FUTURE DIRECTIONS IN COMPARATIVE TAI LINGUISTICS

Preface

This essay, in a shorter version, was originally prepared at the invitation of Professor Russell N. Campbell for presentation before the Thai-Lao group at the University of California, Los Angeles, in May 1967. During that same month a resumé of it also was presented at the University of Washington and at the University of California, Berkeley. Questions and comments on all three of those occasions have been very helpful in the task of revision and expansion of the original paper for publication.

All the data cited were collected during two field trips to Southeast Asia, the first in 1964-65 under the sponsorship of the American Council of Learned Societies and of the Horace H. Rackham School of Graduate Studies and the Center for Southern Asian Studies of the University of Michigan, and the second in the summer of 1966 under a National Science Foundation grant.

Several friends have read the manuscript at various stages. Among those whose criticisms have been especially helpful are William G. Boltz, Robbins

Burling, James E. Dew, Gordon B. Downer, and Thomas W. Gething.

Languages and dialects of the Tai family are found across a broad expanse of Southeast Asia. In Thailand and Laos they are socially and politically dominant; elsewhere, in Burma, Assam, North Vietnam, and southern China, they are spoken by minorities.

In the extreme west, in Assam, we find the important but now extinct literary language Ahom. In Burma there are the well-known Shan dialects, with closely related varieties of speech in adjacent parts of Assam and Yunnan. Lue is spoken in Sipsongpanna, in southern Yunnan; to the north of it there are scattered islands of Tai speech, which do not differ greatly from Shan and Lue.

In Thailand the speech of the central plains and Bangkok forms the basis of the standard official and literary language, which nowadays is widely known throughout the country because of its use in administration, education, the press, radio, and television. Besides this prestige language, however, Thailand has a wide variety of regional and local dialects, which blanket most of the northern, northeastern, and southern parts of the country. In Laos inhabitants of the more populous areas speak dialects of Lao, found also in northeastern Thailand.

Tai languages and dialects are found across the entire northernmost part of North Vietnam, from White, Black, and Red Tai in the northwest to Tho and Nung in the northeast, with intermediate varieties usually designated by place names. Closely related

to the Nung dialects of North Vietnam are the Tai dialects of southern Kwangsi.

In all the areas mentioned above, the ethnic name "Tai," or "Thai," is widely used by the people in identifying themselves and their language, with the exception of a few special local names such as Lue and Lao and Tho and Nung. The two spellings "Tai" and "Thai" derive from the fact that the initial *t* of this word is aspirated in Siamese and Lao, but unaspirated elsewhere. Scholars vary as to which of these two spellings they prefer; one widespread practice, followed here, is to use the spelling "Tai" to designate the family and "Thai" to refer to the language of Thailand, also known as Siamese.

Scholars are agreed in regarding all the languages and dialects mentioned thus far, that is, those found in Assam, Burma, southern and southwestern Yunnan, Thailand, Laos, North Vietnam, and southern Kwangsi, as belonging to the Tai family. For Haudricourt these constitute the Tai family proper.[1] For Li they constitute two of the three branches which make up the Tai family, the Central branch including Tho and Nung and closely related dialects in northeastern North Vietnam and adjacent parts of southern China, and the Southwestern branch including all the others.[2]

There is another important group spoken farther north, in Kwangsi, throughout the southern part of Kweichow, and in the eastern part of Yunnan, with some small spillover across the border into North Vietnam. This group is included by Li in the Tai

family, and regarded as constituting the third of his three branches, designated Northern Tai. Haudricourt, however, regards it as so divergent as to constitute a separate group related to, but not to be included in, the Tai family. It is obvious that the disagreement here is largely a matter of terminology.

Finally, in the extreme east, there are various dialects on the island of Hainan which everyone agrees belong to or are closely related to the Tai family.

Comparative Tai studies, like comparative linguistic studies elsewhere, have these aims: to discover in various presumably related languages and dialects recurrent sound correspondences in cognate items of the inherited vocabulary; to infer from these sound correspondences the regular and systematic sound changes that have taken place; to reconstruct from these correspondences and inferred changes the system of phonological distinctions of the assumed prehistoric language, known as Proto-Tai, of which the modern languages and dialects are divergent continuations; and to arrive at a genetic classification, or "family tree," of the modern languages and dialects.

Research in pursuit of these aims has been in progress for more than half a century,[3] with considerable acceleration in recent years as more data have become available and more scholars have become interested in this field. If we had, as we unfortunately do not, a single published exposition of all that is now known about these matters, it

would be clear to everyone, as it is already clear to workers in this field of study, that our present level of knowledge in comparative Tai linguistics can hardly be matched elsewhere in eastern Asia, and compares respectably with comparative linguistic work in other parts of the world.

The purpose here is not, however, to survey what has been accomplished, but rather to explore the frontiers of our knowledge, to identify problems and puzzles, to speculate upon lines of investigation that seem likely to be fruitful in confirming, clarifying, and expanding our understanding of this field.

The point of view here is that all the certainties that we possess in comparative Tai linguistics have resulted from a strict application of the comparative method, with its assumptions of regular and systematic sound changes resulting in recurrent sound correspondences in cognate forms in genetically related languages and dialects. Although the scientific principles involved are in no way different from those which form the basis for comparative grammar in Indo-European and other families to which the comparative method has been applied, the Tai languages are tonal, and the statements that comparatists make about them differ somewhat in form, though not in principle, from statements made about phonological correspondences and changes in languages of more familiar types. For this reason, and in order to facilitate the statement of the frontier problems and puzzles which are our main subject, it

may be helpful to interpolate at this point a brief exposition of how comparative Tai linguistics works.

The phonological system of any Tai language or dialect can best be described in terms of the distinctions made in various syllable positions. Each syllable has an initial consonant or consonant cluster, a vocalic nucleus, and an optional final consonant, and each syllable has a tone.

As regards initials, there is always a fairly rich variety of consonant distinctions. Most Tai languages or dialects also have some initial consonant clusters; there seem to be almost everywhere[4] at least some clusters of consonant-plus-*w*, and often also of consonant-plus-*y*, and in some languages there are also clusters of obstruent-plus-*l*, or *l* and *r*.

In the vocalic nucleus position there is usually a set of either six or nine distinctive vowels forming the basic pattern, with usually either two or three distinctions in height and three distinctions of front, central, and back. Some languages (Siamese, for example) make a distinction in length for all vowels of the system. Very widespread, however, is a peculiar situation in which only the low central vowel shows a distinction in length. Usually the short *a* is higher phonetically than the long *aa*, so that analysts who dislike ascribing a length distinction to this one vowel sometimes attribute this distinction to quality, so that they may, for example, arrive at a system of seven vowels for a language that others might describe as having six vowels with one of them also occurring long.

Many Tai languages have also, in addition to their basic vowel system, a series of three diphthongs beginning with one of the three high vowels and ending with a centering off-glide, as in Siamese *mia*[1] 'wife', *rɨa*[1] 'boat', and *ŋua*[1] 'ox' (see chart 1).

Chart 1. Typical Tai Vowel Systems[5]

SIAMESE

i, ii, ia	ɨ, ɨɨ, ɨa	u, uu, ua
e, ee	ə, əə	o, oo
ɛ, ɛɛ	a, aa	ɔ, ɔɔ

WHITE TAI			YAY		
i	ɨ	u	i, ia	ɨ, ɨa	u, ua
e	ə	o		ə	
ɛ	a, aa	ɔ	e	a, aa	o

The number of permitted contrasts in syllable-final position is always severely limited. Syllables ending in a vowel, with no final consonant, are very common in all Tai languages, or there may be a final nasal, semivowel, or obstruent. Nasals are *m*, *n*, and *ŋ*; although many Tai languages also have an initial palatal nasal *ɲ*, this is never known to occur finally. The final semivowels *w* and *y* occur in all languages, with limitations as to the vowels that each may follow. Some languages also have a high back unrounded semivowel *ɣ*, usually occurring only after short *a*. Final obstruents are *p*, *t*, and *k*, with in some languages also a final glottal stop.

Each syllable uttered in isolation has a distinctive tone. The maximum number of tonal distinctions, always five, six, or seven, so far as we know, is found on what are sometimes called "smooth" or "free" syllables, that is, those ending in a vowel, semivowel, or nasal. There is always a much smaller number of permitted tonal distinctions on checked syllables, that is, those ending in *p*, *t*, *k*, or *ʔ*. Descriptions sometimes differ as to whether the tones occurring on checked syllables are counted as extras or are identified on the basis of phonetic similarity with tones occurring on smooth syllables, with which they are in complementary distribution; the latter practice is followed in numbering tones in the data cited here.

The tones of a Tai language always show phonetic differences of pitch height (high, mid, low, and so on) and contour (falling, rising, level, and so on), and in many languages (apparently all those in Li's Central and Southwestern areas) there are some tones, usually two but sometimes three, which show an additional phonetic feature of glottal constriction.

The inherited Tai morphemes with which the comparatist has to deal are almost always monosyllables. Most languages also have weak-stressed syllables with neutralized tone occurring in borrowed polysyllabic forms, in complex forms, which have resulted historically from combinations, or in phrasal combinations of monosyllabic morphemes in connected speech. Such weakened syllables, and other more subtle or complicated features of connected speech, often constitute difficult problems in the

exhaustive descriptive analysis of any one language or dialect. Fortunately for the comparatist, at least at the present stage of our work, these more subtle features are for comparative Tai linguistics irrelevant; all that is required for comparative work is an accurate analysis of the distinctions in the various syllable positions of the monosyllabic morpheme pronounced with full stress. The comparatist can likewise disregard more abstruse theoretical phonological questions, which may be of importance in the detailed analysis of any given dialect or language, for example, whether glottal stop can be analyzed out of the system, or whether the number of tones can be reduced by some sort of componential treatment, or whether aspirated stops in a given dialect are to be analyzed as units or clusters. Also, it does not matter whether one operates with classical phonemes or with distinctive features, although for some types of sound correspondences and sound changes it is sometimes necessary to take into consideration phonetic features shared by a number of sounds, and the ultimate aim must always be to account for whole phonological systems, not mere lists of individual phonemes.

For some of the better-known Tai languages, the phonological structure has been well analyzed, so that the comparatist has his data in shape ready for his comparative work. Some minor dialects have also been carefully described by well-trained linguists. But for much of his material the comparatist must do preliminary descriptive work, either in the field, recording and analyzing, or in trying to infer the

phonological structure for unclear descriptions and impressionistic transcriptions in earlier publications.

