overturn the group decisions of some 30 Mien representatives working openly? Second, some felt that he had tried to reopen the divisive issue of religion and tar Orthography 5 with that brush, whereas his own would be neutral. In actual fact, however, there was no such rival script.

Despite the tension and uncertainty over Orthography 5 engendered by these misunderstandings, primers were retranscribed into the new script, and literacy classes got underway in several areas in France and the United States. Furthermore, it was adopted for use in the Phanat Nikhom refugee camp in

⁸ Names have been altered to protect the individuals concerned. No criticism of any of the persons involved is intended. This account is recorded simply for what it contributes to the point of the paper, namely that a non-linguistic factor such as interpersonal relationships can have a profound effect on the development of an orthography.

Thailand by the Consortium (The Experiment in International Living, World Education, and Save the Children). By late 1983, Mr. L and Mr. A had resolved their differences over the orthography and planned to discuss the issue at the Iu Mien New Year's celebration to be held in Portland in January 1984. Early in 1984, however, a new development took place which totally overshadowed previous discussions of orthographies and led to yet one more script.

H. ORTHOGRAPHY 7: ROMANIZED SCRIPT (1984)

Not long after the 1982 orthography conference, three of the Portland leaders sent a letter to China via an American tourist, hoping that it would somehow find its way to Iu Mien people there. In a remarkable way, the letter was taken to the Central Institute for Nationalities in Beijing, and ended up in the hands of the wife of one of the Yao professors there. The letter contained greetings, photographs of Iu Mien in the United States, and a copy of the new 1982 script.

One of the results of this letter and the correspondence that followed⁹, was that a mutual desire for a unified Iu Mien script emerged. Another was that the Chinese side adopted parts of Orthography 5 for use in their literacy classes.

In early 1984, the Portland Iu Mien received an invitation from the China professors involved with the Yao languages (Lakkja, Punu, and Mien) to visit them in order to establish personal contacts, meet Iu Mien and other Yao, and discuss orthography matters. Accordingly, an official delegation comprised of four Mien and me as their linguistic advisor visited China from April 24 to May 11. On May 8, an orthography conference was held in Ruyuan County, Guangdong, to discuss whether a unified script could be developed. In the end, the China side agreed to use most of the U.S. 1982 script except for the initial stops and affricates. They insisted that these 15 consonants follow the Chinese romanized system (*pinyin*) adapted for minority languages. The U.S. side accepted these changes even though the *pinyin* letters have much less transfer value to English. By the end of the day, agreement had been reached on nearly all of the existing differences (see Figure 5).

Upon returning to the United States, the delegation called for a second conference on Iu Mien orthography to discuss the results of the Ruyuan meeting. The conference was held in Portland on July 21 and was attended by some 85 representatives. After careful deliberation, they voted 78 to 7 to ratify the Ruyuan agreement and accept the unified script. Predictably, the area of greatest dissatisfaction was the set of initial stops. For a comparison of the three Romanized orthographies (2, 5, and 7), see Figure 6.

Since the second conference, the literacy primers have been retranscribed once again for use in classes held in a number of places, including the Phanat Nikhom refugee camp in Thailand. In China, literacy classes are being held in several areas, with special emphasis on the training of literacy teachers. Additional materials (e.g., dictionary, folktales, traditional songs, newsletter) in the new script are either being planned or are already in progress.

Nevertheless, some dissatisfaction remains among Iu Mien in the West. The initial stops have been accepted only grudgingly by some. Others are upset

⁹ One letter from China contained a three-page song text in a romanized script which the Portland group took to be a sample of Mr A's. However, it appears to be in a script similar to one developed by Chinese scholars for the Miao (Hmong) some 30 years ago (cf. Ma, 1957).

in general by the numerous changes of orthographies in recent years, particularly by the shift away from the 1982 script. Moreover, one very prominent Mien leader in Thailand was extremely disturbed over the 1984 script, assuming that it was the work of a Westerner, despite having been informed that it had been adopted by Iu Mien in both China and America only after considerable discussion. On the other hand, most are excited by contacts made with their people in the ancestral homeland and see the common orthography as having opened up possibilities for international communication and provided a somewhat more prestigious means whereby their language can be written and their ethnic identity preserved.

Since mid-1984, contacts between Iu Mien in the West and those in China have continued, both through personal visits and, significantly, through correspondence in the unified script. For example, in July 1985, I began to teach an adult literacy class for Mien in Long Beach, California. A few months later, I wrote in Mien to Mr. Pan Cheng Qian, a Mien scholar in Beijing who trains Yao teachers and who was the major spokesman for the China side when the 1984 script was negotiated, telling him what I was doing. Mr. Pan then wrote to my class who, with great excitement, composed their own letters to send to him. In the second round of letters, three of Mr. Pan's students added their letters to his. The letters from China have been copied and circulated among Mien in America, and the American letters, in turn, have been shared with classes in several provinces of China.

One outcome of the correspondence has been that both sides have seen that the script is not yet completely unified. Although the major points had been settled at Ruyuan in 1984, a number of differences have become apparent as a result of each side seeing how the other side actually uses the script. These differences include how tone change is indicated, contractions, punctuation, a variety of specific spellings, and the like. Some Iu Mien are now considering whether a third conference on orthography, attended by representatives from China, might be necessary. Despite these relatively minor differences, however, Orthography 7 is alive and quite well on both sides of the Pacific. The problem now is not how to devise a writing system; it is how to package literacy instruction so that it will be attractive to more people. In the West, motivation to read the unified script is still a problem outside of the Christian groups. It is also a problem for those over age forty, particularly those who are not already literate in Lao, Thai, or English. Literacy programs will need to deal with these and related issues, but the orthography itself appears to be settled.

VI. CONCLUSION

Berry (1958) sounds an appropriate note on which to conclude this paper. He observes that

an alphabet is successful in so far and only in so far as it is scientifically and socially acceptable. The two interests often conflict and it would be a fallacy to assume, as it sometimes is done, that the choice of an orthography can be determined solely on grounds that are linguistically or pedagogically desirable. (p. 737)

The search for an adequate, appropriate, and acceptable practical orthography for the Iu Mien may be nearing an end. In Thailand (except within the refugee camps) the Thai script has won out, but whether the final product will be the present transfer value script or a simplified script (or both!) remains to

be seen. For refugees in Thailand and the West and the Iu Mien in China, the script will be Orthography 7 -- unless the sporadic complaints against it in the West become more widespread and gain a new hearing for a return to Orthography 5.¹⁰

The road to these apparently final solutions has been a rough one indeed. Linguistic factors have not been the major problem, however. All six scripts devised since 1954 have been linguistically sound. Furthermore, at least five of them are pedagogically adequate and feasible in terms of practical factors. Instead, it has been sociolinguistic, cultural, and other factors--including migration and settlement patterns, religion, political developments, and inter-personal relationships--that have militated for so long against any one script becoming *the* Iu Mien orthography.

Figure 1: Mien Phonemes
(adapted from Purnell, 1965)

Consonants:	p	t	ts	c	k	ʔ
	ph	th	tsh	ch	kh	
	b	d	dz	j	g	
	f		s			h
	m	n		ɲ	ŋ	
	mh	nh		ɲh		ŋh
		l				
		lh				
	w			y		
	wh			yh		
Clusters:		Cw-,	Cy-			
Vowels:	i			u		
	e			o		
	æ	a	aa	ɔ		
Diphthongs:	ei,	ai,	aai,	ui,	oi,	ɔi
	iu,	eu,	au,	au,	ou	
	ia,	ua				
Tones:	1. mid	3. high		5. rise		
	2. fall	4. rise-fall		6. low		

(plus grammatically conditioned tone change)

¹⁰ This paper has not dealt with the literacy needs of the Iu Mien in Laos and Vietnam. However, given the present state of international relations in Southeast Asia, one might surmise that a non-romantized script based on Lao for the former and a romantized one based on Vietnamese for the latter would be preferred by the respective governments.

Figure 2. Thai Phonemes

I. CONSONANTS					II. TONES
MID CLASS (1 set of symbols)	p	t	c	k	4 tone marks plus ๑
	b	d			
HIGH CLASS and	ph	th	ch	kh	2 tone marks plus ๑
LOW CLASS (2 sets of symbols)	f	s		h	
	m	n		ŋ	
		l			
		r			
	w		y		

III. *Vowels* are written in 4 positions relative to the initial consonant(s), using either unit symbols as in A, or a combination of symbols as in B. Only long vowels and some diphthongs are shown for purposes of illustration.

