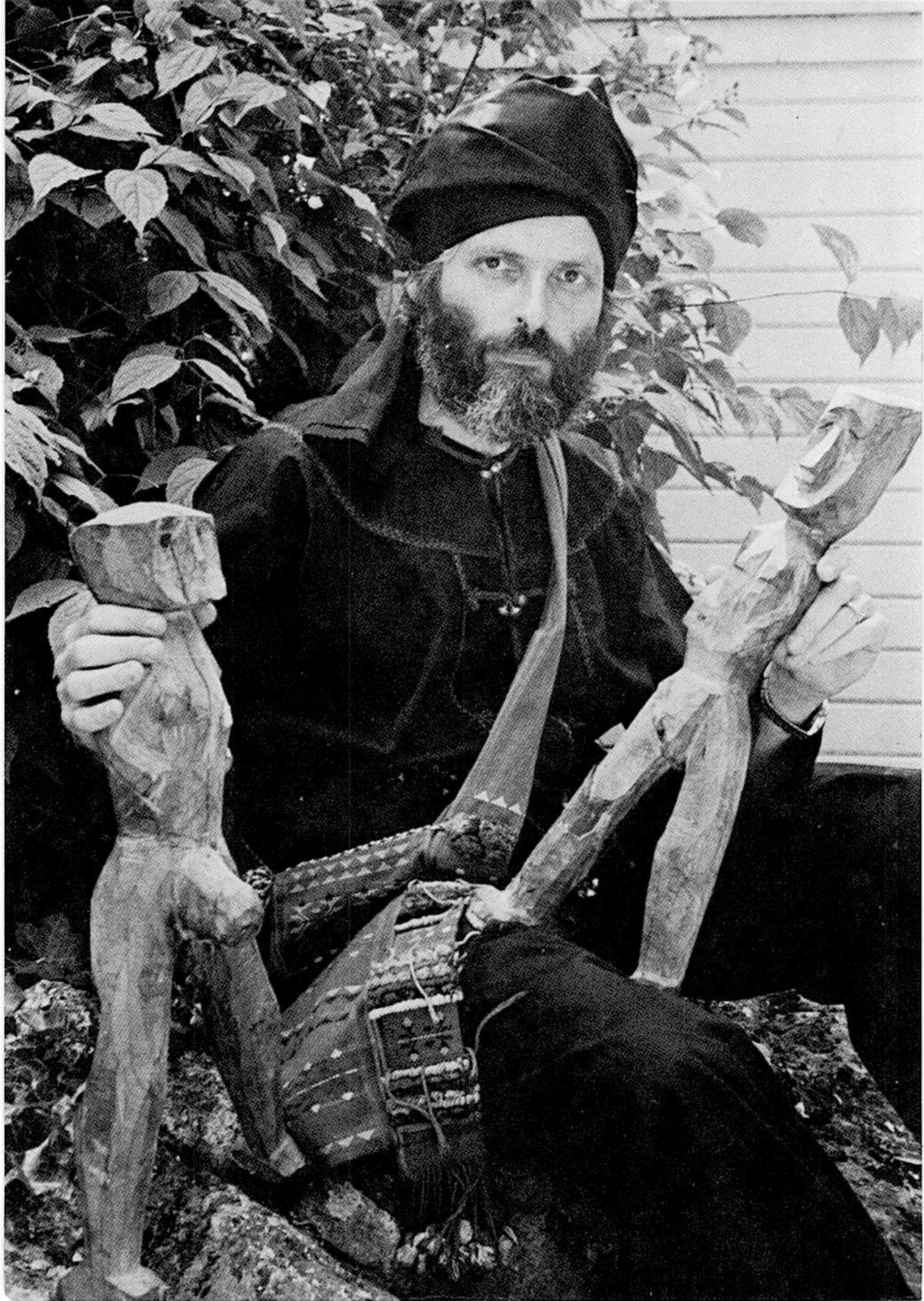




VARIATIONAL SEMANTICS
in
Tibeto-Burman



 Occasional Papers of the Wolfenden Society
on Tibeto-Burman Linguistics / Volume VI 
F. K. Lehman, Editor

VARIATIONAL SEMANTICS
in
Tibeto-Burman

THE "ORGANIC" APPROACH TO LINGUISTIC COMPARISON

James A. Matisoff



A Publication of the
Institute for the Study of Human Issues
Philadelphia

Copyright © 1978 by ISHI,
Institute for the Study of Human Issues, Inc.
All Rights Reserved

No part of this book may be reproduced in any form or by any electronic or mechanical means including information storage and retrieval systems without permission in writing from the publisher, except by a reviewer who may quote brief passages in a review.

Manufactured in the United States of America

Library of Congress Cataloging in Publication Data:

Matisoff, James A.

Variational semantics in Tibeto-Burman.

(Occasional papers of the Wolfenden Society on Tibeto-Burman Linguistics; v. VI)

Bibliography: p.

1. Tibeto-Burman languages—Grammar, Comparative. 2. Tibeto-Burman languages—Semantics. I. Title. II. Series: Wolfenden Society on Tibeto-Burman Linguistics. Occasional papers; v. VI.

PL3553.M38 495'.4 77-28921

ISBN 0-915980-85-1

For information, write:

Director of Publications

ISHI

3401 Science Center

Philadelphia, Pennsylvania 19104

U.S.A.

*For my parents
who always wanted me to be*

*a
doctor*

πολλοῖς γὰρ καὶ ὁ ἐγκέφαλος δοκεῖ μυελὸς εἶναι

Aristotle, *Posterior Analytics*
652^e25)¹

"...the ankle-bone connected to the shin-bone"

Anonymous

*Mais veistes vous onques chien rencontrant quelque
os medulare?....Si veu l'avez: vous avez peu noter
de quelle deuotion il le guette...& de quelle
diligence il le sugce.... Quel est l'espoir de son
estude?.... Rien plus q'un peu de mouelle.*

Rabelais, Prologue to *Gargantua*

FOREWORD

The data for this monograph were collected mostly in the period from January to June 1974, and reported on in a preliminary way at the Seventh International Conference on Sino-Tibetan Languages and Linguistics (Atlanta, October 1974), in my talk "Tibeto-Burman organic relationships: body-part nomenclature and genetic subgrouping." The actual writing was done from February to April 1975, and the first version of the monograph was circulated in dittoed form (viii + 270 pp.) at the Eighth Sino-Tibetan Conference (Berkeley, October 1975).

The response to the dittoed version has been heartwarming and enthusiastic. Many colleagues, students, and friends have taken the time to read the manuscript, and have been kind enough to provide me with a wealth of additional data, references, comments, suggestions, corrections, criticisms, and "personal communications" on a wide range of topics touched upon in this study. A good number of these contributions have been incorporated into the text or the notes of the present version. My warm thanks go to Reinhold A. Aman, Jim Bauman, Alton L. Becker, Madison Beeler, Paul K. Benedict, Stephan Beyer, David Bradley, Jack DuBois, Wallace Chafe, W. South Coblin, Christopher Court, Murray B. Emeneau, Charles Fillmore, Victor Girard, Talmy Givón, Mary R. Haas, André-Georges Haudricourt, E. Hillard, Larry M. Hyman, David Justice, Terrence S. Kaufman, Paul Kay, Chana

Kronfeld, Gari Ledyard, F. K. Lehman, Charlotte Linde, Lorenz G. Löffler, Yakov Malkiel, LaRaw Maran, Susan K. Matisoff, Martine Mazaudon, Boyd Michailovsky, Nai Pan Hla, Marc Okrand, Bambi B. Schieffelin, Edward L. Schieffelin, Axel Schüssler, William Shipley, Purna Chandra Thoudam, Graham Thurgood, Shigeru Tsuchida, Julian Wheatley, and Karl Zimmer.

I would especially like to thank Paul K. Benedict (Briarcliff Manor, New York) and Lorenz G. Löffler (University of Zürich) for their detailed comments and criticisms, enshrined in a series of lengthy letters. I have greatly profited from their expert knowledge, even in those cases where we do not entirely agree. For all blunders and infelicities which remain, I alone am responsible.

It is an equal pleasure to acknowledge the vital practical help of those without whom this monograph would never have seen the light of day. Eleanor Aldrich did yeoman service in the tedious task of transferring lists of body-part words from over 40 languages onto file-cards, where they could be pondered at leisure. Victor Girard, linguist and black-belt typist, beautifully typed both the original dittoed version and the final camera-ready copy (the latter from a difficult, much emended and annotated manuscript). Vic's task was made immeasurably more complicated by the theft of his typewriter in July 1976. I was off on sabbatical in Japan and Thailand at the time, so the final preparation of the manuscript had to wait until my

return to the U.S. in September 1977. Meanwhile, during this trying period, my student Ed Hillard took time out from his busy schedule to supervise the ongoing typing, while Julian Wheatley expertly drew the charts, diagrams, and Chinese characters in India ink, managing to preserve his sanity and good humor even with Figure 19.

Finally I would like to thank Kris Lehman and LaRaw Maran, my colleagues on the editorial board of the Wolfenden Society Monograph Series; David Feingold, director of the Institute for the Study of Human Issues, Inc. ("ISHI"), which has now assumed the sponsorship of the original *Occasional Papers of the Wolfenden Society on Tibeto-Burman Linguistics* ("OPWSTBL"); and Douglas Gordon, Managing Editor of ISHI, who skillfully and patiently saw this monograph through the publication process.

J.A.M.

Berkeley, California

CONTENTS

1.0	Introduction	1
1.1	Strategies for subgrouping Tibeto-Burman in the light of our present knowledge	4
1.11	Phonological criteria for relationship	6
1.12	Morphological and morphophonemic variation and the question of genetic relationship: affixation and compounding, word-families and compound-families	12
1.121	Frequently encountered variational patterns in Tibeto-Burman word-families	21
(1)	Alternations between final homorganic stops and nasals	23
(2)	Alternations between final vowel and final consonant; "parallelism of allofamy"	25
(3)	Changes in position of articulation: late assimilatory developments	29
(4)	Variations of medial glides	33
A.	Intra-lingual -G ₁ - * -G ₂ -	34
B.	Inter-lingual -G ₁ - * -G ₂ -	34
C.	Intra-lingual -G- * -Ø-	35
D.	Inter-lingual -G- * -Ø-	35
(5)	Variations of nuclear vowel	36
A.	Intra-lingual primary vowel alternations	36
B.	Intra-lingual secondary vowel alternations	37
C.	"Inter-lingual secondary vowel alternations"	39

D. Inter-lingual primary vowel alternations	39
(a) * <u>ya</u> - * * <u>i</u> -	40
(b) * <u>u</u> - * * <u>i</u> -	41
E. Miscellaneous but non-phony vowel alternations	43
F. Digression: pseudo-alternations as a product of faulty analysis	45
(6) Variations in syllable-initial position: voicing, aspiration, and prefixation	47
A. Sibilants and affricates in root-initial position	54
B. Resonantal alternation in root-initial position: "lenition" of labial stops	56
(7) Tonal variations in word-families	57
1.122 Compounding and genetic relationship	58
(a) The transparency/opacity continuum; morphans and folk etymology	59
(b) Allocompy and the <i>Urwortschatz</i>	63
(c) Unpredictability of semantic source-fields in compound-formation: illustration from body-part nomenclature	67
(d) Calquing: suspicious similarities in compound-formation	69
(e) Reprise: the compounding/prefixation cycle	71
1.13 Comparative syntax and genetic relationship	72
1.131 Specious syntactic parallelism: chance; universal tendencies; diffusion	75

1.132	Specious syntactic differences	79
1.133	Antonymous changes, in syntax and elsewhere	82
1.134	Cyclicity in syntax	86
1.14	The lexicon and genetic relationship: beyond the lexicostatistical approach	90
	(a) Analytical expertise and "proof of cognacy"	94
	(b) Constancy of replacement rate of core vocabulary	95
	(c) Semantically shifted cognates	99
1.141	Word-family and compound-family scoring problems	106
	(a) Calibrating degrees of phonological/semantic relationship: isofamy and heterofamy	106
	(b) LUNG and BELLY in Tibeto-Burman (morphophonemics)	113
	(c) The problem of setting up "higher-order allofams"—where to draw the line?	127
	(d) Compound-family scoring problems [See 1.122 above].	130
1.142	Patching up the basic vocabulary list: the CALMSEA attempt	133
	(a) Cultural appropriateness	134
	(b) Grammatical appropriateness	136
	(c) Semantic appropriateness: over- and under-differentiation	137
2.0	The organic semantic approach to linguistic comparison	141

2.1	Semantic fields and semantic systems: body-part nomenclature and the Tibeto-Burman Etymological Dictionary	147
2.2	Multidimensional semantic space in body-part nomenclature: classification by association and by opposition	151
2.3	Medical knowledge and "folk usage" of body-part terms	165
	(a) Gross exterior anatomy and obvious physiological facts	167
	(b) Basic physiological facts that are not accessible to direct observation without special instruments	168
	(c) Gross internal anatomy	168
	(d) Subtle internal anatomy and physiology	170
2.4	Semantic vagueness and semantic shifts: classification and diagrammatic representation	173
2.41	Subtypes of intra-field semantic associations in body-part terms	179
	(a) Association via anatomical adjacency	179
	(b) Association via shared physiological function	181
	(c) Association via similarity of relative position	181
	(d) Association via similarity in shape, color, or general appearance	183
	(e) Association via shared symbolic function	185
	(f) Association via metaphorical opposition	186
	(g) Association via synecdoche	186

(h)	Association via euphemism or taboo	187
(i)	Association via figs	188
2.42	Trans-field semantic associations of body-part terms	188
2.43	"Metastatic flowcharts" and their interpretation	193
2.5	Charting the semantic interrelationships among TB internal organs: sketch of a section of a chapter of the Tibeto-Burman Etymological Dictionary	200
2.51	Marrow and its relatives	202
2.52	Heart/liver and their interconnections	202
2.521	The viscera and the emotions	210
2.53	Gastrointestinal relationships	213
2.54	Splenetic forms	217
2.55	Kidneys and their kin	220
2.56	Womb and afterbirth	224
2.57	Composite flowchart of Tibeto-Burman body-parts	228
3.0	Conclusion: genetic relationship, organic semantics, and Sino-Tibetan linguistics	230
NOTES		233
VERMIFORM APPENDIX I		
	Culturally Appropriate Lexicostatistical Model for SouthEast Asia (CALMSEA), or the Matisoff 200-word list, arranged by semantic categories and correlated with some reconstructed roots in STC, TSR, etc.	283
VERMIFORM APPENDIX II		
	List of Source Languages Used in this Study	297
VERMIFORM APPENDIX III		
	Index of New Terms Introduced	301

VERMIFORM APPENDIX IV

Index of Symbols and Abbreviations	303
BIBLIOGRAPHY	308
ADDENDA	327

FIGURES

Figure 1.	Schematic chart of Sino-Tibetan groups	3
Figure 2.	Some overlapping features of special resemblance among the Indo-European languages, conflicting with the family-tree model	5
Figure 3.	Overlapping shared innovations in Lolo-Burmese	11
Figure 4.	Alternative selections from the proto-treasury of morphemes in Tibeto-Burman compound-formation	64
Figure 5.	Yin/Yang and the seeds of antonymous change	83
Figure 6.	A clear-cut family-tree relationship	91
Figure 7.	Some members of the word-family * <u>na</u> * <u>nan</u> * <u>nat</u>	110
Figure 8.	<u>Heart</u> and <u>kidney</u> in the Gnao conceptual scheme	156
Figure 9.	Symbolic oppositions in spatial orientations of the body	159
Figure 10.	The beards of the two faces	160
Figure 11.	Symbolic associations and oppositions between the features of the upper and lower faces	164
Figure 11a.	Taoist interconnections among the viscera, the rest of the body, and the universe	268
Figure 12.	Medical knowledge of "primitive" vs. "modern" man	171
Figure 13.	Medullary matters: the interconnections of <u>marrow</u>	203
Figure 14.	Studia cardiohepatica	204
Figure 15.	The visceral emotions	211

Figure 16. Gastrointestinal goings-on	214
Figure 17. Renal relationships	220
Figure 18. Womb-mates	225
Figure 19. A small part of the Tibeto-Burman body	229

PHOTOS

The Sino-Tibetan body: the author displaying *frontispiece*
male and female fertility statues carved
in an Akha village in northern Thailand
(1970). Photo courtesy of Susan Matisoff.

Lahu men butchering a pig carcass (1965). *following page 168*

1.0 Introduction. The original task I set myself when I undertook to write this monograph was to clarify the nature of "Tibeto-Burman internal relationships." What I thought I would do was come up with some improvement on the Tibeto-Burman subgrouping schema presented in Paul K. Benedict's *Conspectus*,² the most definitive statement of the TB genetic picture that we have.³ After a careful reading of the section "Tibeto-Burman classification" in the STC [pp. 4-11], one is not suffused with any delicious sense of certainty about the fine details of the interrelationships of the hundreds of TB languages. Benedict distinguishes seven "primary divisions" or "nuclei" of TB, but recognizes that "a number of smaller units resist all efforts at taxonomic reduction," largely because of insufficient data. These are his seven primary divisions:

1. Tibetan-Kanauri (Bodish-Himalayish); perhaps also Dzorgai, Lepcha, and Magari.
2. Bahing-Vayu (Kiranti); perhaps also Newari.
3. Abor-Miri-Dafla (Mirish); perhaps also Aka, Digaro, Miju, and Dhimal.
4. Kachin; perhaps also Kadu-Andro-Sengmai (Luish) and Taman.
5. Burmese-Lolo (Burmish); perhaps also Nung.
6. Bodo-Garo (Barish); perhaps also Konyak and Chairel.
7. Kuki-Naga (Kukish); perhaps also Mikir, Meithei, and Mru.

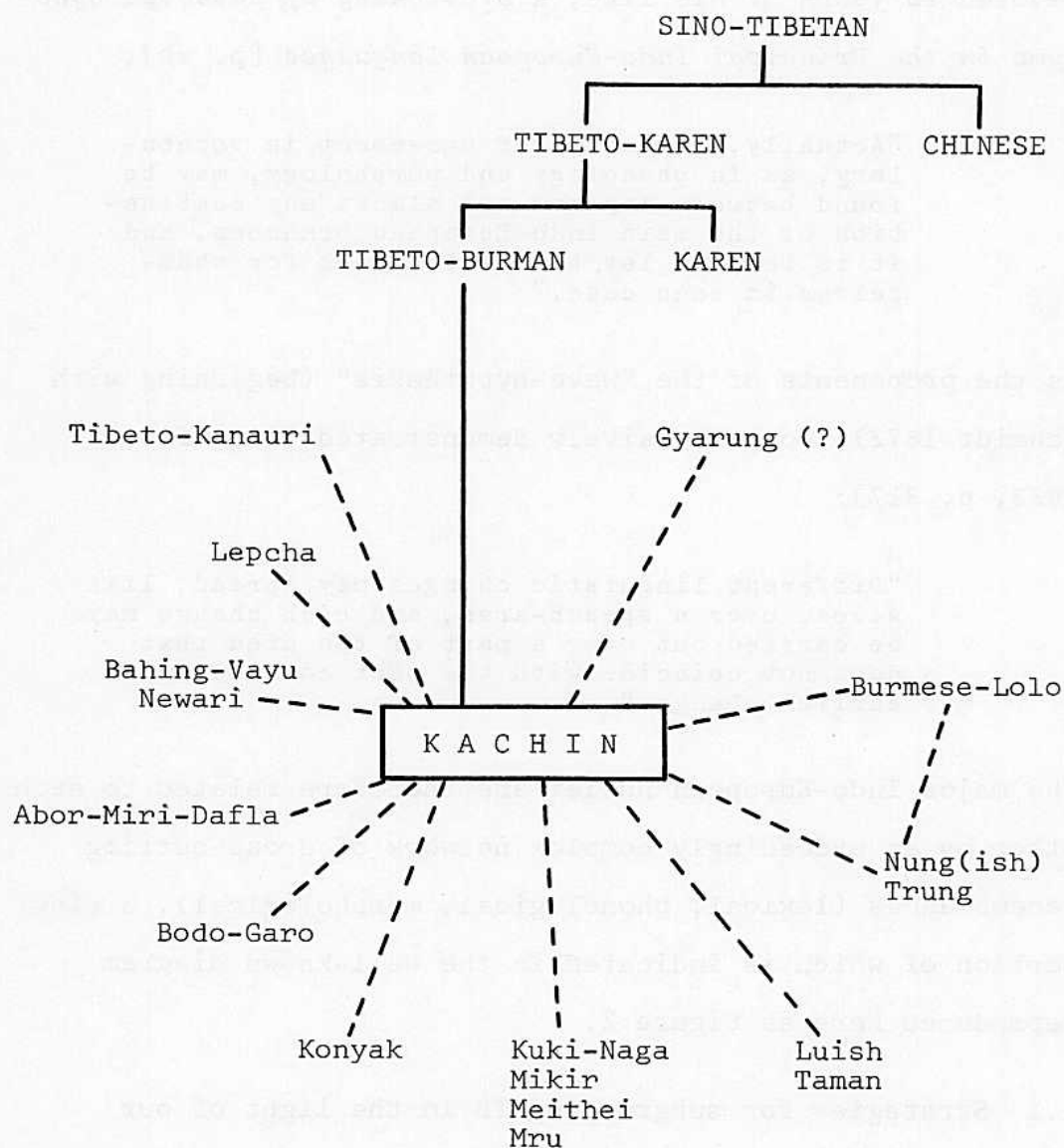
Nuclei 1, 5, and 7 are "supergroups," each with a multitude of languages and dialects. At the other extreme is Kachin (now usually called Jinghpaw or Jingpho), which consists of a single language with only relatively slight differentiation among its dialects. Yet Kachin, spoken in Northern Burma, is of paramount historical significance: it seems to lie at the linguistic (as well as the geographical) crossroads of the entire TB family, showing special lexical and morphological affinities with all the other nuclei simultaneously.⁴

Benedict, with characteristic wisdom, does not offer any pseudo-precise family-tree model of the higher-order taxonomic relationships among his several nuclei. Rather, he conceives of the TB family as an interlocking network of fuzzy-edged clots of languages, emitting waves of mutual influence from their various nuclear ganglia. A mess, in other words. "Supergroups within Tibeto-Burman cannot safely be set up at the present level of investigation" [p. 11].⁵ We are therefore left with a scheme as shown in Figure 1, where everything radiates out of Kachin, the must "nuclear" of the TB nuclei!

Tibeto-Burman is as vast and ramified a linguistic family as Indo-European. Lest anybody feel inclined to smile at the vagueness of Benedict's higher-order groupings,⁶ we would do well to reflect on the multitude of similar uncertainties which beset Indo-Europeanists, even after a century

Figure 1

Schematic Chart of Sino-Tibetan Groups [from STC, p. 6]



and a half of intensive scholarly labor, aided by a wealth of written documentation going back thousands of years. In the words of Carl Darling Buck, the nonpareil Indo-European lexicographer, in the preface to the work to which he devoted 25 years of his life, *A Dictionary of Selected Synonyms in the Principal Indo-European Languages* [p. xv],

"Actually, instances of agreement in vocabulary, as in phonology and morphology, may be found between any two and almost any combination of the main Indo-European branches, and it is best to let the facts speak for themselves in each case."

As the proponents of the "wave-hypothesis" (beginning with Schmidt 1872), so persuasively demonstrated [Bloomfield 1933, p. 317]:

"Different linguistic changes may spread, like waves, over a speech-area, and each change may be carried out over a part of the area that does not coincide with the part covered by an earlier change."

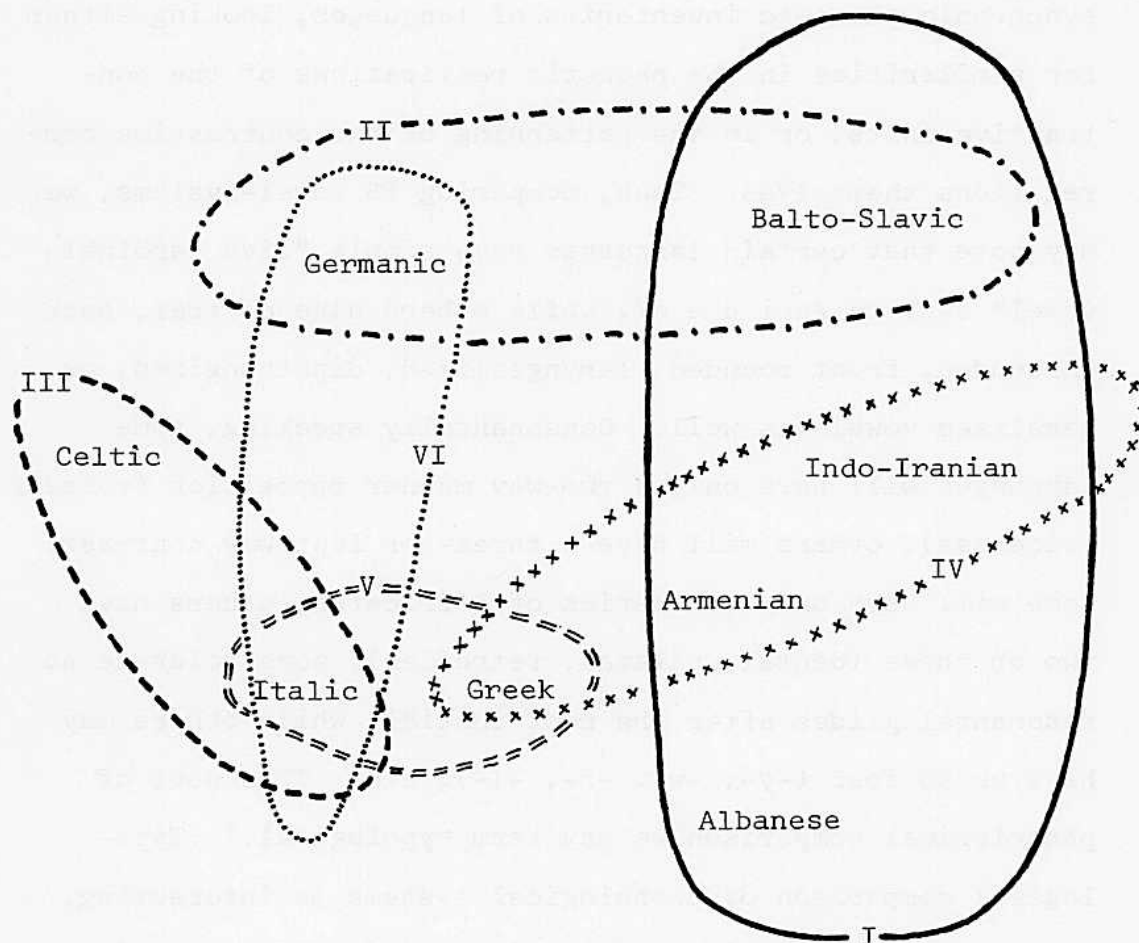
The major Indo-European nuclei are therefore related to each other by an exceedingly complex network of cross-cutting resemblances (lexical, phonological, morphological), a tiny portion of which is indicated in the well-known diagram reproduced here as Figure 2.

1.1 Strategies for subgrouping TB in the light of our present knowledge. Theoretically, any aspect of linguistic structure may be considered to have relevance to genetic

Figure 2

Some overlapping features of special resemblance
among the Indo-European languages,
conflicting with the family-tree model

[adapted from O. Schrader, in Bloomfield 1933, p. 316]



- I. Sibilants for velars in certain forms.
- II. Case-endings with [m] for [bh].
- III. Passive-voice endings with [r].
- IV. Prefix ['e-] in past tenses.
- V. Feminine nouns with masculine suffixes.
- VI. Perfect tense used as general past tense.

classification: phonology, morphology, syntax, lexicon, semantics.

1.11 **Phonological criteria for relationship.** Phonological similarities among languages may be considered from several points of view. On the simplest level, we may compare the synchronic phonemic inventories of languages, looking either for similarities in the phonetic realizations of the contrastive units, or in the patterning of the contrastive correlations themselves. Thus, comparing TB vowel-systems, we may note that certain languages have simple "five cardinal-vowel" systems /a i u e o/, while others have central, back unrounded, front rounded, laryngealized, diphthongized, or nasalized vowels as well. Consonantally speaking, some languages will have only a two-way manner opposition (voiced/voiceless), others will have a three- or four-way contrast; some will have only one series of affricates, others have two or three (dental, palatal, retroflex); some tolerate no resonantal glides after the root-initial, while others may have up to four (-y-, -w-, -r-, -l-), etc. This sort of phonological comparison we may term **typological**.⁷ Typological comparison of phonological systems is interesting, but not very useful in establishing truths about genetic relationships. As is well known, geographically or culturally contiguous languages may come to have extremely similar phonological inventories by "diffusion," even though they are not genetically related at all. Thus Standard Thai and

Lahu have superficially identical 9-vowel systems [Matisoff 1973d, p. 10]. Other non-genetic reasons for typological similarity include **universal tendencies** (e.g., having more tonal contrasts in open than in stopped syllables)—as well as **pure chance**.

One particularly seductive kind of phonological investigation in TB involves the comparison of **tonal systems**, especially in relation to syllable-structure. In the past, when the articulatory and phonological mechanisms of tonogenesis⁸ were less well understood, and the development of contrastive tone was felt to be a strikingly unusual phenomenon, comparatists were led to accept similarities in tone-systems as *prima facie* proof of genetic relationship. Thus the Tai languages, as well as Miao-Yao, were lumped together with Chinese and Tibeto-Burman into the pseudo-stock called "Indo-Chinese."⁹ Vietnamese was lumped together with Tai, instead of with its real Mon-Khmer congeners. But we now know that tone-systems are eminently diffusible among related languages or even across genetic lines; that tones may come and go several times, cyclically, within the history of a given language-family; that even dialects of one and the same language (e.g., modern Tibetan) may some be tonal and others atonal. In short, tonal similarities are in themselves no more useful for establishing genetic relationship than counting the number of vowels in the inventories of the languages under consideration would be.¹⁰

Probably the most reliable way to use purely phonological facts for subgrouping purposes is on the basis of shared phonological rules of innovation leading from the proto-forms of morphemes to the reflexes in the various daughter languages. A celebrated example is the so-called "satem-centum isogloss" in Indo-European [see Area Number I in Figure 2]. Some subgroups of IE developed sibilants from the old PIE velar $*\hat{k}$ (thus $*\text{dek}\hat{m}$ 'ten' > Skt. daśa, Russ. desjat', etc.), while the others retained the velar articulation (Lat. decem, Gk. δέκα, etc.). Similarly in Tibeto-Burman, we might want to exploit for subgrouping purposes any of an indefinitely large number of isoglosses based on particular diachronic innovative sound-changes. Thus we could set off from the others those TB languages that share a certain "occlusivization of sibilant rule," $*\underline{s}- > t(h)-$. This would enable us to demarcate Kuki-Chin¹¹ and Bodo-Garo,¹² where most of the member-languages share this innovation, from the rest of TB, where $*\underline{s}-$ usually remains a sibilant. Yet as Benedict points out [STC, p. 28], a very similar development has occurred (much more recently) in Modern Burmese, where $*\underline{s}-$ has become the interdental fricative θ . We might go on to note that this / θ / is actually articulated by Rangoon speakers as a kind of affricate [$t\theta$]. We could view this affricate as a phonetic missing link between a sibilant \underline{s} and its homorganic voiceless stop, i.e., in articulatory terms a prior stage in the occlusivization

process to what we find in Kuki-Chin and Bodo-Garo, where the reflex no longer shows any trace of its spirantal origin. But does this mean that we must posit any special relationship between Burmese and these other subgroups? Clearly not—all we might say is that the "seeds of sibilant occlusivization" existed already in the proto-language, and were actualized in various subgroups of TB at different times in history.

Interesting and important as these innovation isoglosses are, their interpretation in terms of genetic subgrouping remains an art rather than a science. This is because you can always find cross-cutting and overlapping patterns of apparent "rule-sharing" between any pairs of languages you choose, if you look hard enough. The analyst will implicitly rank these on some subjective aesthetic scale of importance, dismissing some as trivial and fixating upon others as crucial. Let us take a small example from Lolo-Burmese, considering only the three languages Burmese, Lahu, and Lisu. The Proto-Lolo-Burmese (PLB) *voiced series of obstruents, which we may symbolize by its labial representative, *b, retained its voicing in Lisu, but was devoiced to p in Burmese and Lahu:

Rule A: PLB *b > Bs. p, Lh. p, Ls. b.

On the other hand, the PLB rhyme *-ak has retained the centrality of its vowel in the Lahu reflex -a?, while both Modern Burmese and Lisu have developed a front vowel here. All

PLB *-ak words have become -ε? in Mod. Bs., while in Lisu *-ak is similarly fronted to -ae (though only after labials and dentals, remaining -a otherwise.)¹³ We might then write a "pre-velar vowel-palatalization rule," properly constrained in the case of Lisu, but nonetheless setting off Burmese- and-Lisu as a group against Lahu:

Rule B: PLB *-ak > Bs. -ε?, Lh. -a?, Ls. -ae ~ -a.

Finally, we may consider the matter of tonal contrasts in formerly stopped syllables. Here Burmese reflexes of proto-syllables with final */-p -t -k/ are all under the same tone, while Lahu and Lisu show a complex multiple tonal split according to the nature of the syllable-initial consonant group [Matisoff 1972a, Thurgood 1974a]. A small part of this split goes as follows:

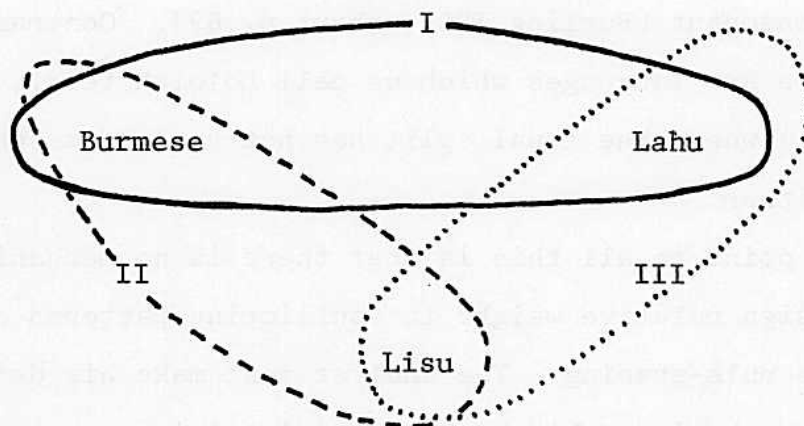
Rule C: { PLB *pak > Bs. pha?, Lh. phâ?, Ls. phâe
 PLB *bak > Bs. pa?, Lh. pâ?, Ls. bâe.

It has therefore not been too difficult to find three phonological innovations that pair up our three languages in three different ways. See Figure 3.

It may be objected that everybody knows what the proper subgrouping of these languages is—naturally Lahu and Lisu, both being Loloish languages, are more closely related to each other than either is to Burmese, which belongs, unsurprisingly, to the "Burmish" branch of the Lolo-Burmese family. Yet suppose we didn't have the testimony of Written Burmese (WB) to guide us. Would it still be so obvious that

Figure 3

Overlapping shared innovations in Lolo-Burmese



- I. Devoicing of PLB *voiced obstruents.
- II. Palatalization of vowel before velar stop.
- III. Tonal split in checked syllables.

isogloss III is to be considered more "diagnostic" than the other two? To be sure, we might dismiss isogloss II as less important than the others, since it only involves a single rhyme, while No. I affects a whole series of obstruents and No. III applies to one of the three primary syllable-types of the family.¹⁴ But what about No. I versus No. III? The devoicing of the old *voiced series took place in other Burmish languages than Burmese (e.g., Atsi and Maru—see Burling 1967), but it also happened in other Loloish languages than Lahu (e.g., Bisu—see Nishida 1966, 1967). The Loloish tonal split in checked syllables is indeed "one of the most striking points of divergence between the Burmish and Loloish branches of Lolo-Burmese."¹⁵ Yet Burmish Maru shows a three-

way phonetic tone-difference in checked syllables—even though this has not yet become "phonologized" since tonal differences are still predictable in terms of the syllable-initial consonant [Burling 1967, chart p. 69]. Contrariwise, there are languages which we call Loloish (e.g., Moso and Na-khi) where the tonal split has not yet become phonologized either.¹⁶

The point to all this is that there is no mechanical way to assign relative weight to conflicting patterns of innovative rule-sharing. The analyst must make his decisions on the basis of his hard-earned intuitions as to what is critical and what is of lesser importance.

[For two alternative subgroupings of the Lahoid branch of Loloish, each of which fixates on a different phonological criterion, see TSR, pp. 21-22.]

1.12 Morphological and morphophonemic variation and the question of genetic relationship: affixation and compounding, word-families and compound-families. The modern Tibeto-Burman languages all show unmistakable effects of morphological processes that operated in the proto-language. Many of the modern languages exhibit more or less productive processes of prefixation and/or suffixation; other languages preserve traces of proto-affixations in more indirect ways (e.g., voicing alternations in initial consonants, alternations between homorganic stops and nasals in syllable-final position,

tonal alternations, etc). To some of the affixes that can be reconstructed at the proto-level it is possible to assign fairly definite meanings (e.g., the prefix *s-, which served as a transitivizer or causativizer of verbal roots). Many of them, however, are semantically quite elusive (e.g. the prefix *r-), and may have been so even at the proto-stage.¹⁷

The fact that a given affixational process is productive in a particular daughter language is by no means a guarantee that it was already productive at an earlier stage. Prefixation is very much a living process in TB. Old prefixes which had been used in a sporadic or unsystematic manner may later be generalized and regularized both in meaning and in "privilege of occurrence." Furthermore, entirely new prefixes are constantly being created through the mechanism of compound-reduction.

For the other great TB morphological process, besides prefixation, is **compounding**—the juxtaposition of semantically related roots to form polysyllabic unitary words. Sino-Tibetan has always been basically monosyllabic, in the sense that morphemes are typically one syllable long. Yet both Chinese and TB have shown a steady evolution towards polysyllabic words, largely as a compensatory device to counteract phonological mergers—or at any rate to increase the redundancy of utterances.

It would be rash to try to decide which is the more

ancient process in TB, affixation or compounding, since the two seem sometimes to stand in a relationship of mutual implication. There are innumerable cases where one can prove that a prefix of the shape C- or Cə- in a daughter language derives from a destressing of a once fully syllabic morpheme in a compound. Thus the WB word pərwak 'ant' clearly comes from *buw-rwak, where the first syllable meant 'insect' (cf. Lahu pú-ǵô? 'ant,' pû 'bug'). On the other hand, there are good examples of the opposite development, whereby a formerly non-syllabic prefix¹⁸ has acquired full syllabic status. Thus PTB *m-ran 'horse' > Jinghpaw (Jg.) gùm-rà ~ gùm-ràn.¹⁹

Compounding, as we shall see below [1.122], is an autonomous process in TB, and operates according to a weird logic of its own. Yet compounding and affixation have worked together in the history of the family. Not only do they shade cyclically into one another in a diachronic sense, but they also "conspire" to produce similar synchronic effects. Both affixation and compounding have contributed to the rich and complex vocabularies of the modern languages.

Any good dictionary of a TB language²⁰ is crowded with sets of variant forms that bear partial phonological/semantic relationships to each other, much like English break-breach-broach; blush-flush; whole-hale-heal-healthy; wring-wrench-wry-wrist-wrest-wrestle-awry; green-grass-grow;

drink-drench; smell-smile-smirk-snout-sniff-snuff-snot-snicker; nose-nozzle-nasal-nuzzle; stand-stable-stagnant-stalk-stake-stage-staff-status; gold-gild-gilt-yellow-yolk-jaundice, etc. As in these English examples, the variant forms can sometimes be shown to reflect a quasi-regular morphological process, living or dead (e.g. causativization [drink/drench]; past-tense formation [drink/drank]; pluralization [tooth/teeth, goose/geese], etc.). But more often than not, the sound/meaning resemblances are more irregular and unstructured, like the English sm- ~ sn- words above, which all refer to the nasal-oral region in some way, and are all probably related etymologically. The judgments of native speakers of English as to whether the words in one of these sets are really "related" or not (either in the synchronic or etymological sense) will vary tremendously from person to person, in a totally idiosyncratic way. Unless the speaker has highly specialized historical linguistic knowledge (and sometimes even then!) he cannot be sure he is not failing to acknowledge a real etymological relationship—or conversely, mistakenly assuming a relationship where there is really none (folk-etymology).

The speaker's problems are equally great whether or not his language happens to be blessed with a long-standing writing system. Thus English speakers will be influenced by the spelling to judge (correctly) that *heal* and *healthy*

contain the same etymon, even though its pronunciation is different in the two words. Yet most people would judge (incorrectly) that *whole* is not related to the other two, on the basis of the different orthographic initial.

Try asking a Lahu whether the following three morphemes are related:²¹ *lâ* (verb) 'to come' [low-falling tone]; *la* (verb-particle) 'motion toward the center of interest' [mid-tone]; *lâ* (verb-particle) 'non-third person beneficiary' [high-falling tone].

In traditional historical linguistic terminology, as applied to, e.g., Indo-European, etymologically related variants of the same proto-morpheme in a given language are referred to as **doublets**.²² The different variants may reflect alternate patterns of borrowing from related languages (e.g. English *royal* < French, *regal* < Latin; *coy* < Fr., *quiet* < Lat.); or result from different morphological elements operating on the same native lexical material; or be survivals of competing forms from different proto-dialects ("dialect mixture").

In Sino-Tibetan linguistics—perhaps due to the greater strength of kinship bonding in the Orient—the traditional label for such groups of variants is **word-family**.²³

This is a fine old term, and should be retained. However, we still need a word to refer to the relationship among the various individual members of the same word-family. For this I suggest the euphonious term **allofam**.

Members of the same word-family are said to be allofams of each other, or to stand in the relationship of **allofamy** with each other.²⁴

Some further distinctions are necessary. (a) We may speak of word-families **within a given daughter language** (like Lahu lâ, la, lâ, above). We are here considering the word-family **synchronically** and **intra-lingually**. All the allofams occur in the same language. (b) We may compare an allofam in one language to an allofam in another daughter language. Thus Lahu qho 'draw water' is an allofam of WB khap 'id.' Yet the two forms are not "directly cognate" to each other. The Lahu form descends from the nasal-finalled variant *kam, while the WB form comes from the homorganic stop-finalled variant *kap. But suppose a Lahu form qhô? also existed, with a meaning plausibly relatable to the drawing of water [see below 1.141(a)]. Such a form would be the exact phonological counterpart of WB khap—i.e., the two forms could be demonstrated to descend from the **same proto-allofam** *kap. We can now give a more precise definition of "cognate": an allofam in Language X is cognate to an allofam in Language Y if the two forms can be shown to derive regularly from the same proto-allofam. Put another way, cognacy is a special case of **inter-lingual allofamy**, such that the cognates can be traced back not only to the same proto-word family, but to one and the same proto-allofam in that proto-family. We

are here thinking of word-families across languages (interlingually) and through time (diachronically).²⁵

In sum, word-families are groups of forms which bear a non-fortuitous phonological and semantic relationship to each other. The sound/meaning relationships among the allofams of a word-family may follow a more or less productive pattern, so that in favorable cases the variations may be traced back to systematic (or at least plausible) alternations in the proto-language itself (often involving proto-affixations). In many cases, however, the synchronically observable intra- or inter-lingual allofamy follows no particular pattern that repeats itself elsewhere. This situation may result from conflicting or overlapping morphological processes that obscure each other's outputs, unsystematic or sporadic increments to roots, interference or contamination from genetically unrelated forms, dialect mixture—or of course it is always possible that the forms in question were never co-allofams at all, and their resemblance is entirely specious.

We shall take a look at some of the more regular variational patterns in TB word-families in a moment [1.121 below]. For now let us just emphasize that we must assume that the proto-language itself was awash with allofamic variation, both systematic and unsystematic. Why should a proto-language be any more monolithically invariant than any living language that we can observe with our

own eyes? No language is ever perfectly regular at any stage of its history. Language change operates on all levels at all times, and the processes responsible for linguistic variation are never "in synch" at any synchronic stage, no matter how far back we venture into the past. (Needless to say, I am not pleading for promiscuity in cognate- or allofam-identification! We must try to be simultaneously conservative and audacious, remaining aware of the full complexity of the question and the limitations of our methodological tools for establishing genetic relationship.)

So far we have been talking mostly about variations whose domain is a single syllable. This is the traditional scope of the term "word-family." Yet the TB languages are full of variations on the level of polysyllabic compound-formation as well. As an analogy, consider the various English compound nouns referring to the jigger that running piped water comes out from: water-tap, water-faucet, water-spigot, water-cock, water-spout.... Different speakers have different favorites. Some speakers use two or more of the forms alternately, others always use the same one.²⁶ Not all speakers of American English have heard of all the compounds ("water-cock" has become impossible to use with a straight face, and is all but obsolete in American usage). Let us call such a group of compounds, which all mean the same thing as units, but some of whose constituent morphemes

are not necessarily related to each other etymologically, a **compound-family**. (Naturally, the members of a compound-family stand in the relationship of **allocompy** to each other.)

The TB languages abound with intra-lingual allocompy—alternative compound-formations that mean the same thing and that coexist in the same language. To take a few random examples from Lahu: khĩ-pè?-kā and khĩ-kĩ?-kā both mean 'cloven hoof;' khĩ-qō-lē and khĩ-mê?-šĩ both mean 'ankle;' chĩ-qwè?, chĩ-qwè?-ni, chĩ-pĩ-qwè?, and chĩ-kε-ne all mean "barking-deer." Speakers differ in their evaluations of the dialectal and semantic status of the various competing allocomps. Some speakers will identify a given compound as belonging to a different dialect than their own, even though they recognize it as "authentic" Lahu. Others might say it is just a seldom-used form in the same dialect. Sometimes a speaker will claim a slight semantic difference between two allocomps (e.g., khĩ-qō-lē may be said to refer to the whole ankle, while khĩ-mê?-šĩ refers only to the projection of the ankle-bone), but other speakers will deny such claims, and say they are perfect synonyms.

As we shall soon see [1.122], not only is allocompy a striking intra-lingual feature of the TB languages, it is also (*a fortiori*) very frequently encountered when one makes inter-lingual comparisons. Different languages make

different selections from the proto-treasury of morphemes in coining their compounds. This greatly complicates the computation of cognate-percentages if one uses the traditional sort of simple-minded lexicostatistical approach [see below 1.141d].

To recapitulate: allofams of a true word-family must show both a phonological and semantic resemblance that is due to a common genetic heritage. Allocomps in a compound-family must all mean more or less the same thing, but they need not have any phonological resemblance to each other, nor even share any particular proto-morpheme in common.

1.121 Frequently encountered variational patterns in TB word-families. In order to provide some justification for our judgments as to whether two given forms are to be assigned to the same word-family or not, we must distinguish between variational patterns that are so frequently encountered and so self-consistent that they may be assumed to reflect genuine proto-variations or proto-morphological processes (i.e. genuine patterns of allofamy), and those variations which are unique, or poorly attested, or not yet explicable in terms of the proto-phonology or proto-morphology (i.e., dubious patterns of allofamy). In this section we shall discuss some of the most important of the "genuine" patterns, as they manifest themselves both within a given TB language and comparatively between different languages or

subgroups of the TB family. Some of these patterns can plausibly be explained without recourse to the notion of dialect mixture — that is, they can be assumed to be the regular reflexes of morphological processes that operated within a single dialect of an earlier stage of the language (e.g., s-prefixation to signal 'causative' or 'transitive'). Other genuine patterns, such as the alternation between medial -r- and -y-, seem never to have had any morphological or semantic significance. Such patterns must either be assumed to be late developments (whose conditioning is totally obscure)²⁷ or else they must be pushed back to different dialects of the proto-language itself.

We may paraphrase the previous paragraph more simply. What is a respectable variation, and what strains the credulity?

Let us introduce the symbol "*" to mean "belongs to the same word-family as" or "is an allofam of." Thus,

(a) Lahu ú- * ố- 'head'

means that both of these Lahu allomorphs for 'head' belong to the same word-family. Since we are here talking intralingually and since the meaning of both variants is the same, the symbol "*" is here equivalent to the traditional alternation sign "~". However, in (b) the meanings of the phonologically variant forms are not the same, so that our new symbol cannot be replaced by the tilde:

(b) Lahu là 'to come' * la 'causative verb-particle'

* lâ 'benefactive verb-particle.'

The new symbol works equally well for word-families considered inter-linguistically:

(c) Lh. qho 'draw water' * WB khap 'id.'

The proto-allofams of the word-family may be listed either with an intervening tilde or with our new symbol:

(d) *kam ~ *kap, or *kam * *kap.

The new symbol may also be used inter-lingually in the special case where the allofams also happen to be direct cognates:

(d) Lh. nè * WB nan 'you.'

However, when the forms being compared are directly cognate, it is usual to list them simply with intervening commas, so perhaps the greatest usefulness of "*" will prove to be in situations like (b) and (c) above.

The PTB syllable-canon had the following shape,

$$(P)(P)C_i(G)V(\cdot)(C_f)(s),$$

where P = prefix, C_i = root-initial consonant, G = glide (-w-, -y-, -r-, -l-), V = vowel, \cdot = vowel length, C_f = root-final consonant, and s = suffixial -s.

We shall divide our discussion of genuine variational patterns according to the particular part of the proto-syllable that was involved.

(1) **Alternations between final homorganic stops and nasals.** This is an extremely common type of allofamic

variation, both within a single language, and across languages. Within an individual language this variation may be exploited in productive morphological processes. Benedict [STC, p. 14] mentions such alternations in the verb paradigms of several Bodish, Bahing-Vayu, and Abor-Miri languages, e.g. Bahing bap-to 'scratch!' (v.t., imperative) * bam-so 'scratch oneself.' The nasal-finalled variant seems to carry an intransitive, stative, or reflexive meaning, while the stop-finalled allofam bears a transitive, directionalized, or causative meaning. Similar alternations occur in modern dialects of Chinese which preserve final stops, as noted most recently by T'sou 1972. Thus, Cantonese kwon 廣 'wide' * kwok 擴 'widen;' saan 散 'be dispersed' * saat 撒 'to disperse (v.t.);' gin 堅 'be solid' * git 結 'to tighten.'²⁸ Alternations like this seem to be generalizations of formerly sporadic metanalytic effects of an old transitivizing or causativizing suffix with a voiceless initial consonant [see next section].

At the other extreme are cases of intra-lingual variation which are low-level, non-productive assimilations to the initial of a following syllable which remains overtly present, like WB yok-ma * yon-ma 'pudding-stick,' where the "basic allofam" had the stop final, as shown by WT yog-po 'poker,' skyogs 'scoop, ladle,' so that the WB nasal-finalled variant is easily explained by the nasal initial

of the second syllable.²⁹

In most cases, however, especially in noun-roots, the final stop/nasal alternation has no identifiable semantic correlate, nor can it be demonstrated to be due to the influence of any following morpheme. In Matisoff 1972a, *The Loloish Tonal Split Revisited* [henceforth "TSR"], I presented evidence for such alternations in about a dozen word-families.³⁰ The patterning of the nasal- vs. stop-finalled allofams is striking in its randomness with respect to the particular Lolo-Burmese languages which exhibit each variant. Sometimes the alternation exists within a single language, in which cases there is usually semantic differentiation between the variants in that language, but not according to any identifiable pattern:

WB mak 'dream' (< *mak) * hman-ca-san' 'somniaambulize' (< *s-man);

Lahu kâ? 'cold (of weather, human sensation, objects)' (< *?krak) * gê 'cool (of objects), peaceful, quiet' (< *Ngran);

WB nok 'space behind, past time,' ?ənok 'the west' (< *nok) * hnon 'come after,' ?əhnôn 'something coming after,' ?əhnon' 'back of a knife' (< *s-non).³¹

(2) Alternations between final vowel and final consonant; "parallelism of allofamy." There is no doubt that we must set up several dental suffixes, /*-t, *-n, *-s/,

for Tibeto-Burman [cf. STC, pp. 98-103], and undoubtedly for Sino-Tibetan as well [cf. STC, pp. 156-159]. These have been discussed repeatedly in the literature for the last seventy years.³² As always, there is no constant semantic content that can be assigned to these suffixes, though in specific instances they can frequently be shown to function as causativizers or nominalizers of verbal roots. Benedict has shown that suffixial **-n* sometimes carried a pluralizing or collectivizing meaning, both in TB [STC, note 284, p. 99] and in Chinese [STC, note 428, p. 157].

Jinghpaw has several striking examples of allofamic alternations between open syllables under the high tone /' / and variants with suffixed *-t* under the low-stopped tone /ˀ/.³³ In four cases the suffixed allofam is a noun derived from the corresponding open-syllabled verb:

<u>khú</u>	'be smoky'	×	<u>?wàn-khùt</u>	'smoke' ³⁴
<u>lэгú</u>	'steal'	×	<u>lэгùt</u>	'thief'
<u>dží</u>	'urinate'	×	<u>džìt</u>	'urine'
<u>šá</u>	'eat'	×	<u>šàt</u>	'food, rice.'

*li 'heavy'
to lit 'load,
burden*

Yet there is a phonologically exact parallel where the semantics of the two allofams are reversed, so that it is the noun which appears as an open syllable:

<u>džú</u>	'thorn'	×	<u>džùt</u>	'be pierced.'
------------	---------	---	-------------	---------------

There is even a case where the same phonological alternation occurs without changing the meaning, so that the two

forms are mere dialectal doublets:

dá 'put, place' * dàt 'put, place'
(Hkauri dialect)

If there is so much semantic inconsistency to this sort of alternation **even within a single language**, one may well imagine how much more anarchic the situation becomes when one proceeds to interlingual comparisons. All six of the Jinghpaw sets just presented have very close phonological counterparts in the corresponding Written Tibetan word-families. For each of the sets WT has both an open- and a dental-finalled allofam—but the semantic/phonological correspondence between the two languages is never quite exact! Thus we have WT du-ba * dud-pa, but both allofams are nouns meaning 'smoke.' For 'steal,' WT has rku, but the derived noun has the nasal- rather than the stop-suffix (rkun-ma 'thief'). For 'urinate,' WT does have an open-syllable variant for the verb (gǎci), but the allofam with dental stop (gǎcid) is still a verb, with the derived noun 'urine' having the nasal suffix (gǎcin). For 'eat,' WT has an open syllable for the verb (za), but a nasal suffix for the derived noun (zan 'food').

In sum, exact semantic/phonological parallelism of allofamy is hard to come by, even in the most favorable cases.

* *

As exemplified in the families for 'steal,' 'urine,' and 'eat' just discussed, a single word-family may comprise allofams belonging to **all three** of the basic TB syllable-final types—open, nasal-finalled, and stop-finalled. Additional examples are readily found:

SWELL UP/SWOLLEN (TSR 92). PTB *s-bwap (> Jg. bòp 'calf of leg') * PLB *Nbwap or *?bwap (> Lh. bú è 'swollen') * PLB *Ø-pwap (> Lh. phò? 'swell up') / PTB *s-bwam or *s-bum (> WT sboṃ 'thick, stout,' Jg. bōm 'swell,' WB phwam 'fat, plump (of young animals),' Lushai puam 'swell, be swollen,' Lh. khî-pē-qu 'calf of leg' / *H-bwa or *baw (> WT hbo 'swell up, rise,' Mru bau 'to swell, swelling').... [See BELLY, below 1.141.]

ILL/IN DIFFICULTY/SPIRIT, DEMON (TSR 136). *nat (> WT nad 'illness,' WB, Jg. nat 'animist deity, spirit capable of causing pain,' Ak. neh 'id.,' Lushai nat 'ache, be in pain') * *nan (> Chinese nān 難 'difficulty, suffering' [STC, p. 159], Lahu nê 'spirit,' Sa. ni 55, Lisu ni 'id.') * *na (> WT na-ba 'be ill,' WB na, Lh. nà 'id.' [STC 80, p. 31; see also Benedict 1939, pp. 227-229, and Figure 7, below].

NOUN PREFIX.³⁵ *a- (> Lh. a- 'vocative prefix') * *aŋ- (> Lh. ṽ- 'noun-prefix') * *ak- (> Lh. á- 'noun-prefix').

In word-families of this complexity, it would be folly to expect to find any universal correlation between the phonological shapes of the allofams and their meanings. The various languages seem to have used their affixational resources idiosyncratically and independently at different stages of their histories.

* *

*

We may mention one further type of alternation between open and stopped syllables—variation between *-w and *-k. In TSR we noted two such cases:

BRAIN (156). *s-nok (> WB û-hnok) * *nuw (> WB nûi); ??

SWALLOW (v.) (137). *m-lyuk (> Jg. məyù?, Akha myo_Λ *
*m-lyuw (> WB myui [mlyui in Inscriptions], Tangkhul

Naga khəməyuy).

This is very reminiscent of the situation in Maru (Burmish group), where the PLB rhymes *uw and *iy have regularly developed secondary stops, becoming Mr. -uk and -it, respectively.³⁶

(3) Changes in position of articulation: late assimilatory developments. So far we have been considering the alternation of homorganic segments and/or the presence or absence of suffixial elements. However, sometimes we find word-families displaying intra- or inter-lingual variation

between final (or initial) consonants occupying different positions of articulation: dentals * labials, dentals * velars, etc. These heterorganic alternations do not have any very ancient etymological status. In most cases they can be shown to be late assimilatory developments in particular "heterorganizing" languages, where the medial glide or the quality of the syllable's vowel has coaxed the consonant out of its proto-position. That is, it can usually be shown that the difference in position of articulation is due to an innovation in the heterorganizing language, and that it does not reflect any such alternation in the proto-language itself.³⁷

Thus, the PTB proto-rhyme *-ik regularly developed into WB -ac, so that the palatality of the vowel was transferred onto the final consonant, the vowel itself becoming depalatalized in the process (e.g. 'joint,' PTB *tsik > WT ts'igs, WB chac). But -ac is the regular WB reflex of the proto-rhyme, and there is no question of assigning any "alternation" -ik * -ac to the proto-language.