Once the scholar has his collection of forms in various Tai languages recorded in a transcription that shows the various distinctions of syllable initial, vowel, final consonant, and tone, he usually finds that for the initials and tones the correspondences from one dialect to another are rather complicated, whereas for vowels and final consonants they are often fairly simple and obvious. The relative simplicity and regularity of the vowel and final correspondences, presumably reflecting a greater stability in the history of these elements, may be illustrated by these examples of the kinds of correspondences one finds in other Tai languages for a few Siamese vowels and final consonants (see chart 2).[6]

It will be noted in chart 2 that uniformity in vowels and final consonants is on the whole more common than diversity.[7] Although, as we shall see later, there are unsolved problems regarding vowels, some of the variations illustrated here are easily explained. For example, White Tai has changed final *-k* to *-ʔ* after a long vowel (the final glottal stop is an automatic feature of the fourth tone in this language and so is not indicated in the transcription); in Lungming and some other dialects of the Nung group high vowels have been diphthongized.

When we come to consider initial consonants and tones, the picture becomes more complex. Comparison of any two or more Tai dialects shows a complicated correlation between initials and tones from which

Chart 2.

	SIAM-ESE	WHITE TAI	LEI PING	LUNG-MING	YAY	SAEK
good	dii^{1}	di^{1}	day^{2}	nay^{4}	di^{1}	dii^{1}
gall bladder	dii^{1}	bi^{1}	dii^{2}	ney^{4}	di^{1}	blii^{1}
older sibling	phii^{3}	pi^{5}	phii^{5}	pey^{5}	pi^{6}	phii^{5}
to have	mii^{1}	mi^{4}	mii^{4}	mey^{4}	mi^{4}	mii^{4}
excrement	khii^{3}	khi^{3}	khii^{3}	khii^{3}	hay^{6}	ɤay^{6}
year	pii^{1}	pi^{1}	pii^{2}	pey^{1}	pi^{1}	pii^{1}
to eat	kin^{1}	kin^{1}	kin^{2}	kin^{1}	kɨn^{1}	kin^{1}
to fly	bin^{1}	bin^{1}	bin^{2}	min^{4}	bin^{1}	bɨl^{1}
tongue	lin^{4}	lin^{6}	lin^{3}	lin^{6}	lin^{6}	liin^{6}
stone	hin^{5}	hin^{1}	hin^{1}	thin^{1}	rin^{1}	riil^{2}
to come	maa^{1}	maa^{4}	maa^{4}	maa^{4}	ma^{1}	maa^{2}
dog	maa^{5}	maa^{1}	maa^{1}	maa^{1}	ma^{1}	maa^{2}
to seek	haa^{5}	haa^{1}	haa^{1}	laa^{1}	ra^{1}	raa^{2}
to kill	khaa^{3}	xaa^{3}	khaa^{3}	khaa^{3}	ka^{3}	kaa^{3}
five	haa^{3}	haa^{3}	haa^{3}	haa^{3}	ha^{3}	haa^{3}
right (hand)	khwaa^{5}	xwaa^{1}	l̥aa^{1}	saa^{1}	kwa^{4}	khwaa^{4}
mouth	paak^{2}	paaʔ^{2}	paak^{5}	paak^{2}	paak^{2}	paak^{6}
fruit	maak^{2}	maaʔ^{2}	maak^{2}	maak^{2}	maak^{2}	maak^{6}
pestle	saak^{2}	saaʔ^{2}	l̥aak^{2}	saak^{2}	θaak^{2}	saak^{6}
otter	naak^{3}	naa^{4}	naak^{5}	naak^{5}	naak^{5}	naak^{5}
root	raak^{3}	haa^{4}	laak^{5}	laak^{5}	raak^{5}	raak^{6}
to vomit	raak^{3}	haa^{4}	l̥aak^{5}	laak^{5}	ruak^{5}	ruak^{5}

scholars have been able to make the following inferences. First, the parent language had a system of three tones (sometimes called A, B, and C) on smooth syllables, and no tonal distinction (sometimes marked D) on checked syllables. Second, at some time after the break-up of the parent speech each daughter language or dialect underwent a number of sound changes involving tonal splits conditioned by the phonetic nature of the initial consonants at the time of the splitting. Third, in the case of checked syllables there was in many cases an additional conditioning factor in these tonal splits, namely, vowel length.

Reconstruction of these changes and of the system that existed before the changes is possible only because each language or dialect made different tonal splits, and is aided by the fact that where subsequent changes (for example, in initial consonants) took place, the changes differed from one dialect to another.

It is helpful in working out the details of the tonal splits in any given dialect to devise a chart showing the tonal system of the parent language and the various phonetic types of initial consonants which conditioned splits in this tonal system in one dialect or another.

Chart 3 provides for a maximum of tonal distinctions in the modern dialects. No one dialect, of course, makes all twenty distinctions, but the system of five or six or seven tones of any given dialect will always be found to divide up the chart in a particular way, usually different in some respects

Chart 3. The Historical Sources of Tones in Tai Dialects

PROTO-TAI TONES / Proto-Tai Initials	SMOOTH SYLLABLES			CHECKED SYLLABLES	
	A	B	C	D-short	D-long
Voiceless friction sounds					
Voiceless unaspirated stops					
Glottal stop and preglottalized sounds					
Voiced sounds					

from the pattern of other dialects. The work of the comparatist, whether he uses or visualizes a chart of this particular form or not, consists essentially in collecting together in one box those morphemes whose cognate forms in the various dialects show the same correspondences in tone and initial consonant. Placement of a morpheme in a particular box implies that it had in the parent language a particular tone, A, B, C, or D, and an initial consonant of a particular phonetic type.

To illustrate something of how tonal correspondences are handled, chart 4 shows how various languages have split the left-hand column of chart 3, which presumably reflects the single tone A of Proto-Tai.

Chart 4.

	SIAM-ESE	WHITE TAI	LEI PING	LUNG-MING	YAY	SAEK
thick	naa^{5}	naa^{1}	naa^{1}	naa^{1}	na^{1}	naa^{2}
dog	maa^{5}	maa^{1}	maa^{1}	maa^{1}	ma^{1}	maa^{2}
to seek	haa^{5}	haa^{1}	haa^{1}	laa^{1}	ra^{1}	raa^{2}
hair of the head	phom^{5}	phum^{1}	phyom^{1}	phyom^{1}	piam^{1}	phram^{2}
three	saam^{5}	saam^{1}	l̥aam^{1}	saam^{1}	θaam^{1}	saam^{2}
	---		---			---
fish	plaa^{1}	paa^{1}	pyaa^{2}	pyaa^{1}	pya^{1}	plaa^{1}
door	-tuu^{1}	tu^{1}	tuu^{2}	tow^{1}	tu^{1}	-tuu^{1}
year	pii^{1}	pi^{1}	pii^{2}	pey^{1}	pi^{1}	pii^{1}
to eat	kin^{1}	kin^{1}	kin^{2}	kin^{1}	kɨn^{1}	kin^{1}
middle	klaaŋ^{1}	kaaŋ^{1}	kyaaŋ^{2}	kyaaŋ^{1}	caaŋ^{1}	tlaaŋ^{1}
				---		
to cough	ʔay^{1}	ʔay^{1}	ʔay^{2}	ʔay^{4}	ʔay^{1}	ʔay^{1}
star	daaw^{1}	daaw^{1}	daaw^{2}	naaw^{4}	daaw^{1}	traaw^{1}
gall bladder	dii^{1}	bi^{1}	dii^{2}	ney^{4}	di^{1}	blii^{1}
stairs, ladder	-day^{1}	day^{1}	day^{2}	nay^{4}	lay^{1}	ray^{1}
leaf	bay^{1}	bay^{1}	bay^{2}	may^{4}	bay^{1}	bəə^{1}
		---	---		---	---
rice-field	naa^{1}	naa^{4}	naa^{4}	naa^{4}	na^{4}	naa^{4}
fire	fay^{1}	fay^{4}	fay^{4}	fay^{4}	fi^{4}	vii^{4}
snake	ŋuu^{1}	ŋu^{4}	vuu^{4}	ŋow^{4}	ŋɨa^{4}	ŋua^{4}
human being	khon^{1}	kun^{4}	khən^{4}	kən^{4}	hun^{4}	hun^{4}
we	raw^{1}	haw^{4}	l̥aw^{4}	law^{4}	raw^{4}	roo^{4}

It will be noted that White Tai and Yay have made a two-way split on the basis of the simple distinction of voiced versus voiceless initial at the time of the split, with subsequent voicing or devoicing of these initials in some cases; many other Tai dialects have behaved in exactly this way. Lungming also has made a two-way split, but here the tone of syllables with original initial glottal stop or preglottalized consonant has fallen together with that of syllables with originally voiced initials. Siamese has made a two-way split on a different basis, treating syllables with originally voiceless friction sounds differently from the rest. Lei Ping and Saek have both made some phonetic features of initials the basis for a three-way tonal split.

The investigator makes a similar study of the tonal splits in each of the other columns of chart 3. The results of this sorting are always most illuminating. When he has placed the sets of cognates in their proper boxes, he will be able to infer, for example, that the word for 'thick' had in the parent language a different kind of initial, perhaps a voiceless **n̥* or a preaspirated nasal **hn*, from the voiced initial **n* reflected in the word for 'rice-field'.