A. Unit symbols :

(2)

ii

(1) ee/oo/ay [Consonant] (3) aa/oo

(4)

uu

B. Combination symbols :

1 + 3: əə, əw

2 + 3: əə, uə

1 + 2 + 3: iə, uə

Figure 3: Orthography 2

Consonants :	p	t	z	J	k	ʔ
	P	T	Z	Q	K	
	B	D	R	F	G	
	f	s				h
	m	n		E	v	
	M	N		H	V	
	w	l		Y		
	W	L		Y		
Clusters :	Cw-	Cu-	Cy-	Ci-	(depending on the next vowel)*	
Vowels :	i			u		
	e		r	o		
	c	a	æ	x		
Diphthongs :	ei, ai, aai, ui, oi, xi					
	iu, eu, au, aau, ru					
	ia, ua					
Tones :	1. (unmarked)		3. -q		5. -j	
	2. -b		4. -d		6. -g	

* A later revision of the script regularized the clusters to Cw- and Cy-.

Figure 4: Orthography 5

Consonants : p t ts jh k -q
 ph th tsh ch kh
 b d dz j g
 f s h
 m n ny ng
 hm hn hny hng
 w l r* y
 hw hl hy

Clusters : Cw-, Cy-

Vowels : i uea* u
 e ue* o
 ae a aa aw

Diphthongs : ei, ai, aai, ui, oei#, oi+
 iu, eu, au, aau, ou
 ie, ia, ua

Tones : 1. (unmarked) 3. -v 5. -x
 2. -h 4. -z 6. -c

Figure 5: Orthography 7

Consonants : b d z j g -q
 p t c q k
 mb nd nz nj nq
 f s h
 m n ny ng
 hm hn hny hng
 w l r* y
 hu hj hi

Clusters : None (treated as Cu-, Ci)

Vowels : i uea* u
 e er* o
 ae a aa or

Diphthongs : ei, ai, aai, ui, oei#, oi+
 iu, eu, au, aau, ou
 ie, uo

Tones : 1. (unmarked) 3. -v 5. -x
 2. -h 4. -z 6. -c

* Indicates letters which are used only to write non-Mien words (e.g., from Lao, Thai, Chinese): *r* = /r/, *uea* = / ʉa /, *ue* and *er* = / ɔ /.

+ Used to write the diphthong / ɔi / which, following the regular pattern, would have had the awkward sequence *awi* in Orthography 5 and *ori* in Orthography 8.

Used to write the rare diphthong / oi / so that the letters *oi* could be used for the much more common / ɔi /.

Figure 6: Comparison of Orthographies 2, 5, and 7

Orthography 2 : mavg tu'q siag nxm zxv siag nxm hxvd c'q, Rwrnj taaib,
 zyrug kxvq Buaj ninb sia'q Bua, Fav Nxi mivb za'q lyavj xij zu'g yetg
 Nxi za'q Tauj Revg.

Orthography 5 : Mange tuqv siac norm tsornq siac norm hornqz aeqv, dzuanx
 taaih, tsyouc kornqv buax ninh siaqv bua, jang hnoi mingh tsaqv lyangx.
 Oix tsuqc yietc hnoi tsaqv thaux dzengc.

Orthography 7 : Mange duqv siec norm zornq siec norm hornqz aeqv, nzuonx
 daaih, ziouc gornqv mbuox ninh sieqv mbuo, njang hnoi mingh zaqv
 liangx. Oix zuqc yietc hnoi zaqv taux nzengc.

REFERENCES

- Berry, J. 1958. "The making of alphabets." In J. Fishman (ed.), *Readings in the Sociology of Language*. The Hague: Mouton, 1968.
- Cushman, R.D. 1975. "The way-ward people: Taoism, ethnic identity and ethno-history among the Yao on the South China frontier." (unpublished manuscript)
- Kaplan, R.B. 1980. *On the Scope of Applied Linguistics*. Rowley, MA: Newbury House.
- Lemoine, J. 1982. *Yao Ceremonial Paintings*. Bangkok: White Lotus Co. Ltd.
- Ma, H-L. 1957. "A script for the Miao people," *China Reconstructs* 6, No. 7.
- Nida, E.A. 1947. *Bible Translating*. New York: American Bible Society.
- , 1954. "Practical limitations to a phonemic alphabet." Reprinted in W. Smalley and others, 1963.
- Pike, K.L. 1947. *Phonemics*. Ann Arbor, MI: University of Michigan Press.
- Purnell, H.C. 1965. *Phonology of a Yao Dialect*. Hartford: Hartford Seminary Foundation Studies in Linguistics 15.
- , 1975. "A survey of secular lu Mien (Yao) songs." (unpublished manuscript)
- , 1986. "The metrical structure of lu Mien secular songs." Paper read at the First International Colloquium on Yao Studies, Hong Kong.
- Savina, F.M. 1926. "Dictionnaire Français-Man, précédé d'une note sur les Man Kim-di-mun et leur langue." *Bulletin de l'École Française d'Extrême Orient* 26:11-255.
- Sjoberk, A.F. 1966. "Socio-cultural and linguistic factors in the development of writing systems for preliterate peoples." In W. Bright (ed.), *Sociolinguistics*. The Hague: Mouton.
- Smalley, W.A. 1976. *Phonemes and Orthography: language planning in ten minority languages of Thailand*. Canberra: Pacific Linguistics Series C, No. 43.
- Smalley, W.A. and others. 1963. *Orthography Studies*. London: United Bible Societies.
- Stubbs, M. 1980. *Language and Literacy*. London: Routledge and Kegan Paul.
- Trung, C.K. (translator). 1932. *Gospel of Mark in Yao*. Bangkok (?): American Bible Society.

THE WORD *nzi* IN GREEN MONG (MIAO)

THOMAS AMIS LYMAN

Professor Emeritus, Chulalongkorn University, Bangkok, Thailand

In view of the current interest in Mong/Hmong studies, I am submitting my translational definition of the word *nzi*¹, a highly productive form in the Mong Njua (Green Mong) language. This form is an extremely difficult one to translate and the present rendering will, therefore, prove of considerable value to Tibeto-Burman Area specialists.

My qualifications as a translator of the Mong Njua idiom are set forth in the following paragraphs.

Field-work with the Mong Njua was first begun in 1949 and later continued during the years 1962-64 and 1966-67 in Naan Province, northern Thailand. The mountain villages where I was able to contact the Mong Njua were mainly Baan Khun Sathan and Kang Ho. The first-named village lies on the mountain Doi Khun Sathan close to the provincial boundary-line between Naan Province and Phrae Province. The second village is situated on the Doi Phuu Wae plateau, District of Pua, northeast of Naan City and close to the Laos border.

After having achieved a thorough knowledge of the language and culture of the Mong Njua, I compiled two dictionaries, wrote one grammar, and published numerous linguistic and ethnographic articles on this folk-group. (I here wish to thank the Wenner-Gren Foundation for Anthropological Research, the American Council of Learned Societies, and the American Philosophical Society for their generous support.)

The Mong Njua (*mɔŋ njuá*) tribe is also known as the 'Green Miao' or 'Green Meo'. The name 'Miao' is derived from Chinese (*miáo*), while the Thai, Yon (Kam Myang), and Lao equivalent is 'Meo' (*เม็ว*).² The Tribe itself, however, uses the ethnic name Mong (*mɔŋ*) to which a descriptive term is added to designate the particular branch of the Mong group. In reference to themselves, the tribesmen used the designator *njuá* 'to be green'.

¹ The Green Mong symbols used in the present article have mainly the phonetic values given them by the International Phonetic Association. For exceptions and modifications, see T. A. Lyman: *English-Meo Dictionary*, 2nd Ed., Bangkok, Thailand, The German Cultural Institute (Goethe Institut), 1973.

² I use the Haas system of Thai romanization. See Mary R. Haas: *The Thai System of Writing*, American Council of Learned Societies, Washington, D.C., 1956.

Given below is my final and current translation of the Mong Njua word **nzi**. The examples of usage are mainly taken from my work *Dictionary of Mong Njua*.³

nzi (tủ) 1. lord (either man or deity), lady (either woman or deity); nobleman, noblewoman; ruler, monarch, king, queen; divine being, god. 2. honorific for 'father'. (NB: The word **nzi** is seemingly only employed in meanings (1) and (2) by the Mong Njua of Khun Sathaan village, Thailand.) 3. (as postposition to nouns) lordly, noble, royal; holy, divine, godlike. 4. (semantic extension of the foregoing) superior, high-ranking, powerful, strong; high (of status); elegant, fastidious. (NB: In meanings (3) and (4), this word is only vaguely analyzed as a free form by most Mong Njua and is considered mainly an integral part of the total noun phrase.) (In meaning (1) = Thai **cảw**)

Examples of Usage:

1. **áo.tủ nzi** two kings; two queens; two deities.
vâng.nzi ruler, monarch; deity. (< **vâng** 'king, queen')
 (NB: This is a tautological compound derived from a Chinese word.) (= Thai **phrá? cảw**)
2. **nzi, kủ sâng náo pả** Lord (Father), I want to eat cake. (NB: This sentence was not given by an informant but was overheard as part of a conversation between father and child in 1949. It is remembered by the author with distinct clarity.)
3. **yâng nzi** the Lamb of God ("godlike sheep"). (NB: This item was given to me by my friend the Rev. Donald Rulison who stated that it was volunteered by a Christian Mong as a translation of the Biblical phrase.)
4. **séng-nzi** the grandchild ("superior descendant" or "descendant of the first rank").
yâng.nzi pattern, example ("superior thing" or "elegant thing").
yéng.nzi drawing, picture, illustration ("elegant markings").