In equally regular fashion, the rhyme *-ik sometimes developed into Lushai -it. Here the palatal vowel remained palatal even after causing the final velar to assimilate into the dental position. Thus, 'scorpion' PTB *(s-)di.k [STC 56] > WT sdig-pa, Lu. ti.t; 'eye' *mik [STC 402] > WT mig, Lu. mit; 'pheasant' *s-rik [STC 403] > Garo grik, Lu. va-hrit.³⁸

A similar innovative development may be found in Jinghpaw, where a secondary rhyme -en has developed from original *-yam: 'fly (v.)' WB pyam, Jg. pyēn; WB khyâm 'cold,' Jg. khyēn * gyēn 'snow, ice.' In some cases, the WB cognate does not show the -y- medial: WB krâm 'rough, coarse,' Jg. grèn 'raw-boned,' tìn-grèn 'rough;' 'sound, voice' WB ʔasam, Jg. nìn-sén * ñsén; WB hlam 'spear,' Jg. məlèn 'sting, as of bee,' pəlèn 'point, as of serpent's tongue.' In these instances we must suppose that WB has secondarily lost the -y- medial, since the simple proto-rhyme *-am developed regularly to -am in both WB and Jg. ('road' *lam > WB, Jg. lam).³⁹

In syllable-initial position a fair number of TB languages show a secondary shift of original palatalized labials into the dental position. Thus *b-yet * *m-yet 'mind, seat of emotions' > Nung mit * nit; *s-myak 'eye' > Nung mɛ * ɲɛ. In Nung, as these examples show, the assimilation is still in progress, so that competing labial and dental allofams still coexist. In the Loloish languages Ahi and Sani, the process has been carried out to completion, so that *my- regularly develops into n-, *py- > t- or the lateral affricate tl-, and *by- > d- or dl-: 'bee' *bya > Ahi do, Sa. dla-ma; 'fly' *pyam > Ahi thö, Sa. tlö; 'monkey' *m-yok > Ahi, Sa. nu; 'eye' *s-myak > Ahi nie, Sa. ne.⁴⁰

The assimilative developments so far discussed are all

quite regular in the particular heterorganizing languages. However, if you poke around in enough word-families, you soon unearth many scattered, unpatterned examples of heterorganic shifts. Naturally these are the most interesting and challenging cases. We shall list a few, but rather than attempt any *ad hoc* explanations let us just leave open the possibility that some of them may reflect real variational pressures that existed in these families already at the proto-stage:

- (a) SUCK/MILK (TSR 73)⁴¹ is an exceptionally complex word-family. We must reconstruct allofams with final stops in all three positions of articulation, labial, dental, and velar, as well as an open-syllabled variant: *ʔcup * *ʔjup (> Atsi suʔ-cʔup); *jut (> WB cut 'suck, imbibe, absorb;'; *cuk (> Jg. tšúʔ); *ʔjwa * *zwa (> WT hdžo-ba ~ žo 'milk'). This wild family also shows -u- * -i- alternation [below Db]: *ʔjip (> WT hdžibs 'suck').
- (b) PIERCE/STAB/THORN (TSR 107). Besides the open-syllable variant meaning 'thorn' (*tsow), the family comprises an allofam with final velar (*Ntsuk * *ʔtsuk > WT hdzugs 'prick,' Jg. tšók 'pierce') and another one with final dental (*Ntsut * *ʔtsut > WT hts'ud 'be put into,' Jg. džút 'be pierced,' šədžút 'cause to pierce').

(c) **SOAK/WATER/EJACULATE** (TSR 109). On the basis of the Loloish evidence so far discovered, we cannot decide whether to reconstruct this set as *Ntit **ʔtit or as *Ntik **ʔtik. Jinghpaw has two allofams, one with a dental final (mədît 'to moisten, wet; be wet') and one with a glottal stop < *-k mədi (mədīʔ 'be moist, wet'). Undoubtedly related is the open-syllabled variant that means 'water,' PTB *m-ti(y) [STC 55]. The newest member of the family I have found is Lahu dī (< *Ndi) 'ejaculate (of a man); moisten due to sexual excitement (of a woman).' ⁴²

(4) **Variations of medial glides.** Liquids and semi-vowels are notorious for their variability and interchangeability into each other. All possible two-way variations among the four sounds /l, r, w, y/ are amply attested in the world's languages: l * r, l * y, l * w, r * y, r * w, y * w. Tibeto-Burman is no exception.

Variations in medial glides are of two basic types: either (a) two or more different glides appear in the word-family, but every allofam has some glide or other, i.e., -G₁- * -G₂-; or (b) one or more allofams have a glide, but one or more do not, i.e., -G- * -Ø-.

Both sorts of variation are well attested in TB, whether we consider word-families within a single language or across languages. Let us confine our discussion to Lolo-Burmese:

A. Intra-lingual -G₁- * -G₂-. Inscriptional Burmese has many forms with medial -l-, but even in the Inscriptions variation among -l- ~ -r- ~ -ly- is common [STC, note 134, p. 41]. The modern written language is full of orthographic doublets where -r- alternates with -y-, especially after velars. A few random examples: 'dove' khrûi * khyûi; 'horn' khruì * khyui; 'boil to a pulp' kyui * krui; 'transverse stay, prop' kyâ * krâ; 'thread' khyañ * khrañ; 'windpipe' khyôn * khron; 'prepare, put in order' pran * pyan. Alternations between -w- and the other glides is rarer, but examples do exist, like the triplet kwap * kyap * krap 'clamp, make fast, secure as by tying.'

Lahu normally keeps its reflexes of PLB *k-, *kr-, and *ky- strictly apart, showing q-, k-, and c- respectively.⁴³ Occasional cases of variation may be found, however: 'red termite hill' khε-pû-lô? (<PLB *kr-) * chε-pû-lô? (< PLB *ky-).

A similar apparent doublet within Lahu might seem to justify setting up a new proto-variant for the important family *m-kri-t 'bile' [STC 412] * *kri(y) 'sour' [STC 413]. Lahu has ṣ-kṣi 'gall-bladder,' which nicely fits into the *-r- prototype, but also has the form ci 'sour,' which makes it look as if we should recognize a new proto-allofam *kyi(y). For evidence to the contrary, see note 260.

B. Inter-lingual -G₁- * -G₂-. There are several cases (recognized already in Matisoff 1968) where WB and Lahu

disagree as to which proto-medial they reflect:

(a) **MANY/NUMEROUS.** WB has myâ, with medial -y-, but standard Black Lahu has mâ. The Lh. form cannot descend from the same proto-allofam, since *-ya regularly becomes Lh. -ɛ (e.g. 'bee' *bya > Lh. pê). In fact, the general TB root is *mra [STC 148], and it is WB which is out of step with the rest of the family.⁴⁴

(b) **MARROW.** WB has khraŋ-chi, but Lahu has ḁ-co-po (Burling 1967 reports a Lahu variant with aspirated initial ḁ-chɔ-pwɛ). The Lahu here reflects a prototype with *-y-, which makes this case the mirror-image of (a) and (c).

(c) **TIME.** WB khyin, Lh. khî. The WB form is from *-y-, and the Lahu is from *-r-. Mru (East Bengal and Arakan) has doublets, rin and khin, the latter possibly a loan < Burmese, but perhaps both reflecting a prototype *krin. [See Löffler 1966.]

C. **Intra-lingual -G- * -Ø-.** Occasionally we can find co-allofams within Lahu, where one variant points to an older medial glide while the other reflects a glideless prototype. Thus, 'the whole foot' khî-to-pa (< *pa)⁴⁵ * khî-to-pe (< *bya); dâ? 'be good' (< PLB *Ndak) * qha-dê? (< PLB *Ndyak)⁴⁶ 'well, properly.'

D. **Inter-lingual -G- * -Ø-.** The word for 'village' reconstructs as *kak in most Loloish languages [TSR 22],

e.g., Lahu qhâ?. But Ahi and Sani have tš'e₄₄, which definitely points to a proto-alternant (not recognized in TSR) that was either *krak or *kyak.

(5) **Variations of nuclear vowel.** The co-allofams in a word family may exhibit alternation in their nuclear vowels. Vowel alternations may be subclassified, according to whether they are observable within a single language (intra-lingually) or across languages (inter-lingually), and according to whether they are "primary" or "secondary." A **primary** vowel alternation is one which existed already in the proto-language. A synchronically observable vowel alternation in one or more daughter languages may also be called primary, if it is the direct reflection of a similar alternation in the proto-language. A **secondary** alternation is one which can be demonstrated to have arisen after the split-off of the daughter-language in question from some higher genetic node in the family, due to the operation of some conditioning factor in the phonological or morphological environment.

A. **Intra-lingual primary vowel alternations.** There is lots of evidence for an alternation between *-u- & *-i- in PTB [see below D(b)]. Some of this evidence is inter-lingual, but we can also find a direct reflection of this primary alternation within dialects of the same daughter language, or even between competing variants in the same dialect. Thus we have WT doublets like pus-mo & pis-mo

'knee' or smyug-ma * smyig-ma 'cane, bamboo.' Similarly with Dimasa (Bodo-Garo group): birin * burun 'bark, call (as an animal),' bithlim * buthlun 'brain.' These pairs of allofams exhibit an intralingual primary vowel alternation.

B. Intra-lingual secondary vowel alternations. A well-known example of this sort is furnished by English words like *foot/feet*, *tooth/teeth*, *goose/geese*. The high vowel in the plural forms arose through assimilation to the old Germanic pluralizing suffix -i. This "umlauting" of the vowel assumed the full burden of signifying plurality when the suffix was apocopated. The synchronically observable vowel alternation in these English words is secondary with respect to Proto-Indo-European, since it arose some time after the split-off of Germanic from the rest of the family.

The verbal morphology of WT displays striking types of vowel alternations, notably a pattern where a verb stem has -a- in the "perfect" and "future," -o- in the imperative, and sometimes -e- or -o- in the "present." Thus,

- (a) -a- * -a- [no alternation, except in the imperative]: Pres. ldag-pa, Pf. bldags, Fut. bldag, Imp. ldog 'lick'
- (b) -a- * -e-: Pres. hdebs-pa, Pf. btap, Fut. gtap, Imp. thob 'throw'
- (c) -a- * -o-: Pres. gsod-pa, Pf. bsad, Fut. bsad, gsad, Imp. sod 'kill.'

Attempts have been made [e.g. Shafer 1940-41, Miller 1956, Pulleyblank 1965] to impute such alternations to some regular morphological process in the proto-language (PTB or PST), i.e. to regard them as **primary**. Benedict, on the other hand, feels that vocalic alternation "played no role in PTB morphology" [STC, p. 126], and explains the WT alternations in terms of secondary influences upon the nuclear vowel of the verb root by older prefixes and suffixes. For Benedict the three alternations just mentioned reflect three different proto-nuclear vowels,

(a) *-a- > -a- * -a-

(b) *-ə- > -a- * -e-

(c) *-â- > -a- * -o-,

which yielded different reflexes in the present tense as they responded differentially to affixial influence.

At a more recent time-depth, modern Lahu has an interesting type of doublet-formation, where some forms having simple back vowels /u o ɔ/ have alternants pronounced with medial -w- plus the front vowel of the corresponding height /i e ε/: -u * -wi, -o * -we, -ɔ * -wε. Thus, ŋâ-ku * ŋâ-kwi 'dried fish,' yê-mí-tō * yê-mí-twē 'bear,' khî-qð? * khî-qwè? 'underside of the knee,' etc. [Matisoff 1973d, p. 19]. Although we must set up some kind of front/back vowel alternation for PTB [see *-u- * *-i-, below], it seems to me that this Lahu phenomenon is a quite late secondary development. For one thing, this doublet-

making is still a living, semi-productive process in Lahu, as proved by the fact that it is frequently applied to loan-words: co * cwe 'era, period of time,' ḥ-lḥ? * ḥ-lwê? 'terraced field,' lḥ * lwê 'swim' (all from Shan). What the source of this alternation may be remains an open question, though it may involve affective sound-symbolism,⁴⁷ since the prelabialized variant sometimes conveys a more colloquial, folksy, or vivid flavor than the allofam with simple vowel.

C. "Inter-lingual secondary vowel alternations." It is a commonplace of comparative linguistics that a given proto-vowel may have different reflexes in two given daughter languages. Thus, the proto-sequence, *-wa(-) is retained as such in WB, but regularly appears as the mid-vowel -o(-) in WT: 'tooth' *s-wa > WB swa, WT so; 'ant' *-rwak > WB pərwak, WT grog, etc. We could call this sort of thing an "interlingual secondary vowel alternation," though this is an unnecessary concept, included here merely for symmetry. There is no reason not to continue calling the relationship between WB -wa and WT -o a "vowel correspondence," and leave it at that.

D. Inter-lingual primary vowel alternations. The most interesting sort of vowel alternations are those which must be assumed to have already existed in the proto-language itself on the basis of comparative evidence from the daughter languages.

So far no one has succeeded in demonstrating any non-chimerical primary vowel alternations in open syllables.

In closed syllables, however, there is strong evidence for two primary variational patterns: (a) -ya- * -i-, and (b) -u- * -i-.

(a) *-ya- * *-i-. The most important root showing this alternation is 'eye,' PTB *mik * *myak [STC 402 and n. 251, p. 84]. Most of the TB languages reflect the -i- variant, with WT mig being typical. However, Karen, Lolo-Burmese, Nung, and Gyarung point to the -ya- variant (with WB myak being typical). Note that WT mig and WB myak do not exemplify a regular correspondence in the sense of paragraph C above. In normal cases, WT -ig corresponds rather to WB -ac, as in 'joint' *tsik > WT ts'igs, WB chac. Conversely, when other TB languages point to *-yak, the normal WT reflex is *-(y)ag, as in 'lick' *lyak > WB lyak, WT ldžags 'tongue (resp.).' We must therefore posit proto-variation in this etymon at least as far back as the PTB level.

It is characteristic of primary variation of this sort that there is no predictable pattern according to which one language will always choose one proto-allofom while another language will always choose the other. Take for example the etymon 'pheasant,' reconstructed as PTB *s-rik * *s-ryak [STC 403, p. 85]. If we consider the WT and WB forms, we find that these languages have retained the

opposite proto-variants from the ones they reflect for 'eye'! Thus, WT has sreg-pa (West Tb. šrag-pa) < *s-ryak, while WB has rac < *s-rik. In tabular form:

	*- <u>i</u> -	*- <u>ya</u> -
'eye'	WT <u>mig</u>	WB <u>myak</u>
'pheasant'	WB <u>rac</u>	WT <u>sreg-pa</u> / <u>šrag-pa</u>

This would seem to indicate that we would be barking up the wrong tree were we to try using patterns of proto-allofam selection as a criterion for higher-order subgroupings of the TB family.⁴⁸

(b) *-u- * *-i-. Alternations between these two high vowels, both intra- and inter-lingually, were already discussed by Wolfenden 1929 (pp. 114-5), and mentioned again in STC (pp. 80-84). Among many examples we could cite, let us take 'sleep,' PTB *yip * *yup. Here Jg. retains both variants, with semantic differentiation: íp ~ yíp 'cover, conceal (information)' * ʔyúp 'sleep.' Inter-lingually, Jinghpaw's -i- allofam goes with PLB *yip 'sleep' [TSR 180] and with WT yib-pa 'hide oneself;' while Jg's -u- allofam is cognate to forms like Miri yup 'sleep.' Similar random patterning of allofams is shown with other inter-lingual pairs like WT k'yim * Lepcha k'yum 'house;' Jg. rīm 'be dusk,' ń-rīm 'evening twilight'⁴⁹ * WT rum 'dark-

ness, obscurity;' Nung dip 'beat' * Jg. dùp 'beat, strike,' etc.

In Matisoff 1972a, I reconstructed a Proto-Lolo-Burmese form for 'sweep/wipe' (#120) as *sut, on the basis of the testimony of WB sut. What I should have done was reconstruct a PLB word-family of the form *sut * *sit, since all the Loloish forms cited point to an allofam with the high front vowel: Lh. šî?, Ak. si[^], Sani sɹ₄₄, Lisu si₂, etc. Similarly, the set 'tear/rip' is reconstructed as *Njut * *?jut in TSR 110, though I should have also recognized an allofam *Njit on the basis of Lahu jè?.

When we speak of "primary" vs. "secondary" alternations, we are of course speaking relatively to a particular point in genetic time. An alternation may simultaneously be primary with respect to some later stage and secondary with respect to some earlier stage. For example, even though we cannot get rid of the *-i- * *-ya- alternation at the PTB level, we might still wonder whether one pole of the variation is more archaic (i.e., "more primary") than the other. In this case, Benedict considers the *-ya- variant to be older, since this is the one that is usually reflected in Chinese (cf. Ch. *miôk/miuk 目 'eye' [GSR 1036a-c],⁵⁰ as well as in early Miao-Yao loans from Sino-Tibetan [STC, p. 84]. This is not conclusive, however, since we still cannot demonstrate that the *-i- variant is somehow derivable

by the operation of regular phonological processes from the hypothetically "older" **-ya-*. Neither can it be shown that any morphological significance attaches to the variation. All we can say is that the conditioning factors for the alternation have not yet been discovered, so that the alternation is "primary" as far back as we can now go.⁵¹

E. Miscellaneous but non-phony vowel alternations. We now enter a gray never-never land, intermediate between the well-established variations we have been discussing and the clearly illusory "pseudo-alternations" we shall consider in the next section.

There are many TB roots displaying irregularities in vowel correspondences which cannot be accounted for in the present state of our knowledge. Some of these will undoubtedly be taken care of with the discovery of additional forms in one or another language that will furnish missing links enabling us to reduce the aberrancies to some more regular pattern. For now, we can just say that in many cases the irregular forms look too good semantically and phonetically to throw out. If it makes us feel better we can talk about "sporadic and irregular vowel gradation" [STC, p. 69], as long as we recognize that we still don't really know what is going on.

(a) *-a- * -u-*. Benedict [STC 405] has a root **sun* 'pleasant smell' (WT *bsun*, Jg. *sun*), where WB unexpectedly has *-a-* (*sân*). To this we may add 'descend' [Matisoff

1974b, No. 363] PLB *zak (see TSR No. 121), where Jg. has a variant with -u-, ?yú?.⁵² Occasionally, we find ablauting compounds involving these two vowels, like WT khrag-khrug 'complicated, confused, like a troop of fighting men or the loose leaves of a book.' This WT form seems to have a direct Lahu cognate, with the order of the syllables reversed: khò?-khà? 'sparse, scattered about.'⁵³

In one interesting word-family referring to a "diffuse organ" of the body, there is strong evidence for **both** the *-u- * *-i- and the *-u- * *-a- alternations: the lateral-finalled root BODY HAIR has (among others) the PTB allomorphs *s-mul (> Lushai hmul, etc.); *s-mil (> Nung mil, Lepcha ămyel, etc.); *s-mal (> Lepcha ă-myal, Magari mi-mhā, and perhaps Chepang myāng). The -y- in the Lepcha doublets is the regular reflex of prefixial *s- (see Benedict 1943), so there is probably no need to posit *-i- * *-ya- alternation in this family.

In STC (p. 75, note 231), Benedict comments on the frequent interchange of -u- and -a- in Lepcha, ascribing this in effect to an earlier proto-variation between PTB long and short medial u, with PTB *-u- > Lp. -a- or -ă-, while PTB *-u*- > Lp. -u-.

(b) *-ow * *-u(w). STC mentions three roots in this category, 'thorn' *tsow * *tsu(w); 'steal' *r-kuw * *kow; and 'hammer' *tu(w) * *tow [p. 69].

(c) *-a- * *-i- and other alternations involving a high

front element. STC sets up a root (404) 'alive' with *-a- * *-i- (*s-rin * *s-ran), implying that this is to be considered as related to the well-established *-ya- * *-i- variation. Benedict (p. 85) also mentions the WB form khraṇ-chi 'marrow,' which has -a- instead of expected -i- (< *r-klin). Actually I now prefer to assign the WB form to a separate prototype, *r-k(y)aṇ * *k-r(y)aṇ.

Other roots showing variations involving -i- or -y- include: 'low, short' *nem [STC 348, p. 74] * *nyam [p. 85]: *-e- * *-ya-; 'father's sister' *ni(y) * *ney [STC, p. 69]: *-i(y) * *-ey; 'near' *ney [STC 291, p. 65] * *na·y [p. 68]: *-ey * *-a·y; 'left (side)' *bay [STC 47, p. 24] * *r-bi(y) [p. 68]: *-i(y) * *-ay.

F. Digression: Pseudo-alternations as a product of faulty analysis. A propos of word-families in Indo-European, Buck observes (p. xiv):

"The uncertainty and speculation which are often involved in the grouping under roots and root extensions (which...are only convenient abstractions of elements common to groups of actual words) are well known."

That goes double for Tibeto-Burman. The analyst must be on his guard lest he be seduced into facile speculation about possible allofamic relationships that are phonetically aberrant and semantically wild at the same time. Roy Andrew Miller has furnished us with several classic examples of this genre, where "alternations" are posited based on faulty cognate identifications or erroneous assignment of un-

related forms to the same word-family. Thus Miller 1956 attempts to justify an alternation among syllable-final *-a * -u * *-i on the basis of forms like WB na 'ill,' nu 'leprous,' and ni 'red.'⁵⁴ These forms are in fact totally unrelated. It is not the semantic difference among them which disqualifies them *per se* from an allofamic relationship. As we shall see below, semantic shifting is unpredictable and often radical. Rather, there is no evidence that the vowels */-a -u -i/ ever interrelate in open syllables in cases where the meanings of the forms in question are identical (or at least close enough so that no special *ad hoc* justification is necessary).⁵⁵ In other words, if a vocalic alternation is well-established in several clear, unimpeachable cases, then we might be willing to accept as genuine relatives forms which showed this same phonetic alternation even if their meanings were quite different [cf. 'sleep/conceal,' above D(b)]. However, if on the one hand the phonetic alternation is **not** well-established **and** on the other hand the semantic relationship is far-fetched, then our argument is like two drunkards supporting each other.⁵⁶ It would be as if some Martian linguist, Yor Werdna Rellim, were to take the two English roots spit and split and attempt to relate them using the following reasoning: "There is a clear phonological relationship between these words. In fact we may set up a proto-form *sp(l)it or *sp-l-it, with the latter showing an -l- infix.

Semantically the two forms are easy to relate. Both refer to an action of separating—just as the saliva is separated from the body by the act of spitting, so is the wood cloven in twain by the act of splitting. The meaning of the infix -l- is, clearly, 'action performed upon an inanimate object.' The form spit, lacking this infix, refers to action performed upon an animate object, namely the body. This analysis is reinforced if we consider two other related forms, slit and shit. Again the words are closely related semantically, since slit refers to the creation of a narrow orifice, and it is well known that the Earthling body expels solid wastes through a similar orifice. Once more we are struck by the fact that the infixed form slit (< *s-l-it) typically refers to action impinging upon an inanimate object, while the morphologically simple form shit (phonetically [šit] < *sit, with secondary palatalization of the root-initial due to the following high vowel) does not, at least in the same sense...."

(6) **Variations in syllable-initial position: voicing, aspiration, and prefixation.** A ubiquitous type of allofamic alternation in syllable-initial position is variation in the voicing of the root-initial consonant. In many cases the [+voice] * [-voice] phenomenon can clearly be demonstrated to be due to the effect of prefixes on the C_i. For example, the old nasal prefix *N- typically voiced a following surd C_i, while the sibilant and glottal prefixes, *s- and *H-,

typically devoiced a following sonant initial. Sometimes, however, the conditioning factor for the voicing alternation remains obscure. A striking case in point is provided in Jinghpaw, whose voiced series of obstruents corresponds randomly in word families both to the PLB *voiced and *voiceless series [Matisoff 1974f, pp. 155-6]. The promiscuous comparatist, faced with such a situation, can always posit proto-prefixes to "account" for the alternation, but in the absence of independent evidence, this is a circular procedure. You can do a lot with prefixes, but they should not be treated as wild cards. In some cases it is best to recognize that the voicing alternation is primary with respect to a given stage of the proto-language that we are reconstructing, i.e. that the proto-language itself displayed some variation [see STC, p. 124].

There has been much misunderstanding on this point, on the part of writers who have a simplistic view of linguistic structure as totally homogeneous. Thus Miller 1974 sarcastically ridicules the *Conspectus* [pp. 17ff] for giving formulae for correspondences of the form,

PTB *k > Jg. k(h) ~ g

PTB *g > Jg. g ~ k(h).

Yet Benedict is merely being honest here. Let us try to understand what he means!

Many TB languages (as well as reconstructed Archaic Chinese) have a three-way contrast between voiceless un-

aspirated, voiceless aspirated, and voiced obstruents. But in most cases one or another of the three series is easily shown to be secondary. Thus in WT the two voiceless series are in near complementary distribution, with the aspirated series occurring in absolutely initial position and after the prefixes m- and h-, and the unaspirated series occurring elsewhere.⁵⁷

In Akha, the two voiceless series are in exact complementary distribution, with the aspirates occurring only under non-checked tones [Lewis 1968].

In Lisu, as first pointed out by Burling (1967), the voiceless unaspirated series is the reflex of the old *glottal prefix, while the aspirates come from the old unprefixated *voiceless series, thus:

PTB * <u>b</u>	>	Lisu <u>b</u> ,
PTB * <u>p</u>	>	Lisu <u>ph</u> ,
PTB * <u>?p</u> * <u>?b</u>	>	Lisu <u>p</u> .

In Lahu, I have shown [Matisoff 1969a and subsequently] how the voiced series /b d j g/ arose secondarily through the influence of the old nasal prefix. The old *voiced series lost its voicing, yielding the modern Lahu voiceless unaspirates, while the old *voiceless series phonologized its presumably always redundantly present aspiration, yielding the new Lahu voiceless aspirated series, thus:

PTB *b > Lahu p,
 PTB *p > Lahu ph,
 PTB $\left. \begin{array}{l} *N-\underline{p} \\ *N-\underline{b} \end{array} \right\} > \text{Lahu } \underline{b}.$

A similar explanation can probably be worked out to account for the clearly secondary WB voiced obstruents, since WB, like Lahu, devoiced the original PTB *voiced series.

In many modern languages, like Burmese and Jinghpaw, there is much synchronically observable alternation between voiced and voiceless initials, occasionally accompanied by tonal alternations as well. There is one kind of Burmese voicing alternation that is predictable in the sense that one can specify a **necessary** (though not sufficient) phonological condition for its occurrence. Roughly stated, a Burmese syllable with voiceless obstruental initial undergoes voicing of the initial if it is preceded "in close juncture" by a syllable under tones "1," "2," or "3" (but not if the preceding syllable is in the "stopped tone" [$< \text{PLB } *-\underline{p} \text{ } -\underline{t} \text{ } -\underline{k}$]). Yet "close juncture" is not definable phonologically with respect to this rule, whose operation obviously depends to a large extent on the **grammatical** relationship between the syllables in question.⁵⁸

The point to all this is that the features of voicing and aspiration in syllable-initial position are one of the least stable areas of Sino-Tibetan phonology. The modern languages are rife with such variation at the pres-

ent day. There is no reason to suppose that the proto-language was any more stable in this regard. There is every reason to follow Benedict in his view that PTB (and PST) should be set up with only a two-way contrast in this area, with the proviso that the contrast was always "bursting at the seams" from the earliest stages of the proto-language that we can now reconstruct. Pressure for the phonologization of the ever-present redundant feature of aspiration was always there.⁵⁹ Interaction with prefixes constantly reshuffled the patterns of contrast in syllable-initial position, as the languages fought to maintain their power to distinguish utterances in the face of disastrous mergers and fusions in the affixial system.

Kun Chang is quite conservative in his view that "we must have a better understanding of this problem [of accounting for anomalies in voicing and aspiration between Chinese and TB] before we can make serious Tibeto-Burman: Chinese comparisons."⁶⁰ After all, "serious" comparisons have been possible now for a long time! Voicing and aspiration are coming to be better understood, both because of our deepened knowledge of the effects of prefixes in TB and because of a more sophisticated understanding of linguistic variation in general.⁶¹ In this same review, Chang several times displays a rigid and mechanistic view of prefixial behavior. For example, he criticizes setting up a noun-prefix *s- with a meaning referring to body-parts or ani-

mals, deriving from the full morpheme *sya 'animal/flesh.'⁶² He calls this "a dubious proposition which fails to account for those nouns with s- which have nothing to do with body parts or animals." This is as if someone were to deny that there is a semantic association between the English initial cluster sm- and sn- and the nose-mouth region (snot, snicker, snuff, smell, smirk, smile, sneeze, snore, snout, snoot, etc., etc.) on the grounds that this observation "fails to account for" words like small, snail, and snapdragon. (Benedict says explicitly [p. 107] that "most TB roots for animals can be reconstructed without this prefix.") It is totally unreasonable to expect a one-to-one correlation between the presence or absence of a particular prefix and the presence or absence of a particular meaning. Things are much less simple than that. We are talking of general (and conflicting) tendencies, not of iron-clad rules. (The behavior of TB prefixes is more reminiscent in its sporadicity of IE "derivation" than "inflection.")

Chang further criticizes the STC (p. 124) for reconstructing the proto-alternation between simplex and causative verb forms in two different ways—sometimes as involving the *s- causative prefix, and sometimes merely involving an alternation between voiced (simplex) and voiceless (causative) root-initial. Chang would like to derive all causative forms from the *s- prefix. This is admittedly an

intricate point, yet one senses that Professor Chang is made uncomfortable by the thought that one and the same grammatical process might possibly be signalled in two different morphological ways.⁶³ Equally well might one be disturbed that some English plurals are signified by zero instead of by a sibilant suffix (one sheep, two sheep, etc.)

As has been stressed repeatedly in the recent literature,⁶⁴ the prefixes interacted with the following root-initial consonant in a bewildering variety of ways, especially when the C_1 was a resonant */w y r l/. Besides affecting the voicing or aspiration of the root-initial, the prefixes could metathesize with it, palatalize it, drive it out entirely ("prefix-preemption"), fuse with it into a single segment, drop altogether, be substituted for by another prefix—and any or all of these activities could be accompanied by an effect on the **tone** of the syllable. There is no way of mechanically predicting what the fate of a given root might be in relation to the prefixes. To this one may either say "alas!" or "hurray!," depending on one's temperament.

Furthermore, prefixation in TB is a living, cyclical process. Fully stressed members of compounds may lose their stress and become "cliticized," ending up as prefixes. A given root whose "original" prefix had somehow become swallowed up, may later get **re-prefixed**, so that not all prefixes in a language at any moment in time are of equal antiquity.

Finally, there is no simple way to relate prefixial behavior to genetic subgrouping. Though one can generalize about the fate of the prefixes in the various nuclear groups of TB [STC, p. 104], when one gets down to individual roots, *chaque mot a son histoire*. Black Lahu and Yellow Lahu are dialects of the same language, as like each other as two (slightly different shaped) peas in the same pod. The Black Lahu word for 'tiger' is lâ, while the Yellow Lahu form is cà-mě. The first syllable is a fusional reflex of a prefixed version of the same root as the Black Lahu form, *k-la. Why did Y. Lh. prefix this root, while B. Lh. did not?⁶⁵

A. Sibilants and affricates in root-initial position.

Another noteworthy type of initial alternation might also be ultimately conditioned by prefixes, and seems to reflect an inherent instability in the PTB phonological structure itself: alternations involving **affricates** and **sibilants**. Many Lolo-Burmese languages have undergone mergers of the *dental and *palatal series of spirants and affricates (WB, Lahu).⁶⁶ In the case of Lahu, the phonetic distinction still survives, in a curious pattern of complementary distribution where the palatal phone appears before all the vowels except /ɨ/, and the dental equivalent appears only before /ɨ/, thus:

/c ch j š y/ → [ts tsh dz s z]/ɨ.

Lahu is thus one of the world's relatively few languages to

lack an /s/-phoneme.⁶⁷

Several TB roots show intra- and inter-lingual variation between dental affricates and sibilants. WT sometimes displays such variation in its verb paradigms (e.g. hdzugs, zug, hdzug 'prick, pierce; enter, begin').⁶⁸ The wide-spread root *dza 'eat' [STC 66] appears with a simple sibilant initial in WT (za-ba) and Kanauri (za). The root for 'urine' [STC 77] reflects an affricated prototype *ts(y)i- in many TB languages (e.g., WT gtśid * gtśi * gtśin, WB chî), but Lolo-Burmese also points to an allofam with a simple sibilant (WB sê < *ziy).⁶⁹ One of the common TB roots for 'hair' must be set up at the proto-stage as *tsam * *sam [STC 73], since both variants are well attested in many languages.⁷⁰ Lahu has a doublet pair for 'hair of the head,' cĕ-khe-mu * yĕ-khe-mu (phonetically [tsĕ * zĕ]—see above). This reflects an older *dz * *z variation, but these forms cannot be directly related to *tsam * *sam, since both the vowel and the tone are wrong.⁷¹ The root for 'child, son' [STC 59 and note 86, p. 27] shows variation between *za and *tsa or *dza, both in TB and in Chinese. Lahu, I believe, preserves both variants, with semantic differentiation: Lh. yâ 'child, son' < *za (PLB tone *2), and câ- 'prefix to male names' < *dza (PLB tone *1).^{71-a} Reconstructive digging in this consonantal area must be carried out with a toothbrush and not a pick-axe. In making comparisons between Jinghpaw and LB, for example, one finds

such a multiplicity of different correspondences in the sibilant/affricate realm that proto-variation is the only plausible explanation [see Matisoff 1974f, pp. 156-7].

B. Resonantal alternation in root-initial position: "lenition" of labial stops. As an example of a variational pattern which has not yet received any attention in TB, we may mention Lahu doublets like $\underline{g}\hat{o}?$ * $\underline{v}\hat{o}?$ 'pick up, hold in the hand' and $\hat{o}-\underline{g}\hat{o}$ * $\hat{o}-\underline{v}\hat{o}$ 'ring' [GL, p. 9]. The modern \underline{g} [γ] * \underline{v} alternation reflects a PLB pattern of * \underline{r} * * \underline{w} . (I like to refer to this as the "wittle wabbit syndwome.") In general Loloish displays much instability in this area. Lisu for example has multiple conditioned reflexes of PLB * \underline{r} : $\underline{\gamma}$, \underline{w} , and \underline{y} [Thurgood 1974a, pp. 33-36].

In several important TB roots ('father,' 'pig,' 'bamboo,' 'leech,' 'hide,' 'left (side),' 'sole,' 'flower'), the modern languages show variations between an initial labial stop and the labial semivowel \underline{w} -. Benedict formerly felt this \underline{p} - * \underline{w} - alternation was due to prefixial influence [STC, p. 23], but has now taken the position that it reflects PST clusters like * \underline{pw} - and * \underline{bw} - [STC, note 78, pp. 23-4]. I have adopted the term **lenition** from Celtic linguistics to characterize the appearance of an initial semivowel in a word-family that also contains members with the homorganic stop. (See the sets for BELLY/STOMACH presented below, 1.141b.) If we accept Benedict's proto-cluster explanation, this "lenition" is really nothing more

than the metanalysis of the original stop-component of the cluster as a prefix, which then was free to drop: *bwak > *b-wak > *wak.

(7) **Tonal variations in word-families.** This is a vast topic which we cannot begin to go into here. All TB languages for which we have reliable and copious tonal information show abundant allofamic variation in this area. Sometimes it is possible to demonstrate a precise origin for this alternation in terms of earlier segmental morphology. Thus Lahu simplex/causative pairs like

dè 'drink' / tɔ 'give to drink,'

câ 'eat' / cā 'feed,'

và? 'hide oneself' / fá 'hide something,'

display regular tonal reflexes of the old *s- (> Proto-Loloish *ʔ-) causative prefix.⁷² Often, however, the tonal alternations (intra- or inter-lingual) show no particular pattern. Examples may be multiplied at will: among those already mentioned in Lahu we have seen yâ 'son' / câ- 'prefix to men's names;' là 'to come,' la 'cisative verb-particle,' lâ 'cisative benefactive particle.'⁷³

Sometimes one can point to similar-looking allofamic tonal variations across languages, without being sure whether this is a common inheritance from some earlier stage or merely parallel independent development. Thus, Jinghpaw, Burmese, and Lahu all have phonologically related pairs of words for 'ear' and 'hear:'

Jg. nā (mid-tone) 'ear' * nà (low-tone) 'to hear,'
 WB. nâ (<PLB Tone *2) 'ear' * na (<PLB Tone *1) 'listen,'
 Lh. nā-pɔ (<PLB *na or *sna, Tone *2) 'ear' * nà
 'listen' (< PLB *na, Tone *1).⁷⁴

As always, care must be taken to distinguish genuine co-allofams from unrelated forms which have "accidentally" come to have the same segmental shape and have similar meanings, but now differ from each other only by tone. Thus, Lahu mu 'high,' resembles mû 'sky' both in sound and meaning; yet mu < *mran 'high,' while mû < *r-muw 'sky.' Similarly, Lh. phu 'money, silver' < *plu, while phû 'price' < *puw.⁷⁵

Some modern TB languages (e.g., Tiddim Chin [Henderson 1967], Lushai [Weidert 1975]) show highly productive patterns of synchronic tonal variation which cut across large form-classes. These are certainly to be considered as late developments, in which formerly sporadic alternational patterns have been analogically regularized. The proto-language had nothing that neat and tidy.

1.122 Compounding and genetic relationship. Compounding is a pervasive TB morphological process. The modern TB languages are still monosyllabic, but only in the sense that the typical **morpheme** is one syllable long. Words (especially nouns, but also verbs) very often have two syllables, and three-syllable nouns are by no means uncommon. In literary, liturgical, or formal style, four-

syllable compounds, both nominal and verbal, abound.⁷⁶ The longest TB words are names for plants and animals, which may run to five or even six syllables: e.g., Lahu pā-pā-qú-ti-ni 'dragonfly,' hɛ-ṣā-tḁ-qú-cè 'plant of the nightshade family,' khâ?-pā-mē-cí-câ-kwî 'racket-tailed drongo,' etc.

(a) **The transparency/opacity continuum: morphans and folk etymology.** As in all languages, there is a multi-dimensional continuum of transparency in TB compounds.⁷⁷ At the most transparent end of the scale are compounds in which each syllable has a clear meaning of its own, such that the meaning of the whole compound is deducible from the sum of its discrete parts. Some examples from Lahu body-part nomenclature will serve: mê?-gǝ 'tear' ("eye" + "liquid"), là?-cí 'wrist' ("hand" + "joint"), mê?-mu 'eyelash' ("eye" + "hair"), nī-ó-qō 'glans penis' (nī "penis" + ó-qō "head"),⁷⁸ etc.

Semantically less transparent are compounds whose constituents have clear meanings individually, but where the meaning of the compound as a whole is not deducible from them. In these cases the language has departed from literality and indulged in a shift or extension or displacement of meaning. When American soldiers speak of their "dog-tags," they mean the metal identification tags they wear around their own necks. When a Lahu speaks of his "dog-shit finger" (phâ-qhê-là?-nɔ) he means his index

finger. (For the Lahu the size and shape of the fore-finger suggest a dog-turd.⁷⁹) The literal meaning of the name for that 'plant of the nightshade family' just mentioned is 'old-field turtle plant' (hɛ-ŋ̃ā 'old-field' [i.e., a mountain field that had been planted to grow rice but which is now reverting to jungle or lying fallow] + tò-qú 'turtle' + cê 'plant.')

Evidently this plant typically springs up in old fields, and perhaps its leaves resemble turtle-shells (or maybe turtles like to eat it?). At any rate a non-Lahu, when presented with several plants, would have a hard time guessing which one of them was associated with turtles in the Lahu mind.

Constituents of a compound frequently acquire a measure of phonological opacity by undergoing a change in pronunciation in combination with the adjacent morpheme. If the phonetic change is not too radical, and provided the meaning of the compound does not undergo any unpredictable shifts, the native speaker will still be able to analyze the compound correctly. In these cases we are justified in speaking of "allomorphs of the same morpheme." The first syllable of Lahu hɛ̃-ŋ̃ā 'jungle-fowl' ('field' + 'chicken') is in the high-rising tone, though the underlying word for 'field' is in the mid-tone (hɛ): hɛ * hɛ̃-. The morpheme for 'head' appears as ó- in ó-qō 'head,' but as ú- in ú-gê 'pillow.' In more extreme cases, a constituent suffers such a change in pronunciation (typically through destress-

ing) that it is no longer segmentable by the native speaker. Well-known English examples are *hussy* (< "house" + "wife") and *waistcoat* /weskit/ (< "waist" + "coat"). The compound is especially vulnerable to semantic shifting once the original meaning of a constituent has been obscured in this way. The ultimate degradation for a constituent is for it to become so phonologically reduced that it is "cliticized," losing its vowel completely and becoming a prefix or suffix [below].

Even without phonetic decay, a syllable in a compound may become totally opaque, in that it has no meaning of its own in isolation. These pitiful entities I have called "morphans" (or "orphan morphs").⁸⁰ The TB languages abound in these morphans, as of course does English. Let us not cite the *cran-* of 'cranberry,' since the example is so trite. Rather let us consider the word *nostril*. Every speaker of English recognizes that the first syllable, though pronounced /nas-/, is morphemically identical to *nose* /nowz/—these forms are allomorphs of each other, in the same sense as Lahu hɛ * hɛ́- 'field' [preceding paragraph]. With *-tril*, it is a different story. Nobody knows (unless he has just looked it up in a good dictionary) that *-tril* comes from Old English *thyrel* 'hole.' The once transparent compound "nose-hole" has become opacified through the obsolescence of one of its constituents. *-tril* is now a morphan. If we did not have documentary evidence

on older stages of English, we would be forced to rely on comparative evidence from other Germanic languages. Suspecting on universal semantic grounds that the second element must mean 'hole,' we would start by looking up the words for 'hole' in other languages, hoping to find a phonological fit. The same game may be played with fun and profit in Tibeto-Burman.

A morphan may be infused with new life by the process known as **folk etymology**. *Asparagus* becomes "sparrow-grass," **grozeberry* (cf. French *groseille* 'currant') becomes "gooseberry," *chaise longue* (Fr. 'long chair') becomes "chezz lounge." The American town named *Chemin Couvert* ('covered road') by early French explorers eventually became known to English-speaking settlers as *Smackover*. "Smackover" doesn't mean much, but at least it consists of two real English words.... As these examples show, loanwords are especially vulnerable to this etymologizing treatment. Speakers of languages where morphemes are typically monosyllabic have a strong tendency to read new meanings into unfamiliar syllables. Lahu borrows English 'cigarette' as śú-kêʔ-lêʔ, identifying the first syllable with the native word śú 'tobacco.'

We might compare a morphan to a hermit-crab, constantly on the make for a new meaning-shell to clothe itself with.

Technological and cultural change may cause the semantic content of an obscure word to be inherited by an

etymologically distinct homophone. Most American children interpret the cheery nautical cry "Anchors aweigh!" as "Anchors **away!**" (that is, "Away with the anchors, let's be off!" rather than "Let us **weigh** [i.e. **heave up**] the anchor!"). Everywhere we look, folk etymologies make headweigh!⁸¹

Small children nowadays are more familiar with dogs than wolves, so that when somebody is gobbling his food they tend to describe the action as "woofing it down" (instead of "wolfing it down"), i.e. "eating in the manner of a dog, the animal that goes woof."

(b) *Allocompy and the Urwortschatz*. I have always liked the German word for 'vocabulary,' *Wortschatz*, literally "word-treasury" or "word-hoard." It is pleasant to think of the proto-TB stock of monosyllabic morphemes as an "Urwortschatz," or "proto-word-treasury"—a bulging treasure-house of jewel-like entities, each with its own idiosyncratic wedding of form to meaning, and each provided with a multitude of invisible hooks for hitching itself into unpredictable combinations with others of its kind. Each TB language, when faced with the necessity of forming polysyllabic compounds, dips into its inheritance from the proto-treasury in its own unpredictable way. Nishida 1967 furnishes us with a beautiful example of this process by listing in tabular form according to proto-root the compound for 'head' in WT and several Lolo-Burmese languages. See Figure 4.

Figure 4

Alternative selections from the proto-treasury
of morphemes in TB compound-formation

[Adapted from Nishida 1967, p. 68]

'HEAD'	Prefix	* <u>bu</u> (> * <u>wu</u>)	* <u>du</u> (k)	* <u>s-ko</u> (ŋ) 'hollow object'	*l(y)am ✕ * <u>lum</u>
Written Tibetan	d-	bu			
Bisu	ʔaŋ-		tù		
Akha		ʔù-	dù		
Written Burmese		ʔû-		khôn	
Lisu		wu ¹ -	dū ³		
Lahu Nâ ⁸² (Black Lahu)		ô-		qô	
Lahu Nyi ⁸² (Red Lahu)	ʔa-		-tù-	kù	
Lahu Shi (Yellow Lahu)	ʔa-			kù	
Maru		ʔáu-			lâm
Lashi		ʔú-			lêm
Atsi		u-			lum

It is clear that each of these closely related languages is a law unto itself, making idiosyncratic selections both from the stock of roots⁸³ and the stock of prefixes. Some languages are satisfied with a single root-morpheme, and prepose a prefix to it for "phonological bulk" (WT, Bisu, Lahu Shi). Several languages select two root-morphemes (Akha, Lisu, Burmese, Black Lahu, Maru, Lashi). Red Lahu has a trisyllabic form, where a prefix is preposed to a sequence of two root-morphemes. The languages which select two root-morphemes (but no prefix) may be further subdivided according to the particular roots selected: Akha and Lisu [*wu + *du]; Burmese and Black Lahu [*wu + *s-kon]; and Maru, Atsi, and Lashi [*wu + *l(y)am or *lum]. The root *s-ko(η) is an intruder from another semantic field, and can be demonstrated to have originally meant 'hollow object' (cf. WT k'un 'hole, pit, hollow, cavity').⁸⁴ Yet in Yellow Lahu (and perhaps in WT mgo) this morpheme has obviously acquired the full meaning of 'head,' since it appears as the only root-morpheme in the form.

The root *du(k) has affinities with both NECK and SKULL. STC #392 sets up a root *tuk * *duk 'neck' on the basis of forms like Jg. dù?, Garo gitok, Mikir tsethok, Lepcha tük-tok. Yet this etymon means 'head' in Abor-Miri a-tuk, Atong dək-əm, Wanang tək-əm, Kaike thoppā (< *thok-pa), etc. Certainly related is Chinese 體 *d'uk/d'uk

[GSR 1224m] 'skull.' Perhaps Chinese 頭 *d'u/d'əu [GSR 118e] 'head' and 脰 *d'u/d'əu- 'neck' are not to be referred to this PST prototype with dental root-initial and final velar stop [e.g. *(s-)tu·k or *(s)-to·k], but rather to the root with labial root-initial (*bu) that has been "pre-empted" by the dental prefix: *d-bu > *du (cf. WT dbu).

The last syllables in the Maru, Atsi, and Lashi forms are probably to be identified with a root meaning ROUND (STC #143 *s-lum or *z-lum), as suggested by WB lûm 'round,' ʔû-lûm 'the whole head.'⁸⁵

Taken all together, this "compound-family" [see above 1.12] demonstrates the futility of trying to correlate selections of particular morphemes in compound-formation with degrees of genetic relationship. The clinching case is provided by the three Lahu dialects. These are varieties of one and the same language (with a high rate of mutual intelligibility)—yet they each went a different route in making up a compound for 'head.'⁸⁶

What is true of 'head' goes for any other part of the body, or for any other compound word. We have no right to expect that any two TB languages, however closely related, will form compounds designating the same object in the real world by using exactly the same lexical material. In other words, there is no way to formulate constraints on the lexical or semantic composition of allocomps in the same com-

pound family.

(c) **Unpredictability of semantic source-fields in compound-formation: illustration from body-part nomenclature.** Consider the human uvula. There it sits in the back of the mouth, hanging down like a stalactite. Obviously it must be there for **some** reason, though it's hard to see what good it does anybody. Let's give it a name. What shall we call it, since we don't really know what its function is? When my elder daughter was about one and a half her mother yawned one day. Little Nadja could hardly restrain her excitement and pointed to the exposed uvula, chirping "nenoo, nenoo!" "Nenoo" was one of Nadja's first words, and meant 'nipple; the breast.' Clearly the shape of her mother's uvula reminded Nadja of the maternal breast. Her excitement was due to finding another one of those marvelous objects in her mother's mouth. For Nadja, a uvula was a "mouth-nipple." When one compares this spontaneously generated infantile Occidental word with the normal adult compound words for 'uvula' in several TB languages, one is struck once again by the truism that human thought-processes are everywhere the same:

Burmese hlya-khan "tongue-knoll," i.e. uvula;

Tangkhul Naga kharok-əməthin "throat-liver," i.e.

uvula;

Tibetan lče-č'uṅ "little tongue," i.e. uvula;

Lushai daṅ-mən "palate-clitoris," i.e. uvula.⁸⁷

Since the true function of the uvula is totally obscure to most people,⁸⁸ anything goes! Any random semantic association may be seized upon in naming it. For the Romans, the uvula looked like a little grape (Lat. *ūva* 'grape,' -*ula* 'fem. dimin. suffix').

François Rabelais, who has gone deeper into body-parts than any other French writer, lived at a time of great excitement with the rebirth of anatomical learning. Rabelais had the Renaissance man's fascination with the sonorous Greek and Latin words coined to name all the new body-parts that were coming to light under the knives of the revived science of dissection. Exuberantly he lists the parts of the body belonging to a grotesque giant called Quaresmeprenant (*Quart Livre*, Chs. 30-31). Each body-part is compared to an object in the outside world with which it seems not to have the remotest connection:

*Qvaresmeprenant, dist Xenomanes, quant aux
parties internes a...la ceruelle en grandeur,
couleur, substance, & vigueur semblable au
couillon guausche d'un Ciron masle....*

Les nerfs, comme un robinet....

La salive, comme une nauette....

Le coeur, comme une chasuble....

Les muscles, comme un soufflet....

Le membre, comme une pantopfle....

Le trou du cul, comme un miroir cristallin....

Les aisselles, comme un eschiquier....

(Quaresmeprenant, said Xenomanes, as far as his internal organs go, has a brain which is

comparable in size, color, substance, and vigor to the left ball of a male mite.... [He has] nerves like a spigot, saliva like a shuttle, a heart like a chasuble, muscles like a slap, a member like a slipper, an asshole like a crystalline mirror, armpits like a chessboard....)

Armpits, to take only the last example, do not seem to have anything much to do with chessboards—yet what Rabelais is telling us is that anything may be compared to or associated with anything else. We may consider some similes to be apter than others—Rabelais is purposely giving incongruous ones, since on one level he is parodying the rhetorical excesses of certain contemporary writers—but there is, in principle, no way to predict which meaning-associations will catch on in a language and which will die at birth.

For some Penutian tribes of California Indians, the human brain is "head-snot."⁸⁹ For the Black Lahu, the index finger is associated with the excrement of dogs.

(d) **Calquing: suspicious similarities in compound formation.** Given the unlimited lexical/semantic possibilities for coining compounds, we might well suspect that when contiguous languages that are not close genetic relatives have semantically identical compounds, this is not mere coincidence, but might be due to borrowing or "diffusion." Since *skyscrapers* originated recently in an English-speaking country, we have enough extralinguistic information to know that French *gratte-ciel* and German *Wolkenkratzer* are modelled on the English form, and not

vice versa. In Southeast Asia, however, there are many old compounds whose semantic structures are remarkably constant across genetic lines, such that it is now impossible to determine in what language the original coinage took place. In the case of semantically transparent compounds like "eye-water" for *tear*, the most reasonable hypothesis is usually parallel independent development. Sometimes, though, the collocation of semantic notions combined in the compound is unusual enough to rule that out, and we are tempted to adopt a diffusionist explanation. Some language coined a striking new compound (like *sky-scraper*) and other languages were so charmed that they adopted it. Purtle 1971 has collected several interesting examples:

"fire" + "tongue" > FLAME (Cambodian, Malagasy, Yao, Jinghpaw, Lahu);

"fly" + "shit" > FRECKLE, MOLE (Cambodian, Mon, Thai, Indonesian/Malay);

"pig" + "crazy/illness" > EPILEPSY (Cambodian, Mon, Thai, Indonesian/Malay, Burmese);

"eye" + "foot" > ANKLEBONE (Indonesian/Malay, Burmese, Lahu; cf. Cambodian "eye" + "cow," Mon "eye" +

"elephant," Vietnamese "eye" + "fish," etc.);

"tooth + "insect" > DENTAL CARIES (Cambodian, Vietnamese, White Meo, Thai, Jinghpaw, Burmese, Chinese);

"finger/hand" + "mother/female" > THUMB (Cambodian, Mon, Vietnamese, Thai, Indonesian/Malay, Burmese, Yao, Lahu, Mandarin, etc.)

But even in these cases one cannot be certain. Japanese

also has "parent-finger" for *thumb* (*oya-yubi*), and this does not appear to be a borrowing from Chinese. Even though we usually use a monomorphemic word for "flame" in ordinary English, we may also speak of "tongues of flame licking up the walls"....

We have seen that one and the same object (e.g., a uvula) may be designated by many different combinations of morphemes. As a final twist, let us note that the converse is also true. Identical combinations of morphemes may mean totally different things in different languages. In Malay, "eye" + "day" means *sun* (*mata hari*). The sun is the eye of the day. In Old English times somebody once looked at a flower with white petals and a yellow center, and called it a *daeges-ēge* or "day's-eye." We still call that flower *daisy*, though phonological reduction has now robbed us of any sense of the original metaphor, and the word is now monomorphemic to all but the most determined antiquarians.⁹⁰

(e) **Reprise: the compounding/prefixation cycle.** In my view both compounding and prefixation are devices which conspire in TB to combat creeping homophony by adding more phonological and/or semantic bulk to monosyllabic morphemes. Both processes feed cyclically into each other, since fully syllabic constituents may be reduced to affixes, affixes may become fully syllabic, and older syllables or prefixes may drop completely at any time and be replaced later. It

is a chicken-and-egg question to ask which process is the more ancient in Sino-Tibetan.

A similar cyclical view is helpful in considering the history of tonal development in the family [Matisoff 1973c] or the question of the number of voicing and aspiration contrasts at various stages of proto-time [above 1.121(6)]. The languages make do with whatever they have, striving to maintain their communicative integrity in the face of constant change.

1.13 Comparative syntax and genetic relationship. Of all the various aspects of linguistic structure, syntax might well be the **least** useful for establishing fine degrees of genetic relationship. For the purposes of this brief discussion, let us take "syntax" to mean the behavior of words in constructions and sentences, especially the inter-relationship between the grammatical category to which words belong and their relative ordering with other words.

Before one can even begin to make cross-linguistic syntactic comparisons, one is forced to make impressionistic decisions about the relative importance of conflicting syntactic tendencies within each individual language to be compared. Suppose for example that you want to compare English genitives to genitival constructions in other languages. The most basic thing to consider would be the relative ordering of the three constituents in a genitive

construction: the possessor nucleus [v_p], the thing possessed (or the "possessed head" [v_h], and the genitive marker.⁹¹ Some languages have the possessor first, some have the thing possessed first. But in English we have constructions of both types:

the razor's edge [v_p + GEN MARKER + v_h]
's

the edge of the razor [v_h + GEN MARKER + v_p].
of

One could argue endlessly about which construction is now the more "basic" in English. Historically the affixial or synthetic genitive ('s) is older, but synchronically the analytic genitive now seems predominant (unless the possessor is animate, where the reverse is true). In German the affixial genitive is even more vigorously maintained than in English, but the usual order has the possessed head coming first, unlike English:

das Buch des Mannes 'the man's book.'

v_h v_p

The reverse order is also possible (*des Mannes Buch*), though this is now poetical or highflown style (*der Früchte Saft* 'the juice of the fruits,'⁹² *der Liebe und des Meeres Wellen* 'the waves of love and the sea').⁹³ Besides, German also has an analytic genitive, with the preposition

von corresponding to English *of*: *das Lied von der Erde*

v_h

v_p

'the song of the earth.' In Yiddish the predominant genitive construction is clearly the analytic one, with the preposition *fun* (cognate to German *von*): *dos bux funem*

v_h

man 'the man's book,' *der pupik fun a floy* 'the navel of

v_p

a flea.' Yet an affixial genitive also exists, though it is restricted mostly to animate possessors: *a mame's harts* 'a mother's heart,' *dem Yidn's simxe* 'the Jew's happiness.'

If these three languages had no written records, a comparatist might well conclude that English and Yiddish were more closely related to each other than either is to German, on the basis of shared similarities in genitive formation (e.g., affixial genitives in both English and Yiddish show a special affinity for animate possessors, while German has no such tendency). The point is, one must know the grammar of a language inside out before one can dare to isolate a given syntactic feature for comparison with a similar feature in another language. Few Tibeto-Burman languages have been described in enough detail to make meaningful syntactic comparisons possible.