This particular conclusion sometimes surprises and troubles beginners in comparative Tai studies, because words like 'thick' and 'dog', for which earlier initials with some sort of voiceless onset are reconstructed, are not found in any modern Tai dialect with anything other than an ordinary voiced nasal initial. The theory that the tones of such

words reflect an earlier voiceless initial was first suggested by the spelling *hn* and *hm* found in languages with archaic writing systems such as Siamese, Lao, Shan, Lue, and White and Black Tai, but it seems likely that the entire hypothesis of tonal splits conditioned by the phonetic nature of initials would, even without such hints from archaic spellings, have led us to our present view, because of the overall power of the theory in explaining the entire tone and initial systems of each dialect. A similar paradox appears in words with initial stops in the bottom row of the chart, where the original initials are believed to have been voiced, as in the word for 'person'. Here we find initial voiceless stops in the modern dialects, in some cases aspirated *ph*, *th*, *kh* and in others unaspirated *p*, *t*, *k*, but again archaic spellings suggest earlier voiced consonants, and the coherence and cogency of the overall hypothesis can leave little room for doubt that the parent language differed from all the daughters. To put it another way, when we find that the initial nasal of words like 'thick' and 'dog' in the top box of column A behaves with respect to tone exactly like such sounds as the fricatives *s*, *f*, *h*, and *x*, and the voiceless aspirated stops *ph*, *th*, and so on, and we note that in such words the initial nasal is spelled *hn* or *hm* in archaic Tai writing systems, we conclude that the initial nasal in such words must have been pronounced with some sort of voiceless friction as its onset at the time of the tonal splitting. Similarly, in the bottom box, the words that have initial *ph*, *th*, *kh* in some dialects and *p*, *t*, *k* in others are

found to show the same tonal development as words having initials that are everywhere voiced, such as nasals, semivowels, and liquids. And, when we confront this evidence with the fact that these initials are spelled in archaic Tai writing systems with letters which represented *b*, *d*, and *g* in the Indic alphabets from which they were borrowed, we conclude that these initials were voiced at the time of the tonal splits, and that they have undergone subsequent alteration in all the dialects, in some to aspirated *ph*, *th*, *kh* and in others to unaspirated *p*, *t*, *k*.[8]

Such inferences are possible, of course, only because for each such correlating correspondence of initial and tone there are numerous examples, sometimes a great many. In the case of fairly closely related dialects it is possible to account for the tones and initials of virtually all of the ordinary vocabulary in this way, with so few exceptions that one usually feels forced to assume that these have undergone interdialectal borrowing or have suffered some sort of internal distortion not related to the regular historic sound changes.

We come now to our real subject of discussion, which is some of the problems and puzzles that lie on or beyond the frontiers of our present knowledge and understanding in the field of comparative Tai. These will be formulated as numbered questions, though in some cases we will find that there is considerable overlap and interdependency.

Question 1. Can the four categories of phonetic types of initial consonants that are assumed to have conditioned tonal splits be further subdivided?

Many Tai languages have made a simple two-way tonal split in all columns conditioned by the voiceless/voiced opposition in initials, as White Tai and Yay have done in column A of chart 3. This pattern is so widespread that some scholars in the field are in the habit of providing for only this distinction in their notation and charts representing the historical changes, so that when they deal with languages that have made tonal distinctions conditioned by other phonetic features of the initials they are compelled to add supplementary statements.

As we have seen, Siamese has made a different tonal split in column A, conditioned apparently by a distinction between originally voiceless friction sounds as opposed to all other types of initials. Siamese has been so well known from the beginning of comparative Tai studies that no one person may be said to have discovered this phenomenon; it was there to account for from the first. We have noted that Saek makes a similar distinction.

The pattern in Lungming, where syllables with original glottal stop or preglottalized consonant have behaved with respect to the conditioning of tonal splits like originally voiced consonants, is found at various isolated points, widely scattered throughout the Tai-speaking domain, including the well-known tone of Chieng Mai in northern Thailand, in the A column, and occasionally (as in Yay) in the C column or elsewhere in the chart. Li first worked out the correct explanation of this,[9] and others have

since spotted the same conditioning factor in other Tai dialects.[10]

The discovery of these phonetic distinctions in initials as the conditioning factor in tonal splits in the various languages and dialects has been so enormously helpful in pointing toward the structure of the consonant system of Proto-Tai that one might wonder if even more such distinctions might exist, waiting to be discovered. Available published descriptions appear to show nothing more of this sort, but I have found one such additional distinction in a dialect of Nung spoken at the village of Bac Va located a little south of That Khe in the extreme northeast of North Vietnam.[11]

In this dialect morphemes that fall into the top box of column A in chart 3, that is, those which have the fifth or rising tone in Siamese, and which are assumed therefore to have had in Proto-Tai tone A and a voiceless friction sound as initial, are divided into two groups, one having low rising tone and the other high rising tone. For example, in this dialect the word for 'leg', cognate with Siamese *khaa*^{5} and White Tai *xaa*^{1}, is *khaa*^{1} with low rising tone, but the word for 'dog', cognate with Siamese *maa*^{5} and White Tai *maa*^{1}, is *maa*^{2} with high rising tone. In this dialect the high rising tone also is found in words of column A having an original unaspirated initial such as *pii*^{2} 'year' (Siamese *pii*^{1}, White Tai *pi*^{1}), and in words with preglottalized initial such as *ʔaw*^{2} 'to take' (Siamese and White Tai *ʔaw*^{1}). Words in column A having an original voiced initial have in this dialect still a third tone, mid level as

in *mɨɨ*^{3} 'hand' (Siamese *mɨɨ*^{1}, White Tai *mɨ*^{4}) or *vaay*^{3} 'water buffalo' (Siamese *khwaay*^{1}, White Tai *xvaay*^{4}).

Thus, in column A this Nung dialect has made, like some other Tai languages, a three-way split, but the horizontal line marking the distinction in the upper part of column A of our chart must be placed even higher than for Siamese.

When we examine the two sets of Bac Va words that correspond to Siamese fifth-tone words in the top box of column A, and compare each of the two sets with the cognates in other Tai languages (leaving languages of the Northern branch out of account for the moment since they present special tonal problems of their own, as we shall see later), we find that in general the words with low rising tone at Bac Va include the sets of cognates that usually are regarded as having had an original voiceless fricative or aspirated stop, or a cluster beginning with such a sound (see chart 5).

Chart 5. Bac Va Words with Low Rising Tone

	SIAM-ESE	WHITE TAI	BAC VA	LEI PING	LUNG-MING
hair of the head	phom^{5}	phum^{1}	phom^{1}	phyom^{1}	phyom^{1}
rain	fon^{5}	fɨn^{1}	phon^{1}	phən^{1}	phən^{1}
wall, lid	faa^{5}	faa^{1}	phaa^{1}	phaa^{1}	phaa^{1}
to dream	fan^{5}	fan^{1}	phon^{1}	phan^{1}	phan^{1}
to arrive	thɨŋ^{5}	thɨŋ^{1}	thəŋ^{1}	thəŋ^{1}	thəŋ^{1}
to plow, a plow	thay^{5}	thay^{1}	thay^{1}	thay^{1}	thay^{1}
tail	haaŋ^{5}	haaŋ^{1}	thaaŋ^{1}	haaŋ^{1}	thaaŋ^{1}

	SIAM-ESE	WHITE TAI	BAC VA	LEI PING	LUNG-MING
(two or more) carry	haam^{5}	haam^{1}	thaam^{1}	haam^{1}	thaam^{1}
testicles	--	ham^{1}	tham^{1}	ham^{1}	tham^{1}
head	hua^{5}	ho^{1}	thuu^{1}	vuu^{1}	thuu^{1}
headlouse	haw^{5}	haw^{1}	thaw^{1}	haw^{1}	thaw^{1}
stone	hin^{5}	hin^{1}	thin^{1}	hin^{1}	thin^{1}
hole, pit	khum^{5}	khum^{1}	khum^{1}	khom^{1}	khom^{1}
to smell bad	khiaw^{5}	khiw^{1}	khew^{1}	khiw^{1}	khiw^{1}
to open	khay^{5}	khay^{1}	khay^{1}	khay^{1}	khay^{1}
bitter	khom^{5}	khum^{1}	khom^{1}	khom^{1}	khom^{1}
hard	khɛŋ^{5}	khɛŋ^{1}	khɛɛŋ^{1}	kheeŋ^{1}	leeŋ^{1}
green	khiaw^{5}	xɛw^{1}	khew^{1}	kheew^{1}	kheew^{1}
arm	khɛɛn^{5}	xɛn^{1}	khɛɛn^{1}	kheen^{1}	kheen^{1}
ginger	khiŋ^{5}	xiŋ^{1}	khiŋ^{1}	khiŋ^{1}	khiŋ^{1}
body hair	khon^{5}	xun^{1}	khon^{1}	khon^{1}	khon^{1}
to crow	khan^{5}	xan^{1}	khan^{1}	khan^{1}	khan^{1}
to sell	khaay^{5}	xaay^{1}	khaay^{1}	khaay^{1}	khaay^{1}
leg	khaa^{5}	xaa^{1}	khaa^{1}	khaa^{1}	khaa^{1}
needle	khem^{5}	xim^{1}	khem^{1}	khim^{1}	khim^{1}
cast-net	hɛɛ^{5}	hɛ^{1}	khee^{1}	hee^{1}	hee^{1}
to yawn	haaw^{5}	haaw^{1}	haaw^{1}	haaw^{1}	haaw^{1}
fragrant	hɔɔm^{5}	hɔm^{1}	hɔɔm^{1}	hoom^{1}	hoom^{1}
to imprison	khaŋ^{5}	chaŋ^{1}	haŋ^{1}	haŋ^{1}	laŋ^{1}
ear	huu^{5}	hu^{1}	huu^{1}	vuu^{1}	low^{1}
cockscomb	ŋɔɔn^{5}	hɔn^{1}	hɔɔm^{1}	hoon^{1}	khoon^{1}

The words for 'eye' and 'to die', which behave in Siamese and many other Tai languages as if they had had in Proto-Tai an initial consonant cluster

beginning with a voiceless unaspirated initial, have in Bac Va fallen into the group represented in chart 5.

	SIAM-ESE	WHITE TAI	BAC VA	LEI PING	LUNG-MING
eye	taa^{1}	taa^{1}	thaa^{1}	haa^{1}	thaa^{1}
to die	taay^{1}	taay^{1}	thaay^{1}	haay^{1}	thaay^{1}

Turning to the set of Bac Va words having high rising tone, these include forms having initial *s* in most languages, represented by *l̥* in some Central dialects (see chart 6).

But the same high rising tone is also found on Bac Va forms in the sets shown in chart 7.

Chart 6. Bac Va Words with High Rising Tone

	SIAM-ESE	WHITE TAI	BAC VA	LEI PING	LUNG-MING
tiger	s̶i̶a̶^{5}	sə^{1}	l̥i̶i̶^{2}	l̥əə^{1}	s̶i̶i̶^{1}
writing, book	-s̶i̶i̶^{5}	s̶i̶^{1}	l̥i̶i̶^{2}	l̥əə^{1}	səy̶^{1}
clear, transparent	say^{5}	say̶^{1}	l̥ay̶^{2}	l̥ay̶^{1}	say̶^{1}
cord, string	saay^{5}	saay^{1}	l̥aay^{2}	l̥aay^{1}	saay^{1}
late in the morning	saay^{5}	saay^{1}	l̥aay^{2}	l̥aay^{1}	saay^{1}
three	saam^{5}	saam^{1}	l̥aam^{2}	l̥aam^{1}	saam^{1}
garden	suan^{5}	son^{1}	l̥uun^{2}	l̥uun^{1}	suun^{1}
two	sɔɔŋ^{5}	sɔŋ^{1}	l̥ɔɔŋ^{2}	l̥ooŋ^{1}	sooŋ^{1}
young unmarried woman	saaw^{5}	saaw^{1}	l̥aaw^{2}	l̥aaw^{1}	saaw^{1}
high	suuŋ^{5}	suŋ^{1}	l̥uŋ^{2}	l̥oŋ^{1}	soŋ^{1}

Chart 7.