³ See T. A. Lyman: *Dictionary of Mong Njua: A Miao (Meo) Language of Southeast Asia*, Mouton and Company, The Hague and Paris, 1974.

SOME CLASSES OF CLASSIFIER IN IU MIEN (YAO)¹

Christopher Court
 Arizona State University, Tempe

Introduction. A classifier (Clf) is a noun (N) of a special grammatical type in Iu Mien, and quite generally in languages of the East Asian and Southeast Asian area, that co-occurs with numerals (Num). Semantically, Clf's serve to specify in some way, either qualitatively or quantitatively, a N (overt or implicit) or a verb (V) (see Figs. 1-3 below). We shall refer to Clf's of N's as N-Clf's, and to Clf's of V's as V-Clf's.² For many of the terms, and for the general plan of the exposition I am indebted to Matisoff (1973: Sec. 3.42). Syntactically, Clf's perform the function of enabling Num's to enter into constituency with N's and V's, while semantically, as specifiers, they individuate, instantiate, or otherwise make precise the reference of a N or V. Especially with N's characterized by homonymy or polysemy Clf's perform a very important disambiguating function, as can be seen from the following examples: (1) *homonymy*:-- /1 nóm diə / 'two pills' vs. /1 tséw diə/ 'two bolts of cloth' (</diə/ #1 'medicine' vs. /diə/ #2 'cloth'); (2) *polysemy*:-- /1 nóm tsyân-hò/ 'two hours' vs. /1 puən tsyân-hâ/ 'two minutes', /1 lít uəm/ 'two liters of water' vs. /1 tîw uəm/ 'two streams/ivers' vs. /1 pəw uəm/ 'two crossings of a stream/river.'³ As the example /1 pəw uəm/ shows, the Clf may not only narrow the sense of a N, but it may also add to the meaning of the NP as a whole -- in this case adding a deverbative meaning. It should also be noted that Clf's may fail to disambiguate: thus /1 tîw cuŋ/ means either 'two dragons' or 'two rainbows', and /1 tîw cāv/, either 'two trails' or 'two methods' (cf. English *two ways*).

As a final general remark on Clf's, it should be noted that finer points of Clf usage are very likely to vary from one locality or subdialect to another.

¹ The present paper is a somewhat reworked version of Court 1986, which is itself an abridgement and slight modification of materials in Court 1985. It has benefited from the input of Martha Ratliff, who was kind enough to read Court 1986 during the Conference at which it was presented, and to make helpful comments. The debt to Matisoff 1973 will be obvious. A great debt is also owed to Lombard 1968, and its editor, Herbert C. Purnell. Finally, a debt is owed to my informants, principally Fou-One Saephan and Mounq-Yoon Chow. Data has been gathered in Berkeley, California, from September 1979 down to the present. The name "Iu Mien," applied to both the language and the people, is commonly shortened to "Mien" by the people themselves, and, following them, by English-speaking people who are familiar with them. They are referred to as "Yao" by the ethnic Thais and Laos, but most linguists now follow Chinese usage in reserving that name, used alone, for a broader ethnic grouping, most of whose languages (or dialects) belong to a common stock, which linguists call collectively Yao, and which forms the Yao branch of the Miao-Yao language family. The subdialect of Iu Mien described in the present paper is substantially the same as that described in Downer 1961 and 1973, Purnell 1965, 1968a and b, and 1970 (s.v. Yao of Chiengrai, or "YCR"), as well as in Mao and Zhou 1962 (s.v. *Mian*), or the translation of the latter, Mao and Chou 1972 (s.v. *Mjen*).

² The idea that Clf's may occur not only in expressions modifying N's but in those modifying V's as well goes back at least as far as Noss 1964:106 s.v. *metric classifier*.

³ Transcription of Iu Mien follows Downer 1961 (which is a form of IPA), except that the palatal series is here transcribed "j", c, ch, ŋ, hŋ", glottal stop is here "+", and the tones (in Downer's and the Chinese linguists' conventional order of listing) (1) unmarked, (2) "ˆ", (3) "ˊ", (4) "ˋ", (5) "ˊˊ", (6) "ˋˋ".

The combination of Num + Clf I shall refer to, following Matisoff op. cit., as a quantity-expression (Q); a Q can quantify a N, as with the N's in the above examples: viz. /diə/ #1, /diə/ #2, /tsyāŋ-hò/, /uəm/ and /cáw/. The resultant expression is an NP of which the N is the head, and the head N (N_H) is, of course, a quantified N_H (N_{qh}). In Figs. 1-3 below are shown the constituent structures of the following sentences respectively: (1) /yie huə puə tâw (mîən) mîŋ/ 'We three go,' (2) /yia bua mîŋ puə tâw (mîən)/ 'We go as a threesome/There are three of us going,' and (3) /nfn dí? dyǎŋ 1 dzùn/puŋ/ 'He/She/It kicks the tree two times/on two sides'.

FIGURE 1. Structure of yie buə puə tâw (mîən) mîŋ 'We three go'. NP₂ and NP₃ are in apposition to one another.

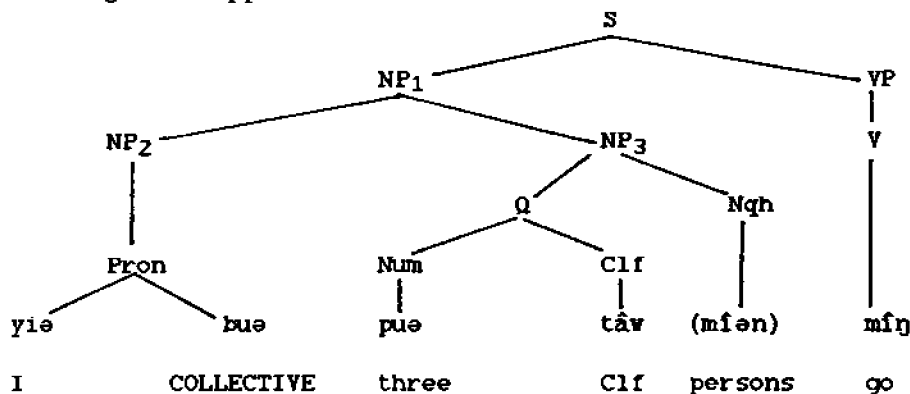


FIGURE 2. Syntactic Structure of yie-buə mîŋ puə-tâw (mîən) 'We go as a threesome (à trois)'.

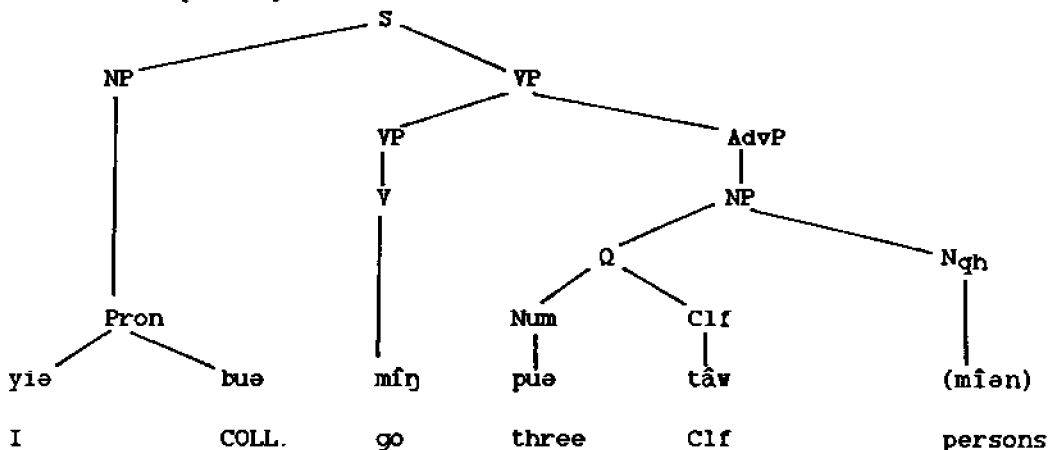
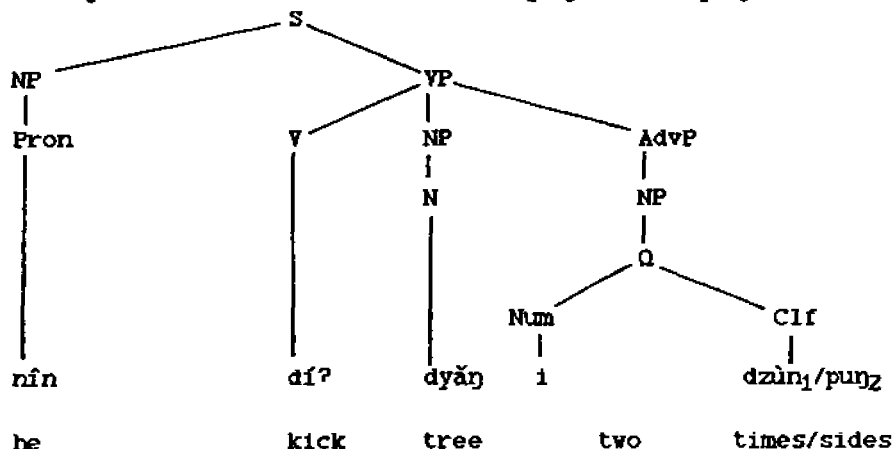


FIGURE 3. Syntactic Structure of *nɪn dɪʔ dyǎŋ 1 dzùn/puŋ*.