Suppose, however, for the sake of argument, that there were available dozens of painstaking descriptions of the grammars of TB languages. Suppose further that in most

cases you could figure out what the dominant tendency of each language was in a given syntactic area. Even then there would still be tremendous difficulties involved in trying to use syntactic criteria to establish degrees of genetic relationship. We may state the difficulty succinctly in terms of two negative propositions: (1) Syntactic similarities are no guarantee of genetic closeness. (2) Syntactic differences are no guarantee of genetic remoteness.

1.131 Specious syntactic parallelism: chance; universal tendencies; diffusion. The most general syntactic relationships involve binary options: either the relative clause comes before its head noun or after its head noun; either the case marker comes before its noun (preposition) or after its noun (postposition); either the object comes before the verb or after the verb; either the numeral-plus-classifier comes before the quantified head or the quantified head comes before the numeral-plus-classifier.

If each alternative were equally likely in human languages, the fact that Language X and Language Y had the same value for some binary syntactic feature (e.g., were both [+postpositional]) would be irrelevant to the question of their genetic relationship. It may turn out that most human languages have a particular value for some feature: e.g., maybe we shall find that 74% of the world's languages

are [+prepositional], but only 21% are [-postpositional], with 5% being [+postpositional] or "difficult to determine" or "category does not apply."⁹⁴ Even so, the fact that Language X and Language Y are both postpositional could still easily be due to **chance** (about one chance in five). Note that this is a far greater chance than that of having any given lexical item of core vocabulary (e.g., dog) show a marked phonological resemblance in any two randomly chosen languages.

Isolated facts of syntactic resemblance may therefore be of merely **typological** interest, no more.

As Greenberg 1963 was the first to make clear, there are valid linguistic universals of at least two types: (1) flat statements that are held to be overwhelmingly true across languages, involving a certain linguistic feature considered in isolation from any other linguistic feature. Universals like this might include: "all languages have a nounlike class of words and a verblike class of words;" "all languages have at least as many oral vowels as nasal vowels," etc. (2) "implicational" statements of the form, "if a language has feature A, it will also tend overwhelmingly to have feature B as well." Thus we might say, "languages that are postpositional tend to have the object coming before the verb, and the relative clause coming before the noun-head; prepositional languages tend to have the object coming after the verb, and the relative

clause following the noun-head."

This means that if we find that Languages X and Y are both postpositional, and then learn that both languages also have the object coming before the verb, this latter piece of information does not constitute **independent** evidence of relationship. The SOV order is a redundant, predictable feature belonging to a whole complex of implicationally related features that one could call the "postpositional syndrome." In any reckoning of overall syntactic similarity, we can "count" any such implicationally related set of features only once.

Maybe you can demonstrate that a given point of syntactic similarity between two languages cannot possibly be due to chance. Furthermore, perhaps you can show that it cannot be due to any universal tendency of human languages—maybe in fact both languages fly in the face of the prevailing trend with respect to this feature. Before jumping to the conclusion that the similarity is due to a common genetic inheritance, one further possibility must be considered: maybe the two languages were "in contact," so that the feature in question was **borrowed** from one language to the other—or, viewed more passively, "diffused" from one language into the other.

It used to be considered a rather daring, not to say naughty idea that a syntactic feature could be borrowed from language to language.⁹⁵ Grammar was supposed to be

the area of linguistic structure most impervious to outside influence. We now know that this is simply not true. If pidgin and creole studies have demonstrated anything, it is that there is no aspect of syntax that cannot be changed due to language contact. It is now widely recognized that languages in a geographically contiguous area tend to develop grammatically (as well as phonologically) along parallel lines, so that it makes sense, for example, to speak of Southeast Asian areal grammatical features.⁹⁶ Among these we may mention numeral-classifier systems; sentence-final particles whose only semantic function is emphasis or the conveying of the speaker's emotional attitude; four-syllable "elaborate expressions" [see 1.122 above and note 76]; complex verb phrases where the meaning of the head verb is modified by up to four auxiliary verbs, sometimes with no intervening markers of this subordination;⁹⁷ the "grammaticalization" of noun- and verb-roots which acquire more abstract meanings as "particles,"⁹⁸ etc.

Sometimes the direction of borrowing is clear. Thus it seems obvious that the Nepali and Bengali classifier systems are due to TB influence, since these languages are in contact with TB, while other Indic languages far removed from the TB area show no signs of developing classifiers. Often, however, there is no way of knowing who has borrowed from whom—just like good neighbors who have borrowed

so many cups of sugar from each other over the years that they can't keep track anymore. The similarities may be quite striking, and may involve very fine details of the grammar. Modern Burmese has a verb nei နေ that means 'live, dwell, be in a place' when it functions as a main verb, but which as an auxiliary verb has the much more abstract meaning of 'continued or progressive action' (sâ nei 'is eating,' literally "eat stay"). Exactly the same roles are played by the phonologically unrelated Lahu verb chê 'live, dwell' (câ chê 'is eating,' lit. "eat stay"). So far this is not surprising—these are both Lolo-Burmese languages, after all.⁹⁹ But look at Thai, which is not Tibeto-Burman at all. Sure enough, there is a verb jùu อยู่ that means 'live, dwell, be in a place' as a main verb, but indicates continuous action as an auxiliary (kin jùu 'is eating,' lit. "eat stay").

1.132 **Specious syntactic differences.** Even closely related languages may come to have very different syntactic structures due to the operation of processes of change internal to each language. Lahu and Lisu are extremely close genetically, with perhaps 80-90% cognacy in core vocabulary. Yet they differ radically in the way they form relative clauses. In Lahu, the normal relative clause precedes its noun-head, and is connected to it by the subordinating particle ve.¹⁰⁰ The relative clause may never

contain a NP (either agent or object) that is coreferential to the noun-head, nor even any pronominal "copy" of such a NP. To say, "the person who smokes opium," you say simply,

- (i) yāphī dò ve cho
 opium smoke REL person
 NP_{obj} V N_{rh}/agt.¹⁰¹

In Lisu, on the other hand, the relative clause follows its noun-head.¹⁰² It is connected to it by the relativizer yí (which I believe to be cognate to Lahu ve). The verb of the relative clause is followed by the "declarative" particle a.¹⁰³ The whole relative clause is then agentively nominalized by the particle ma, "the one who," thus:

- (ii) làthyu yí yàphi thy a ma
 person REL opium smoke "DEC" NOMINALIZER,
 N_{rh}/agt NP_{obj} V

'the person who smokes opium,' literally "the person who is the one who smokes opium." The ma may be regarded as a pronominal copy of the underlying subject of the relative clause.¹⁰⁴

These constructions seem vastly different from each other. If one were to do a serious pan-Loloish study of relative clause formation,¹⁰⁵ I believe the Lahu type, with the relative clause preceding the N_{rh}, could be demonstrated to be more archaic than the Lisu type. This is

because some Lahu RC's, whose structure is simple enough, may be shifted to the right of the N_{rh} . This "right-relative transformation" may operate if the RC has a minimal number of nominal arguments, and if its verb is stative (either a natural adjective or an action-verb stativized by the particle tā):¹⁰⁶

(iii) dā? ve cho → cho dā? ve
 be good person N_{rh}
 V_{adj} N_{rh}

'a person who is good.' Every Lahu RC that appears to the right of its N_{rh} may also grammatically appear to its left—but the contrary is not true: the vast majority of Lahu RC's may not be shifted from their pre- N_{rh} position. This seems to indicate that the prehead position is basic, though there is a tendency to shift to the right. Lisu itself, which has carried the rightward tendency much further than Lahu, seems to show a remnant of the older pre-head position in the form of its agentive nominalizer ma. If ma is really a copy of the underlying deleted NP in the RC that is co-referential to the N_{rh} , maybe we should analyze the relative clause as a structure that is **modifying** ma, at the same time that it is being nominalized by it. Thus to return to (ii) yāphi thyā a 'smoking opium' could be analyzed as a pre-head relative clause modifying ma, so that

yāphi thyɛ a ma
 NP_{obj} V REL N_{rh}

would be exactly analogous to the Lahu clause (i),

yāphī dò ve cho
 NP_{obj} V REL N_{rh}.

When the rightward trend overwhelmed Lisu, the "original" N_{rh} was demoted to a pronominal copy, and this whole structure was then stuck onto the right of the "new" head-noun.

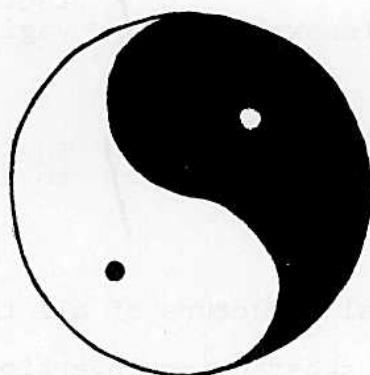
However, even if one can figure out that at some deeper historical level seemingly disparate syntactic structures in two languages derive from an identical source, this does not alter the fact that syntax is just as susceptible to change as any other aspect of the language. In fact, syntax seems to be one of the less stable areas of linguistic structure, diachronically speaking.

1.133 **Antonymous changes, in syntax and elsewhere.** Especially interesting are cases where two poles of a relationship seem to reverse themselves through time, so that A becomes B and B becomes A, or A + B becomes B + A. The attraction of opposites is one of the most ancient and pervasive insights of metaphysical thinking. Nowhere is this expressed more succinctly than in the Taoist emblem that symbolizes the interrelationship of the two organizing principles of the universe, Yin and Yang. See Figure 5.

The white area (roughly, the male principle or Yang) contains a little black circle, just as the black area (Yin; the "female" principle) contains a little white one. Each pole of the opposition carries within itself the seeds of the change into its opposite.

Figure 5

Yin/Yang and the seeds of antonymous change



Lévi-Strauss has provided us with brilliant demonstrations of how mythic motifs repeatedly undergo changes into their opposites as they are transformed from culture to culture. Several South American tribes, not too distant from each other geographically, have myths about the origin of tobacco (Lévi-Strauss 1970, pp. 99-108). In all cases, the plant is said to have first sprung up from the body of a dead person. But many details of the stories are symbolically "inverted" from one tribe to the other. In the Tobe and Tereno versions, the story hinges on

a husband	/	who has a wife	/	(the wife) is des-
(affinal rela-		(jaguar)		tructive through
tionship)				the mouth

of a husband who	/	looking for animals	/	that the wife ought
has climbed a		(birds)		not to eat (but
tree				does).

In the Bororó version, on the other hand, the key figure is

a mother	/	who has a son	/	(the mother) is pro-
(blood rela-		(snake)		tection through the
tionship)				vagina

of a son who has	/	looking for vege-	/	that the mother ought
climbed a tree		table food		to eat (but does not).
		(fruit)		

On some level the formal structure of all these myths is identical, despite the substantive inversion which they display on the surface.

In phonology, and especially in the history of tonal phenomena, we are faced with "flipflops" whose causes have never been explained in anything but a poetical way. Of thirteen Loloish languages and dialects studied with a view to their tonal behavior in stopped syllables (Matisoff 1972a, "TSR"), eleven agree in having one class of words (call them Class H) under a tone which is phonetically higher in pitch than another well-defined class of words (Class L). But two languages, Nasu and Lu-ch'üan Lolo, are just the opposite: the words in Class H appear under

a tone which is **lower** than the words in Class L. Presumably the tonal flipflop in these two languages was an innovation in response to some threat to contrastive power. Maybe the lower tone began creeping upward in pitch, so frightening the higher tone that in its rush to maintain its integrity it plunged as far down as it could, passing the upward-bound lower tone in the opposite direction¹⁰⁷....

In alternating current the negative and positive charges on the two poles are reversed dozens of times per second, yet the output of energy remains steady (as long as the voltage and resistance remain constant).

In politics, the extreme right has much in common with the extreme left.

In semantics, antonymy is an extreme type of similarity, in which the meanings of the two opposed concepts differ from one another only by a value for a single feature.¹⁰⁸ *Black* is much more closely related to *white*, than it is to, say, *brown*, or even *gray*.

Lexically, it may happen through the accidents of phonological change that words of diametrically opposed meanings may come to be pronounced in exactly the same way. English provides us with two classic examples:

let (= permit) < ME *leten* < OE *lāetan* vs.

let (= obstruct, hinder [archaic], as in 'let ball'

[tennis] or 'without let or hindrance' [legal])

< ME *letten* < OE *lettan*;¹⁰⁹

cleave (= adhere) < ME *cleven* < OE *cleofian* vs.

cleave (= sunder) < ME *cleven* < OE *clēofan*.

More to the point are cases where the meaning of one and the same word changes into its opposite through time.

English *candid* (< Lat. *candidus* 'shining white') means 'frank, open, having nothing to hide.' Perhaps its most frequent use nowadays is in expressions referring to unposed photographs ('candid camera,' 'a candid shot,' etc.). However, in order to take a good candid shot of somebody, one often has to use subterfuge, sneaking up to catch the subject unawares so he will not mug for the camera. In the popular mind there is already something sneaky about *candid*, and it is entirely conceivable that a dictionary entry of the twenty-fifth century might look something like this:

- candid* [kædə]¹¹⁰
1. sneaky, surreptitious (as in 'a candid proposition')
 2. taking someone unawares (as in 'a candid videograph')
 3. [archaic] frank, open, having nothing to hide
 4. [obsolete] shining white.

1.134 **Cyclicity in syntax.** In syntax, we have already seen how even closely related languages like Lahu and Lisu can develop a given proto-construction along idiosyncratic lines, so that the daughter structures end up looking like mirror images of each other. When there is greater genetic

distance between the languages involved, the chances for such antithetical developments are correspondingly increased. Among the rich arsenal of Lahu versatile verbs [see note 97], about a dozen occur regularly **before** their head-verb, while all the others come **after** their head. In the Mru language of East Bengal (perhaps most closely related to the Kukish branch of TB), Lorenz Löffler reports that the semantic equivalents (and sometimes the etymological cognates) of Lahu pre-head auxiliaries typically occur **after** the verb-head, and vice versa.¹¹¹

One syntactic relationship that is frequently seized upon and given criterial importance in classifying languages is the position of the verb with respect to its object. The SVO order seems somehow to be fundamentally different from the SOV order, for example. The vast majority of TB languages have the verb phrase coming in clause-final position,¹¹² with its object (and all its other nominal arguments) preceding it:

Clause \rightarrow NPⁿ + VP, where $n \geq 0$.

Karen, however, has the object following the verb (see Jones 1961). This fact has been pounced upon by comparatists as evidence of extreme aberrancy, to the point where Karen is assumed to have split off from the Tibeto-Burman languages before the breakup of common TB into its constituent subgroups. "Tibeto-Karen" is conceived of by

Benedict [STC, p. 6; see Fig. 1 above] as a higher-order taxonomic grouping where Tibeto-Burman and Karen are coordinate branches. Yet lexically Karen shows a marked affinity for Lolo-Burmese, with a very high percentage of shared vocabulary. Furthermore, the tones of Karen can easily be related to those of Lolo-Burmese [STC, pp. 150-152]. Benedict's position is a little unclear here. He explicitly rejects a diffusional explanation for the similarity between the Karen and LB tone-systems [STC, pp. 194-197, note 494]. At the same time he explains the post-verbal position of the Karen object in diffusional terms [STC, p. 4].¹¹³ One would think, therefore, that he would assign greater weight to the genetically-based phonological and lexical similarities of Karen with the eastern branches of TB (especially Lolo-Burmese) than to the diffusively motivated syntactic divergence of Karen from the rest of TB.¹¹⁴

We can gain some perspective on this problem by considering Chinese. Modern Chinese, like Karen, places the object after the verb. Benedict also explains this as due to the influence of neighboring languages [STC, *loc. cit.*], maintaining that the original Sino-Tibetan word-order was SOV, as in Tibeto-Burman. (There is in fact evidence from the oldest Chinese texts that the object sometimes used to precede the verb even after the split-off of Chinese from TB.¹¹⁵) What is fascinating is that modern Mandarin is now

showing strong signs of reversing itself again, so that there is a well-defined construction type where the object does indeed precede the verb! In these cases the object is marked by the preposition baa, as in:

<u>ta</u>		<u>baa</u>	<u>jey</u>	<u>waan</u>	<u>fann</u>		<u>chy</u>	<u>wan</u>	<u>le</u>
NP _{subj}		PREP	Det	Clf	N		V _h	V _{aux}	Aspect Particle
he		OBJECT MARKER	this	bowl	rice		eat	finish	already

NP_{obj}

'He has finished eating this bowl of rice.' Historically, this structure derives from a serial-verb construction, since baa (like many other modern Mandarin prepositions¹¹⁶) descends from a full verb. Baa 把 used to mean 'take, hold': "He, taking the bowl of rice, has finished eating it." However circuitous the route may have been, Chinese seems now to be going back to its Sino-Tibetan roots. Changes in word-order, like changes of all other types, may be cyclical, though the period of oscillation is measured in terms of centuries or millennia.

Even the well-behaved TB languages, where the NP's have always remained to the left of the VP, show great freedom of permutation within the nominal hemistich itself.¹¹⁷ In fact, a Lahu NP may be defined as the "minimum freely permutable element in a nominal hemistich" [GL, p. 189]. Indirect objects may precede or follow direct objects, locative and/or temporal NP's may occur in any

order with respect to the object, as may instrumental NP's, etc. The alternative orderings are concomitant with differences of style or emphasis. In colloquial Lahu it often happens that a NP gets tacked onto the end of a clause as an afterthought:

<u>nà-hi</u>		<u>kô?</u> <u>jâ</u> ,	<u>lâ à?</u>	'We fear (them) very much, tigers.'
NP _{subj}		VP	NP _{obj}	

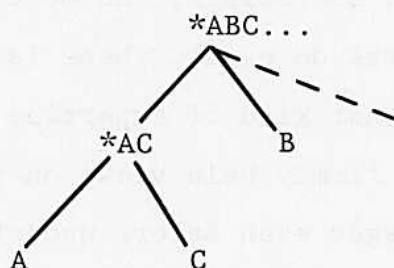
The NP has here burst the bounds of the nominal hemistich entirely and jumped into post-verbal position. Should this construction ever catch on so that it is generalized into "normal" sentences, even Lahu might someday become an SVO language!

1.14 The lexicon and genetic relationship: beyond the lexicostatistical approach. More than any other aspect of linguistic structure, the lexicon has been depended upon as the decisive area in which to establish degrees of genetic relationship. When people try to subgroup a language family nowadays, what they are apt to do is sit down with a bunch of word-lists in the various languages, analyzing them to see how many cognates they can find. The more closely related the languages, *ceteris paribus*, the more cognate sets can be discovered. Let us suppose that a large number of words, carefully selected for their fundamental conceptual importance in the lexicon ("core

vocabulary"), are examined in three languages, A, B, and C. Careful analysis by experts in the language family reveals that languages A and B have 38-40% cognacy in these vocabulary items, as do B and C; while languages A and C have a much higher percentage of cognates, say 85%. In this ideally clear-cut case, everybody would agree that A and C are more closely related to each other than either is to B. According to the familiar family-tree model, the speakers of language B must have split off more or less cleanly from the proto-speech community at a relatively early date, while the ancestors of the speakers of A and C were still in linguistic contact with each other. See Figure 6.¹¹⁸

Figure 6

A clear-cut family-tree relationship



In practice, however, there are many difficulties with this approach, many jokers in the pack. In the first place, there are the defects in the family-tree model itself, with its disregard of diffusional phenomena, contact in-

fluences, borrowings. Two languages may have a very high percentage of shared vocabulary and still be genetically unrelated (e.g., Chinese and Japanese) or at least much less closely related than the percentage would seem to warrant (e.g., English and French). In this latter kind of case, especially when we are dealing with unwritten languages, there is the often severe problem of distinguishing genuine etymological cognates from early borrowings. That is why we specified that the words to be compared should belong to "core vocabulary," and that the job of determining cognacy be performed by "experts" who know the language family well enough to be able to distinguish cognates from borrowings. Yet we are still begging these questions. How do we know what "core vocabulary" is? And suppose there exists no linguist who is "expert" enough in these languages to distinguish *a priori* among genuine cognates, borrowings, and mere chance resemblances? If such experts do exist, there is an opposite problem—anybody with that kind of expertise would be likely to have his own firmly held views on the interrelationships of the languages even before undertaking the investigation, and it would be hard for him to avoid injecting his own preconceived ideas into the question.

In the 1950's, as we all know, the exciting new approach known as lexicostatistics (or glottochronology) was developed in an effort to put linguistic subgrouping on a

less intuitive and more scientific basis, and to make explicit and refine the methodological assumptions involved. The problem of deciding what was "core" vocabulary was solved by making an actual list of English words (first a 100-word list, then one of 200 items), to be translated into whatever languages were to be examined. These words were selected from semantic domains judged to be the most fundamental to human conceptualization of the universe (kinship, body-parts, natural objects, animals, motion, simple activities, basic logical relations like negation and conjunction, etc.). Making such an explicit and universal list seemed like an excellent idea. The key assumption behind the list also seemed eminently reasonable: "core" vocabulary is particularly resistant to replacement by borrowing or through obsolescence, and in fact seems to be replaced at a relatively constant rate through time, say, 20% per 1000 years.¹¹⁹ Thus, if Language A and Language C show 160 cognates out of the 200 items in the core-vocabulary list (i.e., 80% cognacy), they would have separated from their common ancestor about 1000 years ago. But if Language B shares only about 128 core cognates with Languages A and C (i.e., 64% cognacy), it would have separated off from its sister languages about 2000 years ago (i.e., only 80% of the 80% of the original common stock of core vocabulary would have escaped replacement.)

Attractive as this new approach is, many difficulties

of principle and practice make it highly doubtful that it is any more "scientific" than the older intuitive techniques (the "expert-flying-by-the-seat-of-his-pants" methods). Some of these difficulties are inherent in the tenets of the lexicostatistical approach itself; but there are additional problems in applying the method to the Tibeto-Burman languages in particular. These TB-specific problems have already been broached above [1.12]—the fascinating complications introduced by the TB penchants for lexical ganglia that we have called "word-families" and "compound-families."

Let us discuss all these difficulties, and see if there is any escape.

(a) **Analytical expertise and "proof of cognacy."** As the computer scientists tell us, "garbage in, garbage out."¹²⁰ That is, no matter how elegant or precise-looking the computations and calculations to which you subject your data, if the data are wrong, inaccurate or uncertain to begin with, the conclusions and inferences one draws as a result of the calculations must also be wrong, inaccurate or uncertain.

It is often not easy to tell whether a given pair of forms are true cognates or not. The matter cannot always be settled by superficial inspection. In many cases it takes detailed, specialized knowledge even to be able to venture an educated guess. Even professional linguists

are not equipped to decide subtle questions of cognacy involving the very language family to which their native language belongs, unless they have received specific training to that end. Is Latin *dies* 'day' cognate to English *day*? If so, why? If not, why not?¹²¹ How about Latin *habeō* and English *have*? Or does English *have* go rather with Latin *capiō* 'take'?¹²² Often forms which look very much alike are not true cognates at all. Conversely, true cognates which are perfectly regular reflexes of rigorously statable "rules of sound-change" may have no superficial phonological resemblance to each other whatsoever. Everybody knows the thrilling example of Latin *duo* and Armenian *erku* 'two'—perfect cognates! Or Lahu *ḥ* and Burmese *lê* 'four'.¹²³

(b) **Constancy of replacement rate of core vocabulary.**

Although it is certainly true that on the whole core vocabulary is more resistant to replacement than what we might call "peripheral vocabulary," it is easy to find striking examples of lexical replacement in any semantic/conceptual realm. Numerals may be borrowed wholesale; kinship terms may change their referents or disappear due to taboo, euphemism, teknonymy, or social change [see, e.g., Benedict 1941]; body-part words may change for a variety of similar reasons; even words of abstract grammatical function—basic relational particles that mean things like 'and' or 'or' or 'not'—may be replaced by

foreign borrowings.¹²⁴ Furthermore, due partly to the accidents of cultural history and partly to some mysterious differences in their internal linguistic "nature" or "genius" or "essence," synchronically observable languages do seem to differ considerably in their receptivity to borrowings from foreign languages—whether these foreign languages are dead, learned languages like Greek or Sanskrit, geographically contiguous contact languages (whether culturally prestigious or not), or nowadays even culturally important languages spoken on the opposite side of the globe. Japanese borrows many more words from English than English does from Japanese. We can conceive of Japanese borrowing the English word *kiss* (Js. *kissu*) much more easily than our borrowing the Japanese word *seppun* 'kiss,' or even a more culturally indigenous word like *o-jigi* 'a bow.'

The vicissitudes of world cultural history constantly reshuffle the patterns of linguistic dominance and submissiveness with respect to outside incursions on a language's vocabulary. For hundreds of years after the Norman conquest French was accepted as a culturally superior language to English. Both Englishmen and Frenchmen concurred in this view, and the massive flow of borrowed vocabulary was almost totally unidirectional, from French to English. In the 19th century the tide began to turn with the growing political and economic power of the Anglo-Saxon world

and the decline of France's role on the international stage. Words like *revolver* and *redingote* (< Eng. riding-coat) crept into the hallowed French lexicon, as the culture and fashions of the English-speaking world began to become objects of emulation rather than of amused contempt. Now, in the last quarter of the 20th century, the incursions of English into the French lexicon have reached alarming proportions. Purists decry the vile new "franglais" of with-it French youth. English words have invaded all areas of the vocabulary. A French linguist will tell you that he is off to "faire du fieldwork." To this you might answer in good French, "O.K., amuse-toi bien!"

The glottochronologist would have us believe that these cultural vicissitudes "even themselves out" over a long enough period of time. This is an unprovable article of faith.¹²⁵

It is instructive to take a sharper look at English and German, two quite closely related Germanic languages. Despite the fact that they are genetically so close, they seem to differ greatly in the rate at which they are willing to tolerate the replacement of native inherited lexical material.¹²⁶ (We might call this their "replacement tolerance quotient," or RTQ.¹²⁷) The RTQ of English seems much higher than that of German. This difference in RTQ manifests itself in several ways: (A) Coinage of learned vocabulary. While English rushes to Latin and Greek to make up techno-

logical or learned neologisms [*television, tonogenesis, leucorrhoeia, rhinoglottophilia*¹²⁸], German much prefers to use native lexical material [*Fernseher, Tonentstehung, Weissfluss, Nasenstimmbändergegensatzliebe*]. (B) Borrowings of core vocabulary. In words like *dog, pig, skin*, and *animal*, the modern English word is an innovation with respect to Common Germanic. Either English has borrowed its new word from some (often obscure) contiguous Germanic source (*dog, pig, skin*), or from a "higher culture" language (*animal* < French < Latin). German, on the other hand, tends overwhelmingly to preserve the most widespread original Germanic word for such core concepts (*Hund, Schwein, Haut, Tier*). (C) Semantic shifting in core vocabulary. Though this would take a serious study to prove, I would like to claim that English items of core vocabulary are more likely to undergo changes in their meaning through time than are their German equivalents. This would be a smart "strategy" for a language with a high RTQ to adopt! While welcoming innovations in core vocabulary on the one hand, English is reluctant to throw away the older word entirely. Typically the older word is retained, either as a less frequent and stylistically "marked" synonym of the newer word, or else with some change or increment of meaning, so that the two words, old and new, continue to co-exist, each safe within its own semantic domain. This way the language loses nothing, and becomes in fact en-

riched. English cognates to the four German words in (B) above are still very much alive: *hound*, *swine*, *hide*, *deer*. In three of these cases, a semantic specialization has occurred: a *hound* is a particular kind of dog, a *deer* is a particular kind of animal, a *hide* is a particular kind of skin (usually an animal's). *Swine* still means 'pig,' though it now sounds archaic to our ears except in its usage as an insulting epithet ("He's a real swine!") We still cast pearls before the swine, but we don't eat *swine flesh* anymore (even if we do eat pork). A German, on the other hand, thinks nothing of eating *Schweinefleisch* with his dumplings.

In sum, it is not at all obvious that languages do undergo lexical replacement at the same rate, either in core or peripheral vocabulary.

(c) Semantically shifted cognates. It is a fundamental dogma of lexicostatistics that one must not look around too hard for etymological cognates in compiling the basic vocabulary lists in the various languages to be examined. What is wanted is the "usual word in the language at the present time" that is used for the concept expressed by the linguist's gloss. (Let us for the moment assume that it is always possible in principle to decide on one particular "usual" word to translate each of the concepts in the core vocabulary list. Let us make the further assumption that the investigating linguist is using

an English word-list as his point of reference, and that English (being the world's most perfect language) has one single clear and unambiguous word for each of the core concepts.¹²⁹⁾

The lexicostatistical method, in other words, requires that we disregard etymological cognates in cases where the cognate form has been displaced from its central hook-up with the core concept. This displacement may have been due to pressure from a foreign word borrowed from outside, or may have occurred as the result of purely internal pressure for semantic change. To return to our English/German discussion, we cannot count the fact that English *hound* is cognate to German *Hund*. This is not a plus-factor in our calculations. Since the usual English word for *Canis familiaris* is now definitely *dog*, we must register a minus for this item on the list.

It seems to me that this disregard of semantically skewed cognates causes the lexicostatistician to lose a great deal of vital information that might enable him to make his judgments of degrees of relationship much more subtle and refined. Let us take a hypothetical English-German bilingual, somebody who is well-educated but not a professional linguist (and *a fortiori* not a specialist in comparative-historical Germanic studies). Let us suppose that the lexicons of English and German are equally familiar to him, and that he reads novels in both languages for

pleasure. If we were to ask him questions of the form "Can you think of a German word that is related to the English word X?" or "Can you think of an English word that is related to the German word Y?," he would be able to come up instantly with hundreds of correct answers.¹³⁰ The easiest relatives to recognize would be those where the pronunciation, spelling, and meaning of the two forms are identical, or nearly identical: Eng. *finger*/Ger. *Finger*, Eng. *nest*/Ger. *Nest*, Eng. *hand*/Ger. *Hand*, Eng. *hammer*/Ger. *Hammer*, Eng. *house*/Ger. *Haus*, Eng. *mouse*/Ger. *Maus*, Eng. *louse*/Ger. *Laus*, Eng. *shoe*/Ger. *Schuh*....

Only minimally more difficult for our bilingual to identify would be pairs where the meanings of the two forms are still identical, but the pronunciation (and spelling) is "a little different": Eng. *out*/Ger. *aus*, Eng. *foot*/Ger. *Fuss*, Eng. *water*/Ger. *Wasser*, Eng. *day*/Ger. *Tag*, Eng. *eye*/Ger. *Auge*, Eng. *sheep*/Ger. *Schaf*, Eng. *bride*/Ger. *Braut*, Eng. *sheath*/Ger. *Scheide*, Eng. *heath*/Ger. *Heide*, Eng. *thumb*/Ger. *Daumen*....

Sometimes, however, even if there exists a cognate in the other language, and even if the meanings of the two cognate forms are still more or less the same, our bilingual would have a lot of trouble coming up with a correct answer. Given the English word *tail*, he might not be able to cite the cognate *Zagel*,¹³¹ since *Schwanz* is now the usual standard form. Similarly, asked to find English relatives of

Zeichen 'sign, symbol, mark,' *heissen* 'to be named, be called,' *gern* 'willingly,' *Schaden* 'damage,' or *schwarz* 'black, dark,' he would not be very likely to think of *token*, *hight*, *yearn*, *scathe*, and *swart*. English *token* does indeed still mean a mark or sign, and there is nothing archaic or dialectal about the word. When we use a subway token instead of putting real money in the slot, we are using a symbolic substitute for a coin. When we give someone a token of our esteem, we're giving him a sign or symbol of our affection. *Hight* is now obsolete (or poetically archaic at best), but our bilingual has probably encountered it in literature. *Scathe* survives only in the negative adjective *unscathed* 'undamaged,' and in its participial form *scathing* 'hurtful, damaging.' *Swart* is now obsolete, but its derivative *swarthy*¹³² is very much alive, still with its meaning of 'dark, as opposed to light.' One could argue endlessly about the kinds and degrees of differences in meaning that these English/German cognate pairs have undergone. Each pair seems to be a law unto itself. Yet perhaps we can at least agree that the meaning changes, if any, have not been particularly drastic. The relationship of the words is hard for our bilingual to figure out on other grounds—the obsolescent, archaic, or dialectal nature of some of the forms, the fact that the cognate morphemes occur hidden in combination with other morphemes (*unscathed*, *swarthy*), or the considerable surface-

phonological and orthographic divergence of the forms (*to-ken/Zeichen*).

But let us now consider the crucially important class of English/German cognates where one of the languages has indeed undergone a considerable semantic shift. It is my contention that this shift in meaning would not necessarily make it difficult for our bilingual to come up with the correct cognate right away. Asked to find the German cognates to *flesh*, *fowl*, *bone*, and *hound*, would he not quickly supply *Fleisch* 'meat,' *Vogel* 'bird,' *Bein* 'leg,' and *Hund* 'dog'? Granted, there will be cases where the semantic shift has been so great that our bilingual might not be able to find the cognate even if the two forms were phonologically quite similar (Eng. *deer*, Ger. *Tier* 'animal'; Eng. *tidings*, Ger. *Zeitung* 'newspaper'). In extreme cases, both the pronunciations and the meanings of the forms have diverged so much that only the specialist in Germanic linguistics could be expected to know them: Ger. *Schmuck* 'ornament'/Eng. *smug*; Ger. *nehmen* 'to take,' *benommen* 'numb'/Eng. *numb* [i.e. 'taken, seized' < OE *numen*, past participle of *niman* 'take'] and *nimble* [i.e. 'quick in one's bodily movements' < 'quick-witted, quick to grasp things, quick on the uptake']; Ger. *schräg* 'oblique, slanted,' *schränken* 'lay crosswise'/Eng. *shrug*, *shrink*, etc.

The point is this: why should we throw away any information at all that bears on the past history of rela-

tionship of these languages? Whenever we find a cognate, let us take it into account. If the meanings have diverged sharply, this fact should of course be noted. But to ignore a good cognate on the grounds of semantic change is obfuscatory rather than clarificatory. The English words *girl*, *tree*, and *cloud* have, as far as I know, no cognates in modern standard German. Conversely, the German word *Stimme* 'voice' has no modern (or even archaically poetic) English relatives. Surely we must "score" words like these differently from cases like *Vogel/fowl* or *Hund/hound*! Instead of the "all-or-nothing" approach, why not use some more sensitive scoring system that can capture the whole continuum of phonological/semantic relationships among the lexicons of our languages?

Semantic shifting is perhaps the rule in language history rather than the exception, if one takes entire lexicons into account. It is not a rare or sporadically encountered phenomenon, and is just as worthy a subject of investigation as purely phonological change (though a lot harder to deal with). As a start in the right direction, I would advocate a graduated system of scoring when comparing word-lists among languages. Vocabulary items in Language A and Language B that are phonological cognates and that also mean the same thing would be given a perfect score of 10 on our scale from 0 (lowest) to 10 (highest)—e.g. Eng. *louse*/Ger. *Laus*, Eng. *sheath*/Ger. *Scheide*. Items where the phono-

logical correspondence is perfect, but the meanings different, would rank lower and lower on the scale, depending on the degree of semantic divergence. Thus, Eng. *flesh*/Ger. *Fleisch* would score about a 9, while Eng. *bone*/Ger. *Bein* 'leg' would only score about a 7, as would Eng. *dish*/Ger. *Tisch* 'table'; Eng. *beam* (of wood)/Ger. *Baum* 'tree' would score about a 4, and Eng. *nimble*/Ger. *nehmen* 'take' would only score about a 2. Similarly, lower marks would be accorded etyma which only survive in a compound or set expression rather than in isolated root form. Thus, E. *-scathe-* [in 'unscathed,' 'scathing']/Ger. *Schaden* 'damage' would get about a 3. Phonological cognates which hang on only in archaic or poetic language, but which are not known to the average speaker (Eng. *hight*/Ger. *heissen* 'be named'), or which are dialectal rather than standard (Eng. *tail*/Ger. *Zagel*), would rank lower still, say 1 or 2. Forms which have no wholly or even partially surviving relatives in the other language receive a zero (Eng. *tree*/Ger. ____).

Granted, this gets very complicated. Furthermore, it is subjective—we do not have principled ways of measuring degrees of semantic divergence (or phonological divergence either, for that matter [see below 1.141a]). Nevertheless, this approach begins to do justice to the complexity of the problem. It has the virtue of focussing our attention on whole clumps of phonologically and semantically related forms, rather than artificially isolating a single pair of

forms to consider each time.

This sort of fine-tuned scoring is particularly appropriate in attacking the special "family" problems Tibeto-Burmanists encounter in their cognate identifications [below 1.141]. More importantly, it leads us into the "organic semantic" approach of studying meaning-change through time [below 2.0-2.5].

1.141 Word-family and compound-family scoring problems.

We have already [1.12 above] discussed word- and compound-families rather extensively. It is now time to take a closer look at the complications these phenomena introduce into any attempt to apply lexicostatistical methods to the subgrouping of Tibeto-Burman.

(a) **Calibrating degrees of phonological/semantic relationship: isofamy and heterofamy.** A genetically related pair of forms, one from Language A and one from Language B (i.e. a pair of "interlingual allofams"), may stand in any of four phonological/semantic relationships to each other.

[1] Both the phonological and the semantic correspondences may be perfect. Interlingual allofams which display a perfect phonological correspondence we may call **phonological isofams** of each other. (To this relationship we may give the name "phonological isofamy.")^{133, 134}

Interlingual allofams which display a perfect semantic

correspondence we may call semantic isofams of each other ("semantic isofamy").

Thus, English *fish* and Swedish *fisk* are isofamous on both the phonological and semantic planes. Eng. "sh" regularly responds to Sw. "sk" (cf. Eng. *shoe*/Sw. *sko*, Eng. *shirt*/Sw. *skjorta*, etc.), and the forms both mean 'finny, aquatic vertebrate animal.' Similarly with Lahu pā and WB phâ 'frog' (< PLB *pa₂): the correspondence of a Lahu voiceless unaspirate to a WB aspirate is the regular reflex of a PLB *glottalized initial,¹³⁵ and Lahu very-low tone /ˉ/ is the regular correspondence to WB high-tone /ˆ/ for PLB Tone *2 syllables with pre-glottalized initials.

The examples just cited involve isolated etyma that do not form parts of larger word-families. In favorable cases, however, two (or more) co-allofams in Language A may each be demonstrated to be isofamous to particular forms in Language B. In cases like these we are justified in speaking of "parallelism of isofamy." Thus, the simplex/causative verb-pair 'sleep'/'put to sleep' is reconstructible for PLB as *yip (> WB ʔip, Lh. yɛʔ, Sani ji₂₂) and **s-yip > *ʔyip 'put to sleep' (> WB sip, Lh. ɿ, Sani ši₅₅).¹³⁶ Each of these three Lolo-Burmese languages has a surviving pair of intralingual co-allofams, and each member of each pair displays a perfect phonological and semantic isofamy to the corresponding forms in the other languages.

[2] The phonological correspondence may be imperfect,

though the semantic correspondence is exact. Interlingual allofams which display an imperfect phonological correspondence we may call phonological heterofams of each other ("phonological heterofamy").¹³⁷

The Lahu verb qho 'draw water, scoop up water' means the same as WB khap, and the forms are clearly co-allofams. Yet the Lahu form descends from a prototype with nasal final, PLB *kam, while the WB form is from the stop-finalled variant *kap [see above, 1.121(a)]. In other words, *kam and *kap were already differentiated co-allofams in the proto-language, so that the descendants of one must be heterofamous to the descendants of the other.

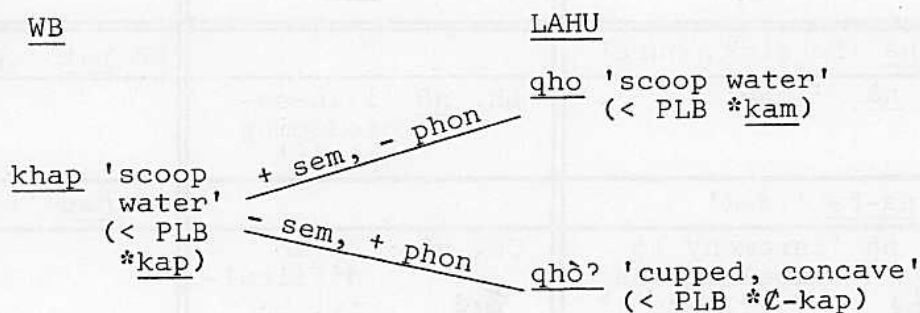
Similarly, the Lahu word for 'needle,' gò?, descends from an unprefixated PLB prototype, *rap. Several other LB forms, however, including WB ʔap, Hani ko₃₃, and Bisu kjaw, descend rather from a proto-allofam with velar prefix, *k-rap [see TSR, set 191].¹³⁸ The Hani and Bisu forms are isofams of each other, but each is heterofamous to the Lahu form.

[3] The semantic correspondence may be imperfect, though the phonological correspondence is exact. Interlingual allofams which display an imperfect semantic correspondence we may call semantic heterofams of each other ("semantic heterofamy").

We have already discussed semantic heterofamy under a different name ("semantically shifted cognates," above

1.14c). Instead of repeating examples presented there (like Ger. *Hund* 'dog'/Eng. *hound* 'hound,' Ger. *Tisch* 'table'/Eng. *dish* 'dish'), let us add a new wrinkle to the LB word-family *kam * *kap 'scoop up water' just mentioned in [1]:

Besides the action-verb qho 'draw water,' Lahu has a stative adjective in the low-stopped tone qhò? that means 'concave, cupped.' This form descends from Proto-Loloish *ɕ-kap.¹³⁹ Phonologically it corresponds perfectly to the WB verb khap 'draw water.' Now the drawing of water is performed either with the cupped hands or with a concave scooping instrument. There seems to be no doubt that both Lahu forms, qho and qhò?, are related to the WB form. Yet both of them are heterofamous with respect to it: the pair Lh. qho/WB khap is semantically isofamous but phonologically heterofamous, while the pair Lh. qhò?/WB khap is semantically heterofamous but phonologically isofamous. We may represent this graphically, using "+" for isofamy and "-" for heterofamy:



[4] Both the phonological and the semantic correspondences may be imperfect. Two forms may still be interlingual allofams even though they are both phonologically and semantically heterofamous. Thus, Lahu nê 'animist spirit capable of causing pain' (< PTB *nan) is ultimately related to WT nad 'illness' (< PTB *nat), even though both the sound- and the meaning-correspondences are not exact.

* *

*

The complex ST word-family *na * nan * nat just mentioned [see also above 1.121(2)] furnishes us with examples of all four types of phonological/semantic relationships. See Figure 7.

Figure 7

Some members of the word-family *na * nan * nat

PST * <u>na</u>	PST * <u>nan</u>	PST * <u>nat</u>
WB <u>na</u> 'be sick, hurt'		WB <u>nat</u> 'spirit'
Lh. <u>nà</u> 'idem'	Lh. <u>nê</u> 'illness-producing spirit'	
WT <u>na-ba</u> 'idem'		WT <u>nad</u> 'illness'
Ch. <u>nâ</u> 'ceremony to expel demons of illness' ¹⁴⁰ (<u>傩</u>)	Ch. <u>nân</u> 'be in difficulty, be suffering' (<u>難</u>)	
		Lu. <u>nat</u> 'ache, be in pain'

The first three forms in the first column are isofamous both in sound and meaning [+phon, +sem]; the two forms in the second column are isofamous in sound, but heterofamous in meaning [+phon, -sem]; Lahu nê and WB nat are heterofamous in sound, but isofamous in meaning [-phon, +sem]; and Lahu nê and Lushai nat are heterofamous both in sound and meaning [-phon, -sem].¹⁴¹

* *

*

Let us return to the problem of how to "score" the various kinds of interlingual allofams for the purpose of calculating degrees of closeness of genetic relationship. The problem divides itself into two parts: (1) identifying the correct proto-allofams that underlie the particular modern forms; and (2) deciding about the internal relationships of closeness and remoteness among the proto-allofams themselves.

In favorable cases it is possible to determine unambiguously which one of a number of intralingual allofams in Language A is the direct cognate of a form in Language B. Thus, WB has two variants for 'gum lac, cochineal,' khrip * khyip. The Lahu form a-kɛ 'pine resin, pitch-pine' is clearly isofamous to the WB variant with medial -r-, and the etymon is reconstructible as PLB *ʔgrip [TSR 46].¹⁴²

In many cases, however, it is not so easy. Imagine a hypothetical form phon 'belly, stomach' in Language A, which we are comparing to two related forms in Language B, pho 'intestines' and phon 'rounded part, swelling.'¹⁴³ The meanings of the latter forms are both heterofamous with respect to the form in Language A and phonologically there is also no *a priori* way of deciding which is closer. Either -o or -on might be the regular correspondence in B to the rhyme -on in A (or maybe neither one is!).

But let us suppose for the sake of argument that the phonological proto-forms underlying all the modern forms to be considered have been perfectly worked out. To do our lexicostatistical analysis we would still need some way of ranking the various forms in Language A on a sliding-scale of overall phonological/semantic similarity with respect to the several forms in Language B. The top and the bottom of the scale are relatively clear. Pairs of forms that are [+phon, +sem] rank highest, and pairs of forms that are [-phon, -sem] rank lowest. Isofamy is more highly valued than heterofamy. But how do we rank [+phon, -sem] and [-phon, +sem] with respect to each other? Are we more impressed by exact phonological correspondence, or by exact semantic correspondence? Most difficult of all are cases where three or more heterofamous meanings or phonological shapes are involved. As yet linguistic theory has provided us with no principled basis for determining degrees of

semantic relationship. Is the meaning 'disease-causing spirit' more "similar" to the meaning 'be sick' than it is to the meaning 'spirit in general'? Even in the phonological realm, where people have been operating with the notion of atomic distinctive features for a long time, we have no unique way of ranking forms on a scale of similarity. Given the three proto-allofams *puk, *pik, and *wuk, can we decide which two are "more closely related" to each other than either is to the third? *puk and *wuk have the same vowel, but their initial consonants differ by the feature [\pm obstruentall]. *puk and *pik have the same initial consonant, but their vowels differ by the feature [\pm frontness].

We shall return to these questions below (c). First, however, it might be a good idea to consider some examples of TB word-families that have not been artificially simplified, so that we may get a feeling for just how complicated their internal morphophonemic structure can be.

1.141(b) LUNG and BELLY in Tibeto-Burman (morphophonemics)¹⁴⁴

LUNG₁

In STC 239, Benedict sets up a TB root *tsywap 'lungs' on the basis of only two forms, Garo kasop and Lushai chuap. My own investigations into TB body-part nomenclature have led me to uncover many related forms in other

languages, which demonstrate that the root is rather to be reconstructed *wap. The various daughter languages have preposed a wide variety of prefixial material to the root, some of which can be positively identified with other, independent TB root-morphemes.

The root *wap itself seems to be an intruder from another semantic field. The key language here is Jinghpaw, which has the verb wóp 'be spongy, soft, yielding' alongside the nouns sîn-wóp 'lung' and məsîn * sîn 'liver.'¹⁴⁵

Let us set up an abstract proto-form *P-wa-p, and gloss it as LUNG₁. "P" stands for one or more prefixes that could precede the root initial in the proto-language. The dash before the final -p indicates that the final consonant was prone to disappear in some of the descendant dialects of the proto-language—perhaps this tendency went back to the proto-language itself.

We may set up "subroots," LUNG_{1a}, LUNG_{1b}, etc., on the basis of the particular proto-variants involved:

LUNG_{1a} [with s-prefix] *s-wa-p: Newari swa; Karen (Bassein Pho pò -θò?¹⁴⁶); also perhaps Bisu ?aŋ-hmaw, Phunoi ?ă-hmap.

Bisu and Phunoi are Loloish languages which show a strong tendency toward interchange of nasals with voiced initials of other kinds. Thus, a regular Bisu reflex of *nasal initials is a voiced stop ('fire' Bi. bî, Lh. mî; 'girl' Bi. bî, Lh. mî; 'black' Bi. da 33c,

Lh. nâ?; 'spirit' Bi. dât, WB nat; 'I' Bi. ga, Lh. nà, etc.)¹⁴⁷ We might therefore surmise that the Bisu and Phunoi forms for 'lung' point to a variant secondary prototype *s-map.

The sibilant prefix *s- is common in TB body-part and animal names, and may be referred to the independent TB root *sya 'flesh, animal' [STC, p. 106]. With this particular root, however, there is also the possibility that the prefix derives rather from the widespread TB morpheme for 'liver' *(m-)sin [STC 234], since there is attested the Jinghpaw disyllabic form just mentioned, sîn-wóp 'lung' (i.e. "spongy liver").

In fact, this interpretation provides us with a better explanation of the Bisu and Phunoi forms than the one just offered. Bisu hmaw and Phunoi hmap reflect a doubly prefixed prototype *s-m-wap, which in turn derives from the full disyllable *sin-wap that we find in Jinghpaw.

The two forms Benedict cites in STC 239 have rather different proveniences. Garo kasop may be referred to a doubly prefixed prototype, *k-s-wap.¹⁴⁸ Lushai chuap, however, is not to be considered the reflex of a triply-prefixed form **t-s-y-wap,¹⁴⁹ but is rather a fused product of an originally disyllabic form *tsi-wap, where the first element is an independent root directly attested in Loloish [see LUNG_{2c}, below].

LUNG_{1b} [with liquid prefix] *l-wa-p *r-wa-p: Jirel
lwa-syakq (2nd syll. = 'flesh'), Abor-Miri â-rop.

Several forms in Himalayan languages point to a doubly prefixed prototype, where a liquid prefix got fused to the root and was in its turn preceded by a secondary velar prefix: *k-l-wa-p > Tamang Risiangku glap-sya (2nd syll. = 'flesh'), WT glo-ba, Gurung kloba, Thakali 'kyahw'. We can only speculate on the morphemic identity of the velar prefix here, but it does seem to be the same entity as the one underlying the Garo form kasop [above]. Benedict [STC, p. 113] points out the appearance of a velar prefix with words for parts of the body in Konyak (Moshang kəmul 'body hair'), Kuki-Chin, and Bodo-Garo.¹⁵⁰

What about the liquid prefix in the words for 'lung'? Liquid prefixes pop up here and there in TB, but as Benedict says [STC, p. 109] they are "of uncertain function." In many cases we shall want to refer them to once independent root-morphemes. With 'lung' such a morpheme is fortunately well-attested in Tangkhul Naga āphār and Mikir ing-phor, which we reconstruct as "LUNG₃" *pwar.¹⁵¹ I suggest that forms like Abor-Miri â-rop might derive by metanalysis from the disyllabic prototype *pwar-wap: *pwar-wap > *pwa-rwap > *pərwap (unstressing of first syllable) > *rwap (loss of unstressed [i.e. now prefixial] first

element) > rop. The â-, of course, is a much more recent re-prefixation (the same element as in the TN form āphār).

LUNG_{1c} [with labial prefix] *p-wa-p: Chepang pop, Lahu ḥ-châ-phô?, Ahi ts144-pu22, Akha shâ-pḍ? (shâ = 'flesh'), Karen (BP pḍ?-ḥḍ?).

The syllables preceding the reflexes of *pwap in the Lahu and Ahi forms, descend from the distinct root that we refer to as LUNG₂ [below].

There is a form in the recently described TB language of Nepal, Kaike,¹⁵² that looks like this: phwa. On the basis of my present knowledge, I cannot tell whether this form is to be ascribed to an open-syllab-
led variant of LUNG_{1c}, that is *p-wa, or rather to the distinct proto-root that I am setting up as LUNG₃, that is *pwar. We return to the question of the relationships between LUNG_{1c} and LUNG₃ below (c).

Three interesting daughter-forms point to a prototype with both a labial and a resonantal element before the root, which we may provisionally set up as *p-r-wap. The forms are Khaling 'swah-prap', Bantawa som-phu-rok, and Chinese 肺 (Old [= 'Archaic'] Ch. *p'iwād, Middle [= 'Ancient'] Ch. p'iwoi- [Karlgren, GSR 501g]. Note that in both the Bantawa and Old Chinese forms the final stop is no longer labial, and I am assuming a **dissimilation of labiality** triggered

*pwar + *wap
LUNG₃ LUNG₁.

This now looks very much like the Bantawa form som-phu-rok (< *s-wam + *p-r-wak < *s-wap + *pwar + *wap)! ¹⁵⁶

Similarly, the Karen (Bassein Pho) form pòʔ-θòʔ clearly reflects the disyllabic prototype *p-wap + *s-wap, i.e., a sequence of LUNG_{1c} + LUNG_{1a}.

It is not at all strange that a language should make up two-syllable compounds consisting of alternate prefixations of the same root. This same fondness for **rhyming compounds** is found all over Southeast Asia, and in many other languages throughout the world. If we can have our Foxy Loxy's and Turkey Lurkey's, why can't they have their swap-pwap's and pwap-swap's? ^{156-a}

LUNG₂

We are now ready to leave LUNG₁ for awhile, and approach a quite separate root to which we give the label LUNG₂. LUNG₂ is, I believe, the original ST root referring to the body-part we call 'lung.' Unlike *wap, which seems to be an intruder from another semantic field (with the original meaning of 'soft, spongy'), no evidence has yet been unearthed to connect LUNG₂ to any extrapulmonary semantic domain. ¹⁵⁷

LUNG₂ has a dental affricate for its root-initial. It is sometimes preceded by a glottal prefix. Its vowel is

always high, but it shows front/back alternation between i and u [see above 1.121(5)]. Some allofams of LUNG₂ are open-syllables, while others end in a dental stop [see above 1.121(2)]. Formulaically, we may set up LUNG₂ as follows:

$$*ʔ-ts^i/_u-t.$$

Following our usual method, we may break down this etymon into "subroots," according to the particular proto-variants involved:

LUNG_{2a} *tsi: Lahu ḥ-chḥ [phonetically tsh1], Sani ts'z₁₁-p'o₃₃-ma₃₃.

The tones of the Lahu and Sani forms correspond regularly, pointing to PLB Tone *2 [see TSR No. 56].

LUNG_{2b} *tsit: Ahi ts1₄₄-pu₂₂.

Tone 44 is characteristic of Ahi syllables belonging to the Loloish high-stopped class [see TSR, p. 11].

LUNG_{2c} *ʔtsit or *ʔdzit: In the archaic Lahu liturgical language there exists an expression meaning 'entrails, viscera in general' of the form cḥ-ṣḥ-mì-cḥ, whose syllables I analyze as meaning 'lung'-'liver'-'tail'-'variety,' respectively. The high-rising tone of cḥ indicates both a glottal prefix and an originally stopped ending [Matisoff 1970], but also a voiced root-initial [TSR, p. 20]. On the other hand, it could be a purely sandhic development from an earlier cḥʔ, which would reflect a *voiceless root initial

(*ʔtsit).¹⁵⁸

LUNG_{2d} *ʔtsut: WB chut (alternatively transcribable as tshut), Atsi tsʔut (Burling), Hayu jot (Michailov-sky 1974a), Lakher pa-chao.¹⁵⁹

Note that most of these forms are from Lolo-Burmese. Yet the Hayu form demonstrates that the root must be set up for PTB. Furthermore, the Lushai form chuap, as we have seen, must be assigned to a disyllabic prototype consisting of a sequence of LUNG_{2a} + LUNG₁: *tsi-wap > *tsywap > chuap.¹⁶⁰

LUNG₃

LUNG₃, as we have seen, is to be reconstructed as *pwar, and is directly represented by Tangkhul Naga āphār and Mikir ingphor, and perhaps by Angami Naga úphié and Kaike phwa as well. *pwar seems also to have been much used as the first syllable in polysyllabic compound formations for 'lung,' combining especially with LUNG₁: *pwar-wap [see above]. This etymon seems to be quite distinct from the labial-prefixed variant of LUNG₁ that we have been calling LUNG_{1c} (*p-wap).

LUNG₄

Finally, to further complicate the picture, we must set up another etymon, LUNG₄, with the shape *pu. This root has semantic affiliations with the notion 'sac, bladder,'

and perhaps ultimately also with one of the subroots for 'belly/stomach,' *s-pu [see below B/S_{2a}].

Our LUNG₄ is directly represented by Lepcha a-bu, Tاراon (= Digaro) a-po, and the second syllable of Ahi ts1₄₄-pu₂₂ and Sani ts'z₁₁-p'o₃₃-ma₃₃. It is possible that the middle syllable of the Bantawa form already discussed, som-phu-rok, belongs here, though the alternative interpretation suggested above is also possible.¹⁶¹ The Lahu form phô 'sac, bladder' (ḁ-jḥ-phô 'urinary bladder' ["urine-sac"] fits perfectly. Since Lahu also shows a distinct reflex of *p-wap (LUNG_{1c}) in its form ḁ-chḥ-phô?, this is good evidence for the separateness of LUNG_{1c} and LUNG₄ at the proto-level.

For further speculations on the interrelatedness of LUNG_{1c} (*p-wap), LUNG₃ (*pwar), and LUNG₄ (*pu), see section 1.141(c) below.