	SIAM-ESE	WHITE TAI	BAC VA	LEI PING	LUNG-MING
dog	maa^{5}	maa^{1}	maa^{2}	maa^{1}	maa^{1}
pig	muu^{5}	mu^{1}	muu^{2}	muu^{1}	mow^{1}
bear	mii^{5}	mi^{1}	mii^{2}	mii^{1}	mey^{1}
thick	naa^{5}	naa^{1}	naa^{2}	naa^{1}	naa^{1}
above, north	n~~ɨa~~^{5}	n~~ə~~^{1}	n~~ɨɨ~~^{2}	n~~əə~~^{1}	n~~ɨɨ~~^{1}
rat, mouse	nuu^{5}	nu^{1}	nuu^{2}	nuu^{1}	now^{1}
skin	naŋ^{5}	naŋ^{1}	naŋ^{2}	naŋ^{1}	naŋ^{1}
many, much	laay^{5}	laay^{1}	laay^{2}	laay^{1}	laay^{1}
nephew, niece; grandchild	laan^{5}	laan^{1}	laan^{2}	laan^{1}	laan^{1}
back	laŋ^{5}	laŋ^{1}	laŋ^{2}	laŋ^{1}	laŋ^{1}
big	luaŋ^{5}	loŋ^{1}	luuŋ^{2}	luuŋ^{1}	luuŋ^{1}
sweet	waan^{5}	vaan^{1}	vaan^{2}	vaan^{1}	vaan^{1}
a comb, to comb	wii^{5}	vi^{1}	vii^{2}	vii^{1}	vey^{1}

Although considerable material on other dialects of the Nung group spoken in northeastern North Vietnam and in southern Kwangsi is available in the literature and in unpublished field notes, the peculiar tonal distinction made at Bac Va has not been noted elsewhere, with a single known exception: Saul (1965) cites data from the Nung dialect of the village of Lang Vo near Lang Sở in which the same distinction apparently occurs. Although the tones

are not described, the phrases cited as examples in her discussion of classifiers show the words for 'rain', 'plow', 'head', 'stone', 'arm' 'leg', 'ear', and 'eye' transcribed with one tone mark, while the words for 'three', 'young woman', 'dog', 'much', 'grandchild', and 'comb' are transcribed with another.[12]

Words like those in chart 7, beginning with 'dog', belong to a large category of Tai morphemes, which have sonorant initials everywhere but behave with respect to tones as if they had had originally voiceless initial consonants. Since they are spelled in Siamese and other languages having archaic writing systems with a preceding letter *h* (for example, *hm*), it has been generally assumed that they reflect earlier preaspirated sonorant initials. The Bac Va evidence suggests, since they behave like words with initial **s* rather than **h*, that perhaps clusters such as **sm-* or **sn-* are involved. These could, of course, have evolved to something like **hm-* or **hn-* in the area and at the period of the establishment of the old Siamese and other archaic writing systems. Or, on the other hand, a reverse process, something like **hm-* > **sm-*, in the Nung area is conceivable. Either assumption, of course, implies subsequent loss of the **h* or **s*. Another possibility is that the reconstruction **hm*, and so on, is after all correct for all dialects, and the special tonal behavior in the first group of words ('hair of the head', 'rain', and so on) was due to the position of the aspiration, with initials like **ph* and **h* behaving differently from these preaspirated sonorants.[13]

Whatever the correct historical explanation of this special tonal distinction in the Bac Va dialect turns out to be, it is striking that the history of comparative Tai studies seems to show increasing differentiation in our knowledge of the types of features in initial consonants that at one place or another conditioned tonal splits. Starting with the simple voiceless/voiced distinction (which is apparently adequate for a similar historical analysis in some other families of tonal languages in Asia), Siamese was found to require a more refined statement. Later it was discovered that for some dialects a distinction had to be made between simple unaspirated stops like *p*, *t*, and *k*, on the one hand, and glottal stop and preglottalized initials on the other. Bac Va now requires a still finer distinction. Each new distinction discovered requires another horizontal line in chart 3. Someday, when we fully understand the historical development of tones in languages of the Northern branch (of which Yay and Saek are examples in our data), it seems likely that still more horizontal lines will have to be drawn in our chart.

Our chart 3 is only one man's preferred way of visualizing what happened between Proto-Tai and the various modern dialects. Others might prefer some different sort of chart to depict the historical changes and correspondences. But, whatever the form of the chart, it seems increasingly likely that the order of the classes of initial consonants represented by the horizontal rows in our chart would

be unalterable because of the interlocking evidence from the various dialects.

Question 2. Is there a systematic phonological explanation for tonal irregularities in languages of the Northern Tai branch?

Li, Haudricourt, and Gedney[14] have commented on a number of words, which in other Tai languages have tones reflecting voiceless initials but in languages of the Northern Tai branch show tones that usually reflect voiced initials, and to a lesser extent vice-versa. Examples cited by these authors include the following (using Yay and Saek as representative Northern Tai languages).

	SIAMESE	YAY	SAEK
person	phuu^{3}	pu^{6}	phuu^{6}
to wear, carry	thɨɨ^{5}	tɨ^{4}	thɨɨ^{4}
bean	thua^{2}	tua^{5}	thua^{5}
bowl	thuay^{3}	tɨay^{6}	thooy^{6}
to arrive	thɨŋ^{5}	taŋ^{4}	thaŋ^{4}
hole, pit	khum^{5}	kum^{4}	khum^{4}
rice	khaaw^{3}	haw^{6}	ɣaw^{6}
ten	sip^{2}	sip^{1}	sip^{6}
to come	maa^{1}	ma^{1}	maa^{2}

To look further into these apparent irregularities we must first set up a table, using the format of our chart 3, displaying the normal or usual (that is, more frequent) tonal patterns of these two languages (see chart 8).

Chart 8. Usual Tonal Developments in Siamese, Yay, and Saek

A	B	C	D-SHORT	D-LONG
thick Si. naa^{5} Y. na^{1} Sk. naa^{2}	four Si. sii^{2} Y. θi^{2} Sk. sii^{6}	five Si. haa^{3} Y. ha^{3} Sk. haa^{3}	six Si. hok^{2} Y. rok^{3} Sk. rɔk^{4}	guest Si. kheek^{2} Y. hek^{2} Sk. heek^{6}
leg Si. khaa^{5} Y. ka^{1} Sk. kwaa^{1}				
eye Si. taa^{1} Y. ta^{1} Sk. praa^{1}	chicken Si kay^{2} Y. kay^{2} Sk. kay^{6}	nine Si. kaaw^{3} Y. ku^{3} Sk. kuu^{3}	liver Si. tap^{2} Y. tap^{3} Sk. tap^{4}	eight Si. peet^{2} Y. pet^{2} Sk. peet^{6}
to take Si. ʔaw^{1} Y. ʔaw^{1} Sk. ʔaw^{1}	young man Si. baaw^{2} Y. baaw^{2} Sk. baaw^{6}	to open (the mouth) Si. ʔaa^{3} Y. ʔa^{6} Sk. ʔaa^{3}	raw, unripe Si. dip^{2} Y. dip^{3} Sk. rip^{4}	to bathe Si. ʔaap^{2} Y. ʔaap^{2} Sk. ʔaap^{6}
rice-field Si. naa^{1} Y. na^{4} Sk. naa^{4}	to sit Si. naŋ^{3} Y. naŋ^{5} Sk. naŋ^{5}	water Si. naam^{4} Y. ram^{6} Sk. nam^{6}	bird Si. nok^{4} Y. rok^{1} Sk. nɔk^{6}	blood Si. liat^{3} Y. liat^{5} Sk. luat^{5}

Thus, for Yay the usual pattern of tones according to their historical sources is

A	B	C	D-SHORT	D-LONG
1	2	3	3	2
1	2	3	3	2
1	2	6	3	2
4	5	6	1	5

with for the most part a simple dichotomy according to an original voiceless/voiced opposition in initials, except that in the C column syllables with an original initial glottal stop or preglottalized consonant have ended up with the same modern tone as syllables with an original voiced initial.

For Saek a similar table takes the following form.

A	B	C	D-SHORT	D-LONG
2, 1	6	3	4	6
1	6	3	4	6
1	6	3	4	6
4	5	6	6	5

We are struck immediately by the fact that Saek has two tones, first and second, in the top box of column A, reminding us of the special distinction that the Bac Va dialect made in this box, but the Saek split turns out to be different from that of Bac

Va. For Saek the normal tone in this box is apparently the second, as in the Saek forms cited earlier: *naa*^{2} 'thick', *maa*^{2} 'dog', *raa*^{2} 'to seek', *phram*^{2} 'hair of the head', *saam*^{2} 'three'. We are prompted to make a list of the Saek words showing the apparently aberrant first tone in this box.

	SIAM-ESE	WHITE TAI	LEI PING	LUNG-MING	YAY	SAEK
needle	khem^{5}	xim^{1}	khim^{1}	khim^{1}	cim^{1}	kim^{1}
arm	kheen^{5}	xen^{1}	kheen^{1}	kheen^{1}	cen^{1}	keen^{1}
horn	khaw^{5}	xaw^{1}	--	--	kaw^{1}	kaw^{1}
leg	khaa^{5}	xaa^{1}	khaa^{1}	khaa^{1}	ka^{1}	kwaa^{1}
to sell	khaay^{5}	xaay^{1}	khaay^{1}	khaay^{1}	kaay^{1}	kwaay^{1}
body hair	khon^{5}	xun^{1}	khon^{1}	khon^{1}	pun^{1}	pul^{1}

At first glance it might appear that Saek has undergone this special tonal development in the A column in the case of syllables that had the original initial that is reflected in White Tai *x*. That this is not the whole explanation is shown by a different development, producing second tone in Saek, in other morphemes having initial *x* in White Tai.