'He/She/It kicks the tree two times.' (Clf₁)

'He/She/It kicks the tree on two sides.' (Clf₂)

Etymologically, several Clf's have been traced back to Chinese, and it is very likely that further investigation would reveal a Chinese origin for quite a few more.

1. **Kinds of Clf's.** On formal or semantic grounds it is convenient to distinguish at least the following sub-types of Clf's.

1.1. **Auto-Clf's. (a-Clf's).**

1.1.1. **Noun auto-Clf's (N-a-Clf's).** Some N's may be their own Clf's. This is often the case in SE Asian languages, but Mien seems to contain remarkably few such. In fact, the present investigator has found only a handful, of which several refer to units of time, e.g. /hnɕy/ 'day' and /hɲǎ:ŋ/ 'year'. Furthermore, they seem to have the syntactic peculiarity that the homophonous N_h must be obligatorily deleted: thus one says, for instance, /1 hnɕy/ and not */1 hnɕy hnɕy/ for 'two days'. This contrasts with analogous expressions in, for instance, Thai, such as /khon sɕɔŋ khon/ ⁴ (N_h Num Clf) 'two people,' where homophory between Clf and N_h is not only permitted, but quite common. Possible reasons for the non-occurrence of expressions with homophonous Clf and N_h in Mien are (a) the two elements are not, as in Thai, separated by the intervention of the Num, so that the result might sound awkward, and (b) Mien, unlike Thai, has the grammatical feature of reduplicated Clf's in the meaning of 'every' (see Sec. 2. below), albeit with tone-sandhi, and homophonous Clf + N_h expressions might be avoided because of their similarity to these reduplicated Clf expressions. From a transformational perspective it is possible that what is going on here is the deletion not of the homophonous N_h but of the Clf, but we

⁴ Transcription of Thai follows IPA usage, except that the tones are marked as in Haas 1964.

will not attempt to settle this question here. Some other N-a-Clf's are /puŋ/ 'direction; side', and /pùŋ/ 'kind; manner; way'.

1.1.1. Familial N-a-Clf's of concerted involvement. These are an interesting kind of N-a-Clf, also found in other languages in the area. As human beings, kinsmen normally take the specific Clf's for human beings (see Sec. 1.1.3. below), but when certain combinations of kinsmen are living or acting together as a group certain kinship terms may be used as a-Clf's.

Thus one says /i mūə/ 'the two siblings (together)' (cf. Thai /sǒŋ pʰiː-nǒːŋ/ '(id.)', literally "two elder sibling-younger sibling"), /i tɔn-màː/ 'mother and son (together)' [literally "two son-mother": cf. Thai /sǒŋ mēːlǔːk/ 'mother and child (together), literally "two mother-child", and its analogue /sǒŋ pʰsː-lǔːk/ 'father and child (together)']. But in order to say in Mien 'father, mother and child together' one resorts to the prototypical familial a-Clf of concerted involvement, /hmwán/. This can be glossed for convenience as 'family', but in order to understand its behavior as an a-Clf, it is better translated as 'family co-members, comprising at least the nucleus of husband and wife': thus, /i hmwán/ ('two hmwán's') means 'married couple; husband and wife (together)'.

This then brings us to the way in which one says 'father, mother, and child (together)', viz. /puə hmwán/ ('three hmwán's'), though one can also say, periphrastically, /i hmwán tshǎw tâv tɔn/ (literally "two hmwán's and [one] Clf son").

1.1.2. Verb-Autoclassifiers (V-a-Clf's): homophonous cognate objects. In some cases, evidence suggests that a V may be its own Clf. Thus, the V 'to kick' is /díʔ/, and as in Fig. 3. above, in order to say 'He kicks the tree two times/He gives the tree two kicks', one may say /nín díʔ dyǎŋ 1 dzùn/ ("He kicks tree two times"); using the Clf /dzùn/ 'time(s) /occasion(s)', one may, on the other hand, also say /nín díʔ dyǎŋ 1 tsǎw-díʔ/ ("He kicks tree two foot-kicks") in order to convey the same meaning. In these two sentences, the word /dzùn/ in the first, and /tsǎw-díʔ/ in the second, have the function of enabling the V /díʔ/ to be numerically specified, and hence are V-Clf's (see Sec. 1.2. below). Now in the second sentence, the syllable /díʔ/ of the V-Clf is an exact echo of the V, but it is used not alone, but in composition with the morpheme /tsǎw/. If it did occur alone, i.e., in the sentence of the same meaning */nín díʔ dyǎŋ 1 díʔ/, it would be an example of a V-a-Clf. The only cases found by the investigator in which the Clf occurs alone and is a replica of V are cases where it is functioning as a measure Clf (see Sec. 1.5. below).

1.2. Verb-Classifiers (V-Clf's). A diagram of the syntactic structure in which V-Clf's occur is given in Fig. 3. above. What is happening is that they are occurring in NP's dominated by adverbial phrases (AdvP), i.e., in NP's that are functioning adverbially. They are in fact what are traditionally known as *cognate objects*, and Chao (1968:312ff) uses this latter term with reference to Chinese. As Chao notes, cognate objects may occur as structures indicating the number of instances of an action, its duration, extent, direction or destination (see also Court 1985: Chap. IV). But we can go beyond cognate objects and argue that the Clf's in Fig. 2. above, which indicate the number of participating *subjects*, might also be viewed as V-Clf's (the possible presence of a N_h is a problem for that view, and we shall not dwell on it here). A few V-Clf's that are

not a-Clf's are: /pəw/, for strokes of an ax; /pəw/, for steps; (cf. /bɪə/ below); /puŋ/ for directions, sides; /bɪə/ for steps or strides, stitches; /dzùn/ for times, occurrences, instances; and /rùŋ/ for manners, ways, modes. The class of V-Clf's, as well as the other classes of Clf's that we are setting up show some overlap.

1.3. Specific Clf's. These are the "classifiers" par excellence, that have, in principle, to be memorized separately for each N in the language. However, to describe them thus is an overstatement for at least two reasons: (a) N's referring to non-discrete entities take *measure* Clf's that are, in many cases, determined not by the identity of the N, but by real-world circumstances, i.e., the object or action used to do the measuring, and (b) specific Clf's tend to refer not so much to particular N's as such as to classes of real-world entities, such as human beings, animals, long thin objects, and so on. Having said that, however, we must, on the other hand, note (a) that some specific Clf's are used for very restricted classes of things, such as /sím/ used only for lengths of thatched mats (Lombard 1968:341; other kinds of mats have the Clf /khwáy/), (b) with other Clf's it is hard to see what the things classified have in common: e.g., /phəw/, for boats and scissors (Lombard 1968:341), and (c) there is a Clf /nəm/, of very broad use (see Sec. 1.4.) that seems to be the unmarked Clf for things in general. Some examples of specific Clf's are /təw/ for human beings, higher animals, ghosts and spirits (in some cases), and doors; /tɪw/ for long narrow objects, animate or inanimate, literal or metaphorical, including snakes, dragons, rainbows, rivers, roads, stories, songs, lives, and minds; /pěŋ/ for pencils, guns, and other long, thin objects (my informants said that the objects classified were not as long as those classified by /tɪw/, and it should be noted that /pěŋ/ also exists as an ordinary N meaning 'handle'); /chəŋ/ for rooms, narrow fields, small spaces; /dzoŋ/ for fields to be cultivated; /pha:n/ for mosquito nets; /tɪw/ for reasons, tasks, Elements (of the Four Elements); /pyáv/ for families, households (also exists as an ordinary N meaning 'house'); /tsun/ for knives, clumps, clusters. Although some of the specific Clf's also exist as ordinary N's, they are not N-a-Clf's, for they are not used to classify those N's: thus the specific Clf for the ordinary N /pyáv/ 'house' is not /pyáv/ but the broad-purpose specific Clf /nəm/ (see following section).