* *

*

Resume of LUNG-etyma. In this section, for ease of reference, we assemble all the proto-forms, both mono- and polysyllabic, that we have reconstructed for 'lung':

LUNG ₁	* <u>P-wa-p</u>
LUNG _{1a}	* <u>s-wa-p</u> * <u>s-m-wap</u> [< * <u>sin-wap</u>] * * <u>k-s-wap</u>
LUNG _{1b}	* <u>l-wa-p</u> * * <u>r-wa-p</u> * * <u>k-l-wap</u>

LUNG _{1c}	* <u>p-wa-p</u>	
LUNG ₂	* <u>ʔ-tsⁱ/u-t</u>	
LUNG _{2a}	* <u>tsi</u>	LUNG _{2d} * <u>ʔtsut</u>
LUNG _{2b}	* <u>tsit</u>	
LUNG _{2c}	* <u>ʔtsit</u> (perhaps * <u>ʔdzit</u>)	
LUNG ₃	* <u>pwar</u>	
LUNG ₄	* <u>pu</u>	

Polysyllables:

*sin-wap: Jg. sin-wóp, Bisu ʔan-hmaw, Phunoi ʔã-hmap

*pwap-swap: Bassein Pho (Karenic) pòʔ-θòʔ

*pwar-wap (> *p-r-wap): Old Chinese *p'iwăd

*swap-prwap (< *swap-pwar-wap): Khaling 'swah-prap

*sin-prwap (< *sin-pwar-wap): Bantawa som-phu-rok

*tsi-wap: Lushai chuap

*tsi-pwap: Lahu ð-chê-phôʔ

*tsi-pu: Sani ts'z₁₁-p'o₃₃-ma₃₃, Ahi tsl₄₄-pu₂₂.

* *

*

The above discussion should convey some idea of the complexity of the morphophonemic and morphological dynamics in ST word- and compound-families. And yet, LUNG is by no means atypical in its complexity. To demonstrate this, we will present some data for one more body-part concept, BELLY/STOMACH. To save time and space, we will not elabor-

ate on every detail of the reconstructions, but merely list the forms in a quasi-tabular way. We shall return to the semantic issues presented by the ST words for BELLY/STOMACH below, 2.53.

BELLY/STOMACH₁ * s-^P/b/_w u/i(.)k

B/S_{1a}: *s-pu.k *s-bu.k [STC 358]

- (a) [voiceless root-initial] Chinese¹⁶² *piôk/piuk 腹
 'belly' [GSR 1034h] (< *s-pu.k) *p'iôk/p'iuk 覆
 'cave' [GSR 1034-L; STC, p. 182] (< *pu.k); WT
 p'ug(s) 'innermost part,' p'ug-pa 'cavern'; Limbu
 səpok 'belly' (also Limbu səpu < *s-pu B/S_{2a} be-
 low), Lushai pu.k 'cave,' WB puik 'pregnancy,'
 wam-puik 'outside of belly,' Mikir pok 'belly,'
 Kabui puk 'belly,' Meithei puk 'belly,' Jyarung
 (Tsa-ku-nao dialect) pog, phog 'guts,' (Chos-kia
 dialect) tipog 'stomach.'

- (b) [voiced root-initial] Chinese *b'iôk/b'iuk 覆
 'cave' [*loc. cit.*], WT bug-pa 'hole,' sbugs
 'cavity.'

B/S_{1b}: *d-puk *d-buk *d-bwak

- (a) [voiceless root-initial] Ao Naga tapok 'cave,'
 tepok 'belly,' Gyarong tēpök 'belly.'

- (b) [voiced root-initial] Lepcha tābak ~ tābok ~ tāfuk
 (< *d-bwak).

(c) [with prefix preemption¹⁶³] Chepang tuk, Magari mi-Tuk.

B/S_{1c}: *wuk [lenited initial—see above 1.121(6B)] Garo ok, Tangkhul Naga wuk, Maring uk, Karen (Moulmein Pho γàu?, Taungthu hố?, Palaychi hù?).¹⁶⁴

The spirant -f- in Lepcha tăfuk represents a stage in lenition intermediate between the stop -p- and the semi-vowel -w-. When a language refuses to tolerate both the lenition of the root-initial and the disappearance of the prefix, it may "choose the strategy" of having the prefix preempt the root-initial position (Chepang, Magari).

In the remaining two allofams of this set, the nuclear vowel is -i-, not -u-:

B/S_{1d}: *ʔ-pi(.)k *ʔ-bi(.)k

(a) [voiceless root-initial] WT p'ig-pa 'make a hole,' Mikir phek 'bowels.'¹⁶⁵

(b) [voiced root-initial] WT hbig(s)-pa 'make hole,' Garo bi-bik 'bowels.'

B/S_{1e}: *ʔwik [lenited initial—see TSR 176] Lahu ð-fɛ-qō 'stomach,' Lisu h'i₆-hchi₆, Sani hI-ma; Karen (Moulmein Sgaw yy?, Bassein Sgaw hy?).

We now leave BELLY/STOMACH₁, and proceed to another set of forms that we choose for the moment to consider as representing a separate proto-group, BELLY/STOMACH₂:

BELLY/STOMACH₂ *(s-)^P/b_w/u

B/S_{2a}: *(s-)pu

- (a) [voiceless root-initial] Limbu səpu 'belly' (*səpok B/S_{1a}); WT p'o-ba 'stomach'; Jinghpaw pù 'bowels,' lùn-pū 'cave'; TN āphurāpham 'belly';¹⁶⁶ Chinese *piu/piu 'the bowels' 腑 [Karlgren, *Analytic Dictionary* 1923].

B/S_{2b}: *(s-)bu

- (b) [voiced root-initial] Boro bi-bú 'guts,' Chinese *b'iu/b'iu 'intestines' 付 [Karlgren, *AD* 1923].¹⁶⁷

B/S_{2c}: *wu [lenited initial] Lahu ḥ-ḡḡ-tê? 'guts,'¹⁶⁸ Akha ḡ-má? 'abdomen,' WB ?u 'intestines,' Lisu wu₄ 'id.'

Note that the semantic affiliations of B/S₂ lie more with 'bowels/intestines/guts,' while those of B/S₁ lie more with the idea 'cave/cavity/hole.' Yet both affiliations are present in both groups in a complicated way that may be graphically (though rather crudely) captured by a "meta-static flowchart" (see below 2.43).

BELLY/STOMACH₃ *pon × *pon

Bisu pòn-ba, Pyen pawng-pawng, Akha baw-ma, Karen (Moulmein Pho yàu?-phòn), Tangkhul Naga āpaiphon.

BELLY/STOMACH₄ *pam × *wam

- (a) [voiceless stop initial]

B/S_{4a}: *pam

Jg. pù-hpam 'stomach,' Tangkhul Naga āphurāpham 'belly' [Note that both Jg. and TN have formed com-

pounds of the shape $B/S_{2a} + B/S_{4a}$.]

B/S_{4b} : *wam [lenited initial] WB wâm 'belly,' Lahu gô-pè,¹⁶⁹ Atsi vâm, Maru wen, wen-tok 'abdomen, stomach,' Mikir vam 'waist, loin,' Lushai von a só 'diarrhea,' Tamlu hwum 'belly,' Lakher a-vy, pa-vy 'stomach.'¹⁷⁰

We could proceed to mention Lushai pùm ~ pùn 'belly,'¹⁷¹ but that would take us somewhat afield to a related but distinct group with the meaning 'swelling, convex, rounded (like a full belly)': PTB *s-bwam [STC 172] > Lushai puam 'to swell,' Jg. bōm 'to swell,' bōp 'calf of leg,'¹⁷² WB phwam 'plump,' WT sbom 'thick, stout'¹⁷³...

1.141 (c) The problem of setting up "higher-order allofams"—where to draw the line? In the present state of our knowledge, many of the fine details of our reconstructions remain somewhat arbitrary. In particular, we face a non-uniqueness of allofamy problem:

HOW DO WE DISTINGUISH "UNRELATED SEPARATE
ROOTS" FROM CO-ALLOFAMS?

From what has already been said, everybody will agree that it is not always easy to decide, when presented with a pair of modern forms, whether they represent two separate proto-roots, or whether they represent co-allofams of the same proto-word family. Yet it is important to emphasize that a similar difficulty exists at the proto-level as well. In

fact, the whole notion of "the same proto-root" as opposed to "unrelated, separate proto-roots" is shrouded in indeterminacy. We are dealing with rarefied abstractions, and just because we can give neat and discrete number-and-letter labels to our various proto-roots ($LUNG_{1a}$, $LUNG_{2b}$, etc.), that does not mean that we have really grasped all the infinitely complex details of their interrelationships. Even at the proto-level, we often have no principled basis for deciding whether two reconstructed forms in a given semantic area are to be interpreted as

*ETYMON₁ vs. *ETYMON₂

or rather as

*ETYMON_{1a} * *ETYMON_{1b}.

Thus, to return to our LUNG example, we set up several "distinct" proto-etyma beginning with a labial stop:

$LUNG_{1c}$ *p-wa-p / $LUNG_3$ *pwar / $LUNG_4$ *pu.

By labelling these proto-forms with separate arabic numerals, we are claiming that they are all etymologically unrelated—i.e., they are "separate roots," separate items in the proto-lexicon, whose phonological resemblance is only accidental. But how can we be so sure? Some smart-aleck could come along and try to lump all three of them into a single "higher-order etymon," say,

$LUNG_{1c/3/4}$ **p-w-a-^P/_r.

In this particular instance, the smart-aleck doesn't have much of a case, and he could easily be beaten down with

arguments like "variation between syllable-final -p and -r is not a well-established phenomenon in Tibeto-Burman."

In other cases, however, there is still room for honest linguists to say they cannot be sure. In our presentation of the data for BELLY/STOMACH, for example, we set up B/S₁ abstractly as

$$* \left(\begin{smallmatrix} s- \\ ?- \end{smallmatrix} \right)^P / b / \begin{smallmatrix} u \\ w \end{smallmatrix} / i (\cdot) k.$$

The various sub-roots under this "single" proto-root, (B/S_{1a}, B/S_{1b}, B/S_{1c}, B/S_{1d}, B/S_{1e}), differ among themselves in the nature of their prefix, initial consonant, and/or nuclear vowel, but they all have a final -k. This final -k was seized upon by me the analyst as the crucial or criterial feature in setting up this "single" proto-root. Yet it is very hard to believe that there is no connection whatever between this etymon and the one that we set up as BELLY/STOMACH₂:

$$*s- \begin{smallmatrix} P \\ b \\ w \end{smallmatrix} / u.$$

The alternation between open and stopped syllables is, in fact, a perfectly respectable variational pattern in TB [above 1.121b], and there can be no principled objection to combining our B/S₁ and B/S₂ into a still more abstract unitary word-family, that could be set up with a hyphen before the final -k to indicate that "some allofams of this family have final -k but others do not," thus:

$$\text{BELLY/STOMACH}_{1/2} \quad * \left(\begin{smallmatrix} s- \\ \gamma- \end{smallmatrix} \right)^{p/b} / \begin{smallmatrix} u \\ w \end{smallmatrix} / i(\cdot) - k.$$

The reason why I didn't do this in the first place was an esthetic one— B/S_1 already looked forbiddingly complicated, and I didn't want to overload everybody's circuits by cramming too much into it. There are "lumpers" in the world, and there are "splitters" in the world. Let a hundred flowers bloom.

1.141(d) **Compound-family scoring problems.** The pervasive TB morphological process of compounding leads to tremendous difficulties for the standard lexicostatistical approach. As we have seen [above 1.122], compound-formation is a highly idiosyncratic process, where even closely related languages or dialects may freely make quite different selections from the "Urwortschatz" in coining a compound word for a given concept.

All sorts of complicated "scoring problems" arise when we try to compare polysyllabic compounds across languages.¹⁷⁴ Suppose Language A has a compound in which one syllable appears to be the direct phonological cognate of one syllable in a compound in Language B. Is this worth more or less than a perfect phonological correspondence between monosyllabic forms in the two languages? (Should we deduct points for non-cognate material elsewhere in the compounds?) The third syllable of one Lahu word for

'lung,' ḁ-chḥ-phô?, is cognate to the second syllable of the Akha form shà-pḁ? (< "LUNG_{1c}" *p-wap). That would be worth some points. Yet Lahu chḥ (< "LUNG_{2a}" *tsi) is totally unrelated to Akha shà, which means 'flesh, meat.' Furthermore, the syllable -phô? is now a bound morpheme of obscure meaning in Lahu, and in fact the usual form for 'lung' is simply ḁ-chḥ. Maybe we should deduct some points on that account?

On the other hand, when **two** syllables of a compound in Language A each correspond to syllables in a compound in Language B, that seems to be worth some extra bonus points, and might even suggest that the languages in question are particularly closely related. Thus, Ahi and Sani, which are indeed two closely related Loloish languages, have forms for 'lung' of the shapes ts₁₄₄-pu₂₂ and ts'z₁₁-p'o₃₃-ma₃₃, respectively. The second syllables both go back to "LUNG₄" *pu, while the first syllables reflect different sub-roots of LUNG₂. (As explained above, the tonal evidence shows the Ahi syllable to derive from *tsit [LUNG_{2b}], while the Sani syllable comes from *tsi [PLB Tone *2], i.e. LUNG_{2a}). But now what about the second and third syllables in the Lahu forms ḁ-chḥ-phô?? The -chḥ- is directly cognate to the Sani first syllable, and it would be nice if the Lahu -phô? were cognate to the second syllables in Ahi and Sani. But the final glottal stop indicates rather that it derives from LUNG_{1c} *p-wap, and not from LUNG₄ *pu. The Ahi and

Sani forms really **are** more closely related to each other than either is to the Lahu form! Needless to say, this could never have been discovered by mere inspection.

Further complications arise when a given language has more than one synonymous compound for a given concept. In Lahu, the succulent quadruped known in English as the "barking-deer" is designated by several names, two of which are chĩ-pĩ-qwè? and chĩ-ke-ne.¹⁷⁵ It is clear that five separate morphemes are represented in these words, to which we may assign arbitrary numbers:

I. chĩ II. pĩ III. qwè? IV. ke V. ne

Let us now take three hypothetical languages, and suppose that their forms for 'barking-deer' are as follows:

LANGUAGE X

LANGUAGE Y

LANGUAGE Z

tsi-pi

ke-ne

pi-nae

That is, Lg. X reflects I + II, Lg. Y reflects IV + V, and Lg. Z reflects II + V. Which of these forms is "most closely related" to the Lahu lexicon? It is hard to make sense of that question. Lg. X shares two syllables with one Lahu form, Lg. Y shares two syllables with the other Lahu form, while Lg. Z shares one syllable with each of the two Lahu forms. To which of the two Lahu forms should we give more weight—the more "common" one? Suppose they are equally common in a given dialect? The problem is similar to what one runs into with a lexicostatistical list. If a language has two equally common near synonyms, like

'little'/'small' or 'start'/'begin,' which one do you use in making interlingual comparisons?

A particularly interesting kind of problem that can arise in this area involves what Benedict calls "split cognates" (1975a, "ATLC"). Split cognates are forms which reflect different syllables of a proto-compound, such that they share no reflexes of any given proto-phoneme. Lest this seem too bizarre, consider words like taxicab. Suppose some future dialect of English retains only a reflex of taxi, say tek, while another dialect retains only a reflex of cab, say kep! Philological evidence will prove to our descendants that there once existed a polysyllabic form taxicab. (Some future dialects may in fact faithfully reflect this, with forms like tekep.) It would be interesting to see how the linguists of the future score such cases. ^{175-a}

1.142 **Patching up the basic vocabulary list: the CALMSEA attempt.** By now it should be fairly evident that I am not a devotee of conventional lexicostatistics. However, even if we have no great faith in glottochronology on methodological grounds, it might at least be worthwhile to try and improve the basic wordlists that are commonly used in these studies, the so-called "Swadesh lists."¹⁷⁶ Even with the "best" list of core vocabulary items that it is possible to draw up, the lexicostatistical method is fraught

with enormous problems and uncertainties. When an attempt is made to apply this method to Tibeto-Burman, these theoretical and methodological difficulties are compounded by the cultural, grammatical, and semantic biases of the standard Swadesh lists. If we remove these biases, we improve the lists. We do not thereby make the lexicostatistical method itself any more reliable—but at least we make its application to Southeast Asian languages a little less arbitrary, and give it a fairer chance of producing meaningful results. We have "eine Sorge weniger," one less thing to worry about.

With this end in view, I have prepared a new wordlist of some 200 items, only about 110 of which are identical to items on the expanded Swadesh list. Our new list, which we might call the "CALMSEA lexicostatistical list" [Culturally and Linguistically Meaningful for Southeast Asia], is given as Appendix I of this monograph. In the Appendix, the items of the CALMSEA list are subdivided according to their semantic category (body-parts, psychological verbs, etc.), and accompanied by all the relevant PTB reconstructions that Benedict makes in the STC.^{176-a}

Let us briefly look at some of the ways in which the CALMSEA list is an improvement over the Swadesh list:

1.142(a) **Cultural appropriateness.** There are two ways in which cultural inappropriateness manifests itself in a wordlist. Either the list contains words which are them-

selves inappropriate to the given cultural area (sins of commission); or it does not contain certain words which are extremely appropriate to the cultural area (sins of omission). In the first category are words like ice and snow (both on the Swadesh list), which are hardly items of core vocabulary in the generally tropical climate of Southeast Asia. In Lahu, for example, if one succeeds in eliciting a word for 'ice' or 'snow' at all, what one gets is neologistic and polymorphemic extensions of words like 'frost' and 'hail' ('ice' = FROST + LUMP), ('snow' = HAIL + POWDER).

Even more serious, perhaps, are omissions of quintessentially appropriate cultural items like monkey, cooked rice, bamboo, village, mortar, banana, pig, etc.

The dyed-in-the-wool practitioner of orthodox lexicostatistics is not impressed by considerations of cultural appropriateness, since he is looking for "comparability of results" across language families throughout the world.¹⁷⁷ Yet I would maintain that true comparability is achievable only if **culturally equivalent** (not referentially identical) lexical items are used cross-areally. The Lahu prize the barking-deer as the most delectable game-animal in the jungle. Let us say that the Nootka prize the seal as the tastiest animal in their environment. Surely we achieve more comparable results if we put 'barking-deer' on our "Southeast Asian list" and 'seal' on a separate "Hyperborean

list," than if we were to insist on including either one of these words on some single "universal basic vocabulary list."

1.142(b) **Grammatical appropriateness.** The Swadesh list contains items of grammatical hardware like the English prepositions at and in. The languages of Southeast Asia certainly have resources for expressing locative relationships, but not necessarily via separate words. Often the locative idea is a semantic component of the verb itself (much as English 'enter' contains INTO, and 'emerge' contains OUT OF). Even if the locative notion is expressed by a separate word (e.g., the Lahu "spatial nouns," or some of the Lahu "noun-particles"), we certainly have no right to expect any one-to-one correlation between the highly abstract meanings of a SEA locative morpheme and an English preposition!¹⁷⁸ To use the glottochronologists' own word, there is no "comparable" Lahu word to English 'at' or 'from.'

Even in the realm of root-vocabulary (nouns and verbs), many concepts which are expressed by a monomorphemic word in English may typically be expressed by a compound in SEA languages. Thus, 'knee' (on the Swadesh list) is very often "LEG + JOINT;" 'feather' (on the Swadesh list) is usually "BIRD + BODY-HAIR;" 'ankle' and 'tear' (not on the Swadesh list) are widely expressed in SEA by "LEG + EYE" and "EYE + WATER," respectively. We have already discussed at length the intricate and idiosyncratic nature of TB compound formation. In any list tailor-made for SEA, poly-

morphemic words must be handled (and scored) with special care.

The Swadesh list contains two separate items for second-person pronouns, thou and ye. This actually represents one of the few attempts Swadesh makes to free himself from Anglocentric bias, since in modern English we of course use a single pronoun, you, whether we mean singular or plural. Ironically, however, the SEA languages are more like modern English than like archaic English or languages like French and German which have two separate unanalyzable etyma for '2nd person singular' and '2nd person plural' (thou/ye, tu/vous, du/ihr). Southeast Asian languages typically do not make the distinction at all, or else (when the communicative situation absolutely demands it), express the plural by adding a morpheme to the singular form [e.g. Lahu nò 'you (sg.)', nò-hĩ 'you (pl.)'].

1.142(c) **Semantic appropriateness: over- and under-differentiation.** Considered in their semantic aspect, many of the examples of incommensurability in the preceding section may be viewed as **over-** and **under-differentiations**. For most SEA languages "thou" vs. "ye" is an over-differentiation. Similarly, since gender distinctions are almost never made in the third person in SEA, it would be a sin of over-differentiation to have both he and she on the list.

(Swadesh has only he; a better gloss for SEA would be 'he/she,' or 'third person pronoun.')

Swadesh does not, with good reason, have two separate items for 'moon' and 'month' on his list, since the mensual unit of time is clearly derived both conceptually and morphologically from the lunar heavenly body in most languages of the world. Yet sun and day appear as separate items on his list, even though in many SEA languages the same morpheme appears in both words (Malay mata hari 'sun' < mata 'eye' + hari 'day'; Lahu ni 'day,' mû-ni 'sun' [with mû- 'sky, heaven']; Burmese nei [tone 1] 'sun,' nei' [creaky tone] 'day,' etc.). As often happens, over- and under-differentiations intertwine with each other. Although the Burmese words for 'sun' and 'day' are not independent etyma (under-differentiation with respect to English), Burmese does make a distinction between 'day as opposed to night' (nei'), and 'full day of 24 hours' (ye?: WB rak), which is over-differentiation with respect to English.

For most modern Western languages it does make sense to have two separate items for animal and meat/flesh. In the hunting cultures of backwoods SEA, these two concepts are often expressed by exactly the same word (e.g. Lahu šā 'animal; meat, flesh;' Jinghpaw shàn 'flesh; deer; large game animal, as used for food').

Cases where the Swadesh list under-differentiates with

respect to SEA languages are particularly striking. One item on his list is the verb 'cut.' SEA languages make a large number of semantic distinctions in this area, according to the nature of the material being cut, the implement used to do the cutting, the kind of motion used by the cutter, the size and shape of the pieces produced, etc. Any given language is bound to have 10 or 12 commonly used verbs in this semantic area. The STC provides reconstructions for no fewer than nine etyma which could receive the English gloss 'cut,' viz. *dan (No. 22), *mrak (No. 147), *tsyat (No. 185), *tsywar (No. 240), *lep (No. 351), *ri·t (No. 371), *kut (No. 383), *tuk (No. 387), *ra-t (No. 458).¹⁷⁹ A wordlist tailored to SEA languages would have to be much more specific in this area, with items like 'cut (as a living upright tree with an axe),' 'cut (as a banana-stalk in small enough pieces for a pig to eat),' etc.

Clearly, if the lexicostatistical method is to have any chance of success in clarifying the degrees of genetic relationship among languages of a given cultural area and typological structure, each item on the wordlist must be subjected to detailed and expert scrutiny by a linguist who has steeped himself in the lexicons and grammars of those languages in advance. Garbage in, garbage out.

Each word has its own history. To maintain, as the orthodox glottochronologists do, that these individual differences "even themselves out" sufficiently so that a

single 200-word list is good enough to figure out the genetic relationships among all the world's languages, is to ask us to subscribe to an unprovable article of faith.

A revised, culturally and typologically specific list is a big improvement. Yet the whole approach of lexico-statistics must give way to something better. In the remainder of this study, we shall outline the beginnings of a new sort of attack on the problem—the "organic semantic" approach to linguistic comparison.

2.0 THE ORGANIC SEMANTIC APPROACH TO LINGUISTIC COMPARISON

The semantic dimension of language was neglected for many years in American structuralist and early generative synchronic linguistics. As everyone knows, that tide has recently been decisively reversed, to the point where semantics is now the chief focus of synchronic theoretical attention. Meaning, as Bloomfield feared, is indeed the most difficult aspect of language to study, impinging as it does on the speaker's and hearer's psyches and the extralinguistic "pragmatic" situation in which their communicative act takes place.¹⁸⁰

It is now time to give semantics its due on the diachronic plane as well. Recent American books on historical linguistics, even the best ones, typically devote scant space to the profoundly important and fascinating topic of semantic change through time.¹⁸¹ Everybody knows that the meanings of words change. However, the factors which cause these changes and the directions these changes take are fantastically complex and unpredictable. This has made most diachronic linguists shy away from the problem. The whole question makes them vaguely uneasy, and it often seems as if they would vastly prefer it if words always meant one and the same thing, forever and ever.¹⁸²

Language change through time is of course intimately connected with synchronically observable linguistic variation.¹⁸³ The same people who are made uneasy by the way

meanings change through time must also be upset by the idea that meanings may be in a state of flux even in a given speech community at a given point in time.

A brilliant exception to the anti-semantic orientation in American diachronic linguistics was the University of Chicago Indo-Europeanist, Carl Darling Buck. His masterpiece, *A Dictionary of Selected Synonyms in the Principal Indo-European Languages* (1949), is a total reorganization of the vast bodies of material to be found in the classic Indo-European etymological dictionaries, like Walde-Pokorny, as well as in the more specialized etymological dictionaries on individual languages and sub-families of Indo-European. These other dictionaries had been organized by the blind tyranny of alphabetical order, whereby utterly random sequences of concepts like KICK, KIDNEY, KILL are juxtaposed to each other. Granted it is possible to kill someone by kicking him in the kidneys. Still, one could hardly claim that these three concepts have any special relationship to each other, except for the fortuitous similarity of their initial letters. Buck's great contribution to historical lexicography was to create a "dictionary of synonyms," a kind of historical Thesaurus, where the various Indo-European etyma are categorized according to their meanings. The advantages of this approach are enormous. First of all, there is the matter—by no means trivial—of sheer interest. Nobody except a dedicated, desiccated

bookworm could conceivably spend hours browsing through Walde-Pokorny's *Vergleichendes Wörterbuch der indogermanischen Sprachen* for the sheer pleasure of it.¹⁸⁴ That is a book one consults for reference, with reverence, perhaps, but not love. It is not a book you take to bed with you. Buck's book makes the Indo-European lexicon spring to life in your mind. As you browse through the various sections—"The Physical World in its Larger Aspects;" "Mankind: Sex, Age, Family Relationship;" "Animals;" "Parts of the Body/Bodily Functions and Conditions;" "Food and Drink, Cooking and Utensils;" "Clothing/Personal Adornment and Care;" "Dwelling, House, Furniture;" "Agriculture/Vegetation;" "Motion: Locomotion, Transportation, Navigation;" "Spatial Relations: Place, Form, Size;" "Time;" "Mind, Thought;" "Warfare;" "Law;" "Religion and Superstition," etc.—you acquire a vivid sense of the real world in which the Proto-Indo-European people and their far-flung descendants have been living for the last several thousand years.

The subsections of each chapter are numbered decimally, with the concepts following one another in a logical order. Thus, a sequence of "Bodily Functions and Conditions" in Chapter IV runs like this: 4.51 BREATHE, BREATH 4.52 YAWN, GAPE 4.53 COUGH 4.54 SNEEZE 4.55 SWEAT 4.56 SPIT 4.57 VOMIT 4.58 BITE 4.59 LICK.... The typical entry begins with a list of 30 or more forms for the concept in

question in a variety of ancient and modern IE languages (in a fixed order usually starting with Greek and ending with Sanskrit or Avestan). Buck then devotes a paragraph or two to the semantic interconnections of the concept. YAWN / GAPE (p. 261) will serve as an example:

"The primary notion of the words listed here is 'open the mouth wide.' Many of them may be extended to 'open wide' in general (of the earth, etc.). Nearly all of them show secondary associations with mental attitudes of which the action is indicative, namely wonder, stupidity, or fatigue. In such secondary associations they may be differentiated, as NE yawn (with fatigue) and gape (indicating wonder or stupidity)...or be restricted to one or another of these notions."

At this point, in separately numbered paragraphs, Buck makes groupings of the forms he had listed above, this time arranging them according to their PIE etymology. The first numbered paragraph of YAWN / GAPE contains the largest block of forms, those which can be derived from the PIE root *ghē(i)-, *ghī-, *ghiā-. Here we learn fascinating facts, e.g. that Greek χάσμα 'open mouth, yawning space, chasm' (whence English *chasm*) is relatable to Latin *hiāre* 'to yawn, gape, be open' (whence English *hiatus* 'a gap or break,' as vowels in hiatus—i.e., with no intervening consonant so that there is a consonantal 'gap')¹⁸⁵ as well as to German *gähnen* and to both English doublets, *yawn* and *gape* (the latter from "an extension of the same root"—i.e., what we would call an "allofam," *ghē-b-). The next bunch of forms, in paragraph 2, is a much more restricted Romance group

from a posited Vulgar Latin **batāre*, whence Late Latin *bataclāre*, French *bâiller*, Italian *sbadigliare*, etc. At the end of the whole entry comes an isolated form with a dubious etymology, Sanskrit *jṛmbh-* [not all problems have been solved yet, even in Indo-European!].

Buck's method thrusts meaning into the central position in linguistic comparison. Semantics becomes the backdrop against which the chaotic-seeming parade of forms is viewed, the stage on which the phonological shapes of the morphemes tumble about for our delectation. Buck has a keen understanding of the dynamics of semantic fluctuation, association, and change. For him, semantic change is the rule, and not the exception, in language history. He is worth quoting *in extenso* on this point:

"In a limited number of words for certain concrete notions the meaning has remained virtually constant for thousands of years, as, for example, those for the numerals and for close family relationship.... But such cases are the exception. *Change is the rule* [italics mine]—sometimes so radical that, without the connecting links, one would be at a loss to see any relation between the earliest known and the latest use. The meaning of most words is not a fixed point, but an area of variable dimensions. It is subject to complex associations, some of them inevitable associations of universal scope, others originating in some particular verbal context, external situation, or among a special social or occupational class...." [pp. v-vi]

"By the study of synonyms, their etymology and semantic history, one seeks to show the various sources of a given notion, the trails of its evolution. With some notable exceptions...a group of synonyms has little resemblance to a group of formal cognates such as we find in the etymological dictionaries...." [p. x]

Words, in short, are endowed with a life of their own. A language, whether a modern vernacular or a hypothetical proto-language, is like an ORGANISM—that is, a tremendously complex object that is **organ**-ized simultaneously in a number of distinct but interrelated systems and dimensions, and that is alive in all its parts. The words of a language are like the cells of the organism, fulfilling many different functions simultaneously, exerting influence upon and receiving influence from their neighbors (both their immediately adjacent neighbors and those far removed in "organic space "). Like the cells in a human body, the words of a language undergo complex changes as the language's life unfolds. Some die and are sloughed out of the lexicon, like dead epithelial cells on the skin. Their demise may or may not be compensated for by replacements. The analogy could be elaborated endlessly—sometimes a language's metabolism seems to change! It may become voraciously hungry for new words, absorbing them from whatever source and assimilating them to itself completely. At other times the language may go on a reducing diet, and cut out massive amounts of lexical fat.¹⁸⁶ It is no accident that one refers to a large amount of textual material from a language as a **corpus**, or 'body.'

By our "organic semantic approach" we are trying to integrate into a single body the head, ears, tusks, back, legs, sides, and tail of that infinite and protean Elephant

that is the object of study of the historical linguist.

2.1 **Semantic fields and semantic systems: body-part nomenclature and the Tibeto-Burman Etymological Dictionary.** The usual type of lexicostatistical list is a hodge-podge of items from a disparate variety of semantic areas—kinship terms, body-parts, names for natural objects, verbs of motion, verbs of manipulation, psychological verbs, artifacts, etc. In sharp contrast, the kind of study we have in mind focuses on one particular area at a time, exploring its internal structure in depth and tracing its interconnections with other areas as necessary.

At this point we should introduce some simple terminology. We use "semantic area" and "semantic field" interchangeably to mean "a relatively well-defined region of semantic space," that is, a group of concepts that share some salient feature of meaning.¹⁸⁷ Thus, English words like *swim, fly, creep, crawl, walk, run, jump, float, hop, move, stagger, swoop* belong to a semantic field /MOTOR ACTIVITY/, since they may all refer to modes of motion by living beings. By "semantic system" we mean a semantic field considered in its relational aspects--relations among the elements within the field, as well as relations between elements within the field and elements from other fields. Returning to our field /MOTOR ACTIVITY/, we may make "intra-field systemic" observations like the follow-

ing: *creep* and *crawl* are more closely related to each other (since they share the feature [+on all fours]) than either is to *run*; *fly* and *swoop* are more closely related to each other (since they share the feature [+through the air]) than either is to *swim*, which contains a paradigmatically opposed feature [+through a liquid], etc. Even when all intrafield relationships have been exhaustively described, we cannot have a true picture of the field's orientation in semantic space until we proceed to make "inter-field systemic" observations as well. Thus, *swim* and *float* are very close to each other **within** the field of /MOTOR ACTIVITY/, sharing the feature [+through a liquid], and differing only by some feature like [+active use of arms and legs] (*swim*) vs. [-active use of arms and legs] (*float*).¹⁸⁸ The extra-field affiliations of the two words are quite different, however. Because of its greater "passivity," *float* is by no means restricted to animate beings. It is just as natural to say "the bottle floated by" or "the corpse floated by" as it is to say "Herman floated by." That is, *float* (like *move*) belongs equally well both to the field /MOTOR ACTIVITY/ and to /INANIMATE MOTION/. *Swim* is much more firmly ensconced in the /MOTOR ACTIVITY/ field. We do not say "the corpse swam by" or "the torpedo swam through the water." When we do use *swim* for inanimate motion ("the planet swam into his ken"), this is felt to be a "shift" or "extension" of its usual meaning. It would take a full-scale study to analyze the inter-field relation-

ships of even these few verbs of motion. One final example: *fly, creep, crawl, run, and move* all have interconnections to the field /PASSAGE OF TIME/. Time flies, time creeps closer, time crawls, time runs out, time moves along—but time does not swim, hop, stagger, swoop, or float....

What does "stagger" us is the fantastic complexity of the whole business. As Buck says:

"Relations are too complex to admit any truly scientific and complete classification; and, even if such were possible, it would have little relation to our instinctive associations.... Yet because the ideal is hopeless, to abandon all such classification would be to sacrifice the obvious advantages of a semantic grouping."
[p. xiii]

The problem then is to find some semantic field that is so well-defined that it can serve as a "test case" or a "pilot project" for the study of semantic relationships in general. By far the best semantic field for this purpose is the /HUMAN BODY/ itself.

What could be more appropriate for an "organic semantic" study than the human body, the most perfect and complex organism in our solar system? In the first place, except for extremely minor anatomical differences between races, the human body is everywhere the same. We therefore have a universal, objective external referent against which to compare the somatic semantics of particular languages.^{188-a} Secondly, the human body (like language itself) is tremendously complex. It has a huge number of nameable component

parts. The latest paperback edition of Gray's *Anatomy* has 1257 pages. Thomas Love Peacock (1785-1866), in his novel *Headlong Hall*, was putting it mildly when he referred to the human body as "osseocarnisanguineoviscericartilagino-nervomedullary."¹⁸⁹ Not only is the number of component parts so large, but they are also organized into a great number of systems and subsystems which interconnect with each other at all points—skeletal, muscular, nervous, circulatory, respiratory, digestive, lymphatic, endocrine, reproductive. Thirdly, most speakers of a language, even a "language of culture" like English or French, let alone speakers of "primitive" languages like Lahu or Eskimo, are only aware of a small fraction of the systemic relationships among the parts of their bodies. This means that all kinds of idiosyncratic taxonomies of body-parts are possible, so that different languages are apt to construct their systems of body-part nomenclature in quite different ways. For the linguist concerned with questions of universalism vs. particularism, this is a boon. Fourthly, the parts of the body come in a wide variety of shapes, colors, and textures, ranging from the smooth white alabaster of a perfectly rounded breast to the gray, wrinkled and squishy convolutions of the brain. This variety of visual and tactile aspects means that the gates are wide open to analogical associations with objects of all shapes, colors, and textures in the outside world. Breasts look like

cantaloupes or pears, brains look like sponges, eyebrows look like willow-branches, etc.¹⁹⁰ It is impossible to predict *a priori* what inter-field semantic associations will be found in a language's body-part nomenclature. Fifthly, the human body is of paramount interest and importance to human beings. As such, its component parts are explicitly or implicitly invested with cosmological significance. The HEAD may be identified with the noble, positive forces of the universe, while the FEET may symbolize all that is base or earthbound in human nature ("feet of clay"), etc. Here is another fertile source of inter-field semantic associations. [On all this, more below.]

It seems to me that there can be no more worthy long-term goal for Tibeto-Burman studies than the compilation of a Tibeto-Burman Etymological Dictionary arranged on organic semantic principles—that is, to do for Tibeto-Burman what Buck did for Indo-European. That ultimate goal is a long way off. The remainder of this monograph is to be regarded as the prototype of one chapter of our projected "TBED."

2.2 Multidimensional semantic space in body-part nomenclature: classification by association and by opposition.

Body-parts are subject to multiple simultaneous classifications. Each taxonomic decision fixates upon one particular function or characteristic of the body-part as criterial,

ignoring all the others. (In Gestalt terms, one particular aspect of the body-part is taken as FIGURE, while the rest becomes mere GROUND.) We may classify Australia as the largest island in the world—or as the world's smallest continent. In the first case, we fixate upon the circum-aqueous surroundings of Australia, categorizing it in the same group as Greenland, Madagascar, Alcatraz, and Coney Island. In the second case, we fixate upon the vastness of the Australian landmass, categorizing it with Eurasia, North and South America, Africa, and Antarctica. Depending on which "key" we are attuned to, a given black note on the piano may be classified either as G-sharp or A-flat. In exactly the same way, we may classify the LUNGS as belonging either to the /RESPIRATORY SYSTEM/—that is, the same class of objects that includes *nose, pharynx, larynx, trachea, bronchi, diaphragm*—or to the /CIRCULATORY SYSTEM/—the class of objects that includes *heart, arteries, veins, capillaries*. Both classifications are equally valid—the role of the lungs is essential to both systems. Some subsystems of the body bear compound names which testify to their multidimensional nature. The /GENITO-URINARY TRACT/ has some members which are urinary but not genital (e.g. *kidneys*), some members which are genital but not urinary (e.g. *testicles*), and some members which are either, depending on what function is fixated upon at the moment.¹⁹¹ The penis is part of the urinary system when it is discharging

urine, but a "member" of the genital system when it is discharging semen.

Thus far we have been discussing classification in terms of **shared** characteristics. But a taxonomic system may also be defined by **opposed** characteristics. Let us return to the LUNGS for a moment. As we have been [1.141b], there are semantic interconnections in Tibeto-Burman between LUNGS and the notion of lightness, softness, sponginess. (The lungs are remarkably light in weight despite their large size, because of the myriad alveoli or air-sacs with which they are honeycombed.¹⁹²) The key root, "LUNG₁" *-wap, seems to have originally referred to this quality, rather than to the organ itself, as demonstrated by the Jinghpaw forms wóp 'spongy, soft, yielding,' phún-wóp 'softwood,' sìn-wóp 'lungs.' The association between LUNGS and SPONGINESS/LIGHTNESS is by no means confined to TB, and in fact is to be found in Indo-European languages as well. In French, the adjective *mou* 'soft' may be used substantively to mean 'lungs of an animal,' as in *acheter du mou pour les chats* ("buy some lungs [e.g. steer's lungs] for the cats").¹⁹³ And even in English, the word *lights* was once in common use to mean 'the lungs, especially of sheep, pigs, etc.'¹⁹⁴ This usage survives faintly in the expression *liver and lights*, meaning 'the entrails in general,' as in "He ran him through, and liver and lights spewed forth." As a matter of fact, the word *lung* itself is an etymological

doublet of *light*! According to the OED, p. 503, OE *lungen* < PGmc **lung-* < PIE **lngh-* 'light' (cf. Skt. *laghu-*, Gk. ἑλαφρός).¹⁹⁵

With the expression "liver and lights" we approach the heart of the question of defining a system by opposed characteristics. For the liver is everything which the lungs are not! It is a massive organ, the heaviest of the viscera, weighing four times as much as the heart (Wilson, p. 112), and almost half again as much as both lungs put together.¹⁹⁶ (The Hebrew word for liver is *koved*, which is the same as the word for 'heavy.')

Let us now take another look at the Jinghpaw form for 'lung,' sìn-wóp. There seems little doubt that the first syllable is a bound form meaning 'liver,' though the usual free form for 'liver' has the *m*-prefix, məsìn (< PTB **m-sin* [STC 234]). The word for 'lung' therefore means 'spongy liver.' This interpretation is further justified by the existence of another word for 'liver,' the compound sìn-jà?, whose second syllable is an adjectival root meaning 'be hard, solid, unyielding.' As Hanson says (p. 588), the sìn-jà? "is regarded as a more solid gland than the sìn-wóp." Note that the Jinghpaw language simultaneously sets up a double relationship of identification and opposition between the liver and the lungs: they are both kinds of 'sìn'—i.e. "conspicuous internal organ"—but the lungs (in the thoracic cavity) are the "spongy sìn," and the liver (in the abdominal cavity)

is the "solid sin."

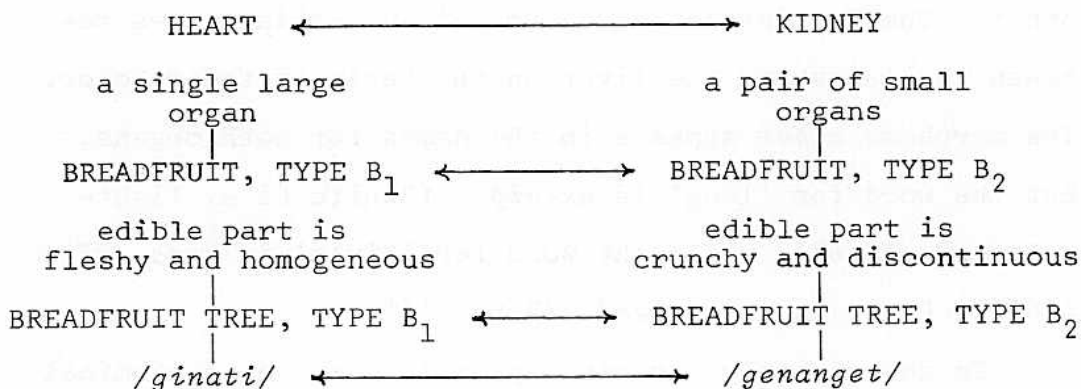
The dialectical opposition between lungs and liver is not predicated solely on their different weights and consistencies, but also on the difference between their colors. The liver, as all hepatophages know, is a dark reddish-brown. The lungs, on the other hand, are a pinkish white at birth, becoming in adult life a bluish gray slaty color, mottled in patches [Gray, p. 975]. It is of course no accident that the English word 'light' may now refer both to weight and to color, even though this is a conflation of two separate PIE roots.¹⁹⁷ The two ideas—small degree of weight/small degree of coloration—mutually reinforce each other. The Turkish language explicitly distinguishes between the lungs and the liver on the basis of their color. The morpheme *ciğer* appears in the names for both organs, but the word for 'lung' is *ak-ciğer* ("white [i.e. light-colored] *ciğer*"), while the word for 'liver' is *kara-ciğer* ("black [i.e. dark-colored] *ciğer*"): ¹⁹⁸

In an extremely interesting article on the anatomical nomenclature of the Gnau people of Papua New Guinea,¹⁹⁹ Gilbert Lewis discusses words which have simultaneous reference both to parts of the body and to objects in the outside world (p. 54). The word *ginati* means 'heart' as a body-part noun, but it is also the name of one of the two large divisions of breadfruit trees, and the name of the fruit that comes from that kind of breadfruit tree

(let's call it B_1). Breadfruit of type B_1 have succulent, edible flesh and no conspicuous stone or pit. Contrasted to *ginati* on many levels is another Gnau word, *genanget*, which means 'kidney' as a body-part noun, but us also the name of the other main class of breadfruit trees (B_2). Breadfruit of type B_2 have inedible flesh, which is thrown away in favor of the many small soft edible nuts nestled inside the flesh. A Lévi-Straussian analysis is irresistible. See Figure 8.

Figure 8

'Heart' and 'kidney' in the Gnau conceptual scheme



The two semantic relationships of association-via-shared-characteristics and association-by-opposition are both exemplified here—entities in the same column are deemed by the language to share a complex cluster of characteristics, which are systematically opposed to the properties of the items in the other column. In this case, Lewis does not

try to decide which semantic field was the "original" referent of the pair *ginati/genanget*. As he says, "the association is clear but its direction not so." The important point is rather that the parts of the human body are rich in metaphorical associations with an unlimited variety of objects and phenomena in the outside world, both vegetable and mineral. Words participate simultaneously in a multiplicity of semantic fields and systems.

The various spatial dimensions and orientations of the body are invested with ideological or cosmological significance in perhaps all cultures. Everybody is aware to some degree of the terrific symbolic importance of the dichotomy between RIGHT and LEFT.²⁰⁰ It would be gauche and maldroit (perhaps even sinister) to dwell upon this here, since the interested reader can be referred to the excellent collection of eighteen articles assembled by Rodney Needham in *Right and Left: Essays on Dual Symbolic Classification* [Needham 1973]. The bilateral symmetry of the body receives explicit linguistic recognition in some languages. In Hebrew, for example, body-parts which come in bilateral pairs do not take a plural affix (-im or -ot), but rather the dual affix -ayim: *ayin* 'eye' / *eynayim* 'eyes,' *ozen* 'ear' / *oznayim* 'ears,' *yad* 'hand' / *yadayim* 'hands,' *regel* 'leg, foot' / *raglayim* 'legs, feet,' etc.^{201, 201-a}

The dimension of FRONT vs. BACK is equally important symbolically, with the positive connotations attaching to

the front, and the negative ones to the back. It is rude to turn your back on somebody if he is speaking to you, even more inexcusable to stick out your backside to him. We place a high value on being "up front." The eyes are in front, and they are the noble "windows of the soul." The back is the sneaky side—it is sinister to have "eyes in back of your head," reprehensible to "talk behind someone's back." We like *pro*-gress, not *re*-gression or *retro*-gression. We like to look forward to things, and we are enjoined to put out best foot—the *right* one, no doubt—forward.

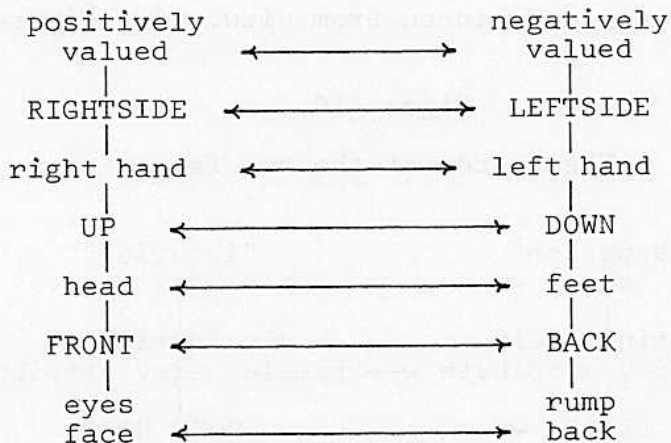
Finally, the vertical dimension of HIGHER vs. LOWER is no less important than the others. The highest part, the HEAD, is the best and noblest, while the FEET (of clay!) are base, in contact with dust and earth. It is a sign of extreme submission to kiss somebody's feet (just as bad in its way as kissing somebody's ass). In Thai etiquette one takes extreme care never to point a foot at anybody, and the ultimate insult is to touch somebody's head with your feet.²⁰² In a beautiful article on the Burmese system of numeral classifiers, Alton L. Becker (1974) convincingly demonstrates that the Burmese classifiers are "a linguistic image of nature," in that the conceptual categories which underlie it have multiple reference both to the human body (especially in its vertical orientation) and to realms of the outside world that are viewed as meta-

phorical reverberations of the body. Thus, the morpheme ʔû means 'head' (*qua* body-part), but also occurs as a polite title (U Nu, U Thant, etc.) and in kin-terms referring to senior relatives (e.g. ʔû-yî 'mother's elder brother'), and is also used as the numeral classifier for worthy or respectable people (like monks, one's elders, etc.)—yahân thôn ʔû 'three monks' ("three headsworth of monks").²⁰³

We may therefore make a diagram of the "good" and "bad" dimensions of the body, as in **Figure 9**:

Figure 9

Symbolic oppositions in spatial orientations of the body

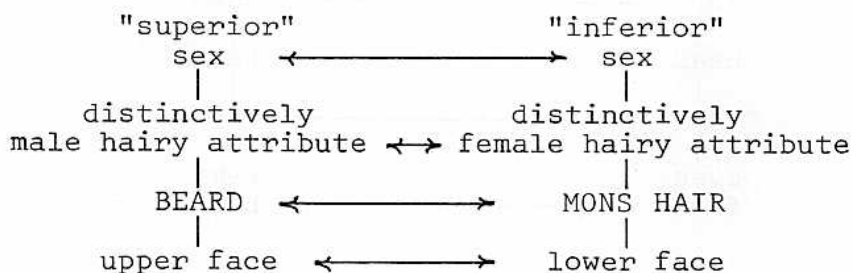


The curious anatomical parallelisms between different areas of the body are packed with tremendous psychological and cosmological significance. Among the "lower animals," the color, shape, and arrangement of the parts of the face often "mimic" the colors and shapes of the organs in the

genital region, constituting a sort of sexual advertisement that has survival value for the species. Thus we find the mandrill (*Papio sphinx*), a ferocious looking baboon of West Africa, with flaming scarlet cheeks that match the vivid crimson of his testes. In Yiddish, the word *bord* 'beard,' in its diminutive form *berdl* 'little beard, nice little beard,' is usable as an affectionate term for the hair on the *mons Veneris*. The beard is the male attribute *par excellence* (ranking at least equal with the penis in Jewish culture). It is situated on a man's face, his "upper face" which is on the front of his noble head. Female pubic hair is situated on a woman's "lower face," a region that is "down" and "back" and hidden from view. See Figure 10:

Figure 10

The beards of the two faces



When a phonetics film on the workings of the larynx is being shown, snickers frequently burst forth from the more irrepressible members of the audience. The flecks of saliva lubricating the lip-like folds of vibrating tissue in the

glottis call forth compelling analogical associations with other areas of the body.

It is no accident that the folds of the vulva are referred to as *labia*, i.e. "lips."²⁰⁴ The *vagina dentata* ("toothed vagina") of many tribal peoples, which expresses the male fear of being harmed by contact with the essence of femininity, carries the metaphor further, supplying a set of teeth to the lower feminine mouth.

In a brilliant essay on Sherpa cultural symbology, "The eyes outnumber the nose two to one," Robert A. Paul (1973) analyzes certain key motifs in Sherpa religious symbolism that owe their vitality to associations with symbolic oppositions between the "upper face" and the "lower face" of a man's body. The eyes are our highest, most intellectual organs of sense. We have two of them, which makes it possible to stand back from things and see them in perspective. The eyes are judgmental (Big Brother is Watching You!); they represent conscious awareness, the Ego, the "Knower" that constantly takes in data from the outside world, focussing now on this and now on that, processing information so that the organism can adapt appropriately to its environment at every instant. In sharp dialectical contrast is the nose, a single protuberant organ below and between the eyes. The sense of smell is our most "animalian" faculty—the one which frequently is more highly developed than sight in the instinct-dominated

lower animals. As Paul says (p. 6):

"Freud went so far as to link the whole passage from nature to culture with the change to upright posture, which shifted emphasis in the area of sexual excitement from the direct smell of the genitals to visual stimuli, [with a] consequent over-valuing of sight and the higher cortex, and the repression of smell in general as a rejection of the transcended past."

The nose represents the passionate, unconscious, subjective, pleasure-seeking, unfocussed, intense, immediate side of our human nature—Freud's *id*, or the "Subject" in Buddhist psychological terminology. The human face presents a metaphorical picture of the triumph of human reason, with the "two conscious eyes victoriously straddling a single and otiose nose" [*ibid.*].

The symbolic meaning of the nose is mightily reinforced by analogy with a similarly salient, protuberant organ in man's "lower face," the penis. It is a cliché of psychoanalysis that the nose frequently stands for the penis in dreams, etc. In various times and places, for a man to have a big nose has been taken as *prima facie* evidence that he was similarly well-endowed down below (cf. *Cyrano de Bergerac*).²⁰⁵ The features of the two faces, in man as in the mandrill, show an uncanny gross-anatomical parallelism. Corresponding to the single nose and the two eyes up above, we have the single penis and the two testicles down below. Corresponding to the mouth, the orifice through which we receive nourishment, is the anus, the orifice through which

we discharge that which contains no nourishment. The cheeks of the face resemble the fleshy buttocks (we even speak of the latter as "cheeks" in English). Facial hair reminds us symbolically of pubic hair.²⁰⁶ Despite this overall parallelism, however, the internal dominance-relationships among the features of the upper and lower faces are diametrically opposite. In our upper face it is the paired organs—the eyes—which are dominant; in our lower face it is the single protuberant organ—the penis—which is dominant. See Figure 11.

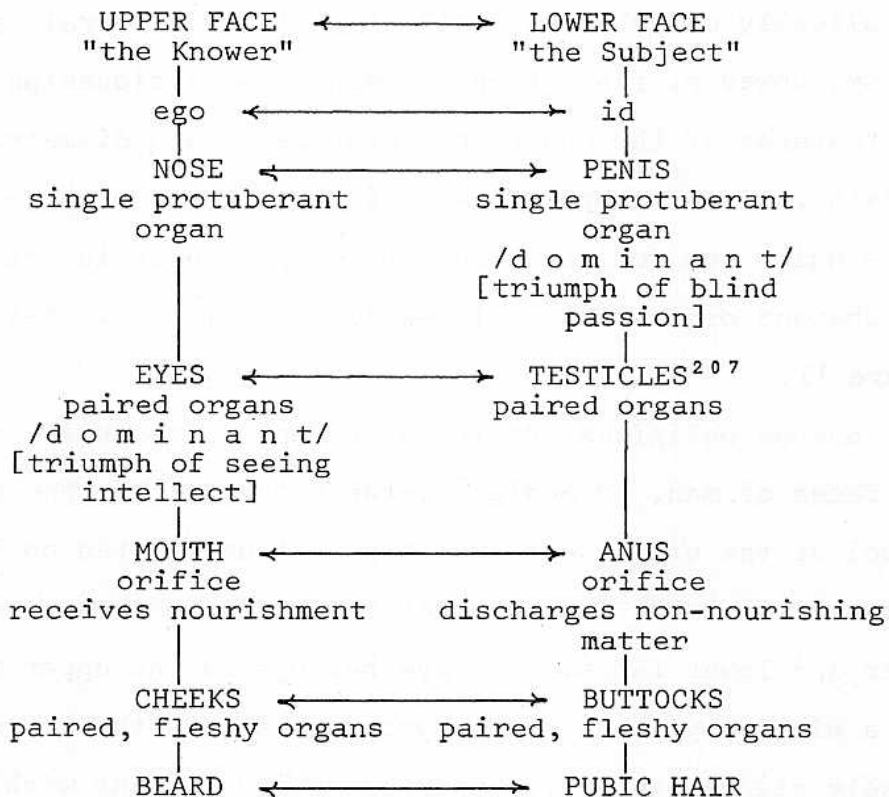
Sherpa religious art is an attempt to reconcile the two faces of man, to mediate between them.^{206-a} The striking symbol of the single eye, which one finds painted on Sherpa stupas or *chorten*, is a sort of compromise between the upper and lower faces. The eye belongs to the upper face, but a single eye "has no perspective, and is therefore more closely tied to subjective purposive involvement with what it sees" (Paul, p. 4).

Widespread use is made of "imagery employing the formal structure of a protuberant figure flanked by a pair of smaller items at its base" (p. 8), with metaphorical connections to the upper and the lower "faces" simultaneously:

"The most important set of offerings on a Sherpa altar...are a large *torma*, a phallic cake of dough representing the major god of the ceremony, with a bowl of fluid on either side.... The main gods in Sherpa temples are three, represented by a large central statue...flanked by two smaller statues...." (*ibid.*)

Figure 11

Symbolic associations and oppositions between the features
of the upper and lower faces



Sherpa religious philosophy, in its most profound acceptance, does not advocate obliteration of man's "lower," instinctive nature in favor of his "higher" faculties:

"Because these 'lower' forces will not go away, ultimate salvation cannot really be achieved by attempting to separate out good from evil, or Reason from Passion. True enlightenment involves integrating both into a higher order synthesis in which the Knower and the Subject become united. In Buddha-knowledge, the limited ego is overthrown and the new conscious awareness, no longer based solely on

sense perception and reason, fully coincides with the true center of one's being."
(p. 11)

2.3 Medical knowledge and "folk usage" of body-part terms

When we read about the quaint medical beliefs of "pre-scientific" peoples, we are inclined to smile. Thus, the Gnau of New Guinea do not think of blood as a tissue or part, but as a **product** of the body. "It lies under the skin, making it smooth and shining. The people list blood not with parts of the body but among its products, with sweat, tears, etc." (Lewis, p. 55). The Tibetans traditionally believed that both the trachea (windpipe) and the esophagus (gullet) were for the passage of food, with the windpipe used for liquid food and the gullet for solid food (Jäschke 1881). The Tibetan word lkog-ma 'gullet,' has been attested in the meanings 'windpipe,' 'throat,' 'larynx,' and 'neck' (Jäschke, p. 18), which indicates a certain lack of conceptual precision in that area.