	SIAM-ESE	WHITE TAI	LEI PING	LUNG-MING	YAY	SAEK
green	khiaw^{5}	xɛw^{1}	kheew^{1}	kheew^{1}	hew^{1}	heew^{2}
ginger	khiŋ^{5}	xiŋ^{1}	khiŋ^{1}	khiŋ^{1}	hiŋ^{1}	hiiŋ^{2}
to crow	khan^{5}	xan^{1}	khan^{1}	khan^{1}	han^{1}	hal^{2}

These are notably different from the previous set in showing initial *h* in Yay.

Although these data suggest that we are on the track of a phonological explanation here, our purpose at the moment is not to solve this curious problem of the extra tonal split in the box of column A in Saek, but rather the general problem of tonal irregularities in languages of the Northern Tai branch, which seem to show original voiced initials where we expect voiceless ones, and vice-versa.

There are so many of these irregularities--far more than anyone has yet set forth in print--that various suggestions have been ventured as to possible explanations. Li[15] has wondered if perhaps the parent language had doublets showing functional morphophonemic alternation between voiced and voiceless initials, with the various daughter languages differing as to which form they preserved. A serious objection to this suggestion is found in the failure of any Tai language anywhere to preserve both forms of any morpheme. Others, who like to believe in an ultimate Malayo-Polynesian connection for the Tai languages, have suggested that these tonal irregularities in the Northern languages, as opposed to the others, might be explained by their assumption that the monosyllabic morphemes often represent conflations or truncations of earlier longer forms, with tonal aberrations on the later monosyllabic forms reflecting earlier initial consonants or longer segments, which were later lost or reduced.

Let us examine the facts. The sets of words in chart 9, from the top horizontal row of chart 3, show

in the Southwestern and Central languages tones that reflect original voiceless initials, but in the Northern languages they have the tones that normally developed from original voiced initials (as the reader will find if he compares the tones with those given in chart 8).

It is apparent that we are dealing with initial consonants of particular phonetic types, including certain (by no means all) words with initial sibilants and what may possibly turn out to be almost all the words with initial aspirated stops in Proto-Tai. Note that velars showing up as initial *x* in White Tai are conspicuously absent here except for the word for 'right (hand)', which is well known for its tendency to have aberrant shapes in many dialects. Saek has exceptionally the tone reflecting a voiceless initial in the word for 'elbow', in which it disagrees with Yay and agrees with the Southwestern and Central languages.

The tonal phenomena exhibited in chart 9 are so characteristic of the languages of the Northern branch that they have been regarded as being among the most important distinctive criteria for classifying a particular language or dialect as belonging to that branch rather than to the Central or Southwestern groups. But a few Nung dialects have been found in southern Kwangsi, which in most other respects (vocabulary, vowel system, treatment of original initial consonant clusters, and so on) clearly belong to the Central group but curiously behave like languages of the Northern group in showing tonal reflexes of original voiced initials in

a few isolated words. For example, the Lei Ping dialect shows this Northern-like tonal aberration in the words for 'male (animal)', 'cooked, ripe', and 'to dig', and Lungming shows it in the word for 'to dig'. There are other Central dialects, for example, those of Ning Ming and Sze-lok, in which examples of this sort are even more frequent, though paradoxically these latter two points are geographically located toward the middle of the Central area rather than on the northern edge near the linguistic boundary between Central and Northern Tai languages.

Chart 9.

	SIAM-ESE	WHITE TAI	LEI PING	LUNG-MING	YAY	SAEK
A:						
eggplant	-khɨa^{5}	khə^{1}	khəə^{1}	khɨɨ^{1}	kɨa^{4}	khɨa^{4}
son-in-law	khəəy^{5}	khəy^{1}	khəəy^{1}	khuy^{1}	kɨay^{4}	khooy^{4}
hole, pit	khum^{5}	khum^{1}	khom^{1}	khom^{1}	kum^{4}	khum^{4}
right (hand)	khwaa^{5}	xwaa^{1}	l̥aa^{1}	saa^{1}	kwa^{4}	khwaa^{4}
to arrive	thɨŋ^{5}	thɨŋ^{1}	thəŋ^{1}	thəŋ^{1}	taŋ^{4}	thaŋ^{4}
bitter	khom^{5}	khum^{1}	khom^{1}	khom^{1}	ham^{4}	ɣam^{4}
B:						
to ride	khii^{2}	khi^{2}	khwii^{2}	khwey^{2}	kɨay^{5}	khooy^{5}
closely spaced	thii^{2}	thi^{2}	thii^{2}	thay^{2}	ti^{5}	thii^{5}
forest	thɨan^{2}	thən^{2}	thəən^{2}	thɨɨŋ^{2}	tɨan^{5}	thuaɨ^{5}
bean	thua^{2}	tho^{2}	thuu^{2}	thuu^{2}	tua^{5}	thua^{5}
C:						
bowl	thuay^{3}	thoy^{3}	thooy^{3}	thuy^{3}	tɨay^{6}	thooy^{6}
person	phuu^{3}	phu^{3}	phuu^{3}	phow^{3}	pu^{6}	phuu^{6}
excrement	khii^{3}	khi^{3}	khii^{3}	khii^{3}	hay^{6}	ɣay^{6}

	SIAM-ESE	WHITE TAI	LEI PING	LUNG-MING	YAY	SAEK
rice	khaaw^{3}	khaw^{3}	khaw^{3}	khaw^{3}	haw^{6}	ɤaw^{6}
D-short:						
male (animal)	thɨk^{2}	thək^{2}	thək^{4}	tək^{3}	tak^{1}	thak^{6}
ten	sip^{2}	sip^{2}	l̥ip^{2}	sip^{3}	sip^{1}	sip^{6}
cooked, ripe	suk^{2}	suk^{2}	l̥ok^{4}	sok^{3}	suk^{1}	suk^{6}
enemy	sɨk^{2}	sək^{2}	--	--	sak^{1}	--
to bite	khop^{2}	khop^{2}	khop^{2}	khop^{3}	hap^{1}	ɤap^{6}
to dig	khut^{2}	khut^{2}	hot^{5}	khot^{4}	hut^{1}	khut^{6}
D-long:						
elbow	sɔɔk^{2}	sɔʔ^{2}	l̥ook^{2}	sook^{2}	suak^{5}	suak^{6}
D-long > D-short:						
cheap; to hit	thuuk^{2}	thuʔ^{2}	--	thok^{3}	tɨk^{1}	thɨk^{6}

Leaving the top horizontal row of chart 3 and turning now to the second horizontal row, reflecting original initials of the type **p*, **t*, and **k*, we find fewer examples of irregularity in the Northern languages. There are no known examples from the two D columns, but we have too little material to judge whether this gap is accidental or systematic (see chart 10).

Some of these, for example, the word 'to become', show reflexes of an original voiced initial also in the Central languages. One might be tempted to explain such extremely frequent phrase-initial words as the classifier for animals and the word 'to become' as having somehow gotten their tone distorted

Chart 10.

	SIAM-ESE	WHITE TAI	LEI PING	LUNG-MING	YAY	SAEK
A:						
wasp	tɛɛn^{1}	tɛn^{1}	pheeŋ^{1}	pheeŋ^{1}	tin^{4}	thiil^{4}
classifier for animals	tua^{1}	to^{1}	tuu^{2}	tuu^{1}	tua^{4}	thua^{4}
to become	pen^{1}	pin^{1}	phin^{4}	pin^{4}	pan^{4}	phal^{4}
to beat	tii^{1}	ti^{1}	--	tey^{1}	ti^{4}	--
deer	kwaaŋ^{1}	kwaaŋ^{1}	--	--	vɨaŋ^{4}	vuaŋ^{4}
B:						
to crawl	taay^{2}	tay^{2}	--	--	raay^{5}, ray^{5}	tay^{6}
C:						
shrimp	kuŋ^{3}	kuŋ^{3}	koŋ^{3}	koŋ^{3}	kuŋ^{2}	kuŋ^{6}
to boil	tom^{3}	tum^{3}	tom^{3}	tom^{3}	tum^{2}	rɔɔm^{6}
wide	kwaaŋ^{3}	kwaaŋ^{3}	kwaaŋ^{3}	kwaaŋ^{3}	kwaaŋ^{2}	--

in weak-stressed phrase-initial position in some branches but not in others. On the other hand, it may not be coincidence that among the few irregular words found in this horizontal row, those in column A shift down to the bottom box and those in column C shift to column B. The hints of a systematic pattern here suggest the possibility that we may be on the track of some phonetic feature that somehow distinguished the initials of these words from the usual unaspirated stops.

The complete lack of examples of tonal irregularties for the third horizontal row of chart 3, which includes forms having an original initial

glottal stop or preglottalized consonant, reinforces our impression that there may be a systematic phonological pattern at work here.

When we get to the bottom row, we find a good many words in the Northern languages exhibiting the tones that would normally reflect original voiceless initials, the reverse of the phenomenon we have just been examining. The morphemes in chart 11 show tonal irregularities in the Northern Tai languages Yay and Saek.

In the A column, we note that both Yay and Saek have the tones associated with original voiceless initials in the words for 'to come', 'man', and 'housefly'. In the other words only Saek shows the variation, not Yay. The clustering of these examples around two types, *r*-words and *f*-words, is striking, with two more examples of *r*-words showing up in the D column. The words for 'older sibling' and 'to defeat, be defeated' show shift between the B and C columns, and may be accidental irregularities in spite of the interesting coincidence in initial consonant.

With regard to the *r*- words, there is another curious puzzle, which seems to have been noticed and which may or may not be relevant here. A great many words that have initial *r* in Siamese have in certain Nung dialects initial *l̥*, which represents a sound more commonly found in the Nung area representing the initial *s*, of whatever historical origin, of other Tai languages. Now, the curious thing about this occurrence of *l̥* for *r* in certain Nung dialects is that it apparently never occurs in words of the type

Chart 11.