1.4. The broad-purpose Clf /nəm/. Like many languages with systems of N-classification, Mien has one specific Clf of very general scope, viz. /nəm/. The functionally corresponding Clf in Mandarin is gè⁵ in Thai /ʔan/, and in Lahu /mə/. In some languages the broad-purpose Clf is fairly freely substitutable for more specific Clf's, but this seems to be much less the case in Mien, and it certainly differs from Mandarin gè and Lahu /mə/ in not being applicable with human beings. Nor is it used with higher animals. But it is used with lower animals, and a whole array of material and immaterial entities: for instance, while it is true that fishes require /təw/, /nəm/ is used with birds, (including chickens) tortoises, crabs, shrimps, ants, flies, mosquitoes, and insects in general, fruits and vegetables (referring to individual pieces of fruit and individual vegetables), medicine (referring to individual pills), ball- and grain-shaped things (perhaps its original sphere of usage), cups, bottles, boxes, bags.

⁵ Transcription of Mandarin forms is in the pinyin system.

sacks, animal horns, teeth, hammers, vehicles, houses, shops, beaches, places, countries, forests, corners, barriers and checkpoints, written characters, names, timepieces, watches of the night, months, seasons (which can also be auto-classified), souls and certain kinds of ghosts (others take /tâw/), customs and traditions, sorrows (there may be some overlap with another Clf. /thǝŋ/, here), etc.

1.5. Measure Clfs. These are used to quantify N_n 's referring to entities that are, or can be viewed, as non-discrete, such as liquids, particulate substances, areas of land, stretches of road, and so forth, i.e., things that are typically designated by mass nouns in English. Except for "abstract" measures, introduced from other languages, such as the Chinese ounce, the liter, the meter, and so forth, they are mostly "concrete," i.e., derived by homonymy from either the N for the *thing* which measures, or the V for the *action* which measures. For instance, /pɿən/ is an ordinary N meaning 'basin.' Now basins are discrete entities, and as (empty) items of equipment they can be counted: /pɿən/ in this case is a N_n and takes as its specific Clf /nɔm/. But when rice is being measured by the basinful, /pɿən/ (or its homonymous derivative) is used as a measure Clf: thus /i pɿən byâw/ 'two basin(fulls) of rice.' It is probable that the name of any container can also serve as a measure Clf (as, for instance, in Thai).

On the other hand we have V's of action from which measure Clfs are homonymously derived. Thus there is a V /phvǎŋ/ 'to scoop up in two hands' from which is derived the measure Clf meaning 'a double handful,' as in /i phvǎŋ byâw/ 'two double handfuls of rice.'⁶ Some measure Clfs that are not "concrete" in the sense just defined are provided by nature itself rather than human intervention, such as /hnɔy/ 'day,' and /hŋǎ:ŋ/ 'year'.

1.6. Group Clfs. Clfs referring to pluralities, aggregates or collectivities of individual entities are what is meant by the term group Clfs. Some examples are /pùəʔ/ for groups of people, /duy/ for piles of things, /kwân/ for herds, crowds, groups, /tɔ/ for bunches or clusters. It seems to be the definitive characteristic for group Clfs that with them the Num /yɿət/ 'one' takes on the meaning of 'all' or 'the whole': thus /yɿət pùəʔ mɿən/ 'the whole group of people.'⁷

1.7. Round-number Clfs. These are the Clfs that refer to powers of ten: viz. /tsɿəp/ 'ten,' /péʔ/ 'hundred,' /tshin/ 'thousand,' /wà:n/ 'ten thousand,' /sɿn/ 'hundred thousand,' /lâ:n/ 'million.' Of these /tsɿəp/ and /péʔ/ are distinguished by taking a special set of Chinese-derived Num's rather than the native stock.

1.8. Quasi-Clfs. There are certain words which fail to meet the criterion for being Clfs, yet resemble them in certain respects: for instance they may be usable only with the Num /yɿət/ 'one,' or they may be reduplicable in the same way as Clfs (see next Sec.), but not be usable with Num's. They require further investigation.

2. Reduplication of Clfs. The only use of Clf reduplication that the investigator has discovered is to convey the meaning of 'every.' The first

⁶ Cf. Thai /kɔ:p/ '(kl.)'

⁷ Actually, other Clfs may take /yɿət/ in this sense, provided that their referent may be viewed as divisible: e.g. /yɿət hnɔy/ 'all day.'

constituent of the reduplicative construction undergoes tone sandhi: e.g., /mêy hnɔy-hnɔy tâ:y náy/ 'You come here every day' (lit. "You day-day come here"). Other examples are /tâv-tâv/ 'everybody,' /nɔm-nɔm/ 'everything.' The reduplicated Clf may be followed by its N_h: e.g., /tâv-tâv mlən/ 'everybody.'

3. Conclusion. This completes our survey of some aspects of Clfs in Iu Mien. Further details about the grammatical structures in which they occur are found in Mao and Zhou 1962 (translated in Purnell 1972: 239-255), and in greater amplitude and technical detail in Court 1985: Sec. 3.11 and passim. A substantial list of Clfs is found in Lombard 1968: 341-343. This list is somewhat expanded, and categorized according to the scheme of the present paper in Court 1985: Secs. 3.4.2.-3.4.218.

REFERENCES

- Chao, Y.R. 1968. *A grammar of spoken Chinese*. Berkeley and Los Angeles, University of California Press.
- Court, C. 1985. *Fundamentals of Iu Mien (Yao) grammar*. UC Berkeley doctoral dissertation available from University Microfilms: Ann Arbor, Mich.
- , 1986. "Some observations on classifiers in Iu Mien (Yao)," paper read the Minority Nationalities of China, Santa Barbara, California, January 26-29, 1986.
- Downer, G.B. 1961. "Phonology of the word in Highland Yao," *Bulletin of the School of Oriental and African Studies* 24.3, 531-41.
- , 1973. "Strata of Chinese loanwords in the Mien dialect of Yao," *Asia Major* 18.1, 1-33.
- Haas, M. 1964. *Thai-English Students' Dictionary*. Stanford University Press.
- Lombard, S.J. 1968. *Yao-English Dictionary*. Data Paper: Number 69. Southeast Asia Program. Department of Asian Studies, Cornell University.
- Mao Zungwu and Zhou Zuyao. 1962. "Yaozu yuyan galkuang" ('Sketch of the languages of the Yao people'). *Zhongguo Yuwen* 114, 141-147.
- , 1972. "A brief description of the Yao language," in Purnell 1972, 239-255 (a translation of Mao and Zhou 1962).
- Mao Zongwu, Meng Chaoji, and Zheng Zongze. 1982. *Yaozu yuyan jianzhi* ('Brief account of the languages of the Yao people'). Beijing: Minzu Chubanshe.
- Matisoff, J.A. 1973. *The Grammar of Lahu*. Berkeley, Los Angeles, London: University of California Press.
- Noss, R.B. 1964. *Thai Reference Grammar*. Washington, D.C.: Foreign Service Institute.
- Purnell, H.C. 1965. *Phonology of a Yao Dialect*. *Hartford Studies in Linguistics* No. 15. Hartford: Hartford Seminary Foundation.
- , 1968a. Editor's introduction in Lombard 1968:x-xvi.
- , 1968b. Appendix A: Numbers, in Lombard 1968:317-320.
- , 1970. *Toward a Reconstruction of Proto-Miao-Yao*. Cornell University doctoral dissertation, unpublished.
- , editor. 1972. *Miao and Yao Linguistic Studies*. Selected articles in Chinese, translated by Chang Yu-hung and Chu Kwo-ray. Data Paper Number 88. Southeast Asia Program, Department of Asian Studies, Cornell University, Ithaca, New York.

**A comparative look at
 Yao numerical classifiers**

Bruce R. Caron
 M.A., University of California, Berkeley

Summary

This paper extends Donald R. Goral's (1978) descriptive analysis of numerical classifiers (CLs) found in several Southeast Asian languages to CLs found in Yao. What is presented is a necessary first step in our understanding of the syntax and semantics of Yao CLs. As is perhaps expected, Yao CLs are in many ways similar to the CLs found in neighboring languages. However, there are certain syntactic and semantic features that are peculiar to Yao. While the syntax of Yao CLs is relatively straightforward, they represent a rich and still not completely explored arena for semantic analysis. The appendices list eighty-one Yao CLs along with main selectional criteria and a beginning attempt at semantic classification. These lists should allow further comparisons between Yao CLs and those of neighboring languages.