Actually of all the TB peoples, the Tibetans had the most highly developed medical science, having been exposed to Sanskrit medical treatises with the introduction of Indian Buddhist learning to Tibet. Along with many valid concepts in anatomy and pharmacology, the Tibetans inherited a number of erroneous notions from Indian medicine. Thus, the Hindu physicians believed that the human fetus was formed through the combination of a man's semen and a woman's

uterine blood. Tibetan medical terminology had a name for this reproductive mixture, k'u-k'rag (k'u- 'fluid, semen,' k'rag 'blood;'; Jäschke, p. 40). As in other medical traditions, Tibetan anatomical classifications revolved around certain "magic numbers," whereby body-parts deemed to form coherent systems were grouped together into classes with a specific number of members.²⁰⁸ Thus we have the don lna, or FIVE MAIN ORGANS (*heart, lungs, liver, spleen, kidneys*; Jäschke, p. 259); and the snod drug, or SIX VESSELS (*gall bladder, stomach, small intestine, large intestine, urinary bladder, seminal vesicles* [man] or *uterus* [woman]; Jäschke, p. 319). The FIVE ORGANS and SIX VESSELS collectively were known as the don-snod.²⁰⁹ Similarly, the body was deemed to have exactly 900 c'u-rgyus, which Jäschke (p. 159) glosses as 'sinews, ligaments, and nerves,' adding "with respect to these, as well as to the veins, Tibetan science seems to be rather in the dark."²¹⁰ Under the entry rtsa 'vein, artery, pulse' (p. 436), Jäschke expresses his frustration in trying to find adequate English translations for the traditional Tibetan terminology in the whole area of blood vessels, nerves and sinews:

"Owing to the imperfect state of Indian and Tibetan anatomy, resulting from inveterate prejudices both of a religious and intellectual nature, great confusion prevails also in the department of angiology [blood vessels and lymphatics], many different vessels of the human body, and even part of the nerves, being classed among the veins, so that it is impossible to find adequate terms for the Tibetan nomenclature."

Nevertheless, if we were to compare the knowledge that "pre-scientific" peoples have of their bodies with the medical knowledge that the average layman in a "highly advanced technological society" has, it is by no means certain that the modern son of science would come out ahead! Suppose we were to stop a man in the street in Modesto, California, and explore his medical knowledge by asking him a number of questions about human anatomy and physiology, then comparing his answers with those of some naked savage in the Amazon basin.²¹¹ From our theoretical armchair, we may distinguish roughly among four kinds of medical knowledge:

(a) **Gross exterior anatomy and obvious physiological facts.**

Both the man from Modesto and our savage from Brazil would predictably score equally well in this area. Both would recognize the parallelism in structure between the upper and lower limbs—the fingers are like the toes, the wrist is like the ankle, the elbow is like the knee. We see with our eyes, hear with our ears, smell with our noses. We get two sets of teeth in our lifetimes. Certain organs serve multiple functions (e.g. the penis). The body is bilaterally symmetrical. Our Adam's apples move when we swallow. Skin turns red when it is slightly burned, blisters when it is more severely burned, and turns black when it is burned to a crisp²¹²....

(b) **Basic physiological facts** that are not accessible to direct observation without special instruments. This is the only area where the man from Modesto is likely to score significantly better than his pre-scientific counterpart. In this category we include basic information that the man in the street could not have found out for himself, but which is more or less "common knowledge" for anyone with a minimum of education in a technological society: the role of the lungs in respiration, the fact that the blood circulates and that the heart is like a pump, that conception occurs when the male sperm fertilizes the female ovum, that the kidneys have something to do with urination, etc. These are facts which have "filtered down" from the medically sophisticated elite until even most laymen know them. Yet we must not take anything for granted—even such basic phenomena as the role of the lungs in the circulatory system are likely to be equally mysterious to our two subjects (see category [d] below).

(c) **Gross internal anatomy.** This is an area where the pre-scientific man is likely to have much more accurate and organized information than his civilized counterpart! In primitive societies people do their own butchering of animals, gaining a good practical knowledge of the 'innards' in the process:

"A considerable familiarity with [internal parts of the body], gained through the dis-



Lahu men butchering a pig carcass (1965). The photo on the reverse of this page was taken moments later.



Indians looking at a dead shark on the beach, 1934.

section of animals for food or sacrifice, is common among primitive peoples and is not surprising for the Proto Indo-European period." [Buck, p. 197]

The Gnau used to dismember the corpses of their parents to get out certain bones which they then kept in their houses (Lewis, p. 51). The "primitive" man knows dozens of facts about the size, shape, color, consistency, and relative position of the principal internal organs, which his civilized cousin would be hard put to come up with. Meat for most Americans is something that comes neatly wrapped and packaged in the supermarket. Chickens have gaping holes where their internal organs should be.

Of course even in primitive societies there are differences in medical knowledge between individuals and social groups. Among the Gnau, it is paradoxically the men who have a better understanding of the womb and placenta than the women (Lewis, pp. 64-5). The women believe that the womb and the placenta are the same thing, and call them both by the word *gelugi*.²¹³ Only women are allowed to handle a human placenta—they therefore believe that the womb/placenta is expelled from the body after each birth (since they have seen it with their own eyes!). The men, on the other hand, refer to the placenta as *gungi wolit* 'blood lake,' distinguishing it from *gelugi* 'womb.' This is because they know from comparing pregnant and unpregnant sows that the *gelugi* remains within the sow after the piglets are born, and that it is only the "blood lake" which is

lost at birth.

(d) **Subtle internal anatomy and physiology.** When it comes to really detailed internal anatomy, or physiological phenomena that involve a sophisticated knowledge of body chemistry, our man from Modesto is no better off than the man in the jungle. If we ask the average American, "What are the functions of the liver?", we will be met with a blank stare. We will **not** get an answer like the following:

"The liver is responsible for the final stage in the treatment of proteins, carbohydrates, and fats. It breaks down proteins into their constituent amino acids, converting any unneeded amino acids into other substances by removing their nitrogen, then sending them to the kidneys for excretion. It breaks down carbohydrates to the simple sugar glucose, then converting the glucose into the starch glycogen, which it stores until the body needs it, when it is converted back into glucose and released into the blood. It also stores fats temporarily, and iron for hemoglobin. In addition, the liver manufactures the proteins fibrinogen, prothrombin, albumin, and globulin. It makes heparin, a mucopolysaccharide acid that keeps the blood from clotting; and bile, stored in the gall bladder and poured into the small intestine to help in the digestion of fats...."214

We may sum up our impressionistic comparison of the medical knowledge of pre-scientific vs. technological man as in Figure 12.

In pre-scientific societies, the combination of a sophistication with respect to gross internal anatomy (the

Figure 12

Medical knowledge of "primitive" vs. "modern" man

	I	II	III	IV
	Gross exterior anatomy/Obvious physiological facts	Basic, but not directly observable physiology	Gross internal anatomy	Subtle internal anatomy and physiology
"PRIMITIVES"	+	-	+	-
"MODERNS"	+	±	-	-

plus in column III) with a naivete with respect to physiological function (the minuses in columns II and IV), favors the naming of the internal organs by morphemes referring to their size, shape, texture, or color, rather than to their (unknown) function: the LUNGS are the "spongies," not the "air-blowers" or the "bellows."²¹⁵ The KIDNEYS are the "breadfruit nuts," not the "blood-cleaners" or the "piss-makers." In the same way, ignorance of the true physiological functions of an organ leaves the way open to the adventitious symbolic associations we have been discussing. Since the true functions of the LIVER are unknown to primitive peoples, they are free to invest it with fanciful ones—since the liver is so big, and in the center of the viscera, and obviously must serve some important function, why then it must be the SEAT OF THE EMOTIONS!

Words serve the needs of their users. For an Eskimo, it might be important to distinguish 87 different kinds of

snow and ice. For the linguist it is well to know the differences among a reflex, a retroflex, and a circumflex. For the cardiac surgeon it is important to have separate words for the aorta, the mitral valve, and the vena cava. In ordinary language we usually have no particular need for terminological precision in specifying parts of our bodies. If we have a pain in the "gut," we can say either, "I've got a belly-ache" or "I have a stomach-ache" (when really the source of our pain might be in the duodenum). If you ask somebody what part of his body he kicks with, he might answer either "My foot, you dummy!" or "My leg, you dummy!" For most purposes we use "arm" whether we mean the stretch from elbow to wrist (the "distal segment" or "forearm"), or the stretch from shoulder to elbow (the "proximal segment" or "upper arm"). Human beings are loath to be more precise than they have to in a given situation. Suppose you nick one of your fingers while chopping onions. You say to your friend, "I hurt my finger—get me a band-aid." At this point it would be highly inappropriate for your friend to ask, "Which finger?" The answer would be, "What do you care, dummy? Just get me the goddamn band-aid!" On the other hand (!), if a surgeon is about to amputate somebody's finger, he had better be sure which finger it is, or he will be open to a malpractice suit.

2.4 Semantic vagueness and semantic shifts: classification and diagrammatic representation. We are interested in the related notions of semantic vagueness and semantic shifting.²¹⁶ We wish to regard the great Tibeto-Burman proto-language as a real human language, spoken by real people. Therefore we are compelled to recognize that already in the proto-language the meanings of words were not fixed and immutable entities, but rather vague, and prone to form new and unpredictable associations with other atoms of semantic stuff. It is often hard to draw the line between a "vagueness of meaning" and a "shift of meaning." In the abstract, however, we may roughly diagram the difference as follows (where M means 'unit of meaning'):

VAGUENESS. An etymon E is vague in a given language at a given point in time if it has ranges of meaning like

$M_1 ; M_1 + M_2 ; M_2 \dots$

SHIFT. An etymon E has shifted in meaning through time if at one stage in the language's history, S_1 , it used to mean " M_1 but not M_2 ," while at a later stage in the language's history, S_2 , it came to mean " M_2 but not M_1 ."

When we find phonological cognates in Language A and Language B, but the meaning of the form in Language A is glossed as "vein" in the best available source, but the meaning in Language B is glossed as "sinew," we have several options open to us:

- (a) We can insist on "exact" correspondence in gloss between the two sources before we admit the possibility of the forms being true cognates at all. In its extreme form, this attitude has been succinctly characterized by Paul K. Benedict as "not knowing enough to go from 'dung' to 'excrement.'" ²¹⁷ Thus, Roy Andrew Miller (1974, p. 208) criticizes Benedict for identifying Chinese hsin 薪 as cognate with the TB root *sin *sik 'tree, wood, firewood,' making the erroneous claim that the Chinese word means 'fuel' and not 'firewood.' In fact, hsin (< Ancient Chinese *siĕn < Archaic Chinese *siĕn) has meant 'firewood' (and **not** 'fuel in general') from the earliest Chinese classical texts to the present (see Karlgren, GSR, p. 109, No. 382n, quoting the *Shi King* as authority).
- (b) We can decide that there was a semantic vagueness already in the proto-language, so that the proto-form should be glossed something like 'vein, sinew' or 'elongated internal tissue of non-apparent function inside the body.' In this case, we would also suspect that the dictionary glosses for the forms in the modern languages are too specific.
- (c) We can decide that the forms in question are indeed etymologically related, but that there has been a real change in meaning from the proto-stage to the later stage. Thus we know that Late Latin *testa* 'pot' changed its meaning in Gallo-Romance, so that now in modern French its

etymological descendant, *tête*, means 'head' and not 'pot.'

But the distinction between decisions (b) and (c) is very tenuous—there must have been a time in Gallo-Romance when the prototype of *tête* meant both 'pot' and 'head,' so that a contemporary dictionary might have had an entry like:

- teste* (n) 1. a pot, as for cooking
2. (slang, vulgar) the head.

A later stage might have been glossed as follows:

- testa* (n) 1. the head
2. (archaic) a pot.

The important thing is that the meanings POT and HEAD have been **associated** in the history of Romance.²¹⁸

The paths that semantic association may take are wildly unpredictable [see above, 1.122b]. As Buck says:

"The associations underlying semantic changes are so complex that no rigid classification...is possible. Many changes may be variously viewed. In a sense, each word has its individual semantic history." [*op. cit.*, p. vi]

In Gnau, the word *basyilape* means both 'grasshopper, cricket' and 'fontanelle' (the "soft spots" on a baby's skull). The association is via the 'cricket-like' jumping pulse of the fontanelle's membrane when the infant coughs or sneezes—it is the mother's duty to try to cover the fontanelle at such times, lest harm befall the baby [Lewis, p. 53].

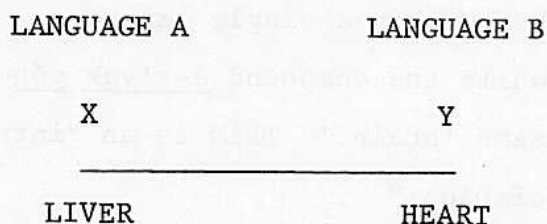
Despite the difficulty of the problem, students of

language have been trying for centuries to impose some sort of conceptual framework on it, as evidenced by the existence of a full arsenal of rhetorician's terms like *metonymy* and *synecdoche* for characterizing different types of semantic association and shift. Any such attempt at classification, however, will still have to reckon with a large "miscellaneous" category.

For our purposes—putting into relief the web of semantic interassociations in the area of Tibeto-Burman body-part nomenclature—we can operate simply with two kinds of semantic association. To these we give the names intra-field association and extra-, inter-, or trans-field association.²¹⁹ To use these terms, even roughly, we must first define which semantic field we are taking as a point of reference. /PARTS OF THE BODY/ is an exceptionally well-defined semantic field, but even here we must deal with fuzzy edges. What do we do with words like *wing*, *tail*, *egg*, *fin*—i.e., words for parts of non-human animal bodies? Do we include in our field words for bodily secretions or products like *sweat*, *spittle*, *urine*, *tears*? How about diseases and pathological conditions, like *goiter*, *leprosy*, *rash*, *fever*? What about verbs of bodily function, like *to vomit*, *to breathe*, *to hear*, *to shiver*? Once we have defined the field to our satisfaction—e.g. /PARTS AND PRODUCTS OF THE HUMAN BODY, NOT INCLUDING BODY-FUNCTIONS OR DISEASES/—we can then look at the pattern of semantic associations involv-

ing the field, and decide which of them are "intra-field" and which are "trans-field."

Suppose we notice that Language A has a word meaning 'liver' (e.g. Lepcha a-byet) while Language B has a phonological cognate that means 'heart' (e.g. Limbu naribet). Since LIVER and HEART are both parts of the human body, this is an intra-field association, which we may diagram by a straight line connecting the two points in semantic space:



[X and Y are in the same semantic field]

Suppose, on the other hand, that Language A has a word meaning 'stomach' (e.g. Tibetan p'o-ba) while Language B has a phonological cognate that means 'cave' (e.g. Jinghpaw lùn-pū). Since the latter concept refers to an inanimate physical object, and not to a body-part, this is a trans-field association, which we may diagram by a looped line connecting the two points in semantic space:

LANGUAGE A

LANGUAGE B

X

Y

a

STOMACH

cave

[X and Y are in different semantic fields]

In both of the above cases, the associations were established inter-lingually, on the basis of data from more than one related language. Associations may equally well be established within a single language. Lepcha a-yón means 'marrow,' while the compound a-t'yak yón (literally "head-marrow") means 'brain.' This is an "intra-lingual intra-field association:"

LANGUAGE A

X

Y

BRAIN

MARROW

In Gnau, as we have seen, the body-part *fontanelle* is associated with the trans-field concept *cricket*, *grass-hopper*. This is an "intra-lingual trans-field association:"

LANGUAGE A

X

Y

a

FONTANELLE

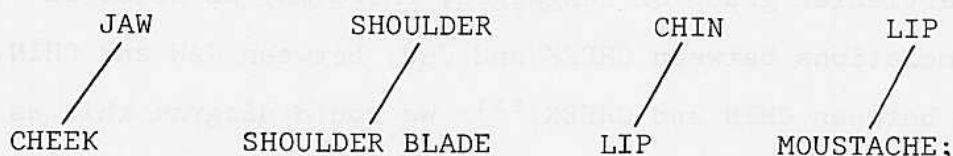
cricket/grasshopper

With these basic categories of semantic associations in

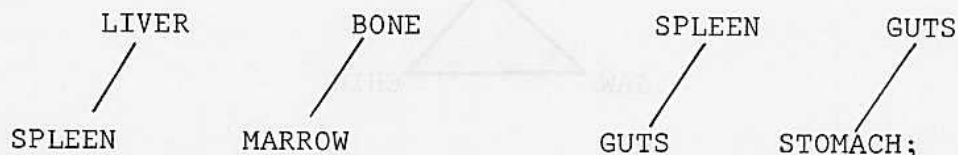
mind, we may proceed to make finer subclassifications, some of which have specific relevance to body-part nomenclature.

2.41 Subtypes of intra-field semantic associations in body-part terms.

(a) Association via anatomical adjacency. Just as the rot may spread from apple to adjacent apple in the barrel, just as each successive falling domino knocks over its neighbor in the row, just as the crazed cells in a neoplasm subvert the functions of adjacent cells, just as the pure radiance of a noble soul inspires good feelings in everybody else in the room, so do words for body-parts move associatively from part to adjacent part of the human anatomy. The association may be from place to place on the exterior of the body:



or from place to place inside the body:

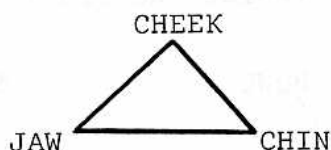


or from a place on the exterior to an adjacent place on the interior:

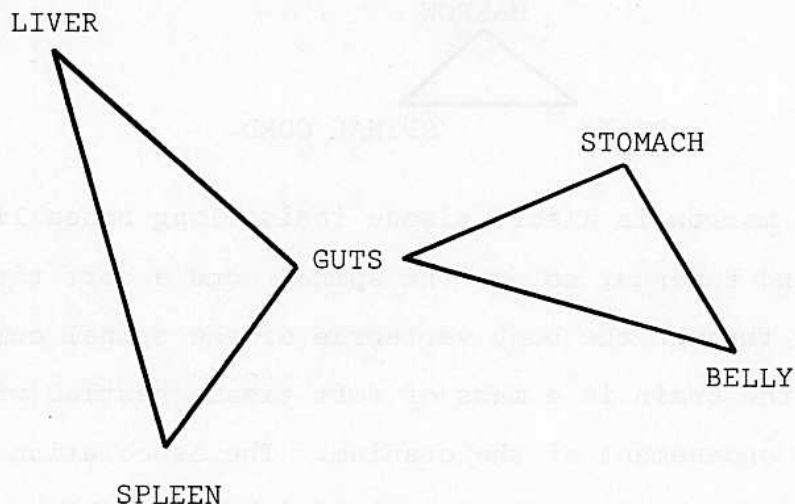


In this type of semantic association, there is not necessarily any shared function between the pairs of adjacent body-parts beyond their "geographical" contiguity. A good name for this phenomenon is *metastasis*, a "standing beyond" or displacement of the meaning from point to point in somatic semantic space.

When considering more than two body-parts, it will usually be the case that a given body-part is involved in a network of "geographical" associations with more than one other N_{bp} at once.²²⁰ Thus, within a single language, or in a particular group of languages, there may be attested associations between CHEEK and JAW, between JAW and CHIN, and between CHIN and CHEEK.²²¹ We could diagram this as a triangle:

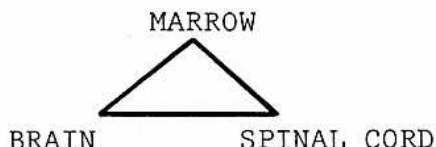


As a slightly more complicated example, the associations LIVER/GUTS, GUTS/SPLEEN, LIVER/SPLEEN, GUTS/STOMACH, STOMACH/BELLY, and BELLY/GUTS may all be represented simultaneously by a geometrical figure like the following:



(b) **Association via shared physiological function.** This category is hard to distinguish from the preceding one, since many anatomically contiguous body-parts also share a systemic physiological function: STOMACH/GUTS, PENIS/TESTICLES, WOMB/PLACENTA, BRAIN/SPINAL CORD, LIVER/GALL BLADDER. We may include here associations between a body-part and a body-product, like EYE/TEAR, ANUS/SHIT, PENIS/SEMEN, NOSE/SNOT, or between a bodily activity and a body product, like FART/SHIT, URINATE/URINE, SPIT (v.)/SPITTLE.²²²

(c) **Association via similarity of relative position.** In this category we include association like FINGER/TOE, WRIST/ANKLE, ELBOW/KNEE—and the following fascinating triad:



Just as the marrow is a soft tissue inside long bones like the tibia and humerus, so is the spinal cord a soft tissue running through the bony vertebrae of the spinal column, while the brain is a mass of soft tissue nestled within the bony encasement of the cranium. The association between BRAIN and MARROW is well-established in Tibeto-Burman (see *r-klīn, STC No. 126), and in fact is found all over the world. In Gnau, the spinal cord is identified with bone marrow (*tepit*), and both are thought to be similar to the brain (Lewis, p. 59). In French, *moelle* means 'marrow,' while *moelle épinière* ("spinal marrow") means 'spinal cord.' The Latin ancestor of the French form, *medulla*, is used in medical terminology not only for the bone marrow, but also for the soft marrow-like center of other organs, like the kidney or adrenal glands, and even for part of the brain itself, the *medulla oblongata* (the lowest or hindmost part of the brain, continuous with the spinal cord). The concepts *brain* and *marrow* have in fact been closely associated throughout the history of Indo-European. Aristotle (quoted in Buck, p. 215), observes in his *Posterior Analytics* that "many think the brain is really marrow."²²³ Words like Russian *mozg* 'brain' are cognate to

forms like English *marrow*.

(d) **Association via similarity in shape, color, or general appearance.** This is the chief type of *trans*-field association [see below, 2.42], but it also operates within a given semantic field. Certain parts of the body are not adjacent to each other, nor are they participants in any common physiological system, yet their general shape or appearance is such that an association may be made on that basis alone. Throughout TB, and in Austroasiatic and Austronesian as well, there is an association between EYE and ANKLE, with the anklebone referred to by compounds meaning "eye of the foot, foot-eye" [above 1.122c].

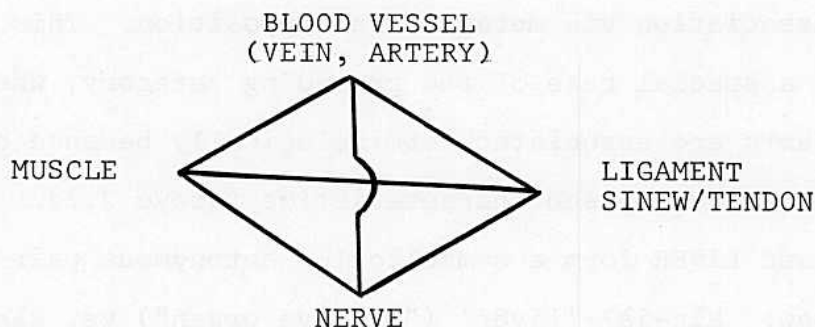
The MARROW participates in associations both with FAT and BLOOD in TB—for the excellent reason that the marrow in some bones is yellow and largely composed of fatty tissue, while the marrow in other bones is red and bloody-looking. As Gray's *Anatomy* puts it (pp. 1096-7):

"In the shafts of adult long bones the marrow is of a *yellow* color, and contains, in 100 parts, 96 of fat, 1 of areolar tissue and vessels, and 3 of fluid.... In the flat and short bones, in the articular ends of the long bones, in the bodies of the vertebrae, in the cranial diploë, and in the sternum and ribs, it is of a *red* color, and contains, in 100 parts, 75 of water and 25 of solid matter ...and only a small proportion of fat...."

The MARROW/FAT association may be established on the basis of the widespread TB root **tsil* [STC pp. 16, 168, 169, 173],

which in some languages means 'fat' (WT ts'il), and in other languages means 'marrow' (Kanauri tsil). Burmese displays both meanings, with chi, ʔəchi 'oil, fat,' alongside khraŋ-chi 'marrow.' As for MARROW/BLOOD, I have proposed a new Sino-Tibetan etymology,²²⁴ whereby the general TB etymon for 'blood,' *s-hwiy [STC No. 222], is not identified directly with Chinese *xiwet/xiwet 'blood' [GSR 410a-c] 血, but rather with Chinese *swia/swig: 'marrow' [GSR 11h] 骨髓 or 髓. A key form in this connection is Jinghpaw ləsəwi 'marrow' [Hanson, p. 380], as opposed to Jg. sài 'blood.'^{225,226}

A gross similarity in appearance, along with an ignorance of their real functions, has caused a terminological "under-differentiation" in TB words for blood vessels, sinews and ligaments, nerves, and muscles. Mikir (Walker 1925) is typical, with the single word artho glossed as 'nerve, sinew, vein, muscle, ligament.' These "string-like" body-parts, to which we may give the name "obscure internal channels and connections," or OICC's for short, occupy for speakers of many TB languages closely adjacent or identical points in semantic space. Speakers of English may be a bit hazy about the difference between a ligament and a tendon or sinew, or between a vein and an artery, but at least it is safe to say that the OICC's are spread over a considerably wider semantic space in English than in Mikir:



(e) Association via shared symbolic function. As we have seen [above 2.2], parts of the body may acquire all kinds of metaphorical connotations, functioning together as common participants in symbolic systems even though they may have no particular anatomical or physiological connection with each other. Thus, the HEART and the LIVER have both been regarded as seats of the emotions in TB, and comparative study reveals many cases where a given phonological etymon means both 'heart' and 'liver' in a given language (e.g. Jinghpaw səlum 'heart; by some the liver or biliary gland,' Hanson p. 609), or else 'heart' in some languages and 'liver' in others.²²⁷ As an example of the semantic interconnections among HEART/LIVER/SEAT OF EMOTIONS, I offer the following new TB etymology:

*b-yet * m-yet. Lepcha a-byet 'liver,'²²⁸ Limbu naribet 'heart,' Tamang Risiangku net 'liver,' Jinghpaw myit 'mind, seat of emotions,' Nung mit ~ nit 'mind, temper,' WT yid 'soul, mind.'

Maru mjit 'think, consult, deliberate'

(f) **Association via metaphorical opposition.** This is really a special case of the preceding category, whereby body-parts are associated terminologically because of dialectically opposed characteristics [above 2.2]. Thus LUNGS and LIVER form a symbolically antonymous pair in Jinghpaw: sìn-jàʔ 'liver' ("massive organ") vs. sìn-wóp 'lungs' ("spongy organ") [*ibid.* See below 2.5(4)].²²⁹

(g) **Association via synecdoche.** Many body-part terms have flexible ranges of meaning, whereby they are usable alternatively either for a more specific or a more general part of the body, either for a part of a given substructure or for the whole substructure.

If asked for a careful definition of ARM vs. HAND, or of LEG vs. FOOT, most English speakers would say that "arm" and "leg" refer to the whole limb, or at any rate from shoulder to wrist or from hip to ankle, while "hand" and "foot" refer specifically to the distal segment of the limb, the part with the fingers or toes. Yet many speakers (including me) often use "hand" (or "foot") when the whole limb is meant—or even if a particular area of the forearm or upper arm (or calf or thigh) is referred to.²³⁰ In TB also, the root *g-lak [STC No. 86] may be glossed either as 'hand' or 'arm,' and *r-kan [STC, pp. 70, 142] as either 'leg' or 'foot.' Emeneau 1974b ("Arm and leg in the Indian linguistic area") points out similar synecdochic interplay

in Indian languages. In Gnau (Lewis, p. 58), the word for 'bone (in general)' may also mean specifically the tibia or shin-bone; the general Gnau word for 'tendon, vein, artery' may mean specifically the hamstrings.²³¹ In the history of German the word for 'knuckle,' *Knochen*, has come to acquire the meaning 'bone in general,' while its diminutive, *Knöchel*, now means 'knuckle;' meanwhile the older Germanic word for 'bone in general' has taken on the specific meaning of 'leg' (*Bein*) [see note 230 and Buck 1949, p. 207].²³²

It is convenient to refer to these processes of semantic generalization (part > whole, specific > general) and specialization (whole > part, general > specific) by the ancient rhetorical term *synecdoche*.

(h) **Association via euphemism or taboo.** Euphemistic avoidance of the plain, unvarnished words for body-parts of sexual or scatological purport is really a cause or motivation of semantic association, rather than a categorial subtype of semantic association itself. Through euphemism, the "shady body-part" is associated with a neutral, anodyne body-part that is anatomically adjacent to it [category (a), above], or indirect and general rather than specific [category (g), above]. Some English speakers (mostly women) do not say "breast" in polite company, but rather the sexually neutral and more general "chest" or "bust," even

when the mammary glands are specifically meant ("Her chest was, you know, kind of small, so she couldn't nurse her baby."²³³)

The use of "stomach" instead of "belly" seems originally to have been a genteelism in English, with "belly" carrying more gross, gluttonous connotations. Everybody is familiar with Victorian euphemisms like the general "limb" for the more specific "leg," or the fatuous compound "gentleman cow" for the too potent and sexual word "bull."

(i) **Association via figs.** By this facetious category label I mean to comprise all the miscellaneous and unpredictable semantic associations that resist classification. The original PIE word for LIVER **y^ǵek^wr/n-* (> Gk. ἥπαρ, ἥπατος, Lat. *iecur*) was replaced in many daughter languages by words connected with the liver as an article of food. In Romance, the practice of stuffing live animals with dried figs to induce liver growth led to the replacement of *iecur* by **fīcātu(m)* "figged" (p. ppl.) < Lat. *fīcus* 'fig,' whence the modern Romance forms (French *foie*, Italian *fegato*, Spanish *hígado*, etc.). [See Buck, pp. 251-252, and Malkiel 1944].

2.42 Trans-field semantic associations of body-part terms.

Almost all trans-field associations involving N_{bp} 's are based upon some similarity in shape, color, texture, or

general appearance between a part of the human body and an object elsewhere in the animal, vegetable, mineral, or manufactured world [see category (d), above].²³⁴

When faced with a trans-field association involving a N_{bp} , which we may symbolize as

$$N_{bp} \xrightarrow{\quad} X,$$

it is always interesting to try to determine the directionality of the association. Was it the body-part term which extended its meaning to the extra-somatic world, or vice versa? Diagrammatically we may distinguish these relationships by pointing an arrowhead either away from the N_{bp} to the outside, or from the outside to the N_{bp} :

$$N_{bp} \xrightarrow{\quad} X$$

$$N_{bp} \xleftarrow{\quad} X.$$

Sometimes it is clear that the word in question was originally a N_{bp} , which later got extended to the outside world. In English, we feel sure that *mouth*, *foot*, *neck*, *tongue*, and *kidney* applied to body-parts before they applied to rivers, mountains, bottles, shoes, or beans (respectively):

$$\begin{array}{ccc} \text{tongue} & \xrightarrow{\quad} & \text{tongue} \\ \text{[of body]} & & \text{[of shoe].} \\ N_{bp} & & \end{array}$$

Similarly in Gnau, Lewis (p. 54) feels sure that *nembalpe* meant 'liver' originally, and only secondarily came to refer

to 'certain begonias with rough liver-red leaves;' *nemenit* first meant 'nose' and only then got extended to mean 'mountain ridge or promontory.'

Just as often, the evidence may point clearly in the other direction, so that we are sure that a given word started out as a member of some other semantic field, but got extended in applicability until it could refer to a N_{bp} as well. The names of the three important bones of the middle ear, the "hammer" (*malleus*), the "anvil" (*incus*), and the "stirrup" (*stapes*), referred originally to artifacts and have only secondarily been applied to body-parts.²³⁵ The word *appendix* originally meant 'anything added onto something else,' and not necessarily the functionally mysterious vermiform appendix. The words *back* and *side* originally had more general spatial meaning, and only secondarily came to refer to parts of the body (Buck 4.19, 12.36). In Gnau, Lewis (pp. 53-54) is sometimes sure that a body-part word had an extra-somatic origin: *wangen* meant 'bowstring' before coming to mean 'Achilles tendon;' *dupe* meant 'kind of large bamboo' before coming to mean 'windpipe;' *dyuelbi* meant 'kind of long puckered scented leaf' before coming to mean 'psoas muscle.'²³⁶

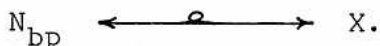
In the case of words for "shady body-parts" the motivation for introducing terms from non-somatic semantic areas is sometimes euphemistic [see 2.41(h), above], but sometimes the effect is vulgar. In Lepcha the penis may be

called a-ṇol, from an adjectival root meaning 'be ringed, surrounded by a ring, as an ornamental pillar.' Mainwaring 1898 characterizes this usage as obscene (p. 74). The testicles may be referred to as eggs, balls, nuts, or stones, with the stylistic value of these names differing from language to language. The word *testis* itself seems to be related to the word for 'witness,' perhaps because they lend "supporting testimony" to the penis (see Buck, p. 257).²³⁷ The word *vagina* originally meant 'sheath' in Latin.²³⁸ By implication, therefore, the penis is a 'sword,' or at any rate some kind of 'weapon' (cf. Old English *wāepēn* 'weapon; penis,' *waepned* 'male' (i.e. "furnished with a weapon").

A special subtype of intrusion into body-part nomenclature from outside is the use of words for parts of the bodies of lower animals to refer to corresponding or similar-looking parts of the human anatomy. Sometimes the use of the animalian word is humorous and good-natured, like the slang English use of *wing* for 'arm, hand.' This same association is found in Sino-Tibetan. The STC (p. 171) identifies PTB *g-lak 'arm, hand' with Chinese *djək/iək 羽翼 'wing.'²³⁹ Often the use of the animal-word is contemptuous or insulting: "Get your filthy **paws** off me!;" "Take your **snout** out of that cookie-jar!;" *ix'l dir gebm a frask in pisk* (Yiddish) 'I'll give you a slap in the muzzle!' A particularly versatile animal-part is TAIL,

which is a common euphemism or vulgarism for "penis" in the world's languages (German *Schwanz*, Yiddish *veydl*, etc.),²⁴⁰ but which can also mean 'buttocks, rear end' ("Get off your tail and get to work!"), and by further extension can even be applied to the female pudenda (especially when used with the classifier 'piece').

Sometimes, however, when the historical linguist confronts a trans-field semantic association, he may not be able to determine which is the "original" sense and which is the derived one. Lewis (p. 54) mentions several cases of this kind in Gnau: *nulape* means either 'scapula' or 'taro-like plant with a few large graceful blade-shaped leaves upon a long slender stem;' *labunat* means either 'thigh, femur' or 'the magnificent rifle bird;' *lauwutag* means either 'labia majora' or 'the loose ridge of skin on the back of a pig's hock.' Associations like these we may diagram with a double arrow:



It can easily happen that a native speaker of a language may have a different notion of the directionality of a semantic association than that held by a historical linguist with specialized knowledge. If we ask the man in the street, "Which is the more basic sense of the word *green*—the color in general, or the condition of unripe fruit?", he would be almost certain to answer that the color-designa-

tion is primary. Yet the Indo-Europeanist knows that the color word is itself derived from the root 'to grow' (the word *grass* also belongs to this word-family) and originally referred to the color of growing vegetation [Buck, pp. vii-viii, 1059].

Finally, we should mention one trans-field association that does not go very far "afield" from body-parts:

body-part ————— ^a ————— body-function.²⁴¹

Thus, English 'ear' is ultimately related to the verb 'hear,' just as Latin *auris* 'ear' (< PIE *aus-) is related to the verbs *audire* and *auscultare* 'listen' [Buck, 4.22, 15.41]. English *hand* is ultimately a doublet of the verb 'hunt' (OE *huntian*), from the notion of grabbing or seizing with the hand [Buck, p. 239].

2.43 "Metastatic flowcharts" and their interpretation. In the last section of this study (2.5), we shall present fairly elaborate diagrammatic representations of the patterns of semantic shifting in a subpart of the system of Tibeto-Burman body-part nomenclature.²⁴² We call these diagrams **metastatic flowcharts**, since they are designed to reflect the complex and unpredictable ways in which the meanings of words "metastasize" from one point in semantic space to the other [see above 2.41(a)].

Before proceeding to the charts, let us summarize and

Lahu:

Mikir:

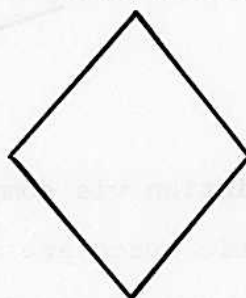
HAND

BLOOD VESSEL



ARM

MUSCLE



TENDON

NERVE

- [b] Inter-lingual semantic metastasis of phonological cognates. There is a form in Language X that is etymologically related to a form in Language Y, but the meanings of these forms are A and B, respectively. Thus, Mikir phek means 'bowels,' but it is related to the second syllable of Lahu ḍ-fá-qō 'stomach' (< PTB *ʔ-pik *ʔ-wik):

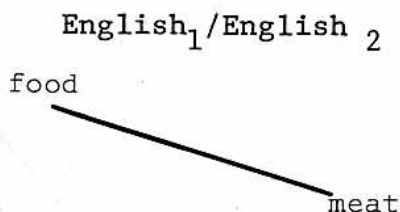
Lahu/Mikir

STOMACH



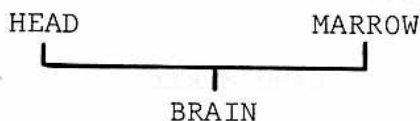
BOWELS

As a special case, X and Y may be earlier and later stages of the same language. Thus, Middle English *mete* and Early Modern English *meat* meant 'food in general,'²⁴³ while at the present time the word means specifically 'flesh of animals used for food:'

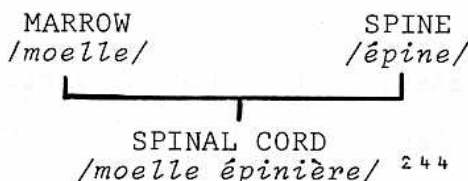


- [e] **Association via compounding.** Three points (A,B,C) in semantic space are related, such that in some language a compound of two morphemes, A+B, has the meaning C. Thus, in Lepcha, a-t'yak means 'head,' a-yón means 'marrow,' but the compound a-t'yak yón means 'brain.' We diagram this relationship by joining the two constituents with a bracket, which is then connected to the point in semantic space which is the combination of their meanings:

Lepcha

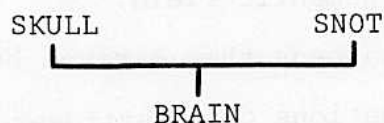


Similarly, in French:

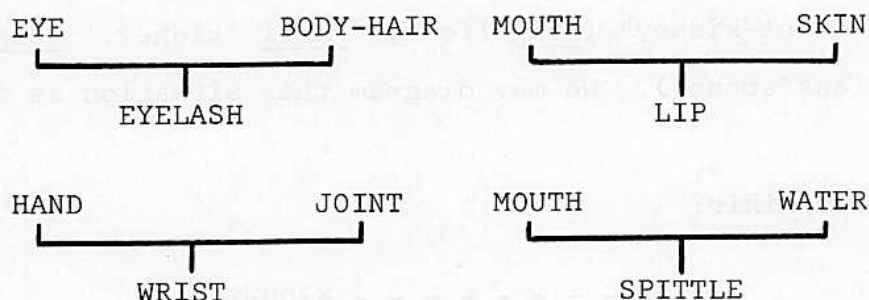


In some California Indian languages of the Penutian stock, the noble *brain* is associated with that vile

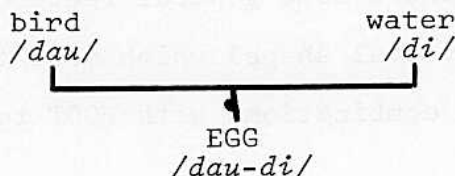
excretion, *snot*, via the compound "skull-snot:"²⁴⁵



Many compound words for body-parts are of obvious semantic structure, such that everybody would admit that the whole is equal to the sum of the parts:



In other cases, however, the compounds may be "semantically exocentric," in that one or more of the constituents may belong to an unpredictably distant semantic field. Thus in Dimasa, as in many other TB languages, the word for 'egg' is a compound meaning "bird-water" [Benedict 1939, p. 225]:

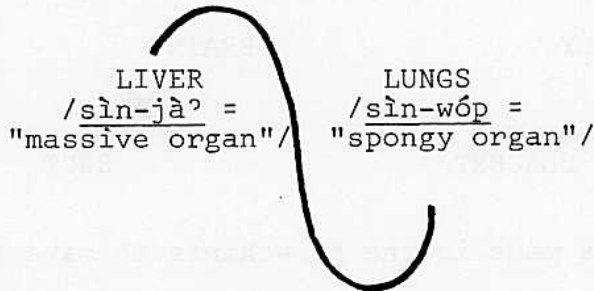


In these cases, we may loop the line connecting the

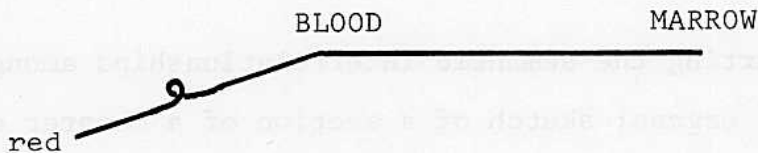
by a convex-concave, hill-valley, Yin-Yang sort of line between the points in semantic space that mark the poles of the opposition,



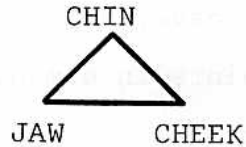
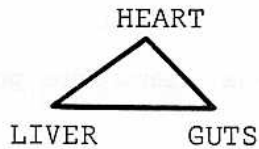
Thus, in Jinghpaw:



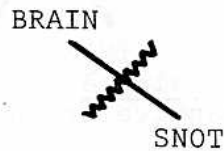
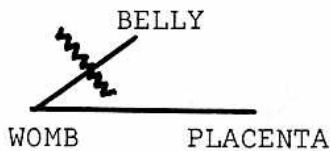
(5) In our metastatic flowcharts, points in semantic space that are body-parts or body-products are labelled in **capital letters**, while points belonging to other relevant semantic fields are labelled in **small letters**, thus:



(6) When three points in semantic space are all connected in a **triangle**, that means that all three dyadic associations (AB, AC, BC) are attested:



(7) A jagged line cutting across the line-of-association between two body-parts, A and B, indicates that they belong to quite different anatomical or physiological subsystems of the body, despite their semantic association:



(8) No attempt is made in the flowcharts to make graphic distinctions among all the various subtypes of semantic association discussed in 2.41 above (e.g., similarity of shape vs. similarity of relative position, etc.). Any given flowchart is liable to embody associations of several different types.

2.5 Charting the semantic interrelationships among TB internal organs: sketch of a section of a chapter of the Tibeto-Burman Etymological Dictionary. A full-scale study of TB body-part nomenclature would naturally have to involve the entire human body. To get a handle on the problem, the parts of the body have to be classified into subsystems, to which I give the non-committal name **somatic areas**. For my own pragmatic purposes I have been working with eight

of these areas, roughly as follows:

- I. Internal organs: heart, liver, kidneys, intestines, etc.
- II. Diffuse organs: bones, skin, hair, blood, etc.²⁴⁸
- III. Mouth and Throat: mouth, teeth, tongue, throat, etc.
- IV. Head and Face: eyes, nose, ears, head, etc.
- V. Limbs and Joints: hands, feet, shoulders, knees, etc.
- VI. Torso: neck, sides, back, hips, chest, etc.
- VII. Reproductive: penis, vulva, womb, etc.
- VIII. Body Products: sweat, spittle, urine, etc.

But by now it should be obvious that these are by no means airtight, mutually exclusive categories. No matter how one tries to draw up a list of somatic areas, many body-parts will remain multiply classifiable.²⁴⁹ This indeterminacy is due partly to the organic structure of the body itself—the breasts are located on the torso (VI), but they are part of the reproductive system (VII). Shall we classify according to the anatomical location, or the physiological function? Partly also, the indeterminacy is due to the unpredictable semantic associations ("metastases") which speakers of a language may make from one part of the body to another, irrespective of any "inherent" connection one could make between them on anatomical or physiological

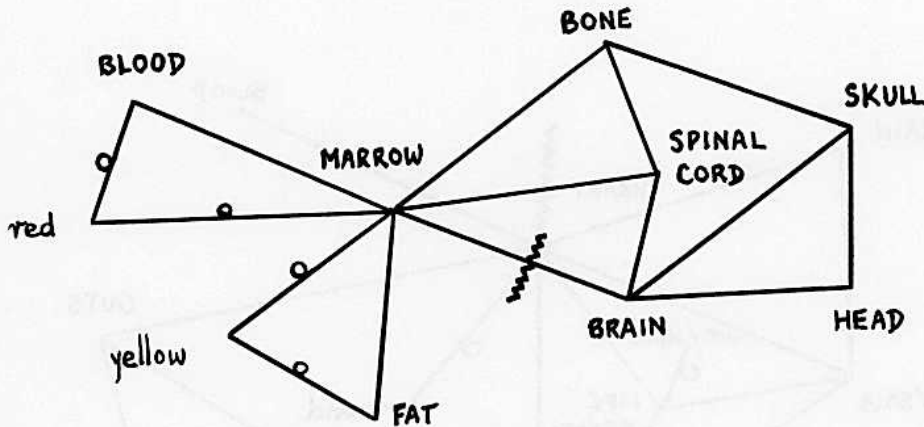
grounds. Thus, adjacent but inherently unrelated inner and outer parts of the body come into semantic contact [see 2.41(a), above]—the kidneys (I) get to be called by the same word as the waist, hips, or small of the back (VI). The brain (I) gets associated with snot (VIII).²⁵⁰

In what remains of this monograph, we shall confine the discussion to the internal and (to a lesser extent) the diffuse organs (I and II).²⁵¹ The data presented here to justify the structure of the flowcharts is only a small fraction of what has been accumulated over the past two years, in the course of my study of the body-part nomenclature of over 40 Tibeto-Burman languages.²⁵²

2.51 Marrow and its relatives. We have already discussed the semantic associations of *marrow* in TB [above 2.41(c) and (d)]. The information is presented in the form of a flowchart in Figure 13.

2.52 Heart/liver and their interconnections. The heart and the liver, the dominant organs in the thoracic and abdominal cavities, respectively, mediate between two conceptual areas of the body. Insofar as they are regarded as internal organs or "viscera," their interconnections are with GUTS, BILE, BLOOD, etc. But insofar as they are considered symbolically, as seats of the emotions or the home of life itself, their associations are rather with MIND, SOUL, BRAIN. Almost all the TB roots that can be recon-

Medullary matters: the interconnections of marrow



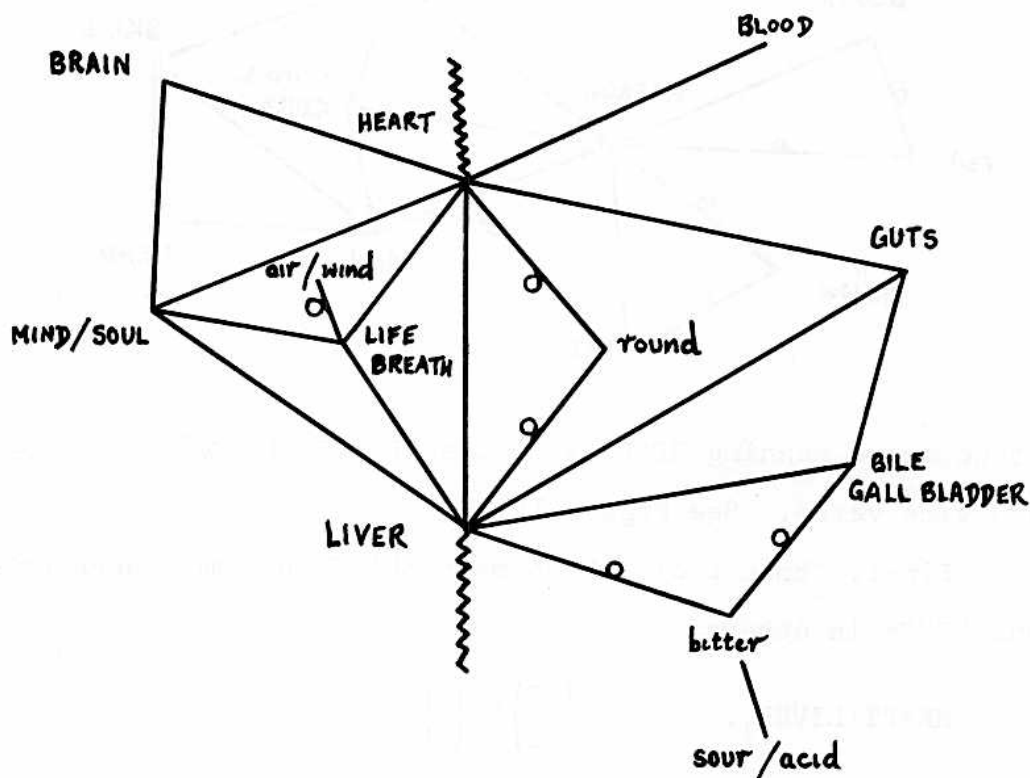
structured as meaning HEART also mean LIVER in some languages, and vice versa. See Figure 14.

First, three roots which mean HEART in some languages, and LIVER in others:

$$\text{HEART/LIVER}_1. \quad * \begin{pmatrix} s- \\ m- \\ b- \end{pmatrix} \text{lu} \begin{Bmatrix} n \\ m \end{Bmatrix}$$

In the STC this root is set up as *m-lun (p. 120), and the claim is made that it is confined to the Kuki-Naga subgroup of TB. Actually this is a general TB root, which could take either the *m- or the *s- "body-part prefixes," and which displays variation between final labial and velar nasals [see above 1.121(c)]. Four or five allofams of this etymon may be set up:

Figure 14
Studia cardiohepatica



HEART/LIVER_{1a}. *lun. Tiddim Chin lun 'heart,' Bantawa lun-mak 'heart, liver,' Sherpa lo-waq 'liver,' Lushai thin-lung 'mind.'

HEART/LIVER_{1b}. *s-lun. Chepang hlung 'heart.'

HEART/LIVER_{1c}. *m-lun × *b-lun.²⁵³ Mru plong 'liver,' N. Khami pālūng [plũ] 'heart,' (corrected form courtesy of L.-G. Löffler), Anal pa-lun 'heart,' Mikir vi-bong

'heart,' Chinbok am-lun 'heart,' Tang-
 khul Naga məluŋ 'heart,' Ao Naga temu-
luŋ 'heart' (with reprefixation).

HEART/LIVER_{1d}. *s-lum.²⁵⁴ Jinghpaw səlum 'heart; by
 some the liver or biliary gland,' WB
hnac-lûm 'heart,' Khaling 'lum 'liver,'
 Hayu thum 'heart' (with prefix pre-
 emptio)

There is a Lepcha form a-lüt 'heart' that looks super-
 ficially like the above forms, but whose true affiliations
 lie elsewhere.²⁵⁵

HEART/LIVER₂. *(s-)din

This root is not recognized in STC. The Kanauri form stin
 (below) is there assigned to PTB *s-nin, which I am treat-
 ing as a separate etymon [see HEART/BRAINS/MIND, below
 2.521]:

Kanauri stin²⁵⁶ 'heart,' Kaike ding
 'heart,' Hayu din 'liver,' Thakali tin
 'heart,' Sunwar au-di 'liver,' Bahing
din 'liver,' Tamang ting 'heart.'
 [Thakali -in for *-in is apparently
 regular; cf. Tamang-Gurung-Thakali
 *syin 'firewood' > Tamang syin, Tha-
 kali sin.]

HEART/LIVER₃. *tywa. This root does not appear in the STC: Thulung Rai thyo-ma bop 'heart,' Khaling 'cō 'heart,' Kanauri zũĩă 'heart,'²⁵⁷ Gurung uiq-dā: 'liver,' Phunoi mî-tchû-sî 'heart,' Thulung Rai dyu-ŋa 'liver,' Kham 'yuh 'heart,' Meithei tha-môî 'heart,' Magari mi-chi-ta 'gall bladder.'

Moving more specifically into the hepatic region of semantic space, we find several roots showing the trans-field association of LIVER/bitter, with the key semantic link being GALL BLADDER or BILE:

LIVER/BILE/bitter₁. *m-sin

This widespread TB root, represented by forms like WT mċ'in-pa, Jinghpaw məsĭn, WB ?əsāñ, etc. [STC No. 234], has been related by Benedict (p. 180) to Chinese *siĕn/siĕn 𪛗 'bitter' [GSR 382a-f].

Two TB forms reflect a variant with dropped initial consonant, *in: Abor a-in 'liver,' Miri pui-ing 'gall bladder.' These forms probably arose through metanalysis, with the original root-initial s- becoming reinterpreted as the body-part prefix s- and then

dropping, something like this: *sin >
 *sə-in [metanalysis] > in [prefix
 loss].²⁵⁸

LIVER/BILE/bitter₂. *ka- $\left\{ \begin{matrix} m \\ n \\ \eta \end{matrix} \right\}$

The important TB root *ka 'bitter'
 [STC No. 8] has a body-part meaning in
 a number of languages within Tibeto-
 Burman: Garo, Kachari bi-ka 'liver,'
 Garo kha-khit 'bile.' The direct Chi-
 nese cognate is *k'o/k'uo 苦
 'bitter' [GSR 49u]. Benedict has shown
 how the Chinese word for 'liver' *kân/
kân 肝 [GSR 139-L] is to be brought
 into the same word-family [STC pp. 154,
 158, 165]. This analysis is somewhat
 reinforced by the fact that related
 forms with final nasal are also to be
 found within TB, e.g. Khaling khang
 'bile,'²⁵⁹ Gurung kă: 'id.,' Thakali
kam 'id.'

LIVER/BILE/bitter₃. * $\left(\begin{matrix} s- \\ m- \end{matrix} \right)$ kri-t

Two allofams are involved here, one
 with final dental and one with an open
 vowel. The forms with final dental

have the substantival meaning of 'bile, gall:' WT mk'ris-pa, Garo kha-khit [< PTB *(m-)krit, STC No. 412]. Related forms with open vowel mean either 'bile, gall, gall-bladder,' or else have the adjectival meaning of 'acid, sour' or 'bitter:' Lepcha kri 'bitter,' Jinghpaw khri 'acid, sour,' Nung səhi 'gall bladder,' WB sāñ-khre 'bile.' To these forms, given in STC No. 413, we may add Jinghpaw səgrì * šəgrì 'bile, gall, gall bladder,' Kaike thi-pā 'bile,' Jirel thi:bāq 'id.,' Chepang hi 'id.,' Tangkhul Naga āthi 'id.,' Boro bi-klê 'liver,' Phunoi bōkê 'bile,' Lisu ji₄ 'id.,' Akha pya.kui^v 'id.,' Magari mi-chi-ṭa 'gall bladder,' and Lahu ḍ-kḥ 'bile, gall bladder.'²⁶⁰ The prefixial s- in the Nung and Jinghpaw forms is clearly a reduced version of the morpheme *(m-)sin (our LIVER/BILE/bitter₁, above), as proven by the WB compound form, where the first syllable remains unreduced.

Several other TB forms may be ac-

commodated in this word-family, even though they have a simple m- initial. In this case the prefixial m- has "pre-empted" the root-initial: Lushai mît 'bile,' Mru müt 'id.,' Mikir amut-alang. These forms display our now-familiar variational pattern between -u- and -i- [above 1.121(c)].

BILE/bitter₄. *kik * kin

This root is less widespread than the preceding. So far I have uncovered only the following forms: Limbu khin 'bile,' khik 'bitter,' Hayu pik-kin * plik-kin 'bile, spleen.'²⁶¹

Both HEART and LIVER have associations with the GUTS or intestines. In Abor-Miri, the forms â-pui, puk-pui, and â-puk all mean 'heart' (Lorrain 1907), but *puk * *wuk is a widespread TB root that usually means 'guts, stomach' [above 1.141(b)]. In Palaychi Karen (Jones 1961), the compound fvúq-tàq means 'liver,' but its constituents mean 'guts' (fvúq) and 'heart' (tàq), respectively.

2.521 **The viscera and the emotions.** In Tibeto-Burman, the intestines do not seem to have been identified with any particular emotion or mental state, though this is a familiar semantic association in, e.g., Indo-European and Semitic languages. In English we use *guts* to mean "courage," and even speak jocularly of "intestinal fortitude." In earlier stages of English, the *bowels* were considered to be the seat of the tender and sympathetic emotions:

"'Twould have moved a Christian's bowels to hear the doubts he stated." Canning 1798 [Quoted in *OED*, p. 1031]

"I am a man that can feel for my neighbors. I have bowels—yes, I have bowels!" Lytton 1832. [*ibid.*]

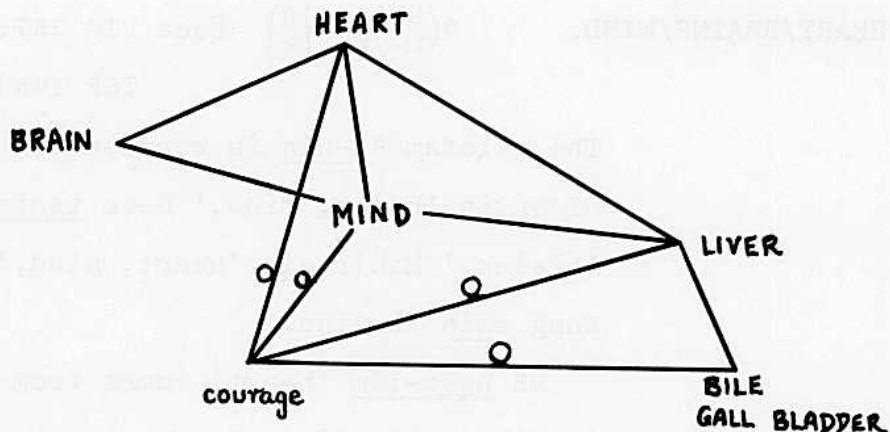
"It has none of the yearnings of the bowels of tenderness." Morley 1873. [*ibid.*]

This English usage was undoubtedly reinforced by the King James translation of the Bible, where the Hebrew word *me'ayim* (Englished as 'bowels') is frequently used to refer to the tender emotions: "Absalom, my son...my bowels are troubled for thee!"²⁶²

In Tibeto-Burman, the internal organs that are associated with emotions like *courage*, and the *mind* in general, are the HEART, the LIVER, and the GALL BLADDER—and to a lesser extent, the BRAIN. See Figure 15.