	SIAM-ESE	WHITE TAI	LEI PING	LUNG-MING	YAY	SAEK
A:						
to come	maa^{1}	maa^{4}	maa^{4}	maa^{4}	ma^{1}	maa^{2}
man	chaay^{1}	caay^{4}	--	caay^{4}	θaay^{1}	saay^{2}
housefly	-wan^{1}	-mun^{4}	--	--	-ɲan^{1}	-ɲɛl^{2}
boat	rɨa^{1}	hə^{4}	ləə^{4}	lɨɨ^{4}	rua^{4}	rua^{1}
day after tomorrow	-rɨɨn^{1}	hɨ^{4}	ləə^{4}	ləɣ^{4}	rɨ^{4}	rɨɨ^{1}
to float	fuu^{1}	fu^{4}	fuu^{4}	fow^{4}	fu^{4}	vuu^{1}
firewood	fɨɨn^{1}	--	fən^{4}	fən^{4}	fun^{4}	vɨl^{1}
to roof	muŋ^{1}	muŋ^{4}	foŋ^{4}	foŋ^{4}	foŋ^{4}	vɔɔŋ^{1}
straw	faaŋ^{1}	fəŋ^{4}	faaŋ^{4}	faaŋ^{4}	fɨaŋ^{4}	viaŋ^{1}
to slash	fan^{1}	fan^{4}	--	--	fun^{4}	val^{1}
B:						
older sibling	phii^{3}	pi^{5}	phii^{5}	pey^{5}	pi^{6}	phii^{5}
C:						
to defeat, be defeated	phɛɛ^{4}	pɛ^{6}	--	--	pe^{5}, pe^{6}	phɛɛ^{5}
D-long:						
root	raak^{3}	haa^{4}	laak^{5}	laak^{5}	raak^{5}	raak^{6}
bedbug	rɨat^{3}	hət^{4}	ləət^{5}	lɨɨt^{5}	rɨat^{5}	ruat^{6}

Chart 12.

	SIAM-ESE	WHITE TAI	LEI PING	LUNG-MING	YAY	SAEK
long	rii^{1}	hi^{4}	l̥ii^{4}	ley^{4}	ray^{4}	ray^{4}
to know	ruu^{4}	hu^{6}	l̥uu^{3}	low^{6}	ro^{6}	rɔɔ^{6}
to leak	rua^{3}	ho^{5}	l̥uu^{5}	luu^{5}	ro^{5}	rɔɔ^{5}
dry field	ray^{3}	hay^{5}	l̥ay^{5}	lay^{5}	ri^{5}	rii^{5}
track, footprint	rɔɔy^{1}	hɔy^{4}	l̥ooy^{4}	looy^{4}	ri^{4}	rii^{4}
house	rɨan^{1}	hən^{4}	l̥əən^{4}	lɨɨn^{4}	raan^{4}	raan^{4}
to vomit	raak^{3}	haa^{4}	l̥aak^{5}	laak^{5}	ruak^{5}	ruak^{5}
chicken louse	ray^{1}	--	l̥ay^{4}	lay^{4}	--	rii^{4}
strength	rɛɛŋ^{1}	hɛŋ^{4}	l̥eeŋ^{4}	leeŋ^{4}	reŋ^{4}	reeŋ^{4}
we	raw^{1}	haw^{4}	l̥aw^{4}	law^{4}	raw^{4}	roo^{4}
to receive	rap^{4}	--	l̥ap^{4}	lap^{4}	--	rap^{6}
bran	ram^{1}	ham^{4}	l̥am^{4}	lam^{4}	ram^{4}	ram^{4}
squirrel	-rɔɔk^{3}	hɔ^{4}	l̥ook^{5}	lok^{5}	rok^{5}	rɔɔk^{5}
light, daybreak	ruŋ^{3}	huŋ^{5}	l̥oŋ^{5}	loŋ^{5}	roŋ^{5}	rɔɔŋ^{5}
to string	rɔɔy^{4}	hɔy^{6}	l̥ooy^{3}	looy^{6}	roy^{6}	rɔɔy^{6}
gutter, trough	raaŋ^{1}	haaŋ^{4}	l̥aaŋ^{4}	laaŋ^{4}	ruaŋ^{4}	--

illustrated above, in which Northern languages show a tonal shift, but only in words that everywhere, in all branches of Tai, including the Northern, always show the tones associated with original voiced intials. Must we posit two types of *r*-sound for the parent language? In the list of words in chart 12 having initial *r* in Siamese, the Lei Ping forms show

initial *l̥*, contrasting with initial *l* in the Lei Ping forms in chart 11.

In our discussion of the evidence that concerns Question 2 we have uncovered a number of unsolved problems, but it seems fair to characterize all this varied evidence as tending to point toward systematic phonological explanations, rather than an explanation based on any morphophonemic alternation in the parent language or on the loss of preceding elements. It seems likely that when all these apparent tonal irregularities in the languages of the Northern branch are understood, we will find that this group has made its tonal splits on an entirely different basis from the other Tai languages, classifying the conditioning features of initial consonants according to some different system that now eludes us.

Question 3. What was the inventory of consonant initials in Proto-Tai?

The general outlines of the consonant system that must be assumed for Proto-Tai have been clear for a long time. There must have been an unaspirated set, **p, t, c, k*, and a corresponding aspirated set, **ph, th, kh*, with apparently a curious gap in the **ch* position. Voiceless fricatives must have included **s, f, x*, and *h*, and there was a series of sonorants with voiceless onset of the types **hm, hl*, and so on. Proof from tonal evidence has established the series **ʔb, ʔd, ʔy*, affecting tones in the same way as **ʔ*.[16] Voiced initials included the obstruents **b, d, j*, and *g*, such fricatives as **v, z*, and *ɣ*, and sonorants such as **m, n, ɲ, ŋ, l, r, w*, and *y*.

Many puzzling irregularities were clarified by Li's historic article on consonant clusters,[17] in which he posited *l* and *r* clusters to account for a great many sets of cognates showing apparently aberrant correspondences in some dialects, especially those of the Central branch.

Some of the material cited above raises doubts, however, as to the degree of completeness of our understanding of the system of initial consonants in Proto-Tai, and there are still other troublesome bits of evidence concerning initial consonants. Consider, for example, the following confusing sets of cognates.

	SIAM-ESE	WHITE TAI	LEI PING	LUNG-MING	YAY	SAEK
civet cat	hen^{5}	hin^{1}	hin^{1}	hin^{1}	ɲan^{1}	ɲɛ1^{2}
to sew	yep^{4}	ŋip^{4}	yap^{4}	yap^{4}	ɲip^{1}	ɲip^{6}
to hear	yin^{1}	ŋin^{4}	yin^{4}	yin^{4}	--	--
to do, make[18]	yet^{4}	yet^{4}	het^{2}	hat^{3}	--	hit^{6}
to shoot, aim	yiŋ^{1}	ɲiŋ^{4}	--	--	ŋiŋ^{4}	ɲiiŋ^{4}

The miscellaneous sets of cognates just cited no doubt involve a number of different original consonants. Probably Proto-Tai initials of the types **y*, **ɲ*, and perhaps also **ŋ*, are to be reconstructed. But we are apparently not yet able to determine which original consonant occurred in each of the sets, or which ones involve aspiration or preglottalization, or to what degree the initial consonants have been affected by contact with original front vowels, or, conversely, which words have undergone vowel fronting

or raising because of the initial consonant. The main outlines of Proto-Tai consonantism have been so widely agreed upon that one feels somewhat mischievous in bringing forth these more intractable examples, but we have to face the fact that so long as such problematic cases as these remain unsolved we cannot feel certain of any overall reconstruction of either the consonant system or the vowel system of Proto-Tai.

Saek, which is a Northern Tai language now located in Southwestern Tai territory, holds special interest for the Tai comparatist because it shows some illuminating archaic features. For example, Saek provides beautiful confirmation of some of Li's reconstructions of initial consonant clusters. In the three sets of cognates presented in chart 13, for example, compare the Saek forms with Li's Proto-Tai reconstructions.

Chart 13.

	SIAM-ESE	WHITE TAI	LEI PING	LUNG-MING	YAY	SAEK
Proto-Tai *phl-, phr-:						
vegetable	phak^{2}	phak^{2}	phyak^{2}	phyak^{3}	piak^{2}	phrak^{4}
hair of the head	phom^{5}	phum^{1}	phyom^{1}	phyom^{1}	piam^{1}	phram^{2}
cliff, rock	phaa^{5}	phaa^{1}	phyaa^{1}	phyaa^{1}	--	phraa^{2}
forehead	phaak^{2}	--	phyaak^{2}	phyaak^{2}	pyaak^{2}	phraak^{6}
thin	phɔɔm^{5}	phɔm^{1}	--	yoom^{1}	pyom^{1}	phrɔɔm^{2}
to burn	phaw^{5}	phaw^{1}	--	phyaw^{1}	piaw^{1}	phraw^{2}

	SIAM-ESE	WHITE TAI	LEI PING	LUNG-MING	YAY	SAEK
Proto-Tai *pr-:						
to expose to the sun	taak^{2}	taa?^{2}	phyaak^{2}	phyaak^{2}	taak^{2}	praak^{6}
to break	tɛɛk^{2}	tɛ?^{2}	pheek^{2}	pheek^{2}	tek^{2}	preek^{6}
cucumber, melon	tɛɛŋ^{1}	tɛŋ^{1}	pheeŋ^{1}	pheeŋ^{1}	tiaŋ^{1}	priaŋ^{1}
bamboo strip	tɔɔk^{2}	tɔ?^{2}	phyook^{2}	phyook^{2}	tuk^{3}	pruk^{4}
Proto-Tai *?bl- or ?br-:						
gall bladder	dii^{1}	bi^{1}	dii^{2}	ney^{4}	di^{1}	blii^{1}
moon, month	dɨan^{1}	bən^{1}	bəən^{2}	mɨɨn^{4}	dɨan^{1}	blian^{1}
flower	dɔɔk^{2}	bɔ?^{2}	byook^{5}	myook^{2}	dok^{2}	blɔɔk^{6}

Saek is not so close to Li's reconstructions in other cases of initial consonant clusters; indeed, in many respects Saek turns out to be no more archaic than most other Tai languages.

When we examine initials that have not been suspected of involving original clusters, Saek sometimes presents problems that current views concerning consonant initials in Proto-Tai cannot handle. Consider, for example, in chart 14, the variety of Saek correspondences for Siamese initial *d*.

Chart 14.

	SIAMESE	SAEK
good	dii^{1}	dii^{1}
to smell (something)	dom^{1}	dam^{1}
earthworm	dɨan^{1}	trual^{1}
star	daaw^{1}	traaw^{1}
to transplant	dam^{1}	tram^{1}
relatives by marriage	dɔɔŋ^{1}	trɔɔŋ^{1}
gall bladder	dii^{1}	blii^{1}
moon, month	dɨan^{1}	blian^{1}
flower	dɔɔk^{2}	blɔɔk^{6}
cockspur	dɨay^{1}	praa^{1}
red	dɛɛŋ^{1}	riiŋ^{1}
sunshine	dɛɛt^{2}	riit^{6}
raw, unripe	dip^{2}	rip^{4}
stairs, ladder	-day^{1}	ray^{1}
bone	-duuk^{2}	rɔɔk^{6}
mountain	dɔɔy^{1}	rɔɔy^{1}
catfish	duk^{2}	rɔk^{4}

We are not in a position at the present state of our knowledge to account for the Saek initials in all these forms, nor, conversely, for all the sources of Saek initial *y* corresponding to several different Siamese initials in such forms as those given in chart 15.