Description of the data

These data were gathered from a native speaker of Yao who grew up in Northern Laos, in the village of Sam Sao, Houa Khong province.¹ This dialect is located in the same region as the Yao who were the sources for Lombard and Purnell's *Yao - English Dictionary*. These people, the Highland Yao, are distributed not only through Northern Laos, but have also migrated into Northern Thailand, and live in their greatest numbers in Southeast China. This dialect has seven tones - five live tones, and two dead.² The following symbols will be used for these tones:

Live Tones:	Dead Tones:
High-level: 55 □	High-level: 5 □
Mid-level: 33 □	
Low-level: 11 □	Low-level: 1 □
Falling: 51 □	
Rising: 15 □	

I SYNTAX

I-A General remarks on Yao syntax

Before we turn to Yao classifiers, some general remarks on Yao syntax are in order. These remarks will be limited to the basic constituent order, and areal syntactic features.

Word order in Yao is unique among the Southeast Asian languages. Like Thai, Yao uses prepositions and noun-adjective constructions, and like Chinese, it puts the head noun after the genitive attribute (see Figure 1 following; cf. Goral, Appendix 1; Purnell 1972, pp. 13-25).

Figure 1:
Sentence word order in Yao, Chinese, and Thai

Yao (and Meo)	SVO	Prepositions	Genitive-N	N-Adjective
Chinese	SVO	Postpositions	Genitive-N	Adjective-N
Thai	SVO	Prepositions	N-Genitive	N-Adjective

The following examples illustrate Yao word ordering:

SVO: y³³iə p¹⁵iin y³³iə n̄³³ɛɛ n̄⁵¹aan
 I{S} hide{V} [I+poss.] money{O}
I hid my money.

Genitival: n⁵¹in n̄³³ɛɛ h⁵¹ee h³³aan
 {he+poss.} shoe lace
his shoelace

Adjectival: c⁵⁵uu k⁵⁵iə
 dog black
black dog

Prepositional: k¹¹a-d⁵⁵iə m⁵⁵iə
 in/under grass
in the grass

Major Southeast Asian areal syntactic features listed in Matisoff (1978:25) are well represented in this dialect of Yao:

1.) Reduplication as an adverbializing process:

c¹5aaw s¹i? red egg
 c¹5aaw s¹i? s⁵i? very red egg

2.) Sentence final particles for expressing "propositional attitudes", i.e. emotions of the speaker. These particles are frequent in Yao, and are quite difficult to fully describe. Particles such as ¹a?, ñ³³εε, and k³³ɔɔ occur at the end of declarative sentences. Their precise function has not yet been fully assessed.

3.) Aspect, not tense, as the primary inflectional category for verbs:

55m k⁵1uu-k¹5oo ts⁵⁵əw c⁵⁵iən nh¹5aan
 the old woman to cook(+prog.) rice
The old woman is/was/will be cooking rice.

4.) Free topicalizability of noun-phrases, with a tendency towards topic prominence. While there is not sufficient space in this piece to fully demonstrate topicalizability in Yao, the following example shows a representative topicalized sentence:

w³³uə 33ii t⁵1aw m⁵1iən y¹iət t⁵1aw [m⁵1iən] ʔ⁵uət 33ii t⁵1aw py⁵⁵aa
 those two CL people one CL {deleted} break two CL sticks
As for those two people, each one broke two sticks.

5.) The nominalization of whole sentences, without their being embedded into any larger unit. The Yao particle ñ³³εε, which appears in genitive and relative clause constructions, may also appear at the end of a sentence, where it has a nominalizing function.

In summary, Yao shows the major areal syntactic features common to Southeast Asian languages, yet its basic word order is unlike that of any other language in the area, except for its sister language, Meo.

I-B. A syntactic definition of Yao numerical classifiers

Goral's syntactic definition of Vietnamese numerical classifiers is useful in describing Yao classifiers. In Yao, a classifier is *any lexical item which can appear in the frame Num + ___ + N, where Num=numeral and N=noun* (cf. Goral: p.6). Yao CLs constitute a subclass of nouns which appear in the above frame to specify a certain attribute of the following noun. Through this explicitly stated attribute the following noun can be counted. Furthermore, without the presence of the CL, the noun cannot be counted. Thus, some CLs not only can appear in the above frame, but must appear

(except for a very few exceptions to be explained below). For this reason, the above definition is useful in predicting the obligatory occurrence of CLs (and it is useful in eliciting CLs). However, this does not fully describe the range of noun-phrase frames where CLs will occur, nor the process of selecting CLs. Of course this definition is a wholly syntactic one, and CLs are primarily semantic in their function. In short, this definition is certainly very useful in eliciting which lexical items can and cannot be used as classifiers. So it approaches descriptive adequacy. It does not, however, define the entire distribution nor the selectional motivation for CLs, i.e. it does not provide explanatory adequacy. Goral's definition does make one thing certain concerning the distribution of CLs: *their obligatory presence in numerative noun-phrases.*

There are two further stages in the analysis of CLs in any language beyond a syntactic definition such as the above. The first stage is a functional semantic description of CLs.³ Studies show that CL systems can be very complex in the interaction between their syntax and semantics, and that the semantics of CL systems can be highly culture-specific. The second stage is more problematic: it is doubtful that a fully adequate description of numerical classifiers can be made without recourse to a satisfying explanation for their existence. Why are they found in such obligatory profusion in certain languages and not at all or vestigially in others? Such an explanation rests uncomfortably on extra-linguistic (psychological, philosophical, etc.) evidence. Until someone uncovers the reason(s) for the existence of complex CL systems and their semiotic function, CLs will continue to evade any explanatorily adequate analysis. This caveat is not given to excuse the shortcomings of the following analysis of Yao CLs, but merely to frame the information provided in the context of the task that lies ahead. For the purposes of this paper, Goral's definition of CLs, and some further notes on the syntactic and semantic features of various Yao CLs, will allow for comparisons between Yao and other Southeast Asian languages.

I-C **The basic structure of the Yao numerative phrase**

The following schema (Figure 2) shows the basic structure of the Yao numerative and/or demonstrative noun-phrase.

Figure 2:
Basic structure of Yao numerative phrase

+ preposition	+ Demonstrative	+ Number	+ CL	N O U N	+ A D J E C T I V E
		- Number	+CL		
- preposition	+ Demonstrative	+Number	+ CL		
+ preposition	- Demonstrative	+ Number	+ CL		
		- Number	- CL		

Notice that Yao preposes demonstratives and numeral+CLs, but postposes adjectives. By this, Yao differs from all other Southeast Asian languages (See Jones p. 6). Thai postposes demonstratives, numeral+CLs, and adjectives. Meo preposes numeral+CLs and postposes adjectives and demonstratives. Modern Chinese preposes demonstratives, numeral+CLs, and adjectives (See Fig. 3 below).

Figure 3: 4
Comparisons of numerative phrase structures

Yao:	Dem	Num + CL	NOUN	Adj
Chinese:	Dem	Num + CL	Adj	NOUN
Meo:	Num + CL	NOUN	Adj	Dem
Thai:	NOUN	Adj	Num + CL	Dem
Lahu:	NOUN	Adj	Dem	Num + CL

The following Yao phrase shows the complete basic numerative phrase structure:

n⁵⁵əy-t¹⁵əy py³³ee t⁵¹aw c⁵⁵uu c⁵¹ə?
 these four CL dog black
these four black dogs

In Yao prepositional phrases, the CL does not obligatorily occur after a demonstrative. A Yao noun phrase can thus exhibit the following structure:

PREP [+Dem.] [-Num.] [-CL] NOUN

In all other NPs, a demonstrative must be followed by a CL. It should also be noted that the number *one*, y¹iət, can be omitted from the noun-phrase, leaving a structure such as the following:

[-Dem.] [-num] [+CL] NOUN

e.g. t⁵¹aw d⁵¹aw-m¹⁵aaw
 CL tiger
one tiger

Apart from the understood number *one*, the head-noun of an NP can never be modified by a classifier without the co-occurrence of either a numeral or a demonstrative.⁵ The role of the CL in prepositional phrases is apparently linked to the definiteness of the phrase and to the emphasis placed on the phrase.

below) that this problem is really apparent. In D1, the CL $dz^{11}ɔm$ is a temporary measure CL derived from the noun $dz^{11}ɔm$, a Yao word for "bowl". One could argue that the second occurrence of $dz^{11}ɔm$ in D1 is, in fact, the noun $dz^{11}ɔm$ and not the CL. The majority of CLs are not homophonous and hyponymous/synonymous with nouns. Yet they do readily occur in similar anaphoric constructions. This supports an argument that perhaps all such anaphora are CLs, and not derived nouns.