The Lahu believe that the timidity of the barking-deer (chĩ-pĩ-qwè?) is caused by the fact that it has no gall-bladder [chĩ-pĩ-qwè? ð-kĩ mâ cð].²⁶³ In many other TB

Figure 15
The visceral emotions



languages, as well as in Chinese and Japanese, a courageous, vigorous person is said to have a "big liver," while a namby-pamby lily-livered coward is said to have a "small liver."²⁶⁴

A newly discovered TB root that unites the meanings HEART, LIVER, and MIND has already been cited above [2.41(e)] in another connection, but we repeat it here:

HEART/LIVER/MIND. *b-yet *m-yet

Lepcha a-byet 'liver,' Limbu naribet
'heart,' Tamang Risiangku net 'liver,'
Jinghpaw myit 'mind, seat of emotions,'
Nung mit ~ nit 'mind, temper,' WT yid
'soul, mind.'

Another widespread root, wherein HEART is associated with MIND and BRAIN (but not with LIVER), may be set up as follows:

HEART/BRAINS/MIND. $\begin{matrix} *(\text{s-})\text{ni-}\{\text{n}\} \\ \text{k-}\{\text{k}\} \end{matrix}$ [see STC 367;
TSR 146].

The allofam *s-nin is represented by
WT snyin 'heart, mind,' Garo tənin
'brains,' Mikir nin 'heart, mind,'
Nung ənin 'brains.'

WB hnac-lûm 'heart' comes from
*s-nik, while Lisu ni₂-ma₃ has been
demonstrated [Thurgood 1974a] on tonal
grounds to derive from PLB *k-nik.

An open-syllabled variant prototype
is reflected by forms like Lahu ni-ma
(< *s-ni), Mikir ne [a doublet form of
Mikir nin, cited above], and perhaps
Thakali 'ngyeh.

This etymon may ultimately be re-
lated to another root that specifically
means BRAIN, reconstructed as PTB
*(s-)nuk in STC 483, and as PLB *nok *
*nok * *nuw in TSR 156.

The heart is sometimes associated with the notion of
LIFE itself, or with the BREATH OF LIFE. Thus in Taungthu

Karen [Jones 1961], sáʔ means 'heart,' while sà means 'breath;' in Moulmein Pho, θàʔ means 'heart,' but θáʔ means 'breath.' [See STC 485 and TSR 123.]

Since the role of the heart in the circulation of the blood is obscure to most speakers of TB languages [above 2.3], the association HEART/BLOOD is not common. Nevertheless, one can find connections like Mikir vi 'blood' (< *s-hwiy, STC 222), vi-bong 'heart.'

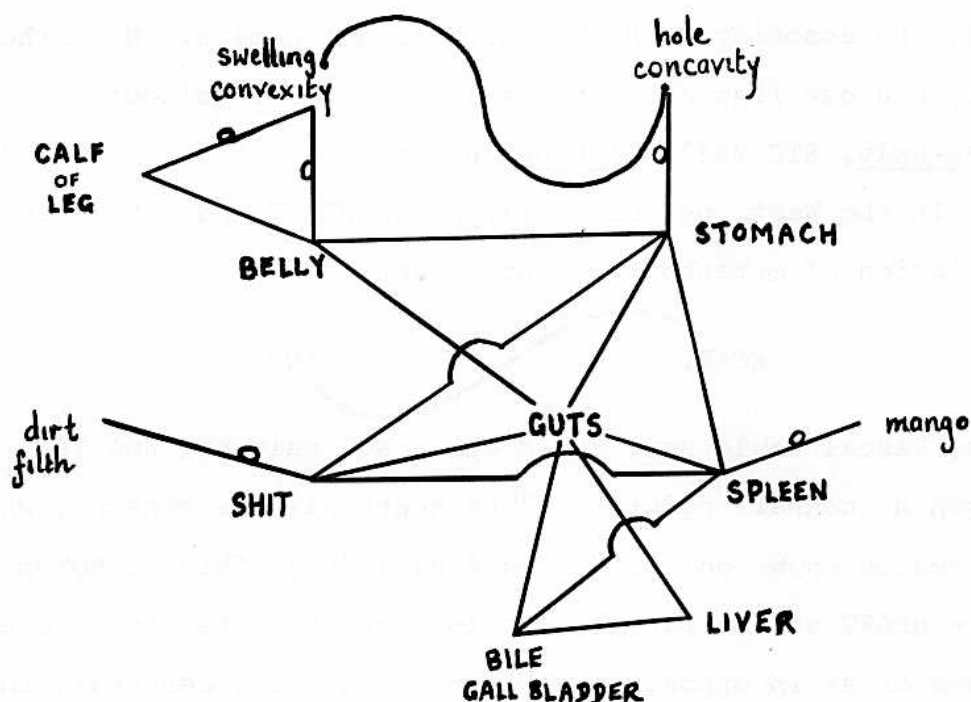
In the West, we frequently place HEART and MIND into a relation of metaphorical opposition.



Thus, Pascal could say, "Le coeur a ses raisons, que la raison ne connaît point." ("The heart has its reasons, which the reason knows not of." *Pensées*, 1670.) This dichotomy, where HEART stands for the emotions or the intuitions, conceived of as in opposition to man's rational, cerebral, intellectual capacities, is alien to Tibeto-Burman thought.

2.53 Gastrointestinal relationships. We have already discussed in detail the most important TB roots for BELLY/STOMACH [above 1.41(b)], pointing out their semantic connections with the GUTS, as well as their trans-field associations with the opposed notions of *swelling*, *convexity* vs. *hole*, *concavity*. In this section we shall add the SPLEEN to our deliberations, we well as a certain amount of SHIT and BILE. See Figure 16.

Figure 16
Gastrointestinal goings-on^{265, 266}



A widespread root whose basic meaning is 'shit' is TB *kliy [STC 125], which Benedict relates to Chinese *śier/śi 屎 [STC pp. 178, 185; GSR 561d]. This etymon also seems to stand in a loan-relationship with a Tai root, represented by Siamese khîi. Besides its nitty-gritty meaning of 'feces,' this word is frequently extended (one could almost say ameliorated) to 'filth in general;' or else, remaining within the somatic semantic field, it sometimes metastasizes to

mean GUTS/INTESTINES, or even STOMACH. Thus, Lepcha təkli is glossed as 'bowels, entrails, guts; the mucus of entrails' [Mainwaring p. 116]. To the forms given in STC 125, we may add the following:

SHIT/GUTS/STOMACH

*kliy

Abor-Miri a-ki 'abdomen, stomach, guts,'
 Lisu h'i₆-hchi₆ 'stomach,' Tangkhul Naga
mon-khui 'abdomen, pelvis,' Sunwar ki.
 'guts,' Chepang yong-kli? 'bowels,'²⁶⁷
 Khaling 'ciskhli 'guts,' Magari me-khe 'guts'
 (me- or mi- is a productive body-part
 prefix in Magari), Hayu so·gali 'belly,'
 Gurung mi-kli 'ashes' (mi 'fire' =
 "fire-shit"), Tamang nak-khli 'earwax,'
mik-khli 'tears' ("ear-shit," "eye-
 shit").²⁶⁸

At least five or six more TB roots (none of them mentioned in the STC) may be set up in this somatic area:

BELLY/STOMACH₅

*g-su(-t)

WT gsus-pa (perhaps from *gsuds-pa),
 Jirel 'su-baq, Sherpa supq [metanalyzed
 from *su-pa?], Hayu so·gali.

It is possible that this root is derived by prefix-preemption from *s-wu

(BELLY/STOMACH_{2c}, above), but there is not enough evidence yet to decide.

GUTS₁*yaŋ

I set up this root for PTB on the basis of two forms, Tاراon (Digaro) kay-ya·ŋ 'bowels' and Chepang yong-kli? 'id.' (The other syllables in these compounds descend, of course, from SHIT/GUTS/STOMACH *kliy, above.)

Benedict [personal communication 1974] has gone on to identify this PTB root with the Chinese *d'ian/d'ian 腸 [GSR 720y], from a PST prefixed variant *d-yaŋ.²⁶⁹

GUTS₂*r-gyu-ŋ

WT rgyu, Tamang kru(ŋ), Jirel gyu-mak (all 'guts'), and perhaps Lakher chhō 'abdomen, belly, stomach, inside of anything.'

GUTS/BELLY/VITAL ORGAN *mak

Akha ù-má? 'abdomen,' Jirel gyu-mak 'guts,' Sherpa ge-mā 'guts,' Bantawa luŋ-mak 'heart.'

GUTS/BELLY

*grwat

WT grod-pa 'belly,' Hayu cot 'guts.'

2.54 **Splenetic forms.** Moving on to the *spleen*, an organ with an "anatomical relationship to the stomach and a physiological relationship to the liver" [Gray, p. 949], we may set up two new SPLEEN roots with phonosemantic interconnections to STOMACH, GUTS, BILE, and SHIT.

SPLEEN₁

*pay *play

Jinghpaw pāi, kānpāi, kūmpāi, sīnpāi
 (all 'spleen'), Tangkhul Naga āpai
 (Pettigrew), pei [i.e. pəy] (Bhat)²⁷⁰
 'spleen,' Mru pai 'id.,' Lahu ḍ-pe
 'id.,' Newari am-pi 'id.,' Kham phis
 'id.,' Abor-Miri tûr-pe 'id.,' Mikir
pli-ha 'id.,' Angami Naga ú-prì 'id.,'
 Akha shi-pyeh 'id.'

The Nepali word for 'spleen' is phiyo, and this is certainly the direct source of Gurung, Khaling, and Sunwar phiyo, but probably not of Kham phis, or Newari am-pi [above], or (*a fortiori*) of any of the other forms cited.²⁷¹

Thulung Rai u-bal and Magari pāl-hi either represent a metathetic variant of this root, *pal, or else (shades of Austro-Thai!) may be taken as evidence that the proto-root should be set up as a disyllable, **pa-lay (> *p-lay).

The Taraon (Digaro) form kleig-ta-plã points to a proto-variant with final nasal. To this form we may perhaps directly compare a group of forms with initial l- (Lep-

cha a-lim, Limbu lem, and Chepang leh), two of which definitely have a final nasal, though the vowel correspondence (Taraon -a vs. front vowels in the other forms) presents a problem. Again we are tempted to invoke Indo-Aryan influence to account for these forms.

This root seems to have phono-semantic associations with STOMACH. Thus Lahu ð-pe 'spleen' is identical except for tone with the second syllable of Lahu gô-pê 'belly, stomach.' Similarly, Tangkhul Naga pəy 'spleen' is very close phonologically to TN pay 'stomach' (as well as to TN páy 'shit').

(From all the above forms, we must sharply distinguish a group of forms with labial stop initial and final palatal semivowel, whose nuclear vowel is -u-, and whose meanings appear to range from 'spleen' to 'stomach' to 'heart:' Abor-Miri tûr-pui, râm-pui 'spleen' [alongside tûr-pe, above], as well as AM puk-pui, â-pui 'heart,' and Lushai pùm-pui 'stomach.' The original meaning of this morpheme is revealed by the Abor-Miri homophones a-pui 'egg (n.); to be round, spherical (v.).' All of these internal organs [spleen, heart, stomach] are roughly oval or rounded in shape. See notes 85, 254.)

SPLEEN₂

*pik * plik

The key forms for establishing this etymon are the Hayu doublets pik-kin *

plik-kin, which mean both 'spleen' and 'bile' [Michailovsky 1974a].²⁷² To this we may compare Chinese *b'iěg/b'jie 脾 [GSR 874h], glossed both as 'spleen' and as 'tripe' in GSR, but as both 'spleen' and 'stomach' in Karlgren's *Analytic Dictionary* (1923).

Benedict [pers. comm. 1974] would like to find Austro-Thai connections for this root, and cites

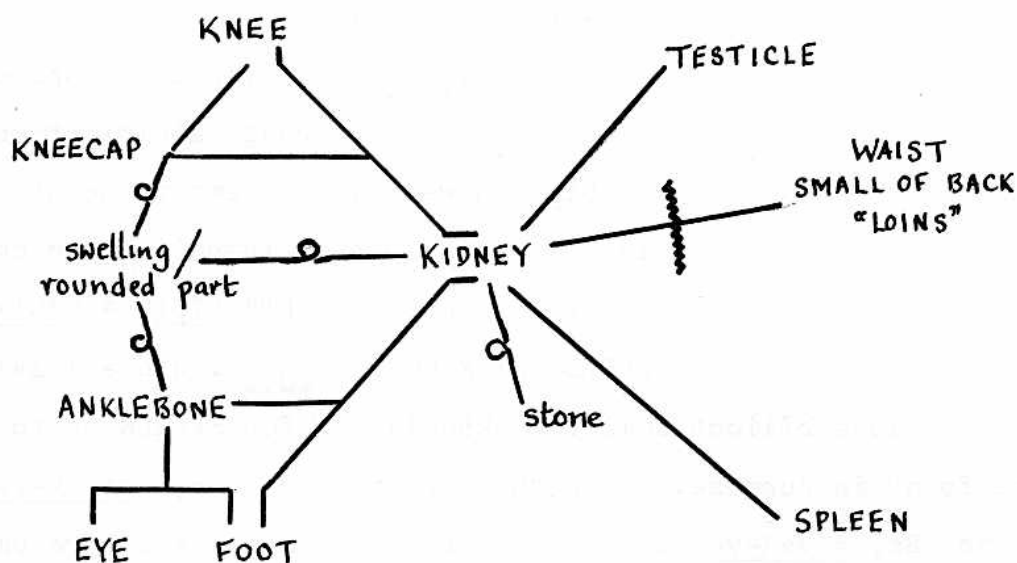
PAT *[1,1]i(m)pa[ak(/paak)] 'spleen' via *[1,1]impik (through assimilation) > *bik. However, it seems to me at least as likely that there is some connection rather with PTB *?pik *?wik [BELLY/STOMACH/GUTS_{1d,1e}, above 1.141(b)].

A less bilious semantic association for SPLEEN is to be found in Burmese, where the word for 'spleen,' sarak-rwak (Mod. Bs. θəye?-ywe?) means literally 'mango-leaf,' presumably because of the organ's shape.²⁷³ Continuing the image, the Burmese word for the pedicle of the spleen is sarak-nywan < nywan 'sprout, young twig or branch' (n.); 'be slender and tapering upwards' (v.).

For the interconnections between SPLEEN and KIDNEY, see the following section.

2.55 Kidneys and their kin. The kidneys interconnect semantically both with their adjacent fellow-internal organ, the spleen, and with the adjacent exterior part of the body, the WAIST, SMALL OF THE BACK, or "LOINS." See Figure 17.

Figure 17
Renal relationships



The STC reconstructs one root in this area:

KIDNEY/SMALL OF BACK₁ *m-kal * s-ga•l [STC 12]

Representing the allofam with the nasal prefix are WT mk'al-ma, Lushai kal,

Tiddim Chin kal, all meaning 'kidney.'
 The variant with voiced initial and s-
 prefix underlies WT sgal-pa 'small of
 back,' Meithei nam-gal * nam-gan 'back,'
 Garo dʒaŋ-gal²⁷⁴ 'id.,' Tiddim Chin
xa·l 'groin,' and probably Jinghpaw kàn
 'put on the back, be on the back.'

To the above forms we may add Che-
 pang gal, Kaike khal and Khaling 'kal
 (all 'kidney'), as well as Tamang and
 Thakali koh 'lower back' and Jirel pi-
koq 'id.' In all these cases, as well
 as with WB khâ 'loins,' we do not yet
 know which proto-allofam to assign the
 forms to.²⁷⁵

A peculiar, and possibly related,
 form is Jinghpaw n-khyun (< *m-kyul?).^{275-a}

Benedict has shown how this root is related to Chinese
 *ḍiēn/ziēn 腎 'kidney' [GSR 368h; STC pp. 173, 175].^{276, 277}

Note that Karlgren glosses 腎 as 'testicles; kidneys'
 in the *Analytic Dictionary* (p. 130), analyzing the character
 as being composed of the elements 取 'strength' and 肉
 'body.' In Chinese medical theory, in fact, the right kid-
 ney was the seat of the imaginary organ known as the "Gate
 of Life" 命門 (Veith 1949, p. 38), which in men was
 thought to be connected with the storage of seminal fluid

and in women with the support of the womb.²⁷⁸

The association KIDNEYS ——— TESTICLES seems to be due basically to the fact that both organs are paired and ovoid in shape [2.41(d)]. From there it is an easy step to transfer the seat of sexual energy from one set of "twins" to the other.²⁷⁹ This complex of associations is not peculiar to Chinese, and is found also in Indo-European, where, e.g. the non-Roman Latin dialect forms *nefrones* * *nebrundines* * *nefrundines* (cf. Greek νεφροῦς, German *Nieren*) meant both 'kidneys' and 'testes'.^{280, 281, 281-a}

KIDNEY/SMALL OF BACK₂ *pil * *pir

This root may well ultimately be relatable to SPLEEN₁ (2.54, above), though several languages (Lahu, Akha, Abor-Miri, Mikir) have distinct reflexes for both:

Akha beh[✓]-si[✓] 'kidney,' Bisu pê-k'à 'id.,' Lahu cò?-pē-ni 'waist, small of back,' Abor-Miri kat-pil * kat-pir 'kidney,'²⁸² Jirel pi-koq 'lower back,' and Mikir ham-bi 'kidney, anklebone, kneecap.'

The astounding Mikir gloss is explicable in terms of "metastasis via compounding" [2.43(3c), above]. The word ham-bi by itself means 'kidney.' The morpheme ham is cer-

tainly related to Hayu and Bantawa ham 'swelling, rounded part,' so that for the Mikir speaker the kidneys are perhaps subliminally associated with roundedness to begin with. When the word kengphu 'knee' is added to ham-bi, the combined meaning is 'kneecap, patella' (kengphu a-ham-bi, lit. "knee's kidney"). When the word keng 'leg, foot' is added, the meaning is 'anklebone' (keng ham-bi, lit. "foot-kidney"). This last compound is remarkable in that for most TB languages (as well as in Austronesian and Austroasiatic) the anklebone is associated with the EYES, not the KIDNEYS.

KIDNEY₃/SPLEEN

*la

This root is not in the STC, but may certainly be set up for PTB. It means sometimes 'kidney' and sometimes 'spleen:'

Lahu ḍ-lâ-šī 'kidney,' Mru la 'spleen' (Löffler 1966),²⁸³ Newari ja-la-si 'kidney,'²⁸⁴ Lushai la 'spleen,' Tiddim Chin la• 'enlarged spleen,' Lakher cha-law 'spleen.'

Some Central Chin languages have special compounds for 'enlarged spleen:' Lushai la a-tla, Lakher cha-law a-tla. At first glance the syllable tla looks as if it might be a prefixed variant of the root *la itself, but atla is actually a totally distinct verbal form meaning "it falls" (i.e.

protrudes) [Löffler, pers. comm.].

An enlarged spleen is a symptom of many pathological conditions. The normal weight of the organ is about seven ounces, but "in intermittent and other fevers it becomes much enlarged, weighing occasionally from 18 to 20 pounds" (Gray, p. 950).

Certain Karen forms with a lateral initial and a front vowel (Taungthu lê, Palaychi lî 'spleen,' Jones 1961), might prove to be directly relatable to the nasal-finalled forms with lateral initial cited above under SPLEEN₁ (Lepcha a-lim, Limbu lem), as well as to Chepang leh.

A trans-field association between KIDNEY and *stone* is perhaps established by Lakher lô 'stone,' lô-po 'kidney.' Again, the connection hinges on the roughly spherical curved shape of the organ.

2.56 **Womb and afterbirth.** It seems appropriate to end our brief introduction to TB body-part nomenclature with a part of the body that has particularly warm and positive connotations: the first home of mankind, the WOMB. The concepts WOMB and PLACENTA/AFTERBIRTH are closely associated in TB, as in Gnau [above 2.3]. There is also evidence for an interconnection between WOMB and BELLY. See Figure 18.

Benedict reconstructs only one PTB root in this area, *s-not [STC pp. 144, 145], with a range of meanings including 'vessel' and 'mouth' as well as WOMB: WT snod 'vessel,'

(a) With medial -a-:

Lushai hlam 'placenta,' Tiddim Chin lam
 'id.,' Tangkhul Naga rəm 'id.,' Lahu
ð-lò 'womb,' Thulung Rai wām 'placenta,'
 Bahing wam 'id.'

The TR and Bh. forms just cited are evidence in favor of a medial *-w- in the proto-form, unless one interprets them as merely being reflexes of *wam 'belly, stomach' (B/S_{4b}, above). See below.

(b) With medial -u-:

WT lhums (respectful) 'womb,' rum 'id.,'
sbrum-pa 'pregnant.'²⁸⁶

The second syllable of WT dku-zlum 'cavity of the abdomen, womb' might also fit here, or might belong with the root ROUND (STC 143; see notes 85, 254).

Karen apparently reflects medial *-a-, even though some of the modern forms have a back vowel: Moulmein Pho dən-lán, Bassein Pho dən-lán?, Palaychi dəq-lò, Sgaw dé-lò. This back-vowel reflex of *-am is also typical of Lahu (cf. ð-lò, above, and many other examples like 'iron' *šam > Lahu šo).

The Mikir forms hem 'womb,' oso-ahem 'placenta' (with oso 'child') show preemption of the initial by the *s- pre-

fix, with the -e- vocalism being a regular reflex of medial *a before *-m (STC, p. 70).

It seems very likely that there is some connection between this etymon and one of the variants for BELLY/STOMACH discussed above [1.141(b)], *pam * *wam. An association between WOMB and BELLY is common in Indo-European languages. French *ventre* may have either meaning, and 'belly' was frequently used to mean 'womb' in earlier stages of English:

"My belly did not blab, so I was still a Mayde."
[William Warner, 1602, *Albion's England*, quoted
in *OED*, p. 789]

In the case of this etymon, it is easy to assign a definite meaning to the *s- prefix: it must be a reduced form of the root *tsa * *za 'child' [STC 59]. WT shows variation between two prefixes (lhums < *s-lum, and zlum < *z-lum), which seems to reflect the voiced/voiceless alternation found in the root for *child* throughout TB.²⁸⁷

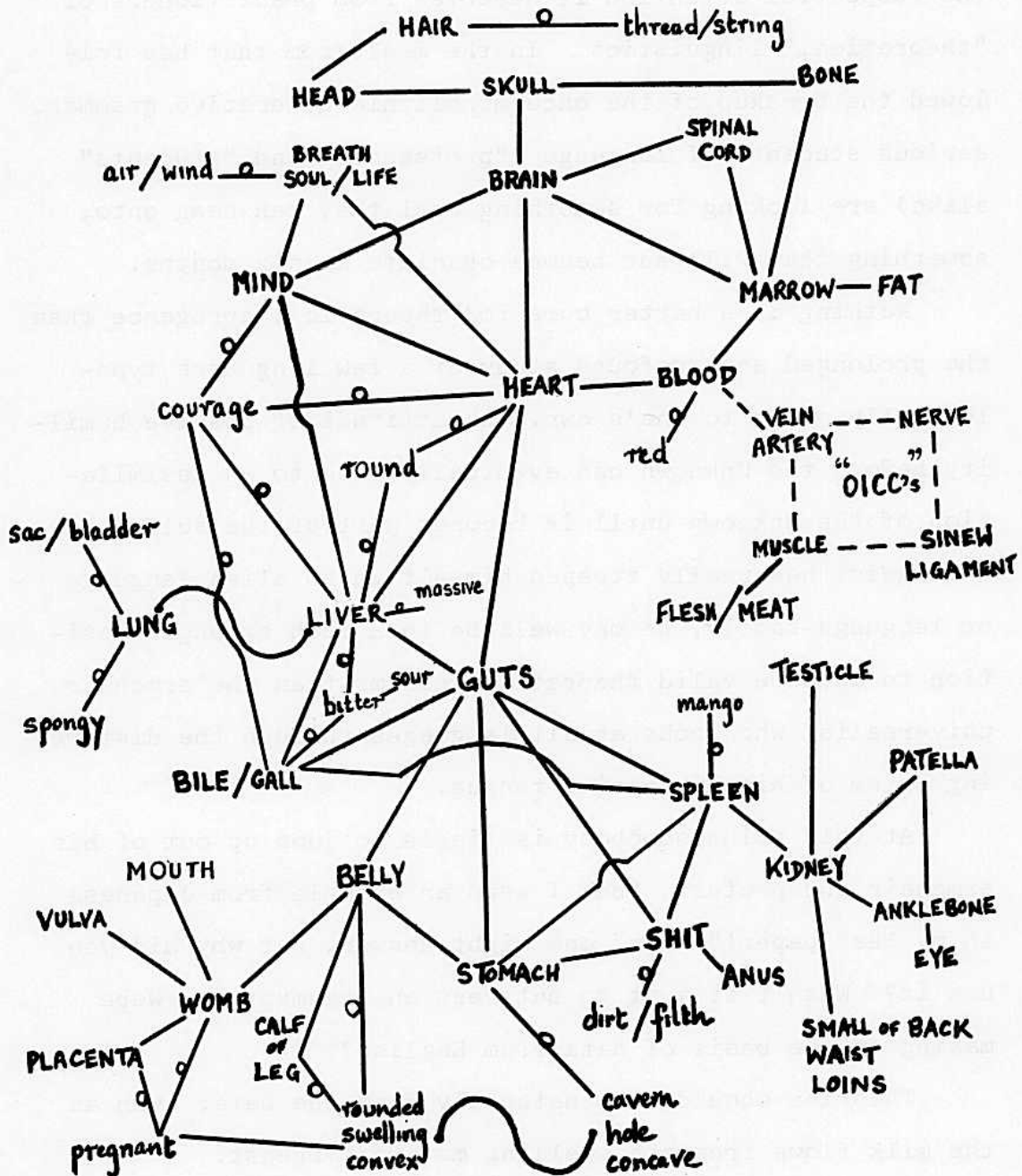
Mru thua 'womb' and Lushai chul 'id.' are, according to Löffler, derivatives of thu/chu 'vulva' (cf. Meithei chu 'vulva,' WT stu 'id.').

In Balinese folk belief there is a cosmological relationship between PLACENTA, AMNIOTIC FLUID, UMBILICAL CORD, and BLOOD, which a man calls his "four brothers," since they were "born" at the same time he was. [Personal communication, Alton L. Becker 1975.]

2.57 Composite flowchart of TB body-parts. The time has come to offer a single chart that simultaneously maps most of the semantic interrelationships among TB body-part terms that we have been discussing in this monograph. Monstrous as it is, the chart reflects only a tiny portion of systematic semantic space. See Figure 19.

Figure 19

A small part of the Tibeto-Burman body



3.0 CONCLUSION: genetic relationship, organic semantics, and Sino-Tibetan linguistics.

The great Sino-Tibetan family is beginning to receive the respectful attention it deserves from practitioners of "theoretical" linguistics. In the maelstrom that has followed the breakup of the once monolithic generative grammar, serious students of language ("professors" and "students" alike) are looking for something real they can hang onto, something that will not become obsolete in six months.

Nothing is a better cure for theoretical arrogance than the prolonged and profound study of a few languages typologically alien to one's own. An attitude of passive humility before the Unknown can eventually lead to an assimilation of the Unknown until it becomes part of the self. If a linguist has really steeped himself in an alien language or language-family, he may well be in a much stronger position to advance valid theoretical claims than the armchair universalist who looks at all languages through the distorting prism of his own native tongue.

At this point somebody is liable to jump up out of his armchair and protest, "But I used an example from Japanese in my last paper!" Aha, one might answer, but **why** did you use it? Wasn't it just to buttress an argument you were making on the basis of data from English?²⁸⁸

Theories should flow naturally from the data, even as the milk flows from the swelling maternal breast.

The Sino-Tibetan languages are of uncommon theoretical interest, for grammarians and phonologists, synchronic and diachronic linguists alike. As the locus of geopolitical power and cultural hegemony shifts from the West to the East, we may confidently look forward to the day when toneless languages like English, with tense distinctions in their verbs and otiose gender distinctions in their third person pronouns, will be treated as exotic curiosities, of interest chiefly for how they compare with familiar languages like Lahu or Chinese.

We have considered various ways to approach the question of genetic relationship—phonological, morphological, syntactic, lexical, and semantic. The difficulties have not been solved by shifting much of the discussion to the semantic plane—but at least a new and shiny arrow has been added to the quiver of the comparatist.

The organic semantic method, pioneered synchronically by inspired amateurs like Roget with his *Thesaurus*, and diachronically by the imaginative Indo-Europeanist, Carl Darling Buck, offers us the chance of penetrating into the deepest mysteries of the Sino-Tibetan proto-lexicon. At the same time it can shed light on the lives, attitudes, values, and material culture of the Sino-Tibetan peoples at various stages of their history and pre-history.

Language is constantly changing, and every word does have its own history. The shifting patterns of semantic

association within a language or a language-family are at least as interesting as phonological changes through time, and may prove to be equally criterial for establishing degrees of genetic relationship. Sounds and meanings change according to separate principles, but they are in the same boat together, and the changes undergone by one subsystem of the language have repercussions all through the other.

The present monograph is only the beginning of a beginning. It deals with only a part of one semantic area in the Sino-Tibetan lexicon. What has been done here for the internal organs can be done for kinship, verbs of cutting, plants and animals, psychological verbs, or anything else. The data are pouring in, freshly digested from old sources and increasingly being generated by new field-work. The surface has barely been scratched, the well has barely been tapped.

NOTES

¹This quote from Aristotle is cited in Buck 1949, p. 215. "For many think the brain is really marrow." See below 2.51, "Marrow and its relatives."

²*Sino-Tibetan: a Conspectus*, henceforth "STC." For details on this and all other references, see the Bibliography. For lists of abbreviations used, see Appendices II and IV.

³See also Shafer 1966-67, for the views of another great Tibeto-Burmanist on this subject.

⁴See Burling 1971; Matisoff 1974f.

⁵For an inconclusive examination of the feasibility of setting up a Jinghpaw/Lolo-Burmese supergroup (called "Ji-bur-ish") on the basis of tone correspondences, see Matisoff 1974f.

⁶Or, *a fortiori*, lest anyone be tempted to sneer at them like Roy Andrew Miller 1974.

⁷Hockett 1955 is a brave early attempt to carry out this kind of investigation on a large scale. The ongoing Language Universals Archiving Project at Stanford University, under the direction of Charles A. Ferguson and Joseph H. Greenberg, is designed to computerize as much typological data as possible about the phonological systems of 150-200 languages.

⁸See Haudricourt 1954; Matisoff 1970, 1973c; Mazaudon 1974c.

⁹See my comments on Li Fang-kuei's "Languages and dialects of China" (Matisoff 1973h). Shafer 1966-67 regards Tai as one of the primary subdivisions of his "Sino-Tibetan." Present-day scholars like Chang Kun still defend

this view, despite the counterevidence put forward for the last 30 years by Benedict (1942; 1967a, b; 1975) that Tai is rather to be grouped with Austronesian and Miao-Yao in a superfamily called "Austro-Thai."

¹⁰On the other hand, once the genetic relationship among a group of tonal languages is well-established on other grounds (as, e.g. in Loloish), the details of the tonal correspondences may then be used to work out the finer sub-grouping of the family, on the basis of "shared phonological rules" (see next paragraph and Matisoff 1972a).

However, when tone-resemblances are themselves taken as the prime criterion for investigating degrees of genetic relationship, we run into trouble. See note 5.

¹¹E.g. *m-sin 'liver' > Lushai thîn; Ngawn, Lai, Laizo, Zotung thin (but Tiddim Chin sin, Chinbok am-sin).

¹²E.g. *sey 'fruit' > Garo the ~ bithe, Dimasa thai 'bear fruit,' bathai 'fruit.'

¹³Thurgood 1974a, p. 49, was the first to explain the conditioning factor behind this double Lisu reflex. The line under the Lisu vowel indicates laryngeal constriction or "creakiness."

¹⁴It is necessary and natural to divide PLB (and PTB) rhymes into (a) open-, (b) nasal-, and (c) stopped- or checked-syllables (e.g. *pa, *paŋ, *pak, respectively). Rhymes ending in semivowels (e.g. *paw, *pay) are best regarded as "open."

¹⁵Matisoff 1972a, p. 3.

¹⁶Partly for this reason, Bradley 1975b puts Na-khi (= Nahsi) somewhat outside of "Loloish proper." To complicate the picture, the Loloish tonal split seems to be directly relatable to a similar two-way contrast in Jinghpaw stopped syllables [Matisoff 1974f]. Either we admit

the possibility of parallel independent development [see Mazaudon 1974c], or we push the split back to the pre-PLB period, claiming that it was Burmish which innovated and not Loloish!

¹⁷See Wolfenden 1929 *passim*, and STC pp. 95-127.

¹⁸By "non-syllabic" I mean "unstressed and with schwa-vocalism."

¹⁹Boodberg 1937 has given the name "dimidiation" to a similar syllabizational process in Chinese whereby a full vowel comes to stand between the elements of an original consonant cluster. For examples, see Yang 1972, 1973, and Boltz 1974. For an alternative explanation of the gùm- in the Jg. word for 'horse' see "Quo Vadimus?" (Matisoff 1973e).

Christopher Court (personal communication, November 1975) provides some interesting observations on vowel epenthesis in Tai. In Siamese, when two consonants meet at morpheme boundary in dissyllabic words (e.g. lûuk-taa 'eyeball' (literally "eye-child")), it sometimes happens in popular speech that the last consonant of the first syllable is doubled after which [ə] is inserted: lûukkətaa. This example is a unitary lexical item, contrasting phonologically with the otherwise homophonous "syntactic nonce-collocation" lûuk taa 'child of maternal grandfather,' where the epenthesis does not occur. For more details on Thai "linker-syllables," see Bee 1975.

²⁰See the dictionaries mentioned in the Bibliography.

²¹I, for one, think they are.

²²Even if there are more than two! It would be too tedious to have to speak of triplets, quadruplets, etc.

²³See especially Karlgren 1933, Wolfenden 1936, 1937. For a readable history of the development of the concept of "word-family" in ST, see Boltz 1974.

²⁴"Allofamy" may be pronounced either /áľlofamy/ or /allóľfamy/ depending on your predilections with respect to "allophony." In fact you might say that ['æľəfaēmiy] and [æ'ľəfəmiy] are themselves dialectal allofams of each other. (Actually, "allofamy" is no uglier than "nibbling," which is now used by anthropologists to mean "niece or nephew.")

²⁵We may also be concerned with tracing allofams in a single language back to variants in older stages of the same language (intra-lingual diachronic study, or "internal reconstruction").

²⁶I personally use "water-faucet" almost exclusively, but "tap" appears in the reversed compound "tap-water" (as opposed to, e.g. "well-water"). Nobody says "faucet-water"—still less "cock-water."

²⁷An example of something similar in IE is the r ~ l in Sanskrit, which cannot be related to any variation elsewhere in the IE family.

²⁸See also STC, p. 156. Benedict (personal communication, 1975) questions the validity of the gin/git example.

²⁹STC, p. 14. Benedict 1972 does not cite WT skyogs (with double prefixation) as belonging to this family, an etymology which I had already proposed in Matisoff 1970 ("GD") #25. (See also Thurgood 1974a, #151.)

However, in Benedict 1975b ("STAL"), fn. 21, he sets up PTB *s-kyok, positing connections with Chinese 勺, 杓, and 鈎 [GSR 1120a, b, c: all Karlgren's *điok/žiak], for which Benedict sets up Archaic Chinese forms like *sgy(i)ok ≠ *skyiok.

³⁰It is all but certain that further research will uncover additional forms that will establish such alternations for many more of the sets presented in TSR.

There follows a list of the sets in TSR that show the alternation. Each reconstructed proto-allofam is followed in parentheses by the names of the languages which reflect that particular variant. Languages that are not Loloish are put in square brackets. Note that in all the sets except DREAM and MAGGOT the Loloish languages are randomly divided with respect to the final stop vs. nasal. Languages which show both variants internally are underlined. DRAW WATER (39) *ʈ-kap ([WB], Akha, Lisu) * *kam (Lh.) COLD (99) *Nkrak * *ʔkrak ([Atsil], [Maru], Lahu, Akha, Lisu, Sani, Hani, Woni, Nasu, Lu-ch'üan) * *Ngran ([Chinese], [WT], [Lushail], [Tangkhul Naga], Lahu). DREAM (144). *s-mak ([WB], Lahu, Akha, Sani, Hani, Lisu, Nasu) * *s-man ([WB], [Jinghpaw], [Chinese]) HEART (146) *s-nik ([WB], Ahi, Sani, Nasu, Lu-ch'üan) * *s-nin ([WT], Lahu, Akha, Hani, Bisu) STEP ON (149). *s-nak (Lahu) * *nan ([WB], [Atsil], Akha) BACK/BEHIND/AFTER (155). *nok * *ʔnok ([WB], Lahu, Sani, Nasu) * *ʔnon ([WB], [Maru], Akha) MAGGOT (186). *k-luk ([WB], [Atsil], [Maru]) * *k-lun ([Lushail], Lahu, Akha, Sani, Hani) TREE (118). *sik ([WB], [Atsil], [Maru], Lahu, Lisu, Ahi, Sani, Woni, Nasu) * *sin ([WT], Hani, Bisu, Lu-ch'üan) STONE (190). *k-lok ([WB], [Lashil], [Maru], Lahu, Ahi, Sani, Bisu, Woni, Lu-ch'üan) * *k-lon ([Magari], [Jinghpaw], [Mikir], Akha, Hani).

³¹These three families also happen to show allofamic alternations in the syllable-initial consonants, though this is irrelevant to the point at issue here. See below 1.121(6).

³²See especially Wolfenden 1936, 1937; Boodberg 1934; and other references in STC (*loci citati*).

³³See Matisoff 1974f, pp. 194-196.

³⁴Lahu has a corresponding allofamic pair of nominal morphemes which show the same phonological relationship but different semantics: mû-qhô 'smoke' < *kuw / qhô?-lá 'ashes' < *ø-kut.

³⁵This set is discussed in GD #71 (p. 32) and in my note 335 in STC (p. 121).

³⁶Benedict was the first (1948) to point this out, positing a parallel development in Archaic Chinese. Burling 1967 rediscovered the Maru phenomenon, for which he was bitterly and incorrectly assailed by Miller 1970, whose alternative explanation does not hold water. See STC, note 193, p. 60.

³⁷For a similar phenomenon in Chinese, see the discussion of the Chinese word for LUNG, below 1.141b (PST *p-wap > OC *p'iwăd), and note 154.

³⁸In a couple of interesting personal communications (Nov. 6, 1975 and Feb. 27, 1976), L. G. Löffler takes issue with the proposed development PTB *-ik > Lushai -it, pointing out the counterexample 'louse,' *srik (STC 439) > Lu. hrik. Instead, Löffler derives Lushai -it from PTB syllables with medial *-y- and a labial or velar final (and similarly, *mutatis mutandis*, for Lushai -in). In the case of 'eye' and 'pheasant,' Löffler's interpretation works, if one assumes that the direct ancestors of the Lushai forms were the PTB allofams in *-ya- and not *-i-: *myak 'eye,' *s-ryak 'pheasant' [see below 1.12 (Da)]. However, the form for 'scorpion' (Lorrain ti·t, Bright trii) would then be unexplained.

In nasal-finalled syllables Löffler's explanation seems convincing. Thus PTB *-in remains Lushai -in if there is no medial -y- ('name' *r-min [STC 83] > Lu. hmin; 'fresh, green' *s-rin [STC 404] > Lu. hrin; 'wood' *sin [STC 233] >

Lu. thin), but *-in or *-im becomes Lushai -in if -y- precedes ('house' *kyim * *kyum [STC 53] > Lu. in; 'sleep' *yip [STC 114] > Lu. it; 'drink' PST *s-k-yim [Ch. 飲 *iəm/*iəm: (GSR 654), WT skyems 'thirst; drink, beverage,' Mru yem, Haka im * lim] > Lushai din, Bawm in).

³⁹See STC, note 171, p. 51; also Matisoff 1974f, note 74, p. 171. Note the similarity between *-yam > Jg. en and Löffler's explanation of Lushai -in/-it as having arisen from a medial *-y- [note 38].

⁴⁰See STC, note 93, p. 29.

⁴¹There is a triple typographical error in the last paragraph of the discussion of this set in TSR (p. 43). Instead of "*jup > WB cut; *jut > Jg. tšú?" read: *jut > WB cut, *cuk > Jg. tšú?.

⁴²I first thought this verb could only be applied to men. Further research (1970 field-trip) clarified the matter. For a semantic association with a different liquid, cf. the contemporary American slang expression "to cream."

The *m- prefix also appears with this etymon in WT mč'i-ba 'tear' ("[eye] its (m-) watering"), analogously to Lahu ("[sex part] its watering"). As we shall see below [1.12(6)], Lahu voiced stops reflect earlier prenasalized initials.

⁴³See Matisoff 1968, 1969, etc.

⁴⁴David Bradley (1975a, p. 475) has just uncovered a Lahu dialectal form which does show vowel fronting, Lahu Bakeo meh[✓]: the long-lost direct isofam of WB myâ!

⁴⁵Or possibly *bra, though there is no outside evidence for any *-r-.

⁴⁶This dà? * dè? example was first pointed out by David Bradley (personal communication, ca. 1972). These Lahu forms are related to the etymon Benedict sets up as PTB *tyak (STC, p. 52). But some of the forms Benedict cites

there also point to a prototype without medial glide (WB tak-tak vs. tyak-tyak 'very;' WT t'ag-pa 'be sure, decided, certain' vs. tig-tig 'certainly'), so the root should already be set up with $-\emptyset-$ * $-y-$ variation at the proto-stage: PTB *t(y)ak.

⁴⁷Wolfenden 1929 (p. 115) offers some speculations on the symbolic value of $-\underline{u}-$ * $-\underline{i}-$ in WT and Jg., and quotes the following interesting example from Lushai: "...with kâl to go, buk-buk is added as a descriptive particle when the action is that of a big heavy man, bak-bak when he is of medium size, bik-bek if of small stature." See also the sound-symbolism in the vocalism of Lushai demonstratives pointed out by Löffler (note 55).

⁴⁸For other examples of PST * $-\underline{i}-$ * $-\underline{ya}-$, see STC, note 251, p. 84.

⁴⁹Erroneously cited as "n-rum" in Wolfenden 1929, p. 115.

⁵⁰We follow the tradition of giving Chinese reconstructed forms in pairs *X/Y, where X is the "Archaic" or "Old" Chinese form (ca. 1200-800 B.C.), and Y is the "Ancient" or "Middle" Chinese form (ca. A.D. 600). See note 162.

We sometimes use Karlgren's tone-marks for the non-checked tones of Ancient Chinese: zero for Level Tone (Tone "A") [pam]; a colon for Rising Tone (Tone "B") [pam:]; and a dash for Departing Tone (Tone "C") [pam-].

⁵¹Very recently (Benedict 1975b, note 16), an attempt is made to get rid of $-\underline{i}-$ * $-\underline{ya}-$ variation on the PST level by setting up two more distinct proto-rhymes, different from both * $-\underline{ik}$ and * $-\underline{yak}$, for the rhymes of EYE [$*-\underline{y\acute{e}k}$] and PHEASANT [$*-\underline{e\cdot k}$], thus:

	'eye'	'pheasant'
PST	* $(s-)\underline{my\acute{e}k}$	* $(s-)\underline{re\cdot k}$
Chinese	[s-]mĭôk/mĭuk 目	səliok/d'iek 羽
WT	mig	sreg

Lepcha	mik	hryak
Jg.	myi?	ri?
WB	myak	rac
Lushai	mit	hrit
Garó	mik	rik.

Here, as elsewhere, the issue is: do we set up more and more separate proto-entities (multiplying the proto-inventory of rhymes), or postulate proto-variation (at some level) between two entities that we already have to set up anyway?

⁵²Also reflecting the -u- variant are Mru yu 'yonder-downward' and Lushai zû(k) [L. G. Löffler, pers. comm.].

⁵³Note that all these examples are in syllables with final velar consonant. For an example of -a- * -u- before final labial, see WOMB/PLACENTA, below 2.56.

⁵⁴Quoted in STC, p. 69, note 217.

⁵⁵To be distinguished from Miller's fantasies are cases where a given language exploits vocalic gradation in a quasi-paradigmatic way in a semantically highly structured set of words whose relationship is beyond doubt. Thus in Lushai, there is a three-way interchange of -i/-a/-u in certain deictic words: khi 'here,' kha 'there,' khu 'yonder' [pers. comm., L. G. Löffler]. Ed Hillard (pers. comm. 1976), based on fieldwork with Lushai informants in the United States, glosses these demonstratives as follows: khi 'that (visible up there),' khu 'that (visible down there),' kha 'that (abstract)' [as opposed to hei 'this' and soo 'that (visible same level)']. (One is reminded of sets like Japanese kore 'this,' sore 'that,' are 'that over there;' konna 'this kind of,' sonna 'that kind of,' anna 'that kind yonder,' etc.)

Also, in **closed** syllables one does find undisputed allofams of the same root that show -a- * -u- * -i- varia-

tion at a relatively recent time depth. Cf. BODY HAIR [section E(a) above] and WARM (STC 381) *-lum * *-lam * *-lim, all ultimately < PTB *s-lum * *g-lum.

⁵⁶As always, there is a Yiddish joke which ridicules this kind of reasoning. See the story entitled Tizal katal imrosi, in Matisoff 1974e (No. 291).

⁵⁷For more details see Li Fang-kuei 1933, and STC, p. 20, where various classes of exceptions to this pattern are explained.

⁵⁸As Christopher Court aptly points out, the well-known phenomenon of "liaison" in French, whereby otherwise silent final consonants are sometimes pronounced before a word beginning with a vowel, also operates under conditions which are not definable phonologically, but only lexically or grammatically. Thus, Mes amis ne font pas comme ça [mezami...] "My friends don't act that way" (mes amis, "my friends" forms a tight syntactic unit) vs. Mais amis ne font pas comme ça [meami...] "But friends don't act that way" (mais amis, "but friends"). Liaison in fact never occurs after mais (not even in such set expressions as mais oui).

⁵⁹See Matisoff 1973c for a similar concept of "ever-present pressure for the phonologization of tone."

⁶⁰Review of STC, 1973 (p. 337).

⁶¹There is little need here to stress the importance that "variation theory" has achieved in current linguistic thinking (both synchronic and diachronic), thanks to the work of scholars like William Labov. See Weinreich, Labov, and Herzog 1968.

⁶²The evidence for this prefix is in fact overwhelming. See STC, pp. 106-108. There is equally strong evidence for a velar prefix which sometimes attached to names of animals

(Matisoff 1969a, 1972a, 1973e; Smith 1975, etc.).

⁶³In TSR (p. 48) I presented evidence that the Proto-Loloish causative prefix *ʔ- was **sometimes (though not always)** opposed to a nasal stativizing prefix *N-.

⁶⁴See especially Matisoff 1968, 1970, 1972a, 1972d, 1973e, 1974f. For a good summary of the results of this line of research, see Thurgood 1974b.

⁶⁵The Yellow Lahu form is not likely to be a loan < Burmese câ (< *kyâ < *k-la), since the former is low tone and the latter is high tone. Furthermore, this word is a famous interdialectal shibboleth among the Lahu themselves (see Matisoff 1969c).

⁶⁶The distinction survives in Lewis' Akha, but Bradley (1969a,b) reports wild variation in this area in other Akha dialects. I conventionally transcribe the "archiphonemic" WB letters ɕ ʈ ʈʰ as /c ch j/, though /ts tsh dz/ would do just as well, and is probably closer to the phonetic truth, since the modern Bs. reflexes are sibilants /s hs z/.

⁶⁷*The Grammar of Lahu*, p. 6 (Matisoff 1973d).

⁶⁸Much of the sibilant/affricate variation in WT can be formulated as "regular" internal Tibetan phonotactic constraints (see, e.g. Li Fang-kuei 1933). Thus WT prefixial h- never occurs before s, z, or ts, but does appear before the voiced and aspirated affricates dz and ts'.

⁶⁹The WB doublets have become stylistically differentiated, so that the affricated allofam chî is polite or medical ("urine"), while the sibilant variant sê is vulgar or familiar ("piss").

⁷⁰Benedict surmises (personal communication, April 1975) that the basic PST allofam of this HAIR root had initial *s-, with the affricated variant arising due to the effect

of the glottal nominalizing or possessive prefix *ʔ- < PST *a-: *a-sam > *ʔsam > *tsam.

This hypothesis is somewhat supported by the fact that Jinghpaw ts- sometimes occurs pre-glottalized, [ʔts] (LaRaw Maran, personal communication, 1963).

⁷¹Proto *-am regularly gives Lh. -o, and the Lh. high-rising tone / ˈ / reflects an earlier syllable with final stop (Matisoff 1970). Lh. cá- * yá- most likely reflects a prototype *ʔdzik *ʔzik, which is yet to be attested elsewhere with the meaning 'hair.'

⁷²See Matisoff 1970; GL 1.64 "Systematic tonal variation" (esp. p. 33); 1975b; 1976b.

⁷³See also GL 1.63 "Tone morphophonemics and sporadic word-families."

⁷⁴See Matisoff 1974f, p. 158. The English words 'ear' and 'hear' themselves stand in an allofamic relationship. See Buck 4.22, 15.41.

⁷⁵*Contra* GL p. 29, where I was too ready to admit a possible genetic relationship between such pairs of forms.

⁷⁶These "elaborate expressions" are a Southeast Asian areal feature. See GL 3.39.

⁷⁷Cf. the extensive discussion of compounding in Lahu, in GL 3.3 (pp. 53-65).

⁷⁸ó-qō 'head' is itself a disyllabic compound, in which the second constituent means "hollow object" [see Figure 4 below]. nī=ó-qō is thus a "2nd-order compound," since one of its constituents is itself a compound. See GL *loc. cit.*

⁷⁹I learned this compound on my very first day in a Lahu village, in February 1965, as I was eliciting body-part words. At first, I couldn't figure out why the assembled villagers were laughing so hard when my informant

solemnly repeated the word three times, as he had already done with many other words on my list.

⁸⁰See GL 3.333.

⁸¹Sometimes even the most plausible etymologies are wrong. "German der Braten 'the roast' and the verb braten 'to roast' are, incredible as it may sound, totally unrelated." (Axel Schüssler, pers. comm., 1975). [That example is probably the world's champion "faux ami"!]

Are the two English verbs 'baste' historically related? Their meanings [(1) 'to moisten during the cooking process,' and (2) 'to sew with long loose stitches'] both involve "home economics." One wonders how many housewives consider them to be the same word.

Folk etymology may involve the conflation of two separate morphemes, such that the combination is homophonous to a **third** lexical item. My daughter Nadja (age 12), flexed the muscles of her forearm the other day and said, "Just look at that forcep!" [forearm + bicep]. She was not even trying to be funny, since she had heard the word forceps before, but was vague as to its meaning.

⁸²Nishida 1967 does not cite the form ó-qō, and erroneously calls the form ?a-tù-kù "Lahu Na." There remains much confusion in the area of Lahu dialect names, since the dialectal subdivisions of the language (however determined) do not necessarily coincide with cultural subdivisions of the Lahu people. See the Introduction to GL, and Walker 1970.

By far the best and most detailed treatment of the interrelationships among the Lahu dialects is David Bradley's SOAS doctoral dissertation, "Lahu Dialects and Proto-Loloish" (Bradley 1975a).

⁸³The four root-morphemes listed here do not begin to

exhaust the TB stock of morphemes that may appear in compounds meaning 'head.' When the data base is expanded to include 40+ TB languages from diverse subgroups of the family [see Appendix II] over a dozen roots may be reconstructed with this meaning. Note the "lenition" in the etymon *bu * *wu [1.121(6B), above).

⁸⁴The -go in the non-respectful WT form mgo might possibly be an allofam of this same root. (WT dbu is an honorific form). Alternatively, mgo might be allofamically related to WT sgo-na * sgon-na * sgon 'egg,' sgon-ba 'make round, globular,' and perhaps (via the testicles) sgo-pur 'foreskin,' as Benedict believes (pers. comm., 1975). For the symbol " $\xrightarrow{\circ}$ " for a "trans-field semantic association," see 2.42 below.

⁸⁵Several other body-part terms have associations with roots that refer to ROUNDEDNESS, including HEART/LIVER_{ld} *s-lum [below 2.52] and BELLY [with a distinct root, SWOLLEN, CONVEX, ROUNDED *(s-)bwam [below 1.141(b)]].

⁸⁶Just as different English dialects developed different compounds for "water-faucet." See above 1.12.

⁸⁷There is a very similar Japanese slang word for uvula: nodo-chinko, literally "throat-penis." I am indebted for this trouvaille to my colleague, Charles Fillmore.

⁸⁸Of course we linguists know that its chief function is to help make uvular r's.

⁸⁹Personal communications (Mary Haas, William Shipley, Marc Okrand, *et al.*), 1974.

⁹⁰To my knowledge, not even the *Sound Pattern of English* tries to set up the underlying form as day's eye.

⁹¹The symbols "v_p" and "v_h" are introduced in my GL, esp. 3.7.

⁹²Phrase from the libretto of Haydn's *Creation*.

⁹³Title of a play by Grillparzer, involving a swimming of the Hellespont for love.

⁹⁴Hypothetical data.

⁹⁵One question put to me during my Ph.D. orals (Berkeley, fall 1964) was "is structural borrowing possible?"

⁹⁶For a pioneering work on areal features in South Asia, see Emeneau 1956. Huffman 1973 is an interesting recent study of some shared grammatical features in Cambodian and Thai.

⁹⁷See the extensive discussion of "verb concatenation" in Matisoff 1973d, pp. 199-265; also Matisoff 1974g.

⁹⁸For example the Lahu locative noun-particle lo (á-qhó lo 'at home; to the house') derives from PTB *lam 'road, way;' the Lahu desiderative verb-particle gâ (câ gâ 'wants to eat') comes from an old full verb gâ 'to think,' which now survives only in fixed elaborate expressions dô-ša-gâ-ša 'pleasant to think and reflect about.'

The history of Chinese has seen the development of many prepositions from old verbs (在 tzay 'in, at' < 'be in a place' (V); 给 gee 'benefactive; for' < 'give' (V); 把 baa 'object particle' < 'hold, take' (V), etc. See Li and Thompson 1973a.

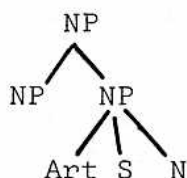
⁹⁹But note that even in these closely related languages the **functionally** corresponding morphemes nei and chê are not phonological cognates. This ties in with Becker's notion (1973) that functional semantic categories are the relatively unchanging grid or framework that actual lexical items move into and out of through time.

¹⁰⁰For an exhaustive (and perhaps exhausting) account of Lahu relative-clause formation, see GL 6.4, pp. 472-503. The particle ve is not only a relativizer, but a **nominalizer**

[GL 6.11, pp. 440-452]. For the connection between nominalization and relativization in TB, and in languages in general, see Matisoff 1972b.

¹⁰¹RC = "relative clause;" N_{rh} = "relative head," or noun-head of the relative clause, here the agent of the higher sentence.

¹⁰²See Hope 1972, esp. pp. 76ff, 101ff. Hope gives as the "surface form of simple relative sentences" the following tree (p. 101):



¹⁰³This is Hope's label. From the limited evidence he presents, I would rather consider this a to be a **general** or **indicative nominalizer** (like Lahu ve or Akha eu or Mod. Bs. te), as opposed to the semantically more marked agentive nominalizer ma (below). As we have seen [note 100] Lahu has a single particle ve that fulfills both of these relativizing (= Lisu yí) and nominalizing (= Lisu a) functions. In Lahu the agentive nominalizing particle pā (= Lisu ma) must follow naked verbs directly (dò pā 'the one who smokes') with no intervening "declarative" particle. How many differences of detail we must recognize in the grammars of these closely related languages!

¹⁰⁴The sister-languages German and Yiddish could also be cited in connection with differences in relative-clause formation. In German an inflected auxiliary comes at the end of the RC, after a past participle,

das Lied, das wir gesungen haben 'the song which we sang,'

N_{rh} REL PPL AUX

and the RC cannot contain a pronominal copy of a deleted

constituent identical to the N_{rh} . But in Yiddish, as in English, the inflected auxiliary comes before the past participle. Furthermore, Yiddish (unlike either German or English) may optionally have a pronominal copy of the RC:

dos	lidl	vos	mir	<u>hobm</u>	(es)	gezungen	lit. 'the song
							which we sang (it).'
	N_{rh}	REL		AUX	PRON _{obj}	PPL	
					(it)		

¹⁰⁵Not a bad idea for a project!

¹⁰⁶See GL 6.49, "Right relative clauses," pp. 490-503. Our sample RC (i) may not undergo the shifting, since it has a non-stativized action verb, dō 'smoke.'