Chart 15.

	SIAMESE	SAEK
to step on	yiap^{2}	yiap^{6}
granary	yaaw^{3}	yiaw^{3}
to be in a place	yuu^{2}	yuu^{6}
to roast	yaaŋ^{3}	yuaŋ^{3}
paternal grandmother	yaa^{3}	yaa^{5}
kind	yaaŋ^{2}	yaaŋ^{6}
lover	chuu^{4}	yuu^{6}
eaves	chaay^{1}	yaay^{4}
sand	saay^{1}	yooy^{4}
to wash (clothes)	sak^{4}	yak^{6}
fishtrap	say^{1}	yay^{4}
straight	sɨɨ^{3}	yɔɔ^{5}

Saek shows initial *kw* in two words, which elsewhere in Tai languages have a velar, with no trace of labial, but which, curiously, turn up with initial *p* in two dialects of Sui cited by Li.[19]

	SIAMESE	SAEK	SUI
leg	khaa^{5}	kwaa^{1}	pa
to sell	khaay^{5}	kwaay^{1}	pe

The word for 'stairs, ladder' shows such aberrant correspondences in initial in the various branches of Tai that Li posited for it a special consonant cluster in the parent language, with only this one example.[20] Confirmation that this word is

somehow unique turns up in the Southwestern Tai dialect of a place called in White Tai *meŋ*^{4} *caaŋ*^{3}, fifty kilometers south of Ha Giang in North Vietnam.[21] There the word has the form *dway*, with an initial cluster for which no other examples have so far been found in this dialect.

Many more miscellaneous examples of forms with puzzling initial correspondences could be cited, especially from the more remote languages of the Northern and Central branches, on which we are beginning to get more data than were available in the past. But perhaps enough troublesome evidence has been presented in the various sets of examples above to make the point that we are farther from having all the answers regarding initial consonants in comparative Tai than might be suggested by the relative transparency of initial consonants in the great majority of words of the inherited vocabulary.

Besides the arduous work of collecting and sorting and struggling over forms involving troublesome problems in initial consonants, are there broader paths open to us that may hold out hope for clarification? Two such paths, at least, seem promising.

For one thing, it may be that in trying to cope with these intractable cases we are in danger of losing sight of the forest because of the trees. It may be helpful to adopt as a tentative hypothesis the view that the consonant system of Proto-tai was highly symmetrical, perhaps even more symmetrical than has been suspected in the past, with stops and nasals in four positions.

p t c k
m n ɲ ŋ

and that these occurred with modifications such as

ph th ch kh
hm hn hɲ hŋ
ʔm ʔn ʔɲ ʔŋ
b d j g

For the preglottalized series, various scholars already have suggested a notation with symbols for nasals rather than *ʔb* and the like;[22] many widely separated dialects show nasal rather than oral stops as reflexes of this series, including the Lungming dialect cited throughout this paper. Glottal stop would then occur in the position occupied by **ʔŋ*. Initial **ch* has commonly been regarded as not occurring in the parent language; perhaps it did occur after all, and may therefore be available to account for some otherwise inexplicable correspondences.

And it may be that in the process of reassessing the material along these lines we will find that we can improve upon Li's reconstruction of initial consonant clusters. Although most of his reconstructions will undoubtedly stand up under further study, he was in some cases uncertain of the exact shape of the original cluster, and there are some sets of correspondences which seem likely to be subject to further subdivision. For example, one might test a hypothesis that in Proto-Tai such a sound as *m* occurred in the same clusters, such as *mr* and *ml*,

regardless of whether it was preglottalized (*ʔm*) or preaspirated (*hm*) or unmodified in any such way.

The possibility that the entire palatal series (*c, j, ɲ*) could be eliminated ought to be examined. For various modern Tai languages analyses of palatals as consisting of dentals plus *y* (*ty* instead of *c*, for example) have sometimes been suggested. The defectiveness of this series in one way or another in most dialects, and its apparent failure to cluster further with *l* and *r*, are suspicious, as is the striking fact that for the other three articulatory positions there are the three corresponding finals *p, t,* and *k* (and also *m, n,* and *ŋ*), but no final palatals.

Thus one possible path that perhaps seems more promising than struggling with individual cases is to attempt a bold reassessment of the entire consonant system.

Another promising line of attack lies in greater use of the distinctive-feature approach and closer attention to articulatory phenomena in formulating a theory of the structure of the parent language and of the changes in the various branches. The most noteworthy contribution along these lines made thus far is found in the brilliant Cornell dissertation of 1962 by J. Marvin Brown,[23] who studied a great number of dialects, mostly in Thailand, and developed an articulatory hypothesis to explain the precise phonetic mechanics of the sound changes in which each dialect, as he puts it, unloaded part of the burden of the initials onto the tones. His approach is from the bottom up, so to speak, using a fine network of rather closely related dialects and working gradually

backward in time. To apply his methods to the entire Tai family we will need data from a much more closely spaced network of geographical points throughout the Tai domain than we now have. It is also conceivable, of course, that Brown's hypotheses and techniques could be applied boldly to the data we now possess, even though for much of the Tai-speaking domain the available data come from rather widely separated places.

It seems possible that as time passes we will find that the major contribution of the present generation of students of comparative Tai has been to collect and organize the data, and that students of a younger generation, more interested in and more adept at newer and more sophisticated approaches to problems of phonological structure and sound change, and less burdened with the arduous work of collecting and analyzing new data, will achieve a coherent explanation of Tai consonantism.

Question 4. What was the vowel system of Proto-Tai?

The vowel system of Siamese, which was displayed in chart 1, is famous for its symmetry. So long as we remain within the Southwestern and Central areas we find the general dimensions of this pattern repeated, though with drastic internal modifications. There is a rather large area, for example, including White Tai, Lue, and Shan, in which the three diphthongs *ia*, *ɨa*, and *ua* are lowered to *e*, *ə*, and *o*. Although many Tai languages have a mid central or mid back unrounded *ə*, we find that Siamese forms containing this vowel turn out either to have no cognates in

other Tai languages or to show derivation of this vowel from some other source, for example, *ɨa*, by secondary processes of one sort or another. Moreover, Siamese long *ee* and long *oo*, and Siamese short *ɛ* and short *ɔ*, are found to be of secondary origin, showing no correlation with vowels in other Tai dialects, so that in these two parts of the vowel system, low front and low back, we are forced to the conclusion that there were earlier only one low front vowel, occurring short and long (reflected in Siamese short *e* and long *ɛɛ*), and only one low back vowel, also occurring short and long (reflected in Siamese short *o* and long *ɔɔ*). Some languages, such as White Tai, coalesce the mid vowels with the corresponding higher vowels (*e* > *i*, and the like) before nasal finals.

But all the variations of the sorts just cited seem, in the Southwestern and Central branches, to take place within the vertical and horizontal dimensions represented in our chart 1 of Siamese and White Tai vowels. When we turn to the Northern languages, however, we find that, although the vowel systems show the same general typological structure as in Southwestern and Central languages, cognate forms often exhibit the wildest variation in the Northern languages; for many words there is complete agreement with other Tai languages, but for many others the variations in vowels are so extreme that we are at a complete loss. Consider the sets in chart 16, which show some of the diversity in Northern vowel correspondences for various Siamese vowels.

Chart 16.

	SIAM-ESE	WHITE TAI	LEI PING	LUNG-MING	YAY	SAEK
sand	saay^{1}	saay^{4}	l̥aay^{4}	saay^{4}	θɨay^{4}	yooy^{4}
to dis-appear	haay^{5}	--	--	haay^{1}	rɨay^{1}	rəəy^{1}
to vomit	raak^{3}	haa^{4}	l̥aak^{5}	laak^{5}	ruak^{5}	ruak^{5}
hungry	yaak^{2}	--	yaak^{5}	yaak^{2}	yiak^{2}	yuak^{6}
to enter	khaw^{3}	xaw^{3}	khaw^{3}	khaw^{3}	haw^{3}	haw^{3}
knee	khaw^{2}	xaw^{2}	khaɣ^{2}	khaw^{2}	ho^{2}	kɔɔ^{6}
empty	plaaw^{2}	paw^{2}	pyaw^{5}	pyaw^{2}	pyu^{2}	pluu^{6}
fire	fay^{1}	fay^{4}	fay^{4}	fay^{4}	fi^{4}	vii^{4}
heart	cay^{1}	caɣ^{1}	--	--	sɨ^{1}	cɨɨ^{1}
new	may^{2}	maɣ^{2}	maɣ^{2}	maɣ^{2}	mo^{2}	mɔɔ^{6}
to dye	yɔɔm^{4}	ɲɔm^{6}	yoom^{3}	yoom^{6}	ɲum^{6}	ɲum^{6}
to sleep	nɔɔn^{1}	nɔn^{4}	noon^{4}	noon^{4}	nin^{4}	nuun^{4}
younger sibling	nɔɔŋ^{4}	nɔŋ^{6}	nooŋ^{3}	nooŋ^{6}	nuaŋ^{6}	nuaŋ^{6}
track, footprint	rɔɔy^{1}	hɔy^{4}	l̥ooy^{4}	looy^{4}	ri^{4}	rii^{4}
flower	dɔɔk^{2}	bɔʔ^{2}	byook^{5}	myook^{2}	dok^{2}	blɔɔk^{6}
bamboo strip	*tɔɔk*^{2}	*tɔʔ*^{2}	*phyook*^{2}	*phyook*^{2}	tuk^{3}	*pruk*^{4}
to have	mii^{1}	mi^{4}	mii^{4}	mey^{4}	mi^{4}	mii^{4}
long	rii^{1}	hi^{4}	l̥ii^{4}	ley^{4}	ray^{4}	ray^{4}
straight	sɨɨ^{3}	sɨ^{5}	--	səɣ^{5}	θo^{5}	yɔɔ^{5}
navel	-dɨɨ^{1}	bi^{1}	--	--	dɨa^{1}	dua^{1}