I-E The issue of repeaters and CL+CL constructions

In Yao there are no repeaters or partial repeaters. Repeaters and partial repeaters are plentiful in languages such as Thai and Lahu, where the noun precedes the Numeral+CL. In fact, none of the languages of the area which place the Numeral+CL before the noun contain repeaters or partial repeaters. As Goral (*op cit*) posits, "in the languages with numerative phrases of the form Num+CL+Noun: Vietnamese, Chinese, and Indonesian, [and also Yao,] there are no repeaters, supporting the hypothesis that CL+CL sequences, or just sequences of identical morphemes, are generally discouraged" (p. 55). While Yao has no repeaters or partial repeaters, it does permit CL+CL constructions. In Yao, CLs are reduplicated to provide the meaning "every". For example:

$\bar{n}h^{51}a\bar{a}ŋ \bar{n}h^{51}a\bar{a}ŋ$	<i>every year</i>
$t^{51}aw t^{51}aw$	<i>everyone</i>

The reduplicated CLs are in close juncture, producing tone sandhi on the tone of the first syllable(s) of the reduplicated pair(s). The only tones permitted for the first syllable of a reduplicated CL are: falling, if the syllable is live, and low, if the syllable is dead. For example:

$nh^{51}ɔy nh^{33}ɔy$	<i>every day</i>
$t^{51}ɔy t^{11}ɔy$	<i>every generation</i>
$dy^{1ep} dy^{5ep}$	<i>every drop</i>

This tone sandhi dissimilates many of the reduplicated pairs of CLs, but dissimilation is not the underlying factor here, as falling live tones and low dead tones are reduplicated without dissimilation. So Yao does not support the hypothesis that "CL+CL sequences, or just sequences of identical morphemes, are generally discouraged" (*op. cit.*).⁶

I-F Other uses of certain CLs

Certain CLs occur in constructions other than the basic NPs outlined above in I-C. In particular, certain CLs are used to quantify actions, i.e. verbs or deverbal nouns. A few nouns serve as their own CLs, and many nouns can serve as temporary measure CLs.

Yao has a CL dz¹¹un, which is used to count events or repetitions of an action. The semantic difference between an "event" and a "repetition" is the following. An event is a *discrete occurrence*. An event may include a repetitive action. Repetitions are re-occurrences of an action (which may itself be complex) within a *single event*. A syntactic distinction is found between these two uses of dz¹¹un. The following Yao phrases illustrate this:

b ⁵ ɔʔ ʒʒii dz ¹¹ un	<i>hit twice</i>	2 EVENTS
ts ¹⁵ əw ʒʒii dz ¹¹ un b ⁵ ɔʔ	<i>make two hits</i>	REPETITION
h ⁵ op ʒʒii dz ¹¹ un	<i>drink twice</i>	2 EVENTS
ts ¹⁵ əw ʒʒii dz ¹¹ un h ⁵ op	<i>take two sips</i>	REPETITION

To "drink twice" is to drink on two separate occasions, while to "take two sips" is to repeat an action in a single event. As these examples show, when the CL dz¹¹un is used to count events, it follows the verb it modifies, as do other adverbs in Yao. When used to count repetitions, the CL precedes a deverbal noun, which it modifies, e.g. dz¹¹un b⁵ɔʔ "two hits."⁷

Certain nouns in Yao serve as their own classifier. This means that they occur in numerative phrases without a(nother) CL. These nouns are few, and are restricted semantically to a set of temporal nouns. The only self-classifying nouns that I have found in Yao are the following two: nh³³ɔy "day", and nh⁵¹aag "year", e.g.

ʒʒii nh ³³ ɔy	<i>two days</i>
ʒʒii nh ⁵¹ aag	<i>two years</i>

These noun/CLs are distinct among temporal nouns as they measure quantities of time which are empirically salient and non-conventional. Minutes, hours, weeks, etc., are standardized measures which subdivide natural units of time: days and years. Such arbitrary measures require classifiers. That certain units of time should be perceived as natural, and therefore require no further numerical classification, brings up the question of the need for classifiers to unitize objects such as houses, persons, trees, etc. Many languages take the single instance of an object as its natural measure, i.e. these create a subset of singular nouns which are then unmarked when classified as instances ("two trees"="two instances of trees"). Yao and many other languages do not have an unmarked set of nouns, except for these few self-classifying temporal nouns. Perhaps then,

the *single instance* of an object is not preeminently salient as a unitizing category in Yao (although it is in English). Only in the few cases where the single instance *is the only empirically relevant unit*, e.g. "day", "year", does this unmarked (or self-marking) appear in the Yao numerative phrase.

Many nouns can serve as temporary measure CLs. They are simply *ad hoc* measures for things. For example:

ʒʒi tʂiə sʒəw
 two table/CL book
two tables of books

The use of these CLs is determined closely by the context. Temporary measure CLs demonstrate the creative potential for CLs in Yao. Perhaps many Yao CLs were originally temporary measure CLs and have subsequently become conventionalized.

I-G Summary of a Yao CL syntax

In Yao, CLs occur in the frame "Num + CL + Noun." They serve anaphorically in conjunction with at least one determiner. CLs can be reduplicated to mean "every." Some CLs can modify verbs. Some nouns are self-classifying. Nouns can serve as temporary measure CLs. Finally, there is some still-unanalyzed connection between CLs and the definiteness of the NP (and between the definiteness of the NP and its possible topicalization).

II. SEMANTICS OF YAO CLS

II-A The semantic structure of Yao CLs

There is no simple semantic ordering to be found for Yao CLs. As Goral (p. 31) warns, "the evidence suggests that CL systems are combinations of structured subsystems and isolated, idiosyncratic relations." This is very much the case with Yao.

These features demonstrate the complex patterning of CLs within even general semantic domains. The semantic complexity of Yao CLs probably prevents an attempt at a componential analysis of the whole system. Even general componential features, such as "animate" or "human," cannot be imposed on the entire set of Yao CLs. For example, tʂ¹aw is a CL for humans, [+animate, +human]; for animals, [+animate, -human]; and for doors, [-animate]. So tʂ¹aw can classify nouns which *are or are not* human and/or animate.

It seems imprecise to even describe Yao CLs as forming a single, internally-ordered semantic system. Yao CLs are at most a set of lexical items with certain defining syntactic properties, expressing a variety of semantic relations. The central function of these relations is to provide a unit reference for nouns. Secondarily, CLs provide definiteness to nouns.

and they emphasize topically salient features for the referents of the nouns they modify. CLs also function to disambiguate potential homophones, a particularly useful function in monosyllabic languages such as Yao.

If Yao CLs do not themselves constitute an ordered system, it should be remembered that Yao CLs, as a set, refer to the entire range of countable nouns in Yao, and that as a whole, these do not form a single, semantically-ordered system. Furthermore, unlike Burmese (see Becker), where CLs provide an order for nouns, Yao CLs are attached to their nouns with varying degrees of conventionality and meaningfulness. They do not order nouns as much as they are disordered by the nouns they modify. While Yao CLs do not collectively systematize the nouns they modify, many Yao CLs individually describe subsets of nouns which are ordered according to some common salient feature.⁸ A few remarks on the selectional criteria for CLs, and the study of CLs in general, will conclude this paper.

II-B The selectional criteria for Yao CLs

In Yao there are a variety of criteria for pairing any one CL with a particular noun. Some selections are semantically unmotivated and purely conventional. These CL-Noun pairs are simply memorized (e.g. #67 in Appendix I). Other CLs are matched with nouns according to some topically salient (physical) feature of the noun's referent object (e.g. #17 in Appendix I). Then there is $n^{33}om$.

The CL $n^{33}om$, used for birds, hours, and weeks, serves also as a general CL for any non-human object which is not properly modified by another CL. All objects which are alien to Yao culture can be modified by $n^{33}om$. There is some evidence that $n^{33}om$, as an unmarked CL, is beginning to replace other CLs in this dialect of Yao, or at least that $n^{33}om$ can be used when the appropriate CL cannot be remembered.

Because CLs often focus on a single, topically salient feature of the modified noun's referent, CLs are sensitive to the topic of discourse. For example, the same two people might be referred to in the following ways:

w ³³ uə ʒʒji t ⁵¹ aw m ⁵¹ iən	<i>These two people [are people]</i>
w ³³ uə ʒʒji ty ⁵⁵ aa m ⁵¹ iən	<i>These two people [belong to the same family]</i>
w ³³ uə t ¹⁵ ɔy m ⁵¹ iən	<i>These [two] people [form a couple]</i>
w ³³ uə h ⁵¹ əw m ⁵¹ iən	<i>These [two] people [are a married couple]</i>
w ³³ uə mh ^w ʒʒaŋ m ⁵¹ iən	<i>These [two] people [comprise a family]</i>

The selection of the CL serves to foreground certain attributes of the modified noun's referent. Sometimes there is only one possible CL for a given noun. In many cases, however, the choice of the CL is determined by pragmatic considerations, such as the topic of discourse or stylistic constraints.