¹⁰⁷The genial Dr. Seuss has a children's story (1961) which serves as an excellent illustration of the dynamics of the social motivations of antonymous change. One class of sneetches (hypothetical large birds that live on beaches and seem incapable of flying) have stars on their bellies, and consider themselves vastly superior to another class of sneetches whose bellies are devoid of ornament. But when the starless sneetches finally acquire stars (through some means which need not concern us), and clamor for recognition as social equals, the **originally** starful sneetches find their distinctiveness threatened and contrive to remove their stars in order to maintain contrastive distance. At this point the two classes of sneetches oscillate rapidly between the starful and starless phases, until in utter exhaustion, they realize that opposites are illusory and that sneetchdom is one and indivisible.

For a similar point of view, see the Heart Sutra (Prajñāpāramitā).

¹⁰⁸See Matisoff 1969b; also GL 4.331C(2), p. 234, "antonymous pairs [of verbs] are mutually exclusive on 'similarity' grounds."

¹⁰⁹This particular pair of antonymous meanings wedded to a single phonological form is so intolerable to English speakers that the difficulty has been resolved by folk etymology [above 1.122a]. Most adult speakers of American English, and probably 100% of the new generation of children, have substituted the semantically appropriate but quite innovative expression "net ball" for the now meaningless "let ball"—after all a let ball does graze the top of the net!

We might also mention the proverb "The exception proves the rule," which everybody now takes to mean "An exception or two is to be expected for every rule, and in fact these exceptions themselves demonstrate the general validity of the rule." This rather bizarre proverb used to have a much more sensible, practically opposite meaning, i.e. "It is the exception which tests the rule, which probes the rule." Since prove now almost always means "demonstrate conclusively" in American English (rather than "put to the test, probe"), the original meaning of the proverb has been lost. (Both prove and probe are of course doublets descending from the same etymon, with probe having been borrowed directly from Latin probāre, while prove entered English via the intermediary of French (modern prouver, Old French prover). Cf. *Thessalonians* I, 5-12: "Prove all things; hold fast that which is good" [quoted in the Introduction to GL, p. 1 (= Roman numeral 50)].

¹¹⁰We can expect considerable phonological deterioration by the 25th century.

Another example is the word obbligato, which used to mean "an obligatory line in the music (as opposed to ad libitum)," but which now usually means "a subordinate, extra, supernumerary part that embellishes the main line of the music."

¹¹¹Personal communication, 1968. See GL, p. 55a, note

92. Perhaps this ordering difference is correlated with the fact that in Mru the object follows the verb [see next note].

¹¹²In STC (p. 4), Benedict states too categorically that "all TB languages without exception place the object before the verb." But Löffler 1966 (p. 157) reports that the object follows the verb in Mru (East Bengal) and Bodo-Garo.

Löffler speculates that the Mru SVO order may be directly related to Mon and Karen SVO. (Mru has a good number of vocabulary items in common with those languages.) The SVO order would thus be an areal feature dating from the period before the intrusion of the Burmese. In the case of Garo, Löffler surmises that the SVO order is due to Austroasiatic influence from Khasi.

¹¹³"...Karen word-order has been influenced by that of contiguous stocks (Thai, Miao-Yao, Mon-Khmer), all of verb + object type."

¹¹⁴The whole question of the genetic position of Karen is of crucial theoretical importance because of the juggling act that one must perform in evaluating the various factors involved—lexical, phonological, and grammatical. I intend to devote a full-scale study to this problem before long.

¹¹⁵Charles Li and Benjamin K. T'sou, personal communications, March 1974.

¹¹⁶See note 98 above, and Li and Thompson 1973b.

¹¹⁷The useful term "nominal hemistich" was introduced in GL (p. 40) to mean "all the NP's of a clause taken collectively." See esp. 3.10.0 "Sequences of NP's within the nominal hemistich," pp. 185-191.

¹¹⁸In this simple diagram, *ABC... symbolizes the proto-language at the earliest reconstructed stage. *AC is a later speech community with respect to *ABC, but is still

a proto-language with respect to the attested modern languages A and C. An intermediate entity like *AC is currently called a "mesolanguage" by many diachronic linguists, whether or not they are rigorous practitioners of lexicostatistics. Witness, e.g., the lively discussions about setting up mesolanguages for Austronesian at the First International Conference on Austronesian Linguistics, Honolulu, January 1974.

¹¹⁹This figure was arrived at empirically, at least in part, through calculations in language families whose history is particularly well documented. This is not the place for a serious account of the development of lexicostatistical theory. The reader is referred to Lehmann 1962 for an elementary presentation and for the actual 100- and 200-word "Swadesh lists."

¹²⁰Or, as my grandmother used to say, "You can't make a silk purse out of a sow's ear."

¹²¹They are not. Eng. d < PIE *dh, which gives Latin f- in initial position (cf. Eng. do, Lat. facio). Latin d < PIE *d, which gives English t- (cf. Eng. ten, Lat. decem; Eng. two, Lat. duo; Eng. tooth, Lat. dens, dentis).

¹²²Have does indeed go with capiō and not habeō. PIE *k > Germanic x (part of "Grimm's Law") > Eng. h. In Latin PIE *k is retained as a velar stop (Eng. hundred, Latin centum; Eng. horn, Lat. cornū; Eng. heart, Lat. cors, cordis)

¹²³See Matisoff 1968, 1969a where I first discussed 'four' and several other parallel sets with similar vocalic and consonantal correspondences between Lahu and Burmese.

¹²⁴The extremely important Lahu particle le, which conjoins both NP's ('and') and clauses ('suspensive'), is probably a borrowing from Tai (cf. Siamese lé?). See GL, Index, p. 665.

¹²⁵"Unprovable" in both senses (see note 109): not only is it **undemonstrable**, but it is also **untestable**.

¹²⁶The following discussion of English vs. German is admittedly impressionistic. It would be interesting to undertake a larger-scale investigation of the question.

¹²⁷Löffler suggests "Ersatzduldungsrate" as the German translation of RTQ!

¹²⁸See Matisoff 1975c.

¹²⁹These assumptions will of course be challenged presently.

¹³⁰Notice that we have to assume that our bilingual does have some educated layman's notion of what it means to say that a word in one language "is related to" a word in another language.

¹³¹Now only surviving in the dialects of East Carpathia known as "Siebenbürgish," according to *Der Sprach-Brockhaus*, Wiesbaden 1962.

¹³²A direct German cognate of this form would be the non-existent *schwarzig.

¹³³The traditional scope of the term "direct cognate" seems to me to be equivalent to "phonological isofam," regardless of the exactness or inexactness of the semantic correspondence.

¹³⁴We have chosen the term "isofam" rather than "homofam," since "homofamy" sounds too homophonous with "homophony"!

¹³⁵See Burling 1967, Matisoff 1969. No evidence has yet been uncovered for setting up a voicing contrast in PLB preglottalized nonstopped syllables (i.e., *?pa vs. *?ba). In **stopped** syllables, on the other hand, both series must be reconstructed to account for the tonal

developments (e.g. PLB *ʔpak > Lh. pāʔ, but *ʔbak > Lh. pá). See Matisoff 1972a.

¹³⁶See TSR #180. The numbers next to the Sani forms are tone-marks. "22s" means "low-mid level constricted," "55" means "high level."

¹³⁷Note that the terms isofam and heterofam only apply across languages ("interlingually"). The term allofam, on the other hand, may refer either to related forms within a given language ("intralingual allofamy") or across languages. See 1.12 above.

¹³⁸The unusual Burmese reflex ʔ- can perhaps most easily be explained as due to the preemption of the syllable-initial slot by the velar prefix.

¹³⁹The symbol "Ø-" stands for a hypothetical voiced prefix that pushed the syllable into the LOW-stopped tone class, despite the voicelessness of the root-initial. See TSR p. 33, note 8. This proto-entity is irrelevant to Burmese, which has no tone contrast in stopped syllables. See above 1.12, where we entertained the possibility of finding a Lahu isofam to WB khap (our hypothetical Lahu form there showed no evidence of the Ø- prefix).

¹⁴⁰Actually this Chinese form does not display the regular PST *-a > Chinese -o shift, so it is either a loan (Benedict, pers. comm.), or else the reflex of a syllable with final liquid (-r or -l: suggested by L. G. Löffler). For another Chinese open-syllabled allofam that **does** reflect PST *-a, see 怒 [next note].

¹⁴¹Benedict (pers. comm., April 26, 1975) suggests several additional Chinese members of this word-family which reflect the *s- prefix: PST *s-na > 怒 'anger, angry; violent, vigorous' [GSR 94a' *no/nuo B/C], reconstructible for Archaic as *sno because of the n- * sn- alternation

in GSR series 94. The semantic association between 'bewitched' and 'angry' is further attested in the word 疾 [GSR 494a: *dz'iət/dz'iět], 'sickness, pain; injure; aggrieved, to hate; violent; evil; urgent, active, energetic'] < *s-nat (the metathetic development *s^ən- > *zn- > *zd- > dz- is regular) [see Bodman 1969 and STAL]. The same semantic association is evident in the homophonous pair of words 憤 'angry' [GSR 593j; dz'iər/dz'i^A] and 瘡 'disease' [GSR 593k: dz'iər/dz'i^{A/B}].

In many primitive societies (e.g. the Kaluli of Papua New Guinea) the normal reaction one is expected to display to injury or bereavement is **anger** at the evil spirit which caused the trouble. See Schieffelin 1976, Ch. 7: "Anger, reciprocity, and the rhythms of experience."

Löffler reminds me that in Judson's dictionary (1893/1966), p. 558, it is suggested that Burmese nat 'spirit, demigod' is a borrowing from Sanskrit nath 'master, husband, lord.' However, this seems to me quite unlikely in view of the well-attested variants meaning 'spirit, evil spirit' elsewhere in TB, which reflect either a nasal final (e.g. Lahu nê) or an open syllable (e.g. Proto-Karen *hna 'witch' < *s-na).

¹⁴²PLB [*velar + r] clusters give Lahu simple velars, while [*velar + y] clusters give Lahu palatal affricates. See Matisoff 1968, 1969a. A phonologically isofamous Lahu form for WB khyip would be the non-existent *cɛ̃.

¹⁴³We are assuming that the historical phonology of languages A and B is not as well understood as that of Lahu and Burmese!

¹⁴⁴The material in this section was informally presented at the Seventh ST Conference in Atlanta, October 1974 (Matisoff 1974c), and will be incorporated into a larger study of TB words for the internal organs of the body (see below, 2.5).

¹⁴⁵We return to a consideration of the semantics of the sets for LUNG and BELLY below, 2.2, 2.5.

¹⁴⁶The syllable θoʔ is glossed 'light, loose, not compact' in Wade's dictionary of Sgaw Karen (1896), which further nails down the basic 'spongy' meaning of *wap. See below 2.2.

¹⁴⁷Cf. Matisoff 1973e (*Quo Vadimus*), 3.3a. The details of the Bisu developments remain to be worked out. But now see Bradley 1975a.

¹⁴⁸Prefixation is a cyclical process in ST. Once an original prefix becomes phonetically and/or "psychologically" fused with the root-initial, the word is eligible for reprefixation. Many TB languages retain multiply prefixed words to the present day, often with a protective schwa intervening between the prefixes. Cf. Tangkhul Naga forms like aməthin 'liver' (< *a-m-sin), khəməlek 'to lick' (< *k-m-lyak), etc. See Matisoff 1972d, "Tangkhul Naga and comparative Tibeto-Burman."

¹⁴⁹*Contra* my Atlanta handout, p. 9 (Matisoff 1974c).

¹⁵⁰Ultimately we might want to identify this "body-part *k-" with the "animal-prefix *k-" that has been so widely discussed in the recent literature: STC, note 301, p. 107; Matisoff 1969a, 1973e. The probable Mon-Khmer origin of this prefix is discussed in Matisoff 1973b ("The Mon-Khmer substratum in TB"), Shorto 1973, and Smith 1975.

¹⁵¹Perhaps Angami Naga úphié also goes directly with the Tangkhul and Mikir forms.

¹⁵²See the work of Tagla Tsering Buhda and James F. Fisher, in CSDPN.

¹⁵³This is similar to the "glottal dissimilation" that plays a role in Loloish tonal developments (Matisoff 1970).

¹⁵⁴The "dentilabialization" of OC/MC *p > Mandarin f is a regular development in "third division" syllables (i.e. syllables with a palatal medial glide -i̯-), where the OC/MC syllable had two labial segments: OC/MC *piw- > Mand. f-. See Karlgren 1954, pp. 269-270.

For a phonetic series which displays a very similar -p > -t dissimilation after medial -w-, see GSR 695: 入 'enter' [695a: *ńi̯əp/ńzi̯əp], but 訥 'slow of speech' [695j: *nwet/nuət].

However, as Axel Schüssler points out (pers. comm., November 1975), Archaic Chinese does permit syllables of the type piwap, piwām: see GSR 625a 凡 *biwām; 642k piwāp 法; 641a biwāp 乏. Since the Ancient Chinese form is in the chiuh-sheng (Tone "C"), which arose, according to Haudricourt's famous theory (1954b), through the influence of an *-s suffix, perhaps we should set up the pre-Archaic form for LUNG as *p'iwāps, explaining the -p > -t dissimilation as having been "helped along" by the apicality of the suffix.

¹⁵⁵The Bantawa (East Himalayish) case is not as clear as the Chinese. The second syllable of som-phu-rok might actually go with our LUNG₄ *pu (see below), rather than descend from *p-r-wak < *p-r-wap < *pwar-wap.

Forms in several other Himalayan languages show final velar stops: Tamang phoksa, Magari and Sunwar phokso, Kham pho·sa. But these are undoubtedly loans from Nepali phokso.

¹⁵⁶Alternatively, of course, the first syllable of the Bantawa form, som-, could also derive from *sin 'liver; conspicuous internal organ' (discussed above), with assimilation of the -n to the following labial.

¹⁵⁷But now see Benedict's new suggestion relating LUNG₂ to a root meaning 'breathe, exhale' (note 160).

¹⁵⁸For a discussion of sandhic variation between Lahu ၵ and ၶ, see GL 1.631, p. 28.

¹⁵⁹According to Löffler, Lakher -ao is the reflex of *-o + stop, so it looks as if it goes with the Hayu form.

¹⁶⁰Benedict, who now finally accepts my analysis of the Lushai form, has "retaliated" by the excellent suggestion that LUNG_{2b} *tsit is also to be related to Garo raʔn-sit 'breathe, exhale' and WT sid-pa 'whistle' (cf. also PLB *sit 'whistle' [TSR 119]). This would then be another word-family displaying sibilant * affricate alternation [above 1.121(b)], and would establish the semantic association

LUNG — o — airflow
(respiration, whistling). See note 215.

¹⁶¹See notes 155, 156.

¹⁶²We are still observing the Sinological convention of citing Karlgren's reconstructions for Archaic (= Old) and Ancient (= Middle) Chinese on either side of a slash, with the simple label "Chinese," thus "Chinese *m-liôk/m̥iuk" means "Archaic Chinese *m-liôk and Ancient Chinese *m̥iuk, as reconstructed by Karlgren." (See note 50, above.)

¹⁶³For a discussion of the phenomenon I call "prefix preemption," whereby the prefix drives out the root-initial, see Matisoff 1972d, 1973e. Thurgood 1974b has discovered special **tonal** consequences of this process in Loloish.

The Chepang and Magari forms might conceivably be related to Maru wen-tok (see B/S_{4b}, below).

¹⁶⁴Benedict (pers. comm.) sets up the Proto-Karen initial underlying these forms in γ- and h- as *hr-, but concedes that this might reflect an earlier *n- prefixed form, *r-wuk.

¹⁶⁵STC 35 groups Mk. phək and Garo bi-bik together, but does not identify them with the WT forms, which are given in note 237 to set 358 (p. 77). For the semantic shifting

involved,

{cavern
cavity → stomach → intestines}, see below 2.53.
hole
belly

¹⁶⁶The r in the TN form may be trying to tell us that this root should really be set up as *pur—but more evidence is needed to establish this point. (Perhaps TN has sometimes developed an epenthetic r intervocalically.)

¹⁶⁷The Chinese allofams 腑 and 付 are different tonally as well as in the voicing of their initial consonants. 腑 is tone B, while 付 is tone C.

¹⁶⁸Lahu g̃ [ɣ] comes from *r, while *w gives Lahu v. However, the syllables vu and vo do not occur in modern Lahu—i.e., the reflexes of both proto-resonants *w and *r have merged to ɣ before modern back rounded vowels.

¹⁶⁹See note 168. Detailed knowledge is required to correctly assign similar-looking forms to their proper prototypes. Thus the -g̃ù- in Lahu ð-g̃ù-tê? 'guts' goes with WB ʔu (PLB tone *1) < *wu (B/S₂b), while the g̃ô- in Lahu g̃ô-pê 'belly' goes with WB wâm (PLB tone *2) < *wam (B/S₄b).

¹⁷⁰According to Löffler (pers. comm.), the vocalism of Lakher vy reflects a development *wam > *vom > vy. The Lakher prefix pa- (as in pa-vy) appears in several other important body-part words (pa-khu 'knee,' pa-lei 'tongue,' pa-lia 'navel,' pa-thi 'liver,' pa-tia 'lower abdomen,' pa-zô 'urine.') Benedict has explained this Lakher prefix as being the regular reflex of PST prefixial *m- (STC, p. 119).

¹⁷¹Bright (n.d.) cites pùn as a "rare" variant of pùm (the latter being the only form to appear in Lorrain and Savidge 1898).

¹⁷²The same association between BELLY and CALF OF LEG

has occurred to the Western mind as well. The name of the prominent superficial muscle that forms the greater part of the bulge of the calf is the gastrocnemius (< Gk. gastēr 'stomach'). See Gray, p. 436.

¹⁷³See above 1.12(2) for more information on this etymon.

¹⁷⁴We are of course speaking of "interlingual semantically isofamous" compounds, i.e. compounds in different languages which have the same meaning (whether or not any of their constituents are relatable phonologically).

¹⁷⁵None of the morphemes in these Lahu words has anything to do with "barking." The English compound refers to the short, sharp cry of the animal, which sounds something like a dog's barking.

¹⁷⁶We have especially in mind the "expanded" list of 206 items. Since the list is so well known, we shall not reproduce it as an appendix, but rather refer the interested reader to Lehmann 1962, pp. 112-113.

¹⁷⁷Isidore Dyen, personal communication, First International Conference on Austronesian Linguistics, Honolulu, January 1974.

¹⁷⁸See GL 3.89, pp. 167-168: "Illustration of the incommensurability of cross-language semantic-grammatical categories: how to say 'from' in Lahu."

¹⁷⁹I plan soon to write an article "TB verbs of cutting," where I hope to explore the semantic interrelationships of these (and other) roots in this conceptual area.

¹⁸⁰See Matisoff 1974e, "Psycho-ostensive expressions in Yiddish."

¹⁸¹Thus Hoenigswald 1960 devotes only a few terse pages of his Chapter 4 to semantic change (pp. 29-30, 43-47), out of a book of 168 pages.

¹⁸²See for example the reviews of *STC* by Chang (1973), and Miller (1974), and my replies thereto (Matisoff 1973f, 1975a).

¹⁸³See Weinreich, Labov, and Herzog 1968 (e.g., pp. 165-183).

¹⁸⁴*Pace* all dedicated Walde-Pokorny fans in the audience.

¹⁸⁵Buck does not mention English 'hiatus,' since space is at a premium and he cannot endlessly expatiate on each form.

¹⁸⁶As when German purged itself of foreign accretions in its learned technological vocabulary in the *entre deux guerres* period, replacing such "foreign matter" as *Telefon* with home-grown neologisms like *Fernsprecher*.

¹⁸⁷I am well aware that the concept of "semantic field" has been around for a long time. As Christopher Court, reacting to the first version of this monograph, wrote me (November 1975): "The systematic or structural approach to semantics, including the metaphor of the human body, has quite a long history—entirely non-Bloomfieldian, needless to say. It goes back at least as far as Saussure, runs through R. M. Meyer, Jost Trier, G. Ipsen, Charles Bally, Helmut Hatzfeld, Georges Matoré, Hans Sperber, Melin-Milleron, Guiraud...." (To this galaxy of [mainly German and French] semanticists we may add E. Coseriu, Karl Jaberg, Porzig, L. Weisgerber....)

The concept of lexical field, semantic field, or *Wortfeld* has in fact been endlessly thrashed out in subtle scholarly discussions over the past 40-odd years in Europe. (An entertaining and erudite account of some of this activity is to be found in Malkiel 1974.) *Wortfeldtheorie*, like generative grammar, has evolved its own abstruse jargon over the years:

"Aside from *Feld* and *Grossfeld*, there has sprouted—among ardent followers of Ipsen, Trier, Porzig, and Weisgerber—a whole bunch of technical and semitechnical terms involving the 'field,' down to *feldfreundlich* vs. *feldfeindlich* 'friendly vs. hostile to the field theory.'" [Malkiel 1974, p. 271]

Among the characteristic neologisms of this school are Trier's *Begriffsblock*, *Feldganzes*, *Sinnbezirk*, *Wortdecke*, *Wortmantel*, *Zeichenfeld*, *Begriffsfeld*, *sprachliches Feld*, and of course *Wortfeld*; and Weisgerber's *Sprachzugriff*, *Zwischenschicht*, *Zwischenwelt* [Malkiel, *op. cit.*, p. 272].

Apparently this terminological exuberance has gotten somewhat out of hand. Reinhold A. Aman (pers. comm., December 1975), refers to H. Bergenholtz 1975, "*Zur Wortfeldterminologie*," as "the latest attempt to clean up their terminological mess," by introducing five subclasses of semantic field: (1) morphological field; (2) lexeme field; (3) lexical field; (4) associative field; and (5) environmental field [translation of terms by Aman].

Apparently also, certain theoreticians of the *Wortfeld* have been considerably influenced by romantic notions of the ineffable expressive power of the German language, ideas which were congenial to the *Zeitgeist* of pre-World War II Germany. Works with titles like *Von den Kräften der deutschen Sprache* were written in an overblown and bombastic style—so that more sober scholars rebelled against the "intuitive" nature of much *Wortfeldforschung* and strove to make it more "technical and operational" (Malkiel 1974, p. 277).

As one interesting sidelight on the immediate success of Trier 1931, Malkiel notes [p. 274] that Einstein's *Feldtheorie* had just burst like a bombshell on the world of physics (1929), so that "fields" of any kind were all the rage at the time.

I have not yet read any of this vast and intricate literature, and I do not presume to set forth any "semantic field theory" of my own. All I am trying to do at the moment is to apply some elementary notions of semantic structure to the study of a proto-lexicon and its development through time.

¹⁸⁸Actually it is more complicated than this. It is possible to use the arms and legs actively even while floating on the back. The feature [face upward] vs. [face downward] will not do either, since we speak of the 'deadman's float,' wherein the floater is lying passively and face down. To distinguish 'swim' and 'float' properly we really need both features:

+active limb motion	-active limb motion	+active limb motion
+face down	+face down	-face down
SWIM	DEADMAN'S FLOAT	FLOAT

¹⁸⁹*Guinness Book of World Records*, revised new Bantam edition 1970, 5th printing, p. 157.

¹⁹⁰See the passages quoted from Rabelais, above 1.222(c).

¹⁹¹Actually KIDNEYS and TESTICLES are sometimes found in semantic association, but the connection seems to be due more to the fact that they are both paired and rounded organs, rather than because of any fancied functional relationship. See below 2.55.

¹⁹²"The substance of the lung is of a light, porous, spongy texture; it floats in water and crepitates when handled, owing to the presence of air in the tissue." Gray, *Anatomy*, p. 975.

¹⁹³Personal communication, Martine Mazaudon (1974).

¹⁹⁴In butcher-parlance even today the lungs are referred to as "lights" (Asimov 1963, p. 145).

As late as the 17th century, "lights" was in common use even by physicians referring to the lungs of a human being. See William Salmon, *Synopsis Medicinae, or a Compendium of Physick* (1671), I.xliii.94 "The Difficulty of Breathing shews the Lights [to be affected]." Quoted in the *Oxford English Dictionary* (Compact Edition), Vol. I, p. 280.

¹⁹⁵Buck, p. 1073, sets up the PIE root as *leg^wh-, *leng^wh- (which accounts better for the -ph- in the Greek form), but does not adduce English lung in this connection.

¹⁹⁶According to Gray, p. 975, both lungs together weigh about 42 ounces, while the liver (in the male) may weigh up to 60 ounces (p. 933).

¹⁹⁷'Light (of color)' is from the same root as the noun light (< PIE *leuk- [Buck, p. 60], while 'light (in weight)' is from *leg^wh- (see note 195).

¹⁹⁸Personal communication, Karl Zimmer (1974). In Indonesian (Alton L. Becker, pers. comm., 1975), color intensity is associated with **relative age**: colors are either muda 'young' (i.e., "light") or tua 'old' (i.e. "dark"), as in mera muda 'light red, pink,' mera tua 'dark red.'

¹⁹⁹Gilbert Lewis, "Gnau anatomy and vocabulary for illnesses," *Oceania* 45.1, pp. 50-78 (1974). I am grateful to Edward L. and Bambi B. Schieffelin for calling my attention to this article (March 1975), and for supplying me with vivid verbal descriptions of the two opposed kinds of bread-fruit (see below).

²⁰⁰Perhaps the importance of the RIGHT/LEFT opposition partly hinges on the paradox that the two vertical halves of the body **look alike** (bilateral symmetry), even though the hands and feet of most people have **very different** "dexterity" due to the neurological process of cortico-

lateralization, in which one hemisphere of the brain comes to "dominate" the other. See Ornstein 1972.

²⁰¹Curiously, the Hebrew word for teeth is also morphologically dual, šeynayim (perhaps the opposition historically was between the teeth of the upper jaw and those of the lower?). The common word for 'testicles' beytsim has a plural suffix, but that is because it is an intruder from another semantic field, with the original meaning 'eggs.'

²⁰²As one socially acceptable way of showing ritual submission to the King of Thailand one may lay a handkerchief in his path for him to step on, and then put the handkerchief on top of one's head.

²⁰³In contrast, the English classifier 'head,' as in "three head of cattle" has no particular positive connotations, but is merely synecdochic, referring to the whole by a salient part [see below]. This demonstrates that the symbolic value of body-part terms is not necessarily identical across languages, despite the striking universal tendencies we have been noticing.

²⁰⁴As always German has a graphic compound here: *Scham-lippen*, literally "shame lips;" probably a calque on Latin *pudenda*.

²⁰⁵As the German saying has it, *Wie die Nase des Mannes, so ist sein Johannes*. (Courtesy of Reinhold A. Aman).

²⁰⁶My guru Benedict (personal communication) has kindly supplied me with a reference in this connection, an American jingle of the 1920's that goes

"If skirts get any shorter,
Said the flapper with a sigh,
I'll have two more cheeks to powder
And a lot more hair to dye."

²⁰⁷In Bali, according to Pete Becker, a man has a different strategy for crossing a street heavy with vehicular

traffic than he does in the West. We are apt to wait for the light to change, or at least look both ways: "Stop, look, and listen / Before you cross the street. / Use your eyes and use your ears / Before you use your feet!"

But in Bali the main thing is to protect your vital spots and plug up your orifices for protection, so that one races into traffic with eyes tightly shut, the right hand covering the eyes and mouth and the left hand covering the crotch, with constricted anus—with the Upper and Lower Faces thus protected, one need fear no evil and need not "look both ways" before stepping off the curb.

²⁰⁸This numerical precision undoubtedly had numerical/cosmological motivation, but it also must have had a practical mnemonic value, helping in the training of medical students, etc. Even modern Western medical students must memorize such numerical statistics as that the body has 206 bones.

²⁰⁹Veith 1949/1973 is a fascinating translation with commentary of the Chinese medical work, *The Yellow Emperor's Classic of Internal Medicine*. (My thanks to Talmy Givón for calling it to my attention.) According to this work, man has five "viscera" 五臟 and six "bowels" 六腑. These are substantially identical to the Tibetan don-snod. Each of the viscera participates in a rigid and elaborate constellation of Taoist interconnections both with other parts of the body and with objects and qualities in the rest of the universe, as in the chart compiled from Veith, pp. 21-24. See Figure 11a, pp. 268-269.

²¹⁰It is only fair to remind ourselves that there is still much about the human body with respect to which even Western medical science is still "rather in the dark." When a modern physician gives the diagnosis "essential" or "idiopathic" hypertension, he is merely saying in effect that this high blood pressure has no known cause (idio-pathic = "self-diseased"). Not so different from Molière's doctor who explained how a sleeping potion worked because of its "dormitive virtue."

²¹¹This seems to me to be not at all a bad idea for an actual research project!

²¹²Victor Girard points out that serious burns have the opposite effect on the skin of black people, turning it pink-white due to the destruction of the melanin.

²¹³Gelugi is itself an intruder from another semantic realm, with the basic meaning of 'stringbag, which women make and all men carry.'

²¹⁴See Wilson 1959, pp. 112-115; Kappas and Alvares 1975. For the American man in the street of the present day, the inside of the body is a mysterious region which he conceives of, if at all, something like the schematizations of a Bufferin or Dristan ad on TV, where the medicine is shown going through a system of internal pipes directly to the "CD zone" (congestion and distress), where it swallows up the animated little men with hammers who are beating on the victim's sinuses from within....

²¹⁵However, Benedict does seem now to have established such a connection by suggesting that our root LUNG₂ was associated with the notion 'breathe, exhale.' See note 160.

²¹⁶See above 1.14c. Semantic shifting goes by many different names. In his pioneering article, Benedict 1939 calls it "semantic differentiation." In his article on

Figure 11a. Taoist interconnections among the viscera, the rest of the body, and the universe

	YANG	YANG	[NEUTRAL]	YIN	YIN
Viscera	liver 肝	heart 心	spleen 脾	lungs 肺	kidneys 腎
Bowels	gall bladder	small intestine	stomach	large intestine	3 burning places (三焦) and bladder
"Tissues"	ligaments	arteries	muscles	skin + hair	bones
Orifices	eyes	ears	nose	mouth	"lower orifices"
Secretions	tears [<eyes]	sweat [<pores]	"saliva" [<lips + mouth]	mucus [<nostrils]	"spittle" [<teeth]
Rules over other viscera	liver > lungs	heart > kidneys	spleen > liver	lungs > heart	kidneys > spleen
Rules over spiritual resources	>soul 魂	>spirit 神	>ideas 意	>inferior animal spirit 魄	>will 志
Emotions	anger	joy	sympathy	grief	fear
Sounds	shout	laugh	sing	weep	groan
Odors	rancid	scorched	fragrant	rotten	putrid
Flavors	sour	bitter	sweet	pungent	salty
Colors	green	red	yellow	white	black

Animals	fowl	sheep	ox	horse	pig
Grains	wheat	glutinous millet	millet	rice	beans/peas
Climate	wind	heat	humidity	dryness	cold
Seasons	spring	summer	late summer 季夏	fall	winter
			long summer 長夏		
Directions	east	south	center	west	north
Numbers	8	7	5	9	6
Musical notes	<u>chiao</u> 角	<u>chih</u> 徵	<u>kung</u> 宮	<u>shang</u> 商	<u>yü</u> 羽
Elements	wood 木	fire 火	earth 土	metal 金	water 水
Planets	Jupiter	Mars	Saturn	Venus	Mercury

Gnau, Lewis calls it "referent slippage." American Indian-ists have traditionally called it "gloss shifting." The latter term has the merit of recognizing that the meanings given in dictionaries are only the constructs of the lexicographer's mind, and do not necessarily exactly mirror the meanings "in the minds of the native speakers," whether they are now long dead or still alive.

For our use of the term 'metastasis' for semantic shifting, see below 2.4-2.5.

²¹⁷Personal communication 1974.

²¹⁸Some of these "potheads" survive in present-day counterculture circles.

²¹⁹"Trans-field" is perhaps the best of these three terms, and we use it henceforth.

²²⁰From now on we use the handy abbreviation "N_{bp}" for 'body part nouns,' defined broadly to include body products (sweat, urine) as well.

²²¹For an analysis of the interconnections of jaw/cheek/chin/face in Romance, see Kahane 1941.

²²²If we wish to exclude bodily activities from the field /Body parts & products/ we would regard pairs like spit/spittle as trans-field associations (2.42 below).

²²³Cf. my talk "For many think the brain is really marrow," Matisoff 1974a (Berkeley), and the epigraph to this monograph.

²²⁴Which Benedict by and large accepts (personal communication 1974).

²²⁵Note that the association between marrow and blood is based entirely on appearance, not on the fact (which could not have been known to speakers of Proto-ST) that the marrow manufactures blood cells!

Jinghpaw lēsəwi, Benedict now suggests, reflects in its first syllable a reduced form of PTB *lak 'hand/arm/limb,' so that MARROW would be analyzable as "limb-blood."

²²⁶A propos of Karlgren's Archaic reconstruction *swia, Axel Schüssler comments as follows (November 1975):

"Most authorities agree, and I concur, that Karlgren's distinction between *-a and *-ar is untenable; these two final categories must have been identical. I agree with those who postulate *-r or *-l for Karlgren's *-a and *-ar. Therefore, I would write *sywal for 'marrow.' Karlgren's vocalic medial -i- in his *swia (and the whole *-a category) is also an unfortunate choice: it consistently behaves like his -i-, not the -i- (as in im xiwet), in the other rhyme categories. Another possibility for 'marrow' would be Arch. Ch. *sywe (Karlgren would write *siwəg), since these finals (*-yal and *-ye = Karlgren *-iəg, both Ancient Ch. -iə), are somehow terribly mixed up in phonetic series. There are quite a few doublets or cognate words where one is written with a phonetic that points to *-yal, the other with a phonetic that indicates *-ye (= Karlgren's -iəg). This is, incidentally, one of the problems which I am at the moment trying to disentangle."

²²⁷For a list of all such heart/liver roots discovered to date, see the supporting examples for the "visceral flow-charts" adduced below 2.52.

²²⁸The -y- in this Lepcha form is from *y and **not** from prefixial *s- (see Benedict 1943 and 1.12 E(a), above). We do not yet know enough about the details of Himalayish phonology to be able to decide between *-et and *-it as the rhyme of this root.

²²⁹A striking example of **interlingual** antonymic association is the Austro-Thai root set up as *(m)pu(N)qi [Benedict 1975a, p. 417], which in some Formosan languages (Ami) means VULVA (< *puki < *puNqi) but in other Formosan languages (Tsou) means PENIS (< *buki < *mpuNqi). As

Benedict adds in an informative parenthetical note, "penis [is] (often associated with vulva)."

²³⁰In my own case, the use of 'foot' for 'entire lower limb' may be due to a Yiddish substratum. The Yiddish fus means either 'foot' or 'leg,' according to context. German distinguishes between Fuss 'foot' and Bein 'leg,' with the latter representing a semantic shift from the original Germanic meaning of 'bone' (Eng. bone, Yid. beyn). (According to R. Aman, in the Bavarian dialect of German, 'foot' may be used, as in Yiddish, to mean the whole leg.)

²³¹The shin-bone is a particularly long and salient bone, with very little flesh covering it, as anyone who has ever received a kick in the shins will testify. Similarly, the hamstrings are particularly salient tendons. One could also claim that the hands are more salient than the arms (since we are constantly manipulating objects and usually have the hands in our field of vision).

²³²The original meaning 'bone' survives only in compounds like Elfenbein 'ivory.'

²³³The class of people who use "chest" for mammary glands may well turn out to be roughly coextensive with the class of people who believe that breast size has anything to do with lactational capacity. (So far the term "chest-feed" has not been attested, to my knowledge.)

²³⁴For some discussion of extended meanings of body-part terms in American Indian languages, see Lehrer 1974.

²³⁵Nevertheless, it is conceivable that a modern urban American child could acquire the "stapes" sense of stirrup before the original sense, since he would be likely to learn about the bones of the middle ear in school before he ever had to think much about the trappings of a horse's saddle.

²³⁶A muscle lying on the internal aspect of the pelvic bone, attached to the upper end of the femur.

²³⁷My own hypothesis is that the association TESTICLE — *e* — witness derives from ancient oath-taking practices. In the Old Testament one swore an oath to a man by placing a hand *taxat matnav* "under his loins." By swearing on his interlocutor's most precious possessions, the essence of his virility, the oath-taker signified his willingness to be exposed to the utmost wrath of the other man should the oath be broken. See Genesis 24:2, Genesis 47:29, Ezekiel 17:18; also Simons 1975, p. 200.

In French schoolboy slang of an earlier era, classically educated French youths would refer to their testes as "mes témoins" (Buck, *loc. cit.*)

²³⁸The German word *Scheide* 'sheath' is in fact now used mostly to mean "vagina" in current German, so that when "sheath" is specifically meant, one must say *Messerscheide* ("knife-sheath") [R. Aman, pers. comm.].

²³⁹Karlgren (GSR 954d) reconstructs the Old Chinese form with a velar, *giək, but Benedict prefers the dental because of the morpheme 趯 *t'iək/t'iək 'the sound of marching' in the same phonetic series (GSR 954g-h). Chang 1973 attacks this ST etymology because the "semantic matching is far-fetched" (p. 336). See my forthcoming paper, "Winging it: arm and wing in Tibeto-Burman," to appear in LTBA. It is clear that the Chinese word for 'armpit' *ziäk/iäk (written variously 腋, 掖 or 腋 [GSR 800a, 1, m]) also belongs in this word family.

²⁴⁰The Latin word *pēnis* itself of course means 'tail.'

²⁴¹This is similar to, but not identical to "body function — *e* — body product" (e.g. spit (v.) — *e* — spittle), above 2.41(b).

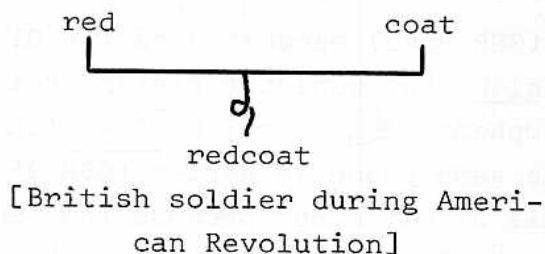
²⁴²I claim no particular geometrical sophistication for my diagrams, which contain many arbitrary features that might make a topologist wince. I intend them only as mnemonic devices.

²⁴³This archaic meaning survives faintly in expressions like "meat and drink" (i.e. food and drink), "sweetmeat," "mincemeat," etc. Similarly for the German cognate *Mett*-, which survives only in *Mettwurst* 'chopped meat sausage' (Löffler, pers. comm.). See Hoenigswald 1960, pp. 43-44.

²⁴⁴Thanks to Martine Mazaudon for this example (pers. comm. 1974).

²⁴⁵Personal communication, Mary R. Haas *et alia* (1974).
See note 89.

²⁴⁶To take a hallowed example, we would similarly diagram the compound "redcoat":



²⁴⁷I used to think that FOOT + EYE = ANKLEBONE was unique to East Asia. However, Jack DuBois tells me (1976) that in Quiché Maya the expression for ANKLE also has the same semantic structure:

u- 'poqoč r- aqan 'his ankle' ("the eye of his foot").
3 poss eye 3 poss foot/leg

²⁴⁸By "diffuse" I mean "not localizable to any particular single spot in the body."

²⁴⁹Dr. Benedict points out that from a strictly anatomical/physiological point of view, MOUTH and THROAT go with the intestinal tract, and even with the ANUS, the "bottom mouth." He cites PMY *hñou 'intestines/mind/seat of emotions,' whose reflex is beautifully defined for Highland Yao in Lombard 1968 (p. 263) as 'heart, mind; center of the personality; (physiologically conceived) the canal running from the mouth through the throat and intestines to the anus.' [The form in the Lombard/Purnell transcription is "Hruq," "H" standing for the voiceless palatal nasal.]

²⁵⁰It is this sort of semantic association **across** "somatic area" that we symbolize with an exclamatory jagged line in our flowcharts [above 2.43(7)].

²⁵¹In fact this monograph narrowly escaped being called "Blood and Guts in Tibeto-Burman." (I **never** considered the title "Bowels and Consonants in Tibeto-Burman.")

²⁵²For a list of the languages being considered in this ongoing study, see Vermiform Appendix II. For the sources used, see the Bibliography.

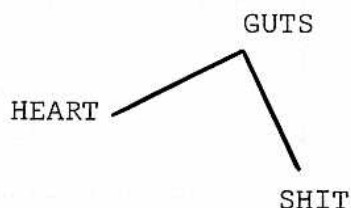
²⁵³The Mru, Khami, and Anal labial stops might be the regular innovative reflexes of PTB prefixial *m-. However, Mikir seems clearly to respond to PTB *m- by in- (many examples in STC, pp. 118-120), so that on the whole it seems preferable to set up a *m- *b- alternation here.

In the Mikir form the labial prefix has preempted the lateral root-initial.

²⁵⁴This allofam is probably to be associated with the root for ROUND *s-lum or *zlum [STC 143] (see note 85), though at the moment I see no way to "prove" that it does not also belong with the velar-finalled forms in HEART/

LIVER_{1a,b,c}. Mru lüm-la 'soul (of the dead)' may perhaps be brought in here, but cf. also Lushai hlum 'dead' (Löffler).

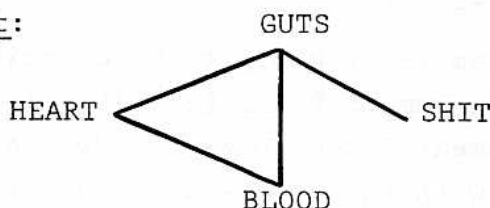
²⁵⁵The Lepcha form a-lüt 'heart' seems to reflect a separate root altogether. Benedict (personal communication) tries to relate it to Chinese 脬 *liwət/liuět (< *s-liwət) [GSR 498h] 'fat around the intestines' and to WT lud 'manure, dung,' thus confirming an association



that is already established by other etyma (see Figure 19).

On the other hand, Wulff 1934 has plausibly suggested a relationship between Chinese 脬 and Thai l̥yad 'blood,' and posits an allofamic relationship within Chinese between 脬 and 血 *xiwet/xiwet 'blood.' In fact, as I pointed out in Matisoff 1976a, the word 脬 is glossed as 'meat for a blood-offering' in Ueda's *Dai-jiten* (p. 1823), with 'fat around the intestines' given only as the secondary meaning. [These blood-offerings (Japanese *chi-matsuri*), according to *Kōjien*, were a practice in ancient China, when a sacrificial animal was slaughtered before a battle to enlist the help of the war god in the fray.] Since all of these interconnections seem valid, I believe we are dealing with an association like the following:

PST *s-lut:



²⁵⁶Bailey cites another Kanauri form dil 'heart,' which Benedict provocatively associates with Lushai di·l 'the inside part of a bamboo or cane; the heart-wood' (personal communication). However, it has been pointed out to me independently both by L. G. Löffler and Boyd Michailovsky that the Hindi word for 'heart' is also dil (ult. < Persian). Since there seems to be a strong Hindi component in the Kanauri vocabulary (see next note), we are probably dealing with a loanword. Perhaps the Lushai form is also ultimately borrowed from Indo-Aryan.

²⁵⁷Bailey 1910-11 cites the Hindi ji as a possible source of the Kanauri form, but this looks like an excellent native TB root to me.

²⁵⁸Benedict (pers. comm.) has just come up with a more general explanation of these curious Abor-Miri forms. Consider the following sets:

	PTB	pre-Abor-Miri	Abor	Miri
'three'	*g-sum	*?a-sum	a-um	a-um
'fruit'	*sey	*?a-se	a-ye	a-ye
'liver'	*(m-)sin	*?a-sin	a-in	a-śin
'son'	*za	*?a-so	a-o	a-o
'die'	*si	*si	śi	śi
'tree/wood/ firewood'	*sin	*ă-sin	$\begin{cases} \text{e-śin} \\ \text{śi} \end{cases}$	$\begin{cases} \text{e-śin} \\ \text{śi} \end{cases}$ (shortened form)

What apparently happened was that the (non-distinctive) glottal stop in the prefix *?a- drove out or "preempted" the root-initial sibilant, in a manner very reminiscent of the "glottal dissimilation" I have discussed in connection with Loloish tonal developments (Matisoff 1970, 1972a), where s- and ?- have similar dissimilation-provoking properties.

Benedict surmises that the Abor-Miri prefix e- is from the unstressed PTB *a- prefix, which was not perceptibly preglottalized, so that PTB *s- was retained as a spirant in 'tree, wood' [palatalized to ś- before i, as in 'die']. The Miri form for 'liver' is unexplained.

²⁵⁹The final -ŋ in the Khaling form is perhaps to be explained as due to assimilation to the velar initial.

²⁶⁰Lahu ci * ce 'sour' probably belongs to a separate etymon along with WB khyañ 'sour, acid,' perhaps from PLB *kyan. See Nishi Yoshio 1974, p. 026.

Note the development of the *(s-)kr- cluster to retroflex (Jirel) or dental (Kaike, Tangkhul) stops in some languages.

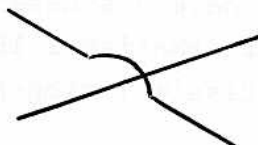
²⁶¹For the interconnections of SPLEEN, see below 2.54.

²⁶²The expression *me'ay homim* 'my bowels are troubled' recurs in the Song of Songs and the Book of Job (pers. comm., Chana Kronfeld).

²⁶³Only the grammar of English nominalizations obliges me to say "by the **fact** that + S." In point of fact, the barking deer does indeed have a gall-bladder (pers. comm., Gordon Young, ca. 1966).

²⁶⁴Cf. Japanese *kimo ga futoi* 'be brave, daring' ("liver is thick"); *kimo ga chiisai* 'be cowardly' ("liver is small").

²⁶⁵The topological complexities of this flowchart lead us to use the "line-crossing convention" current in electrical wiring diagrams, where a bulge in a line where it apparently crosses another line indicates that the two lines are "not really touching" at any point:



This is to be distinguished from our "looped line," which indicates a transfield semantic association:



²⁶⁶For a far-ranging romp through the viscera of the world's languages, see Rahder 1964.

Among all the English words available to refer to the intestines, "guts" is the simplest and most manly. "Bowels" now occurs mostly in the expression "move one's bowels," or figuratively ("bowels of the earth," etc.). "Entrails" calls up images of augurs plunging their hands into steaming bird-carcasses, and "viscera" is quite medical. "Intestines" is the most *salonfähig* of these words, being the most colorless ("the things inside"). Here again small children, who have only the foggiest notion of the insides, fall easy prey to folk etymology. My daughter Nadja, who as a three-year-old was more familiar with daytime TV giveaway shows than her viscera, once called them her "contestants" ("If you eat fast, it's no good for your contestants, right?")

²⁶⁷The tone of the Lisu form -hchi₆ and the glottal stop in the Chepang -kli[?] point to a stop-finalled allofam of the root.

²⁶⁸In the first version of this monograph I guessed that Tiddim Chin gil 'abdomen' was a reflex of this root "with metathesis," but Löffler says it is from *ril, cognate with Lushai ril, Mru ria 'guts.' (If only we had an Anal form [Old Kukish] to help us decide the matter!)

Solnit 1978

²⁶⁹Benedict also points out that Karlgren 1933 had hypothesized a word-family connection between 腸 'intestine' and 長 *d'ian/dzian 'long, tall' [GSR 721a], because of the great length of the intestines.

In STAL, Benedict makes a further connection with

Ch. 藏 *dz'ân/dz'ân [GSR727g'] 'conceal, store, treasure; intestines' < pre-Archaic *sæg'yân < PST *s-g-yân. The semantic association would flow from the intestines' role as the "keepers" of ingested food. Cf. West Tibetan sgyon-ba, pf. bsgyans 'to put into (the pocket, a box, a coffin); to keep, lock, or shut up.'

²⁷⁰Bhat 1969 misglosses this word as 'liver.' See my review of Bhat 1969, Matisoff 1972c. The real Tangkhul word for 'liver' is āmathin (< *m-sin LIVER/BILE/bitter₁, above 2.52).

²⁷¹However, the Sanskrit word for spleen is plīhan, and the Bengali is biliḥ, which look suspiciously like the Mikir form! Perhaps we have here uncovered a new Sino-Tibetan/Indo-European Wanderwort! Skt. plīhan is, of course, ultimately related to Gk. and Latin splēn, and to the variant Latin form liēn. Strikingly enough, the p- seems to have been separable from the following l- both in IE (Latin) and in TB (Lepcha, Limbu)!

²⁷²For the second syllable, see BILE/bitter₄, above 2.52.

²⁷³My thanks to Marc Okrand for first pointing this out to me.

²⁷⁴The first syllable of the Garo form represents a separate root from LOWER BACK (not in the STC), *džan, to which we also assign Magari jāng 'lower back,' Newari jā 'id.,' and perhaps Chepang ging 'id.' In this latter form the palatality seems to have transferred from the initial consonant to the vowel. Alternatively, of course, we could set up the Chepang form as representing an original *gin, with the Garo, Magari, and Newari forms now explained as having shifted the palatality from the vowel to the consonant, in much the same way as PTB *-ik > WB -ac.

²⁷⁵Benedict (STC, note 66) does cite WB khâ, and flatly states that it comes from *s-ga·l, though I don't see how he can be so sure.

²⁷⁶Yang 1975 relates the Chinese form rather with WT rked-pa * sked-pa 'waist, loins,' an alternative which must be taken seriously.

²⁷⁷Graham Thurgood points out (pers. comm. 1975) that the WB form for 'kidney,' kyok-kap, has as its first element a reflex of the Loloish root for WAIST set up in TSR #6, PLB *gyok * džok.

²⁷⁸In Veith (p. 26) there is a drawing of the kidneys and an explanation of their nature, reproduced from a medical work of the Ch'ing dynasty, the *Ling Shu Wên Chieh Yao Ch'ien Chu*, compiled by Ch'ên Hsiu-yüan (alias Ch'ên Nien-shih). The drawing makes the kidneys look very much like testicles, and the text says: 左為腎右為命門。
命門者, 男子以存精, 女子以繫胞。

²⁷⁹Cf. Greek δίδυμοι 'twins; testes' (whence English epididymis).

²⁸⁰Palmer 1955 (p. 61) quotes the grammarian Festus in this connection: "*sunt qui nefrendes testiculos dici putent, quos Lanuvini appellant nebrundines, Graeci νεφρους, Praenestini nefrones.*" Palmer adds, "Here we have a dialect word for 'kidneys' cognate with German *Niere*, for which Roman Latin used rēnes."

²⁸¹Chana Kronfeld reports (pers. comm. 1976) that in modern Israeli Hebrew slang the word *klayot* 'kidneys' is used approvingly in the sense of 'brains, good sense, *seykhl*': *Yesh lo klayot* 'He really has brains!' Perhaps here too there is some connection with the notion of vital (ultimately sexual) power.

²⁸²The kat- in the AM forms might be related to *kal (KIDNEY₁).

²⁸³Löffler 1966 (p. 137) derives the Mru form from *t-hla, but this is too complicated, since *la will do equally well.

²⁸⁴The first syllable of the Newari form perhaps means 'back' (see note 222), though it is not nasalized in this compound (CSDPN Vol. IV, p. 65).

²⁸⁵For some discussion of the relatively rare but well-established TB alternation *-a- *-u-, see above 1.121(e) "Miscellaneous but non-phony vowel alternations."

²⁸⁶The association of rum and sbrum is made in Simon 1975, p. 250.

²⁸⁷See STC, note 86, p. 27. In Tsangla, both the voiced and voiceless allofams of this root co-occur in the compound za-sa 'child, baby.'

Although the usual WT word for child is bu, PTB *tsa *za is reflected in the Tibetan forms btsa-ba 'to bear children,' and ts'a-bo 'grandchild, nephew.' To these forms, we may perhaps also add WT śa-ma 'placenta.'

For another example of a voiced *z- prefix from this etymon, see 'son-in-law' *z-mak (TSR #153; see also the note on p. 71).

²⁸⁸See the Introduction to *The Grammar of Lahu*, pp. xliv-1.

VERMIFORM APPENDIX I

Culturally Appropriate Lexicostatistical Model
for SouthEast Asia

[CALMSEA]

or

"The Matisoff 200-word List"
arranged by semantic categories and cor-
related with some reconstructed roots in
STC, TSR, etc.

A. Bodyparts (N_{bp})

- | | |
|----------------------|---|
| 1. *belly (exterior) | STC *pu•k (358) [see above
1.141(b)] |
| 2. *blood | STC *s-hwiy (222) |
| 3. *bone | STC *rus (6) |
| 4. *ear/hear | STC *g-na (453) |
| 5. *egg | STC *twiy (168) |
| 6. *eye | STC *mik * *myak (402) |
| 7. *fat/grease | STC *ryak (204), *tsil (pp. 16,
168, 173), *sa•w (272) |
| 8. *foot | STC *kriy (38) |
| 9. *guts | STC *pik (35) [see above
1.141(b)] |
| 10. *hair (head) | STC *s-kra (115); *ney (292);
*tsam (73) |
| 11. *hair (body)† | STC *mul (2), *tsam (73) |
| 12. *hand/arm | STC *lak (86) |
| 13. *head | STC *m-gaw * *s-gaw (490), *d-bu
(p. 117) |
| 14. *heart | STC *s-nin (367) |
| 15. *horn | STC *kruw (37), *run (85) |
| 16. *liver | STC *m-sin (234) |

*words which also appear on the Swadesh 200-word list are asterisked.

**words which further differentiate a single item on the Swadesh 200-word list.

†corresponds to Swadesh's 'feather.'

17. *mouth STC *ak (106), *m-ka (468)
18. *neck STC *ke (251), *liŋ (96), *tuk
(392)
19. *nose STC *s-na ~ *s-na•r (101)
20. *skin/bark STC *s-graw (121), *kok (342)
21. *spit STC *m-ts(y)il (231), *m-tuk ×
*s-tu•k × *s-du•k (pp. 58,
75, 126, 132, 146), *twiy
(168)
22. *tail STC *r-may (282)
23. *tongue STC *m-lay × *s-lay (281)
24. *tooth STC *s-wa (437)
25. *wing JAM *s-ton; *(s/k-)l-w-ak
[cf. ARM #12]; *(s/p-)lyap
[cf. FLY #149]
26. *nail/claw STC *m-(t)sin (74)

In Swadesh, but not in our list:

- back STC *s-ga•l (p. 18), *s-nun
(354);
- knee STC *du (p. 21), *(m-)ku•k (pp.
120, 159, 182); *put (7);
- leg STC *(r-)kaŋ (pp. 70, 142).

The following are not in Swadesh, but are on our list:

27. finger/toe STC *m-yun (355)
28. palm STC *pwa or *b-wa (418)
29. penis STC *li × *m-ley (262)

30. vagina or breast/milk STC *nuw (419)
31. brain STC *nuk (483)
32. navel STC *la•y, *s-tay (299)
33. shit STC *kliy (125), *n(y)ik (235),
*s-baŋ (p. 21), *r-kyak *
*s-kyak (pp. 26, 146), *e•k
(pp. 26, 146)
34. piss STC *ts(y)i * *ziy (77)
35. sweat PLB *?krwiy₂
36. snot STC *s-nap (102)
37. vomit PLB *Npat (TSR 38), STC *on (343)
38. marrow STC *klin (126)
39. breath/life STC *sak (485); TSR *C-sak (123)

B. Pronouns/kinship terms/nouns referring to humans (N_{hum})

40. *person/human being STC *r-mi(y) (pp. 107, 119, 158)
41. *thou STC *na(ŋ) (407)
42. *I STC *ŋa (406) * *ŋay (285)
43. *child/son STC *za * *tsa (59)
44. grandchild (nephew) STC *b-liy (448)
45. son-in-law STC *ma•k (324), TSR *z-ma•k (153)
46. *name STC *r-miŋ (83)

C. Foodstuffs (N_{food})

47. peas, beans STC *be (253), TSR *s-nok (140)
48. poison [antifood] STC *duk * *tuk (472)
49. mushroom / fungus STC *g-muw (455) * (JAM *s-muw)
50. liquor STC *yu(w) (94)

51. plantain/banana STC *s-ŋak (477)
 51a. medicine / juice STC *tsiy (65)
 51b. rice (in fields) STC *b-ras (pp. 17, 123)
 51c. rice (cooked) PLB *haŋ₂ (Bradley 1975a,
 #281b)

D. Animal names or animal products (N_{anim})

52. meat/*animal STC *sya (181)
 53. *bird STC *bya (177), TSR *s-ŋak (141)
 54. *dog STC *kwiŋ (159)
 55. *fish STC *ŋya (189)
 56. *louse STC *s-rik (439), *sar *śar
 (pp. 15, 53, 84, 147, 172,
 189)
 57. *snake STC *b-ru•l (447)
 58. frog STC *s-bal (15, 21, 107)
 59. insect, bug, vermin STC *buw (27)
 60. bee PLB *bya₂; STC *kwa•y (157),
 *was (17), *taŋ (494) [cf.
 BIRD #53]
 61. dove STC *m-kruw (118), *kuy
 'pigeon' (495)
 62. monkey STC *mruk (pp. 43, 112), *woy
 (314); PLB *m-yuk (JAM
 1973e)
 63. pig STC *pwak (43); PLB *wak (TSR
 168)

64.	fowl	STC *k-rak (pp. 88, 107, 187-9)
65.	otter	STC *s-ram (438)
66.	horse	STC *s-ran ~ *m-ran (145)
67.	ant	STC *rwak (199)
68.	bear	STC *d-wam (461)
68a.	leech	STC *rpat
68b.	water leech	STC *m-li·t
69.	rat/rodent	TSR *k-r-wak (188); STC *bwiy (173) 'bamboo rat,' STC *b-yuw (93) 'rat, rabbit'

E. Natural objects or phenomena; the inanimate landscape;
vegetable and mineral kingdoms (N_{nat})

70.	*ashes	STC *pla (137)
71.	*cloud	PLB *C-tim ₁ (Bradley 1975a, #320-2)
72.	*earth	STC *r-ka (97), *mliy (152)
73.	*fire	STC *bar × *par (220), *mey (290)
74.	*flower	STC *ba•r (1)
75.	*fruit	STC *sey (57)
76.	*grass	STC *mrak (149)
77.	*leaf	STC *la (486), *lap (321), *pak (40)
78.	*moon	STC *s-la × *g-la (144)
79.	*mountain	PLB *kaŋ ₁ (Bradley 1975a, #312)
80.	*rain	STC *r-wa (443)

81. *river/valley · STC *klu·ŋ (127), *kor (349)
82. *road STC *lam (87)
83. *root STC *bul × *pul (pp. 166, 173),
*r-sa (442)
84. *salt STC *g-ryum (245), *tsa (214)
85. *sky STC *muw (488)
86. *smoke STC *kuw (256)
87. *star STC *s-kar (49)
88. *stick PLB *da₁ (cf. Lahu á-tà)
89. *stone STC *brak (134) 'rock,' *r-lun
(88), PLB *k-lok × *k-lon (TSR 190)
90. *sun/day STC *nam (48), *niy (81),
*tsyar (187)
91. *tree/wood STC *siŋ (233); PLB *sik (TSR
118)
92. *water STC *twiy (168), *m-téril (p.
30)
93. *wind STC *g-liy (454)
94. branch STC *ka·k (327), *ku·ŋ (359)
95. silver STC *d-ŋul (pp. 15, 173)
96. bamboo STC *g-pa (44)
97. shade/shadow STC *g-rip × *s-rip (p. 113)
98. joint STC *tsik (64)
99. thorn (prick) STC *tsow (276)
100. *night STC *ya (417)
101. iron STC *si·r (372), *syam (228)

102. field

102a. swidden PLB *hya₁ (Bradley 1975a, #414)

102b. irrigated paddy field PLB *ʔdan₃; *C-mi₁ (Bradley 1975a, #415)

F. Artifacts and social organization (N_{art})

103. arrow STC *m-da (pp. 96, 111-112, 118),
*b-la (449)

104. needle STC *kap (52); PLB *k-rap (TSR 191)

105. house STC *kim * *kyim * *kyum (53)

106. bow STC *d-liy (463)

107. boat STC *m-liy (474)

108. mortar STC *tsum (75)

109. village STC *r-wa * *g-wa (444), TSR
*kak (22)

G. Spatial/directional (N_{spat})

110. *leftside STC *bay (47)

111. *rightside STC *g-ya * *g-ra (98)

112. *far (v.) STC *dzya•l (229), *wiy₂ (PLB)

113. *near (v.) STC *ney (291)

114. *year STC *nin (368)

H. Numerals and quantifiers (N_{num})

115. twenty/score STC *(m)-kul (397)

116. *one STC *it (pp. 94, 162), *kat (p. 94).
*t(y)ik * *(g-)tyik (pp. 84, 94,
169, 189)

117. seven STC *s-nis (5) [cf. TWO #120]
 118. ten STC *gip (16)
 119. hundred STC *r-gya (164)
 120. *two STC *g-nis (4)
 121. *three STC *g-sum (409)
 122. *four STC *b-liy (410)
 123. *five STC *l-na * *b-na (78)
 124. six STC *d-ruk (411); PLB *C-krok (TSR 35)
 125. eight STC *b-r-gyat (163)
 126. nine STC *d-kuw (13)
 127. (be) many (v.) STC *mra (148)

I. Verbs of utterance, body position or function (V_{body})

128. be born STC *bran (135), *krun (382)
 129. *sleep/lie down STC *ip (114), *mwi (196),
 *n(y)it (236)
 130. weep STC *krap (116), *nuw (79)
 131. *laugh STC *m-nwi(y) (191), *rya-t
 (202)
 132. *die STC *siy (232)
 133. awaken STC *m-sow (295); PLB *nuw₂ *
 *s-nuw₂
 134. cough STC *su(w) (423)
 135. *stand STC *g-ryap (246)
 136. *sit STC *tu•η (361)

J. Verbs of motion (V_{mot})

137. *fall (from a
 height) STC *kla (123)

138. climb, ascend STC *l-tak (pp. 52, 110, 123),
 *syar 'rise' (n. 90, p. 28)
139. descend STC *yu(w) (n. 289, p. 101);
 *ʔ-yuk [above 1.121(5E)];
 *zak (TSR 121)
140. fly STC *pur * *pir (398), *pyam
 (n. 93, p. 29; n. 171, p. 51),
 *pyaw 'fly, swim, float' (176)
141. hide STC *p(w)ak (46); PLB *wak *
 *ʔ-wak (TSR 178)
142. run/flee STC *plon (140)
143. emerge STC *twak (p. 17), *pro (248)

K. Verbs of emotion, cognition, perception (V_{psych})

144. *fear/frighten STC *grok * *krok (473), *kri(y)
 (416), *b-ray (450)
145. *know STC *m-kyen (223), *syey (182)
146. ashamed STC *kyen (162), *s-rak (431),
 *g-yak (452)
147. forget STC *b-la•p (335)
148. dream STC *man (82); PLB *s-mak * *s-man (TSR 144)
149. *see STC *mran (149)
150. *smell STC *m-nam (464)

L. Stative verbs with human patients ($V_{\text{hum stat}}$)

151. *thin STC *ba (25)
152. *old STC *r-ga (445)
153. alive STC *krun (382), *s-rin *
 *s-ran (404)

171. *thick STC *r-ta-t (426), *tow (319),
 *tu•k (356)
172. *new STC *sar (pp. 147, 172, 189)
173. *sharp STC *s-ryam (pp. 53, 171, 189),
 *tak (STC p. 87; TSR 41)
174. lightweight STC *r-ya•ŋ

N. Action verbs with human agent ($V_{\text{hum act}}$)

175. *eat STC *am (481), *dza (66)
176. *drink STC *am (481); PLB *Ndan₁ (TSR
 p. 15)
177. *give STC *biy (427), *pek (pp. 101,
 149)
178. *tie STC *du-t * *tu-t (421), *kik
 (484)
179. steal STC *r-kuw (33)
180. lick STC *m/s-lyak (211)
181. *bite STC *gam (491), *hap (89), *ku•k
 (388), *wa (424)
182. *scratch/scrape STC *hyak (230); *krak * *Nkrak
 (TSR 96); *kret * *Nkret
 (TSR 97); *kut (STC 383),
 *pruk (STC 391)
183. cook/boil STC *klak (124), *prut (131),
 *pryo (250)
184. grind STC *krit (119)
185. *wash STC *kruw (117), *m-s(y)il (492)

186. *dig STC *klaw (269), *r-ko-t (420),
 *la•y (288), *tu (258)
187. let go; set free;
 loosen STC *g-lwat (209)
188. extinguish STC *mit (374)
189. *blow STC *s-mut (p. 75)
190. buy STC *b-rey (293)
191. *sew STC *d-rup (456), TSR (63) *(?-) grup
 * *ʔdrup; *byar * *pyar
 (STC 178)
192. *kill STC *g-sat (58)
193. weave STC *tak (17)
194. *rub STC *nu•l (365), *s(y)wiy (180)
195. *squeeze STC *nyap (192), *tsyur (188)
196. shoot STC *ga•p (219)
197. kick PLB *tek (TSR 14)
198. sell STC *par (p. 35), *ywar (pp.
 15, 51, 89)
199. put, place STC *s-ta (19)
200. drive/*hunt PLB *rak * *Ngak (TSR 162)
- 200a. *burn STC *bar * *par (220), *plon
 (139), *tsow (275), *ka•n
 (330), *duk * *ʔduk (TSR 62),
 *put (TSR 8)
- 200b. *cut STC *da•n (22), *lep (351),
 *mrak (147), *ra-t (458),

*ri•t (371), *tsywar (240),
*kut (383), *tsyat (185),
*tuk (387); TSR *Ntök *
*?tök (101), *?twap *
*C-dwap (69)

VERMIFORM APPENDIX II

List of Source Languages Used in this Study

I. Major sources

The following languages furnished the bulk of the data on which this monograph is based. For the subgroup affiliations of these and other TB languages, see Matisoff 1974b.

After the name of the language, and its abbreviation, the author of the principal data-source is given in parentheses. Full references to these sources are to be found in the Bibliography.

Abor-Miri	AM	(J. H. Lorrain 1907)
Akha	Ak.	(Lewis 1968)
Angami Naga	Ang.	(Burling 1960; Vikuosa Nienu [native informant, Berkeley 1974-75])
Bantawa (= Bontawa)	Bt.	(Michailovsky 1974)
Bisu	Bi.	(Nishida 1966-67)
Boro (= Bodo)	Bo.	(Burling 1959; Bhat 1968)
Burmese	Bs.	
Written Burmese	WB	(Anonymous ca. 1941; Judson 1893, 1956)
Chepeng	Cp.	(CSDPN: Bhobikan/Caughley 1973)
Chinese	Ch.	(Benedict 1972; Chou 1972; Karlsgren 1923, 1933, 1954, 1957)
Old Chinese	OC	(= "Archaic Chinese")
Middle Chinese	MC	(= "Ancient Chinese")

Garó	G.	(Burling 1961)
Gurung	Gu.	(CSDPN: Deu Bahadur/Glover 1973)
Hayu (= Vayu)	Hy.	(Michailovsky 1974)
Jinghpaw (= Kachin)	Jg.	(Hanson 1906; Maran [in preparation])
Jirel	Jir.	(CSDPN: Olak Bahadur, Strahm, Maibaum 1973)
Jyarung (= Gyarong)	Jyar.	(Chang and Chang 1974)
Kaike	Kk.	(CSDPN: Tagla Tsering Buhda/Fisher 1973)
Kanauri	Knr.	(Bailey 1911)
Karen	K.	(Jones 1961)
Palaychi	(K _{Pal})	Bassein Sgaw (K _{BS})
Bassein Pho	(K _{BP})	Moulmein Sgaw (K _{MS})
Moulmein Pho	(K _{MP})	Taungthu (K _{Tg})
Khaling	Kl.	(CSDPN: Sapta Man/Toba 1973)
Kham	Kham	(CSDPN: Hasta Ram Buhda/Watters 1973)
Lahu	Lh.	(Matisoff 1965-1976; A. R. Walker 1970; D. Bradley 1975)
Lakher	Lk.	(R. A. Lorrain 1951)
Lepcha	Lp.	(Mainwaring/Grünwedel 1898)
Limbu	Lb.	(Iman Singh Chemjong, n.d.)
Lisu	Ls.	(Fraser 1922; Hope 1972; Thurgood 1974a)
Lushai	Lu.	(Lorrain and Savidge 1898)

Magar(i)	Mg.	(CSDPN: Krishna Bahadur Thapa/ Shepherd 1973)
Meithei (= Manipuri)	Me.	(Primrose 1888)
Mikir	Mk.	(G. D. Walker 1925)
Mru	Mru	(Löffler 1966)
Newari	Nw.	(CSDPN: Thakurlal Manandhar/ Hale 1973)
Phunoi	Phu.	(Bradley 1973)
Sherpa	Sh.	(CSDPN: Ang Nyima Lama/ Schoettelndreyer 1973)
Sunwar (= Sunawari)	Sun.	(CSDPN: Gyen Singh, Bieri, Schulze 1973)
Tamang	Tm.	(CSDPN: Karna Bahadur, Taylor, Everitt 1973; Mazaudon 1973, 1974a)
Tangkhul Naga (= Tankhur)	TN	(Pettigrew 1918; Bhat 1969)
Taraon (= Digaro, Digaru)	Tar.	(Chakravarty <i>et al.</i> 1963)
Thakali	Tk.	(CSDPN: Narendra Garchan/Maria Hari 1973)
Thulung Rai	TR	(Agam Sing Dewsa Rai 1944)
Tibetan	Tb	
Written Tibetan	WT	(Jäschke 1881)
Tiddim Chin	TC	(Henderson 1965)
Trung	Tr.	(Lo 1945)

II. Supplementary sources

A. Bodman ca. 1965-69.

Brief wordlists of the following languages: Adi,
Ao Changki, Chang, Chungli, Dafla, Jögli (= Yögli),
Kamhau Tiddim, Konyak, Lhota, Limbu, Lungchang,
Lungmi, Meyöl, Miju, Moshung, Mon-pa/Tsangla,
Mongsen, Nokmung, Rawang, Rongrang, Tikhak.

B. Burling 1959.

Atong, Bodo, Garo, Kachari, Wanang.

C. Löffler 1964.

Kadu, Lui, Sak.

D. Matisoff 1972a.

Ahi, Ch'iang, Hani, Lu-chüan Lolo, Moso/Nakhi, Nasu,
Sani, Woni.

E. Nagaland Bhasha Parishad.

See Bibliography for list of 18 dictionaries published by the NBP.

F. Ono 1965.

Anal, Chinbok, Kamhau Tiddim, Khumi, Lai, Laizo,
Ngawn, Zotung.

VERMIFORM APPENDIX III

Index of New Terms Introduced

Unless otherwise stated, these terms were first introduced in this monograph.

{allocomp {allocompy	1.12
{allofam {allofamy	1.12
compound family	1.12
diffuse organs	2.5
{heterofam {heterofamy	1.141(a)
heterorganizing	1.121(3)
{isofam {isofamy	1.141(a)
{metastasis (of meaning) {metastatic flowchart	2.43
morphan	Matisoff 1973d: 3.333
nominal hemistich	Matisoff 1973d: 2.0, 3.48, 3.611c, 3.10.0
obscure internal channels and connections (OICC's)	2.41d
organic semantics	2.0 <i>et seq.</i>
preemption (prefixial)	Matisoff 1972d, 1973e
replacement tolerance quotient (RTQ)	1.14b

semantic field	2.1 [see note 187]
semantic system	2.1 [see note 187]
somatic area	2.5
tonogenesis	Matisoff 1970, 1973c
trans-field association	2.42
wittle wabbit syndwome	1.121(6B)

VERMIFORM APPENDIX IV

Index of Symbols and Abbreviations

A. Phonological, grammatical, semantic

$X \times Y$	X and Y are co-allofams (1.12)
$X \text{---} Y$	there is an intra-field semantic association between X and Y (2.4)
$X \text{---}^e Y$	there is a trans-field semantic association between X and Y (2.4)
$X \text{~} Y$	there is an antonymic opposition between X and Y (2.4)
C_f	final consonant
C_i	initial consonant
\emptyset -	voiced obstruent prefix hypothesized to account for shift of syllables into HIGH-checked tonal category of Loloish [TSR p. 33].
G	glide /-w-, -y-, -r-, -l-/
N_{bp}	body-part noun
N_{rh}	relative head; noun that is serving as the head of a relative clause [GL 6.42]

v_h	head nucleus of a genitive construction [GL 3.7]
v_p	possessor nucleus of a genitive construction [GL 3.7]
P	prefix
RC	relative clause
RTQ	replacement tolerance quotient (vis-à-vis loanwords) [1.14(b)]
V	vowel; verb

B. Names of Proto-languages and Language-families

[for abbreviations for names of individual languages,
see Appendix II]

AA	Austroasiatic
PAA or pAA	proto-Austroasiatic
AN	Austronesian
PAN or pAN	proto-Austronesian
AT	Austro-Thai [Benedict 1975a]
PAT or pAT	proto-Austro-Thai
IE	Indo-European
PIE or pIE	proto-Indo-European
LB	Lolo-Burmese
PLB or pLB	proto-Lolo-Burmese
MY	Miao-Yao
PMY or pMY	proto-Miao-Yao
ST	Sino-Tibetan
PST or pST	proto-Sino-Tibetan

TB	Tibeto-Burman
PTB or pTB	proto-Tibeto-Burman

C. Names of works cited

ATLC	<i>Austro-Thai Language and Culture</i> see Benedict 1975a
CSDPN	<i>Clause, Sentence, and Discourse Patterns in Selected Languages of Nepal</i> see Hale, ed. 1973
GD	"Glottal dissimilation and the Lahu high-rising tone: a tonogenic case-study" see Matisoff 1970
GL	<i>The Grammar of Lahu</i> see Matisoff 1973d
GSR	<i>Grammata Serica Recensa</i> see Karlgren 1957
LSI	<i>Linguistic Survey of India</i> see Grierson 1903-1928
OED	<i>Oxford English Dictionary</i>
STAL	"Sino-Tibetan: another look" see Benedict 1975b
STC	<i>Sino-Tibetan: a Conspectus</i> see Benedict 1972
STL	<i>Sino-Tibetan Linguistics</i> see Shafer and Benedict, 1939-41

TSR *The Loloish Tonal Split Re-visited*

see Matisoff 1972a

D. Names of scholarly journals

AA	<i>American Anthropologist</i>
ALH	<i>Acta Linguistica Hafniensia</i> (Copenhagen)
AO	<i>Acta Orientalia</i> (Copenhagen)
BIHP	<i>Bulletin of the Institute of History and Philology, Academia Sinica</i> (Taipei)
BMFA	<i>Bulletin of the Museum of Far Eastern Antiquities</i> (Stockholm)
BSLP	<i>Bulletin de la Société de Linguistique de Paris</i>
BSOAS	<i>Bulletin of the School of Oriental and African Studies</i> (London)
GK	<i>Gengo Kenkyū</i> (Tokyo)
HJAS	<i>Harvard Journal of Asiatic Studies</i>
IAE	<i>Internationale Archive der Ethnographie</i> (Zürich)
IJAL	<i>International Journal of American Linguistics</i>
JAOS	<i>Journal of the American Oriental Society</i>

JAS	<i>Journal of Asian Studies</i>
JCL	<i>Journal of Chinese Linguistics</i> (Berkeley)
JRAS	<i>Journal of the Royal Asiatic Society</i>
Lg	<i>Language</i>
LTBA	<i>Linguistics of the Tibeto-Burman Area</i> (Berkeley)
NBP	<i>Nagaland Bhasha Parishad</i> [<i>Linguistic Circle of Nagaland</i>] (Kohima, Nagaland, India)
OPWSTBL	<i>Occasional Papers of the Wolfenden Society on Tibeto-Burman Linguistics</i> (formerly Ann Arbor and Urbana; now Philadelphia)
TAK	<i>Tōnan Azia Kenkyū</i> (Kyoto)
TG	<i>Tōyō Gakuhō</i> (Tokyo)
ZDMG	<i>Zeitschrift der deutschen morgenländischen Gesellschaft</i> (Wiesbaden)

BIBLIOGRAPHY

- Allen, Nicholas J. 1975. *Sketch of Thulung Grammar with three texts and a glossary*. Cornell University East Asia Papers, No. 6. Ithaca, New York.
- Aman, Reinhold A. 1973. *Bayrisch-österreichisches Schimpfwörterbuch*. Munich.
- Anonymous. ca. 1941. *Rhyming Dictionary of Written Burmese*. Compiled under the supervision of Paul K. Benedict. Unpublished typescript, 91 pp. Published in *LTBA* 3.1, x and 93 pages (1976-77).
- Anonymous. 1959. *An Outline of Lisu Grammar* [in Chinese: *Li-su yü yü-fa kang-yao*]. Publication of the Minority Language Bureau of the Chinese Academy of Sciences. Scientific Publishing Co., Peking.
- Aristotle. ca. 340 B.C. *Posterior Analytics*. Cf. Richard McKeon, ed., *Introduction to Aristotle*. Modern Library, New York.
- Asimov, Isaac. 1963. *The Human Body: its Structure and Operation*. Illustrated by Anthony Ravielli. Reprinted as Signet Science Library paperback (New American Library). Hardcover edition published by Houghton Mifflin Co.
- Bailey, T. Grahame. 1911. *Kanauri Vocabulary in Two Parts: English-Kanauri and Kanauri-English*. Royal Asiatic Society Monographs, vol. XIII. London. [Also appeared in *JRAS* (1910), pp. 695-705; (1911), pp. 315-364.]
- Becker, Alton L. 1973. "Person in Kawi: exploration of an elementary semantic dimension." Paper presented at First International Austronesian Conference, Honolulu.
- . 1974. "A linguistic image of nature: the Burmese numerative classifier system." Mimeo, 19 pp.

- Bee, Peter. 1975. "Restricted phonology in certain Thai linker-syllables." In Jimmy R. Harris and James R. Chamberlain, eds., *Studies in Tai Linguistics in Honor of William J. Gedney* (Bangkok), pp. 17-32.
- Benedict, Paul K. 1939. "Semantic differentiation in Indo-Chinese." *HJAS* 4.3-4, 213-229.
- . 1941. "Kinship in Southeastern Asia." Unpublished doctoral dissertation, Harvard University, Dept. of Anthropology. 526 pp.
- . 1942. "Thai, Kadai, and Indonesian: a new alignment in Southeastern Asia." *American Anthropologist*, n.s., 44.576-601.
- . 1943. "Secondary infixation in Lepcha." *Studies in Linguistics* 1.19.
- . 1948. "Archaic Chinese *g and *d." *HJAS* 11.1-2, 197-206.
- . 1967a. "Austro-Thai studies: material culture and kinship terms." *Behavior Science Notes* 2.203-244.
- . 1967b. "Austro-Thai studies: Austro-Thai and Chinese." *Behavior Science Notes* 2.275-336.
- . 1972. *Sino-Tibetan: a Conspectus*. ["STC"]. Contributing editor, James A. Matisoff. Princeton-Cambridge Studies in Chinese Linguistics. Cambridge University Press, Cambridge (U.K.).
- . 1975a. *Austro-Thai Language and Culture, with a glossary of roots*. ["ATLC"]. Human Relations Area Files Press, New Haven.
- . 1976. 1975b. "Sino-Tibetan: another look." ["STAL"]. *JAOS* 96.2, pp. 167-197 (1976).
- Bergenholtz, Henning. 1975. "Zur Wortfeldterminologie." *Muttersprache* 85 (July-August), pp. 278-285.
- Bhat, D. N. Shankara. 1968. *Boro Vocabulary, with a grammatical sketch*. Deccan College Postgraduate and Research Institute, Poona.

- . 1969. *Tankhur Naga Vocabulary*. Deccan College Postgraduate and Research Institute, Poona.
- Bloomfield, Leonard. 1933. *Language*. Holt, Rinehart, and Winston, New York.
- Bodman, Nicholas C. ca. 1965-69. [Brief wordlists from 21 languages of the Kuki-Chin-Naga, Abor-Miri, Dafla, Nungish, and Bahing-Vayu groups. See Appendix II.]
- . 1969. "Tibetan sdud 'folds of a garment,' the character 卒, and the *ST-hypothesis." *BIHP*, Academia Sinica, Vol. 39, Part 2. Taipei.
- . 1971. "Some phonological correspondences between Chinese and Tibetan." Paper presented at Fourth Sino-Tibetan Conference, Bloomington, Indiana.
- . 1973. "Some Chinese reflexes of Sino-Tibetan s-clusters." *Journal of Chinese Linguistics* 1.3, 383-396.
- Boltz, William G. 1974. "Studies in Old Chinese word families." Doctoral dissertation, University of California, Berkeley. 166 pp.
- Boodberg, Peter. 1934. "The morphology of final -n and -t." *Notes on Chinese Morphology and Syntax, III*. Berkeley.
- . 1937. "Some proleptical remarks on the evolution of Archaic Chinese." *HJAS* 2.3-4, 329-372.
- Bradley, David. 1969a. "Two Akha dialects in Thailand: Aru and Saenchai." 5 pp. ms.
- . 1969b. "Problems in Akha phonology: synchronic and diachronic." 42 pp. ms.
- . 1970. "Comparison of Nishida's and Lewis' Akha materials." 20 pp. ms.
- . 1973. "Phunoi today." Paper presented at Sixth Sino-Tibetan Conference, San Diego. To appear in T'sou, ed., forthcoming.
- . 1975a. *Lahu Dialects and Proto-Loloish*. University

- of London (School of Oriental and African Studies) doctoral dissertation. 541 pp.
- . 1975b. "Nahsi and Proto-Burmese-Lolo." *LTBA* 2.1, 93-150.
- Bright, William. n.d. "An English-Lushai word list." Mimeo, 7 pp.
- Buck, Carl Darling. 1949. *A Dictionary of Selected Synonyms in the Principal Indo-European Languages: a contribution to the history of ideas*. University of Chicago Press, Chicago.
- Burling, Robbins. 1959. "Proto-Bodo." *Language* 35.3, 433-453.
- . 1960. "Angami Naga phonemics and word list." *Indian Linguistics* 21, pp. 51-60.
- . 1961. *A Garo Grammar*. Deccan College Monograph Series, No. 25. Poona.
- . 1967. *Proto-Lolo-Burmese*. *IJAL* 33.2, Part II. Issued simultaneously as Publication 43 of Indiana Publications in Anthropology and Linguistics, Mouton and Co., 1967.
- . 1971. "The historical place of Jinghpaw in Tibeto-Burman." *OPWSTBL*, Vol. II, 1-54. Urbana, Illinois.
- Chakravarty, L. N., et al. 1963. *A Dictionary of the Taron Language*. Produced for the use of officers of the NEFA administration by the Philology Section, Research Department, Northeast Frontier Agency, Shillong.
- Chang Kun. 1973. Review of Paul K. Benedict, *Sino-Tibetan: a Conspectus*. *JAS* 32.2, 335-337.
- Chang Kun and Betty Shefts Chang. 1974. "Gyarong historical phonology." Paper circulated at Seventh Sino-Tibetan Conference, Atlanta.
- Chemjong, Iman Singh. n.d. [ca. 1961?]. *Limbu-Nepali-English Dictionary* [in Nāgarī]. Kathmandu. [4 pp. of body-part words, transliterated by Boyd]

- Michailovsky.]
- Chou Fa-kao. 1972. "Archaic Chinese and Sino-Tibetan."
[in Chinese: "Shang-ku Han-yü ho Han-Tsang-yü].
*Journal of the Institute of Chinese Studies of the
Chinese University of Hongkong* 5.1, 159-237. [A
review article on Benedict 1972. Contains an
important glossary of Benedict's Tibeto-Burman/
Chinese comparisons.]
- Douglas, M. 1971. "Do dogs laugh? A cross-cultural
approach to body symbolism." *Journal of Psycho-
somatic Research*, Vol. 15, pp. 387-390.
- Egerod, Søren. 1973. Review of Paul K. Benedict, *Sino-
Tibetan: a Conspectus*. *Journal of Chinese Linguis-
tics* 1.3, 498-505.
- Emeneau, Murray B. 1956. "India as a linguistic area."
Language 32, 3-16.
- . 1974a. "The Indian linguistic area revisited."
International Journal of Dravidian Linguistics 3.
- . 1974b. "Arm and leg in the Indian linguistic area."
19 pp. ms.
- Fraser, J. O. 1922. *Handbook of the Lisu (Yawyin) Language*.
Rangoon.
- Geckeler, Horst. 1971. *Zur Wortfelddiskussion: Unter-
suchungen zur Gliederung des Wortfeldes 'alt-jung-
neu' im heutigen Französisch*. Wilhelm Fink Verlag.
Munich. [Reviewed in Malkiel 1974 (q.v.).]
- Gray, Henry. 1901. *Anatomy, Descriptive and Surgical*.
Reprinted 1974 by Running Press, Philadelphia
(paperback, 1257 pp.).
- Greenberg, Joseph H. 1963. "Some universals of grammar,
with particular reference to the order of meaning-
ful elements." In J. H. Greenberg, ed., *Universals
of Language*. MIT Press.
- Grierson, Sir George A. 1903-1928. *Linguistic Survey of
India* ["LSI"]. Esp. Vol. III, Parts 1, 2, 3:

- "Tibeto-Burman Family." Calcutta. Reprinted by Motilal Banarsidass 1967 (Delhi, Varanasi, Patna). *Guinness Book of World Records*. 1970. Revised new Bantam edition, Fifth printing.
- Hale, Austin, ed. 1973. *Clause, Sentence, and Discourse Patterns in Selected Languages of Nepal*. ["CSDPN"]. Four volumes: I. *General Approach*. II (with David Watters) *Clause*. III. *Texts*. IV. *Word Lists*. [Word lists contain forms from Nepali, Gurung, Tamang, Thakali, Magar(i), Kham, Kaike, Jirel, Sherpa, Sunwar, Khaling, Newari, Chepang.] Publication of the Summer Institute of Linguistics. University Press, Tribhuvan University, Kathmandu.
- Hamp, Eric P. 1975. "A functional view of bodily functions." In *Papers from the Parasession on Functionalism*. Chicago Lgc. Society, pp. 209-212.
- Hanson, Ola. 1906. *A Dictionary of the Kachin Language*. Rangoon. Reprinted by Baptist Board of Publications, Rangoon 1954. [This work is now being revised and enlarged by LaRaw Maran.]
- Haudricourt, André-Georges. 1954a. "De l'origine des tons en vietnamien." *Journal asiatique* 242.68-82.
- . 1954b. "Comment reconstruire le chinois archaïque." *Word* 10,2-3, pp. 351-364.
- Henderson, Eugénie J. A. 1965. *Tiddim Chin: a Descriptive Analysis of Two Texts*. Oxford University Press, London.
- Hertz, Robert. 1909. "La prééminence de la main droite: étude sur la polarité religieuse." *Revue philosophique*. Translated by Rodney Needham, in Needham, ed. 1973.
- Hockett, Charles F. 1955. *A Manual of Phonology*. IJAL 21.4, Part II.
- Hodgson, B. H. 1847. "Essay on the Kocch, Bódó, and Dhimál Tribes." Calcutta. Reprinted in *Essays Relating*

- to *Indian Subjects*, Vol. I, Section 1 (esp. Part I, "Vocabulary," pp. 7-8). Trübner and Co., London 1880.
- Hoenigswald, Henry. 1960. *Language Change and Linguistic Reconstruction*. University of Chicago Press, Chicago.
- Hope, Edward R. 1972. "The deep syntax of Lisu sentences: a transformational case-grammar." Ph.D. thesis, Australian National University, Canberra. To appear in *Pacific Linguistics*, Series B, No. 34, Canberra (1974).
- Huffman, Franklin E. 1973. "Thai and Cambodian: a case of syntactic borrowing?" *JAOS* 93.4, 488-509.
- Hunn, Eugene. 1975. "Cognitive processes in folk ornithology: the identification of gulls." Working paper No. 42, Language Behavior Research Laboratory, University of California, Berkeley. 73 pp.
- Huttar, George L. 1975. "Sources of creole semantic structures." *Language* 51.3, 684-695.
- Jäschke, H. A. 1881. *A Tibetan-English Dictionary*. London. Reprinted by Routledge and Kegan Paul Ltd., London 1958.
- Jones, Robert B. 1961. *Karen Linguistic Studies: description, comparison, and texts*. University of California Publications in Linguistics 25, Berkeley and Los Angeles.
- Judson, Adoniram. 1893. *Burmese-English Dictionary*. Rangoon. Reprinted by Baptist Board of Publications, Rangoon 1966.
- . 1956. *English and Burmese Dictionary*. 9th edition. Baptist Board of Publications, Rangoon.
- Kahane, Henry R. 1941. "Designation of the cheek in the Italian dialects." *Language* 17, pp. 212-222.
- Kappas, Attallah and Alvito P. Alvares. 1975. "How the liver metabolizes foreign substances." *Scientific American* 232.6, pp. 22-31.

- Karlgren, Bernhard. 1923. *Analytic Dictionary of Chinese and Sino-Japanese*. Paris.
- . 1933. "Word families in Chinese." *BMFEA* 5, 5-120.
- . 1954. *Compendium of Phonetics in Ancient and Archaic Chinese*. *BMFEA Bulletin* 22, Stockholm, pp. 211-367.
- . 1957. *Grammata Serica Recensa* ["GSR"]. *BMFEA Bulletin* 29, Stockholm. 332 pp.
- Kronfeld, Chana. 1976. "Body-part terms in Yiddish: a tentative list." 7 pp. ms. UC Berkeley course paper.
- Lehmann, Winfred P. 1962. *Historical Linguistics: an Introduction*. Holt, Rinehart and Winston, New York.
- Lehrer, Adrienne. 1974. "Extended meanings of body-part terms." In Notes and Reviews section of *IJAL* 40.2, 135-137.
- Lévi-Strauss, Claude. 1964. *Le cru et le cuit*. Librairie Plon, Paris. (Translated as *The Raw and the Cooked*. Harper Torchbook TB 1487, Harper and Row. New York and Evanston 1970).
- Lewis, Gilbert. 1974. "Gnau anatomy and vocabulary for illnesses." *Oceania* 45.1, 50-78.
- Lewis, Paul. 1968. *Akha-English Dictionary*. Data Paper 70, Southeast Asia Program, Cornell University, Ithaca, New York.
- Li, Charles N. and Sandra A. Thompson. 1973a. "Co-verbs in Mandarin Chinese: verbs or prepositions?" Paper presented at Sixth Sino-Tibetan Conference, San Diego.
- . 1973b. "Historical change of word-order: a case study in Chinese and its implications." 26 pp. ms.
- Li Fang-kuei. 1933. "Certain phonetic influences of the Tibetan prefixes upon the root initials."

- BIHP 4.135-157.
- Lo Ch'ang-p'ei. 1945. "A preliminary study of the Trung language of Kung Shan." *HJAS* 8.3-4, 343-348.
- Löffler, Lorenz G. 1964. "Chakma und Sak: ethnolinguistische Beiträge zur Geschichte eines Volkes." *Internationale Archive der Ethnographie* 50.1, 72-115. [Contains forms from Sak, Kadu, Lui.]
- . 1966. "The contribution of Mru to Sino-Tibetan linguistics." *ZDMG* 116.1, 118-159.
- Lombard, Sylvia. 1968. *Yao-English Dictionary*. Ed. by Herbert C. Purnell, Jr. Data Paper 69, Southeast Asia Program, Cornell University, Ithaca, New York.
- Lorrain, J. Herbert. 1907. *A Dictionary of the Abor-Miri Language*. Shillong.
- and F. W. Savidge. 1898. *A Grammar and Dictionary of the Lushai Language (Dulien Dialect)*. Shillong.
- Lorrain, Reginald Arthur. 1951. *Grammar and Dictionary of the Lakher or Mara Language*. Shillong.
- Mainwaring, G. B. and A. Grünwedel. 1898. *Dictionary of the Lepcha Language*. Berlin.
- Malkiel, Yakov. 1944. "The etymology of Portuguese iguaria 'liver'." *Language* 20.1, 108-130.
- . 1974. Review of Geckeler 1971 (q.v.). *Foundations of Language* 12, 271-285.
- Maran, LaRaw. [in preparation]. *A Dictionary of Modern Jinghpaw*. [A revised and enlarged version of Hanson 1906 (q.v.)]
- Matisoff, James A. 1957. "Rabelais est-il obscène?" Harvard College term paper, Seminar on Rabelais. 86 pp. ms.
- . 1968. Review of Burling 1967 (q.v.). *Language* 44.4, 879-897.
- . 1969a. "Lahu and Proto-Lolo-Burmese." *OPWSTBL*, Vol. I, 117-221. Ann Arbor, Michigan.
- . 1969b. "Verb concatenation in Lahu: the syntax and

- semantics of 'simple' juxtaposition." *Acta Linguistica Hafniensia* (Copenhagen) 12.1, 69-120.
- . 1969c. "Lahu bilingual humor." *Acta Linguistica Hafniensia* (Copenhagen) 12.2, 171-206.
- . 1969d. Review of Paul Lewis, *Akha-English Dictionary*. *JAS* 28.3, 644-5.
- . 1970. "Glottal dissimilation and the Lahu high-rising tone: a tonogenetic case-study." ["GD"]. *JAOS* 90.1, 13-44.
- . 1972a. *The Loloish Tonal Split Revisited*. ["TSR"]. Research Monograph No. 7, Center for South and Southeast Asia Studies, University of California, Berkeley. 88 pp.
- . 1972b. "Lahu nominalization, relativization, and genitivization." In John Kimball, ed., *Syntax and Semantics*, Volume I, pp. 237-257. Seminar Press, New York.
- . 1972c. Review of D. N. S. Bhat, *Tankhur Naga Vocabulary*. *Language* 48.2, 476-9.
- . 1972d. "Tangkhul Naga and comparative Tibeto-Burman." *Tōnan Azia Kenkyū* (Kyoto) 10.2, 1-13.
- . 1973a. "Reconstructed Tibeto-Burman roots from STC, arranged by semantic category: I. Nouns." Handout for Linguistics 245, University of California, Berkeley. 9 pp.
- . 1973b. "The Mon-Khmer substratum in Tibeto-Burman." Talk presented at First International Conference on Austroasiatic Linguistics, Honolulu.
- . 1973c. "Tonogenesis in Southeast Asia." In Larry M. Hyman, ed., *Consonant Types and Tone*. Southern California Occasional Papers in Linguistics, No. 1, pp. 71-96. Los Angeles.
- . 1973d. *The Grammar of Lahu* ["GL"]. University of California Publications in Linguistics 75, U.C. Press, Berkeley and Los Angeles. li + 673 pp.

- . 1973e. "Problems and progress in Lolo-Burmese: Quo Vadimus?" Paper presented at Sixth Sino-Tibetan Conference, San Diego. To appear in T'sou, ed.
- . 1973f. "KIKE and CAKRAJAAN, or 'Why is a Jew like the Thai word for bicycle?': semantic shifts in language history." Handout distributed at Sixth Sino-Tibetan Conference, San Diego.
- . 1973g. "The annual Sino-Tibetan conferences: the first five years, 1968-1972." *Journal of Chinese Linguistics* 1.1, 152-162. [Contains bibliography of all papers presented during those five years.]
- . 1973h. "Notes on Fang-kuei Li's 'Languages and dialects of China.'" *Journal of Chinese Linguistics* 1.3, 471-4.
- . 1974a. "For some think the brain is really marrow." Talk presented at SEAPO [Southeast Asia and the Pacific Organization], University of California, Berkeley.
- . 1974b. "Alphabetical list of Tibeto-Burman languages, with their genetic affiliations." 41 pp. ms.
- . 1974c. "Tibeto-Burman organic relationships: body-part nomenclature and genetic subgrouping." Paper presented at Seventh Sino-Tibetan Conference, Atlanta.
- . 1974d. "Verb concatenation in Kachin." *Linguistics of the Tibeto-Burman Area* (Berkeley) 1.1, 186-207.
- . 1974e. "Psycho-ostensive expressions in Yiddish." Appeared in Hebrew translation as "Sifrut ha-bituy b'Yidish," in *Ha-sifrut* (Tel-Aviv), No. 18-19, 181-223. The English original is to appear under the title *Blessings, Curses, Hopes, and Fears: psycho-ostensive expressions in Yiddish*, Institute for

the Study of Human Issues, Philadelphia.

- _____. 1974f. "The tones of Jinghpaw and Lolo-Burmese: common origin vs. independent development." *Acta Linguistica Hafniensia* (Copenhagen) 15.2, 153-212.
 - _____. 1975a. "Benedict's Sino-Tibetan: a rejection of Miller's *Conspectus* inspection." *LTBA* 2.1, 155-172.
 - _____. 1975b. "A new Lahu simplex/causative pair: study/train." *LTBA* 2.1, 151-154.
 - _____. 1975c. "Rhinoglottophilia: the mysterious connection between nasality and glottality." In Ferguson, Hyman, and Ohala, eds., *Nasâlfest: Papers from a Symposium on Nasals and Nasalization*, Stanford University, Stanford, California, pp. 265-287.
 - _____. 1976a. "Austro-Thai and Sino-Tibetan: an examination of body-part contact relationships." In Mantaro J. Hashimoto, ed., *Genetic Relationship, Diffusion, and Typological Similarities of East and Southeast Asian Languages*, pp. 256-289. (Proceedings of 1st. Japan-U.S. Joint Seminar on East and SEA Linguistics, Shizuoka, Japan, July, 1976.) Japan Society for the Promotion of Science, Tokyo.
 - _____. 1976b. "Lahu causative constructions: case hierarchies and the morphology/syntax cycle in a Tibeto-Burman perspective." In Masayoshi Shibatani, ed., *The Grammar of Causative Constructions*, Academic Press, New York, pp. 413-442.
 - _____. [forthcoming]. "Winging it: arm and wing in Tibeto-Burman." To appear in *LTBA*.
 - _____. [in preparation]. *Lahu-English and English-Lahu Dictionary*.
- Mazaudon, Martine. 1973. *Phonologie tamang*. Société d'études linguistiques et anthropologiques de France "SELAF." Paris.
- _____. 1974a. "Body parts in the Risiangku dialect of Tamang."

- 2 pp. ms.
- . 1974b. "Notes on tone in Tibeto-Burman." *Linguistics of the Tibeto-Burman Area* (Berkeley) 1.1, 27-54.
- . 1974c. "Tibeto-Burman tonogenetics." 173 pp. ms. Published in *LTBA* 3.2 (1977: 123 pp.).
- Michailovsky, Boyd. 1974a. "Hayu body parts, in phonemic notation." 2 pp. ms.
- . 1974b. "Bantawa body parts." 2 pp. ms.
- and Martine Mazaudon. 1974. "Bahing body parts: Rāngadip village." 3 pp. mimeo.
- Miller, Roy Andrew. 1956. "The Tibeto-Burman ablaut system." *Transactions of the First International Conference of Orientalists in Japan*. Tōhō Gakkai. pp. 29-56.
- . 1958. "The Tibeto-Burman infix system." *JAOS* 78.3, 193-204.
- . 1970. Review of Robbins Burling, *Proto-Lolo-Burmese*. *Indo-Iranian Journal* 12.2, 146-159.
- . 1974. Review of Paul K. Benedict, *Sino-Tibetan: a Conspectus*. *JAOS* 94.2, 195-209.
- Miyaji, Atsuko. 1976. "From kokoro to shinzō: changes in body-part vocabulary." [In Japanese.] Author's English title: "On the change of usages of the words denoting heart." *Kokugogaku* 104.3, pp. 33-44.
- Mundhenk, Norman Arthur. 1968. "Words and reconstructions: Proto-Lushai-Chin." 11 pp. ms. [Compares forms from Henderson 1965 (Tiddim), Stern 1963 (Sizang), and Bright n.d. (Lushai).]
- Murray, I. MacKay. 1969. *Human Anatomy Made Simple*. Doubleday & Company, Garden City, N. Y.
- Nagaland Bhasha Parishad [Linguistic Circle of Nagaland]. ["NBP"]. (18 dictionaries of Kuki-Chin-Naga and

Barish languages, including Angami, Ao, Chakhesang, Chang, Garo, Kheza, "Konyak," "Kuki," Liangmai, Lotha, Mao, Phom, Pochury, Rengma, Sangtam, Sema, Yimchungrü (Yimchurr), and Zeliang.)

Needham, Rodney. 1973. *Right and Left: Essays on Dual Symbolic Classification*. University of Chicago Press, Chicago and London.

Neighbor, R. E. 1878. *A Vocabulary of English and Mikir*. Calcutta.

Nishi, Yoshio. 1974. "The Proto Lolo-Burmese and Old Burmese sources of Written Burmese -ac" [in Japanese: "Biruma bungo no -ac ni tsuite"]. *Tōyō Gakuhō* 56.1, 01-043.

Nishida, Tatsuo. 1966, 1967. "A comparative study of the Bisu, Akha, and Burmese languages, Parts I and II" [in Japanese: "Bisu-go no keitō," literally "The lineage of Bisu"]. *Tōnan Azia Kenkyū* 4.3 (Dec, 1966) and 4.5 (March 1967).

Okall 7 Ono, Tōru. 1965. "The reconstruction of Proto-Kuki-Chin (I): initial consonants" [in Japanese: "Kyōtsū Kuki-Chin-go no saikōsei (I): gotō shi-in."]. *Gengo Kenkyū* 47.8-20. [Contains forms from Tiddim Chin, Ngawn, Lai, Laizo, Anal, Zotung, Khumi, Chinbok.]

Ornstein, Robert. 1972. *The Psychology of Consciousness*. W. H. Freeman and Co. San Francisco.

Oxford English Dictionary. 1971. Compact Edition, 2 volumes. Complete text reproduced micrographically. Oxford University Press. ["OED"].

Palmer, L. R. 1955. *The Latin Language*. 2d. ed. Faber & Faber. London.

Paul, Robert A. 1973. "The eyes outnumber the nose two to one." Paper prepared for the symposium "Images of Body States," Annual Meeting of the American Anthropological Association, New Orleans. 12 pp.

- ms. + 4 figures.
- . 1975. "The Sherpa temple as a model of the psyche: symbolic analysis of Tibetan Buddhist iconography." Talk presented at University of California, Berkeley (November 10).
- Pettigrew, W. 1918. *Tāngkhul Nāga Grammar and Dictionary* (Ukhrul Dialect). Assam Secretariat Printing Office, Shillong.
- Primrose, A. J. 1888. *A Manipuri Grammar, Vocabulary, and Phrase Book*. Shillong. [Manipuri = Meithei]
- Pulleyblank, Edwin G. 1965. "Close/open ablaut in Sino-Tibetan." *Lingua* 14, 230-240.
- Purtle, Dale. 1971. "Some w(v)- roots in the languages of East and Southeast Asia." Paper presented at Fourth Sino-Tibetan Conference, Bloomington, Indiana.
- Rabelais, François. 1542. *La vie tres horrifique du grand Gargantua, pere de Pantagruel*. François Juste, Lyon. See *Les oeuvres de Maître François Rabelais* ..., Ch. Marty-Laveaux, ed. Vol. I. Alphonse Lemerre, Paris 1868.
- . 1552. *Le quart livre des faicts et dicts Heroiques du bon Pantagruel*. Michel Fezandat, Paris. See *ibid.*, Vol. II. Alphonse Lemerre, Paris 1870.
- Rahder, J. 1964. "Words for abdomen and entrails." *Zeits. für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 17.6, 609-20.
- Rai, Agam Sing Dewsa. 1944. *Thulung Rāī Bhāsā*. [The Thulung Rai Language]. In Nepali. Darjeeling. [For further details on this work, see Allen 1975, p. 2.]
- Roget, Peter Mark. 1852. *Thesaurus of English Words and Phrases*.
- Schieffelin, Edward L. 1976. *The Sorrow of the Lonely and the Burning of the Dancers*. St. Martin's Press, N.Y.

- Schmidt, Johannes. 1872. *Die Verwandtschaftsverhältnisse der indogermanischen Sprachen*. Weimar.
- Sčur, G. S. 1971. "Об основных теориях поля в языкознании" ["Ob osnovnyx teorijax polja v jazykoznanii" - On fundamental field theories in linguistics]. In N. E. Ul'janov, ed., *Вопросы филологии* [Voprosy filologii], pp. 81-126. Siberian Omsk State Pedagogical Institute. Cited in Malkiel 1974, p. 284.
- Seuss, Dr. 1961. *The Sneetches and Other Stories*. Random House, New York.
- Shafer, Robert. 1940, 1941. "The vocalism of Sino-Tibetan." Part I: *JAOS* 60.302-337. Part II: *JAOS* 61.18-31.
- . 1966, 1967. *Introduction to Sino-Tibetan*. Parts I and II. Otto Harrassowitz, Wiesbaden.
- and Paul K. Benedict. 1939-1941. *Sino-Tibetan Linguistics* ["STL"]. 14 ms. volumes. Unpublished bound typescript produced on a Work Projects Administration Grant, University of California, Berkeley.
- Shorto, Harry L. 1973. "Mon-Khmer contact words in Sino-Tibetan." Paper presented at First International Conference on Austroasiatic Linguistics, Honolulu.
- Simon, Walter. 1975. "Tibetan initial clusters of nasal and R." *Asia Major* 19.2, 246-251.
- Simons, G. L. 1975. *Simons' Book of World Sexual Records*. Pyramid Books, New York.
- Smith, Kenneth D. 1975. "The velar animal-prefix relic in Vietnam languages." *LTBA* 2.1, 1-18.
- Sprach-Brockhaus, Der: deutsches Bildwörterbuch für jedermann*. 1962. F. A. Brockhaus, Wiesbaden.
- Stern, Theodore. 1963. "A provisional sketch of Sizang (Siyin) Chin." *Asia Major* 10.228-278.
- Thomas, Lewis. 1975. *The Lives of a Cell: notes of a*

- biology watcher*. Bantam Books. Published by arrangement with Viking Press.
- Thurgood, Graham. 1974a. "Lisu and Proto-Lolo-Burmese." 95 pp. ms. To appear in *Acta Orientalia*. [Contains 320 cognate sets.]
- . 1974b. Special Field Examination in Tibeto-Burman Linguistics. U.C. Berkeley, Department of Linguistics.
- Trier, Jost. 1931. *Der deutsche Wortschatz im Sinnbezirk des Verstandes*. Vol. I: *Von den Anfängen bis zum Beginn des 13. Jahrhunderts*. Heidelberg.
- . 1951. "Lehm: Etymologie zum Fachwerk." *Münstersche Forschungen* 3, Marburg.
- . 1952. "Holz: Etymologien aus dem Niederwald." *Münstersche Forschungen* 6, Münster.
- . "Venus: Etymologien um das Futterlaub." *Münstersche Forschungen* 15, Köln.
- T'sou, Benjamin K. 1972. "From morphology to syntax: developments in the Chinese causative." Paper presented at Sixth Sino-Tibetan Conference, San Diego.
- , ed. [to appear]. *Proceedings of the Sixth International Conference on Sino-Tibetan Languages and Linguistics*.
- Veith, Ilza. 1949. *Huang Ti Nei Ching Su Wên: The Yellow Emperor's Classic of Internal Medicine*. Chapters 1-34 translated from the Chinese, with an introductory study. 2d printing, 1st paperback edition (1973). University of California Press, Berkeley, Los Angeles, London.
- Wade, J. 1896. *A Dictionary of the Sgau Karen Language*. Rangoon.
- Walker, Anthony R. 1970. *Lahu Nyi (Red Lahu) Village Society and Economy in North Thailand*. 2 volumes. xxv + 580 pp. ms. Chiang Mai.
- Walker, G. D. 1925. *A Dictionary of the Mikir Language*,

- Mikir-English and English-Mikir.* Shillong.
- Watters, David and Nancy. 1973. *An English-Kham Kham-English Glossary.* Summer Institute of Linguistics, Institute of Nepal and Asian Studies, Tribhuvan University, Kathmandu.
- Weidert, Alfons. 1975. *A Componential Analysis of Lushai Phonology.* *Amsterdam Studies in the Theory and History of Linguistic Science*, IV. John Benjamins B.V., Amsterdam.
- Weinreich, Uriel, William Labov, and Marvin I. Herzog. 1968. "Empirical foundations for a theory of language change." In *Directions for Historical Linguistics: a Symposium*, pp. 97-195. University of Texas Press, Austin.
- Wilson, Mitchell. 1959. *The Human Body: What it is and How it Works.* With illustrations by Cornelius DeWitt. Golden Press, N. Y. (First limp bound printing, 1973).
- Wolfenden, Stuart N. 1929. *Outlines of Tibeto-Burman Linguistic Morphology.* London.
- . 1936. "On certain alternations between dental finals in Tibetan and Chinese." *JRAS* (July), 401-416.
- . 1937. "Concerning the variation of final consonants in the word families of Tibetan, Kachin, and Chinese." *JRAS* (October), 625-655.
- Wulff, Kurt. 1934. *Chinesisch und Tai.* *Det küniglige Danske Videnskabernes Selskab, Historisk-filologiske Meddelelser* 20.3.
- Yang, Paul Fu-mien. 1972. "Prefix kə- in modern Chinese dialects and Proto-Chinese." Paper presented at Fifth Sino-Tibetan Conference, Ann Arbor.
- . 1973. "Initial consonant cluster KL- in modern Chinese dialects and Proto-Chinese." Paper presented at Sixth Sino-Tibetan Conference, San Diego.

- . 1975. "Prefix *s- in Proto-Chinese." Paper presented at Eighth Sino-Tibetan Conference, Berkeley.

ADDENDA

ADDITIONAL NOTES

71-a I now realize that Lahu has still a third allofam of this root, ḥ-ṣā 'nephew or niece; sibling's child' < PLB *sa₂. The tonal alternation between /[^]/ in yā (< *za) and /^ˊ/ in ṣā (< *sa) is perfectly regular for PLB Tone *2 syllables.

156-a A similar example is the newly coined English compound house-husband (meaning a man who stays home and takes care of house and children while his wife goes off to work). The first element in the word husband itself means 'house' (< OE hús), so that the first two syllables of house-husband are co-allofams of the same etymon.

175-a Other examples of "split cognates" in English are furnished by hypocoristic forms of long proper names. Among the innumerable diminutives of Elizabeth, some descend from the first part (Eliza, Liz, Liza, etc.), some from the last part (Beth, Betty, Betsy, etc.) and some from both (Libby). Similarly, Alexandra may be diminutivized in favor of the initial part (Alexa), the final part (Sandra), or the middle (Lexa).

176-a Semantically ordered wordlists have in fact been in use in recent years in Austronesian linguistics. A well worked out list where the vocabulary is graded according to semantic area is to be found in Zorc 1974 [see Additional Bibliography].

Zorc's whole discussion is excellent, and my thanks to Shigeru Tsuchida for calling it to my attention (July 1976).

188-a In an early study [1929; see Additional Bibliography], L. Homburger seeks to shed light on the genetic interrelationships of more than 50 African languages by examining their body-part terminologies. As he says (p. 3):

"Si nous avons choisi les noms des parties du corps c'est parce que ceux-ci ne s'empruntent guère et qu'ils désignent des choses concrètes qui ne varient pas et se retrouvent chez tous les êtres humains, donc que toutes les langues désignent d'une façon ou de l'autre."

I am indebted for this reference to A-G. Haudricourt.

201-a The Hebrew word me'ayim 'bowels' (see p. 210) is also morphologically dual, perhaps referring to the pair LARGE INTESTINE/SMALL INTESTINE?

206-a Geertz 1973 [see Additional Bibliography] has a beautiful chapter on the Balinese cockfight, which he also analyzes in terms of a synthetic reconciliation of opposites. Balinese men are preoccupied with the care and handling of their fighting cocks, and obviously regard them in some sense as extensions of their own masculinity ("detachable penises", as Bateson and Mead called them). In Balinese (as, strangely enough in English) there is a transfield association PENIS — rooster, and punning on these notions is frequent. Yet cocks are also negatively experienced in Bali as incarnations of animality, which is reviled in the culture. (Even eating in Bali is performed

also
animal
ji-ba

furtively, because it is viewed as a grossly animal function; babies are not allowed to crawl because that suggests animal movement, etc.) In the cruel drama of the cockfight are fused "man and beast, good and evil, ego and id, the creative power of aroused masculinity and the destructive power of loosened animality."

My thanks to Bambi B. Schieffelin for referring me to this essay.

275-a I have recently (October 1976) found a cognate to this hitherto isolated Jinghpaw form in the Southern Loloish language Mpi: ŋ⁴kjo⁵ . See Matisoff 1976c, p. 15 [Additional Bibliography].

281-a According to Bambi Schieffelin, in Australian English slang kidneys is used in the opposite sense to what one finds in Israeli Hebrew, to mean "the part of the body where one thinks up half-assed ideas", as in He's really using his kidneys today (i.e. 'He's not thinking very intelligently today').

ADDITIONAL BIBLIOGRAPHY

- Andersen, Elaine. 1978. "Lexical universals of body-part terminology." In J. Greenberg, C. Ferguson, and E. Moravčik, eds., *Universals of Human Language*. Stanford University Press, Stanford, Calif.
- Benveniste, Emile. 1969. *Le vocabulaire des institutions indo-européennes*. Vol. I: "Economie, parenté, société." Vol. II: "Pouvoir, droit, religion." Les éditions de minuit, Paris.
- . 1971. "Semantic problems in reconstruction." In *Problems in General Linguistics*, pp. 249-264. Translated from the French by Mary Elizabeth Meek. University of Miami Press, Coral Gables, Fla.
- Brown, Cecil. 1976. "General principles of human anatomical partonomy and speculations on the growth of partonomic nomenclature." *American Ethnologist* 3.3, pp. 400-424.
- . 1977. "Folk anatomy and language change." 49 pp. ms. De Kalb, Illinois.
- Geertz, Clifford. 1973. *The Interpretation of Cultures: Selected Essays*. Basic Books, New York.
- Homburger, L. 1929. *Noms des parties du corps dans les langues négro-africaines*. Société de linguistique de Paris, Collection linguistique XXV. Paris.
- Matisoff, James A. 1976c. "Mpi and Lolo-Burmese microlin-

guistics." Paper presented at Ninth Sino-Tibetan Conference, Copenhagen. 34 pp. ms.

Ultan, Russell. 1975. *Descriptivity Grading of Finnish Body-Part Terms*. Arbeiten des Kölner Universalien-Projekts, #16.

———. 1976. *Descriptivity in the Domain of Body-Part Terms*. Arbeiten des Kölner Universalien-Projekts, #21.

Zorc, R. David. 1974. "A tentative Philippine wordlist: the qualitative use of vocabulary in classifying languages." Paper presented at First International Conference on Comparative Austronesian Linguistics, Honolulu.