II-C Yao CLs and theories of classification

As was mentioned above, it is inappropriate to describe Yao CLs as forming a semantic system. This means theories which describe systems of classification cannot readily be applied over the complete set of Yao CLs. The following is a summary of descriptive statements which can be applied to Yao CLs:

- 1) There are no discernible componential features which operate over the entire set of Yao CLs in any non-trivial way. Any set of descriptive semantic features, such as those in Appendix II, will not operate over the entire set of Yao CLs in a hierarchical fashion.
- 2) The selection among alternate possible CLs seems to be made on the basis of (a) the topic of discourse, or (b) stylistic or metaphorical considerations.
- 3) The notion that shape is an organizing feature for CLs (see Adams and Conklin) applies only to a subset of CLs in Yao. Some Yao CLs are organized according to function instead, and still others are not organized according to any prototypical shape or function. Such CLs are paired with nouns by convention, and the resultant Cl-Noun pairs are learned.
- 4) Yao CLs unitize objects in three ways:
 - PARTICULARIZING: describing a unit which is neither a set of objects, nor part of a larger whole.
 - COLLECTIVIZING: describing a unit as a set of objects.
 - FRACTIONALIZING: describing a unit as a part of a larger whole.
- 5) Some Yao CLs carry diminutive meaning; some modify abstract nouns.

In sum, taken as a complete set, Yao CLs are not amenable to description by current theories of semantic classification. Any complete description of Yao CLs, particularly if it aims at explaining the selectional criteria, will require a thorough analysis of the semantics of Yao nouns. Prototypes and componential analysis can be used to describe various subsets of Yao CLs, however. Such an analysis is provided in Appendix II.

Notes

1. Several linguists working with one or another Southeast Asian language have been able to describe in increasing detail the intricate workings of complex CL systems. Placzek's (1978) work with Thai, and Becker's (1975) work with Burmese, show that perseverance will lead the linguist to a more precise description of a CL system.

2. Adopted from Jones, p. 6. Lahu information from James Matisoff (personal communication).
3. Thai also permits the numeral one to be omitted (Matisoff: personal communication).
4. Lahu also allows for the approximate reduplication of CLs. (Matisoff: personal communication).
5. Rather like English "strike twice" (two events) and "strike two blows" (two repetitions).
6. The same types of nouns are self-classifying in other South East Asian languages, notably Chinese (Goral: p.29). It is curious that the term for "lunar month" (lh¹⁵aa) is not self-classifying while the terms for "day" and "year" are.
7. Appendix I shows a large selection of Yao CLs. Standard measure CLs such as l⁵it for "liter," etc. have been omitted. In Appendix II, the CLs from Appendix I are described according to some general semantic features. A finer feature analysis of Yao CLs is, of course, possible, and would no doubt uncover some of the more complex selectional criteria for individual CLs. Such an analysis would require a more highly sophisticated understanding of Yao lexical semantics, and of Yao culture in general, than this author now possesses.
8. The Yao consultant replaced CLs occurring in Lombard and Purnell's Yao-English Dictionary with n⁵⁵om five times exclusively, and several other times as an acceptable alternate CL.

BIBLIOGRAPHY

- Adams, Karen L. and Nancy Faires Conklin. 1973. "Toward a theory of natural classification." *Papers from the Ninth Regional Meeting of the Chicago Linguistic Society*, 1-10.
- Allan, Keith. 1977. "Classifiers." *Language* Vol. 53.2:285-311.
- Becker, A.L. 1975. "A linguistic image of nature: the Burmese numerative classifier system." *Linguistics*. Vol. 165:109-121.
- Denny, J. Peter. 1976., "What are noun classifiers good for?" *Papers from the 12th Regional Meeting, Chicago Linguistic Society*. 122-132.
- Goral, Donald R. 1978. "Numeral classifier systems: a Southeast Asian cross-linguistic analysis." *Linguistics of the Tibeto-Burman Area*. Vol. 4.1:1-72.

- Jones, Robert B. 1970. "Classifier constructions in Southeast Asia." *Journal of the American Oriental Society*. Vol. 90.1:1-12.
- Lombard, Sylvia J. and Herbert C. Purnell, Jr. 1968. *Yao-English Dictionary*. Data Paper 69, Southeast Asia Program, Department of Asian Studies, Cornell University.
- Matisoff, James A. 1983. "Linguistic diversity and language contact in Thailand." (In Wanat Bhruksasri and John McKinnon, eds., *Highlanders of Thailand*, pp. 56-86.)
- Placzek, James Anthony. 1978. "Classifiers in Standard Thai: a study of semantic relations between headwords and classifiers." Unpublished M.A. Thesis: University of British Columbia.
- Purnell, Herbert C., Jr., ed. 1972. *Miao and Yao Linguistic Studies*. Data Paper 88, Southeast Asia Program, Department of Asian Studies, Cornell University.

APPENDIX I

Yao Classifiers

Note: ♪ indicates a cross-reference to a classifier in Appendix E of Lombard and Purnell's *Yao-English Dictionary*.

♪	1.	c ¹ 5aa	Occasions, periods of time
♪	2.	c ⁵ 5aw	Groups of travelers travelling at different times. Cf. #28 kw ³ 3an
♪	3.	ch ¹ 5oŋ	Small fields and small rooms
♪	4.	c ¹ 5iəw	Words or phrases
	5.	c ⁵ 1oy	Bunches/clusters of fruit
	6.	d ¹ 1aaŋ	Spatial directions
♪	7.	dy ⁵ ep	Drops of liquid
	8.	d ⁵ 5om	Hollow objects
♪	9.	d ¹ 5?	Cubits
	10.	d ^w 33ii	Piled up objects
	11.	dz ¹ 1aan	Letters of the alphabet
	12.	dz ¹ 1om	Sides, e.g. sides of house
♪	13.	dz ³ 3oŋ	Fields to be cultivated
	14.	dz ⁵ op	Clusters of large plants; trees, forests, bamboo groves. Cf. #50 p ³ 3om
♪	15.	dz ¹ 1un	Times or occurrences
♪	16.	g ¹ 5an	Chunks, parts of a whole
	17.	g ¹ 1ən	Long round objects
♪	18.	gw ¹ 1aa	Branches and limbs
	19.	h ⁵ 1əw	Married couples

ㄣ	20.	h ⁵⁵ iŋ	Opposite poles of time or space
	21.	j ¹¹ un	Bolts of cloth
ㄣ	22.	k ⁵⁵ əw	Sections, divisions of an object
ㄣ	23.	kh ⁵⁵ uu	Needles, pins, small instruments
ㄣ	24.	kh ⁵⁵ ɔʔ	For measuring dried fruit, bark, etc.
ㄣ	25.	khw ⁵⁵ ay	Flat, thin things; a general CL; e.g. doors, paper, planks, pages
	26.	k ¹⁵ un	Lumps of things, e.g. sugar
	27.	k ⁵⁵ ɔp	Small boxes, e.g. match boxes
ㄣ	28.	kw ³³ an	Groups; a general CL; e.g. herds of animals, crowds of people, flocks of chickens
ㄣ	29.	ky ^{1a} ʔ	Hands and feet
ㄣ	30.	l ⁵¹ aan	Unrelated adult humans
ㄣ	31.	l ⁵⁵ ɛŋ	Pairs; used for shoes
ㄣ	32.	lh ¹⁵ am	Sections of bamboo
ㄣ	33.	lh ¹⁵ ɛŋ	Slices of fruit
ㄣ	34.	lw ¹⁵ ij	Litters of animals
ㄣ	35.	m ¹⁵ ən	Parts, edges, sides
ㄣ	36.	mhw ⁵⁵ aŋ	A family as a group
ㄣ	37.	n ¹⁵ aam	Hand spans
ㄣ	38.	n ³³ ɔm	Objects in general: birds, hours, weeks; general CL used for all objects alien to Yao culture
	39.	n̄ ³³ aw	Fistfuls of objects
ㄣ	40.	n̄ ³³ ɔm	Bundles of paddy
	41.	n̄ ¹¹ uŋ	For counting <i>kinds</i> of things
	42.	ŋ ^{5a} ʔ	Lengths of bamboo
ㄣ	43.	p ^{5a} ʔ	Bundles of things
ㄣ	44.	p ¹⁵ ɛŋ	Pencils, pens; general CL for long, thin objects
ㄣ	45.	ph ³³ aan	Mosquito nets, floor mats
ㄣ	46.	ph ¹⁵ əw	Ships, scissors
ㄣ	47.	phw ⁵⁵ aŋ	Double handfuls
ㄣ	48.	p ⁵¹ iən	Basins of rice
ㄣ	49.	p ⁵¹ ip	Dry measures
ㄣ	50.	p ³³ om	Plants growing in clumps, small stands of bamboo, clumps of grass, etc.; this CL is used for clusters of plants which are smaller than clusters classified by dz ^{5op} .
ㄣ	51.	pw ⁵⁵ ən	Books
ㄣ	52.	pw ³³ ən	Hours, minutes
ㄣ	53.	pw ¹¹ ən	Families, households
	54.	pw ³³ ij	Basketfuls (used with a certain type of basket which is carried on the back)
	55.	py ¹¹ en	Round, doughnut shaped things, e.g. wheels
	56.	py ⁵¹ oŋ	Clusters of things