	SIAM-ESE	WHITE TAI	LEI PING	LUNG-MING	YAY	SAEK
snake	ŋuu^{1}	ŋu^{4}	vuu^{4}	ŋow^{4}	ŋɨa^{4}	ŋua^{4}
to know	ruu^{4}	hu^{6}	l̥uu^{3}	low^{6}	ro^{6}	rɔɔ^{6}
crab	puu^{1}	pu^{1}	puu^{2}	pow^{1}	paw^{1}	paw^{1}
child	luuk^{3}	lu^{4}	lok^{5}	lok^{5}	lɨk^{1}	lɨk^{6}
high	suuŋ^{5}	suŋ^{1}	l̥oŋ^{1}	soŋ^{1}	θaaŋ^{1}	saaŋ^{2}
salt	klɨa^{1}	kə^{1}	kəə^{2}, cəə^{2}	kyɨɨ^{1}	kua^{1}	tlua^{1}
rope	chɨak^{3}	cə^{4}	səək^{5}	cɨɨk^{5}	saak^{5}	saak^{5}
silver	ŋen^{1}	ŋɨn^{4}	ŋən^{4}	ŋən^{4}	ŋan^{4}	ɲɛn^{4}
red	dɛɛŋ^{1}	dɛŋ^{1}	deeŋ^{2}	neeŋ^{4}	diŋ^{1}	riiŋ^{1}
mother	mɛɛ^{3}	mɛ^{5}	mee^{5}	mee^{5}	me^{5}	mee^{5}
to laugh	hua^{5}	--	vuu^{1}	luu^{1}	riaw^{1}	ruaw^{2}
to leak	rua^{3}	ho^{5}	l̥uu^{5}	luu^{5}	ro^{5}	rɔɔ^{5}
garden	suan^{5}	son^{1}	l̥uun^{1}	suun^{1}	θɨan^{1}	suan^{2}
deaf	nuak^{2}	no?^{2}	nuuk^{2}	nuuk^{2}	nuk^{3}	nuuk^{6}
mountain stream	huay^{3}	hoy^{3}	khooy^{3}	luy^{3}	vi^{3}	rii^{3}
dog	maa^{5}	maa^{1}	maa^{1}	maa^{1}	ma^{1}	maa^{2}
grass	yaa^{3}	ɲaa^{3}	yaa^{3}	yaa^{3}	ɲia^{3}	ɲua^{3}
crossbow	naa^{3}	naa^{3}	--	--	nɨa^{3}	nua^{3}
cloud	faa^{3}	faa^{3}	phaa^{3}	phaa^{3}	vɨa^{3}	via^{3}

In these examples, the vowels of the Southwestern and Central forms, though varied, are readily explained along the lines suggested above, except for an occasional isolated aberrant form such as the Lei Ping word for 'knee'.

No one, however, has yet published a hypothesis as to the vowel system of Proto-Tai, which would handle the deviant forms, exemplified here by Yay and Saek.

Li has suggested that where the Northern languages have, for example, long *aa* corresponding to a diphthong in Siamese, as in the word for 'house', an original post-initial semivowel may have been involved.[24] This seems plausible, and a number of types of deviation might be handled by such an approach, though not all. There also may have been influences, differing from one branch to another, of other initials, and perhaps also finals, on vowels. Obviously it will be necessary to collect all sets of cognates showing the same deviant vowel correspondences, and then to search for similar types of deviation, to reveal whether any such conditioning factors can have been at work in one branch or another.

Ultimately it seems clear that someone must reconstruct a vowel system that will account for the vowel system of the Northern Tai languages on the one hand and of the Southwestern and Central Tai languages on the other, and it seems likely that this will have to be daringly different from anything immediately suggested by the vowel system of any of the modern languages.

With respect to vowels, the Northern branch is markedly different from the others, but, as in the case of consonants, we find a few isolated instances of spillover of Northern-like features into some dialects of the Central group. For example, Ning Ming in southern Kwangsi has *phɨɨ*^{3} 'cloud' and *ŋɨɨ*^{4} 'snake'; the vowel *ɨɨ* in this dialect regularly reflects an earlier **ɨa*, so that this Central dialect has Northern-like vocalism on these two words.

The Central dialect of Lungming has a peculiarity that may be explained finally as a local secondary development. Or, on the other hand, it may turn out to have a much older and more significant origin. In this dialect many words having medial short *e* in Siamese show medial *-ya-*, producing all manner of initial *y*-clusters which would not otherwise occur. A few examples will illustrate the principle.

	SIAMESE	LUNG-MING	LEI PING
duck	pet^{2}	pyat^{3}	pit^{2}
fishhook	bet^{2}	myat^{3}	bit^{2}
spicy hot	phet^{2}	phyat^{3}	phit^{2}
to pick up	kep^{2}	kyap^{3}	kip^{2}
see also:			
ten	sip^{2}	sip^{3}	l̥ip^{2}

In some of the Central dialects of southern Kwangsi, including Lungming, high vowels have been diphthongized, for example, *ii* > *ey*, and in some dialects of the same group diphthongs beginning with a high vowel have been simplified to monophthongs,

for example, *ɨa* > *ɨɨ*. The details of these changes vary from dialect to dialect, but are in every case so transparent as to suggest a relatively recent date for the shifts. Diphthongization of high vowels similar to that found at Lungming also has occurred in some Tai languages of the Northern branch in the same area, for example, Wu-ming, and in some non-Tai languages even including Cantonese. (Can Cantonese, which is the culturally and politically dominant language in this area, have triggered this change in the minority languages?)

But clearly much more ancient, and much more problematical, are sporadic cases of variation from one dialect to another in certain morphemes between *ii* and *ay*, *ɨɨ* and *ay̆*, and *uu* and *aw*, apparently without pattern, since any given dialect is found to have some diphthongized forms and some nondiphthongized ones (see chart 17).

The underlined forms in chart 17 reflect the diphthongs *ay*, *ay̆*, *aw* (in some languages subsequently altered, for example, *ay̆* > *əə* in Saek). Forms not underlined reflect the three high vowels *i*, *ɨ*, and *u* (in some languages subsequently altered, for example, *i* > *ey*, *ɨ* > *əy̆*, *u* > *ow* in Lungming). While we may admire the phonological symmetry displayed here, we are at a loss to account for the apparent freedom of any particular dialect to choose between the two available forms, one monophthongal and the other diphthongal. It can hardly be a case of original doublets, since any given dialect has always only one of the two forms, never both. This phenomenon is apparently to be found throughout the Tai family, in

Chart 17.

	SIAM-ESE	WHITE TAI	LEI PING	LUNG-MING	YAY	SAEK
good	dii^{1}	di^{1}	day^{2}	nay^{4}	di^{1}	dii^{1}
dirt	khlay^{1}	--	--	lay^{4}	hi^{4}	ɣii^{4}
thread	may^{5}	may^{1}	may^{1}	may^{1}	may^{1}	mii^{2}
excrement	khii^{3}	khi^{3}	khii^{3}	khii^{3}	hay^{6}	ɣay^{6}
fire	fay^{1}	fay^{4}	fay^{4}	fay^{4}	fi^{4}	vii^{4}
closely spaced	thii^{2}	thi^{2}	thii^{2}	thay^{2}	ti^{5}	thii^{5}
long	rii^{1}	hi^{4}	l̥ii^{4}	ley^{4}	ray^{4}	ray^{4}
dry field	ray^{3}	hay^{5}	l̥ay^{5}	lay^{5}	ri^{5}	rii^{5}
chicken louse	ray^{1}	--	l̥ay^{4}	lay^{4}	rii^{4}	rii^{4}

	SIAM-ESE	WHITE TAI	LEI PING	LUNG-MING	YAY	SAEK
heart	cay^{1}	cay̶^{1}	--	--	si̶^{1}	cii̶^{1}
correct	chay^{3}	--	chɔɔ^{5}	cɔy̶^{5}	si̶^{5}	--
nine	kaaw^{3}	kaw^{3}	kaw^{3}	kaw^{3}	ku^{3}	kuu^{3}
paternal grand-father	puu^{2}	pu^{2}	--	--	paw^{2}	--
crab	puu^{1}	pu^{1}	puu^{2}	pow^{1}	paw^{1}	paw^{1}
empty	plaaw^{2}	paw^{2}	pyaw^{5}	pyaw^{2}	pyu^{2}	pluu^{6}

all three branches. If we were including data from additional dialects in this study we would be able to add other examples; for instance, many dialects have the i̶ vowel in the word 'to give' rather than the diphthong found in Siamese *hay*^{3}, White Tai *hay̶*^{3}, and

Yay *hay*^{3}. This problem, which appears not to have been noted previously, will constitute a special challenge to any scholar attempting to work out comparative Tai vocalism, because, whereas other vowel problems, though complicated, involve regularity in correspondences from one language to another, in this case the alternation between monophthongs and diphthongs appears to be sporadic and unpredictable.

Question 5. How are the main branches of the Tai family to be defined?

There is general agreement that the Northern Tai languages form a separate branch. Two important questions arise. First, what are the defining features of the Northern branch? Second, do the Southwestern and Central groups constitute a single branch as opposed to the Northern, or are there definite criteria for distinguishing between Southwestern and Central languages, and if so, where is the geographical boundary? To put it another way, does the Tai family have two or three main branches?

Languages of the Northern branch have been shown by Li and Haudricourt to share a number of lexical items not found in other Tai languages. The Yay and Saek forms of some of their examples, to which I have ventured to make some additions, are given in chart 18.

Perhaps more important for defining the Northern branch are phonological criteria, for which the special tonal developments and the deviations in vowels described earlier are basically important. The difficulty here is that we are beginning to obtain data

Chart 18.

	YAY	SAEK
spirit	faaŋ^{4}	maaŋ^{4}
thorn	ʔwan^{1}	ʔɔn^{1}
on, above	kɨn^{4}	khɨn^{4}
tiger	kuk^{3}	kuuk^{6}
iron	fa^{4}	maa^{4}
lazy	cik^{3}	tliik^{6}
grass mat	bin^{6}	biin^{3}
sky	bun^{1}	bɨn^{1}
girl	bɨk^{3}	bɨk^{4}
wife	pa^{4}	phaa^{4}
young shoot	raaŋ^{4}	naaŋ^{4}
wing	fɨat^{5}	viat^{5}
finger	ɲiaŋ^{5}	ŋiaŋ^{5}

on various Nung dialects spoken in southern Kwangsi, which clearly belong to Li's Central branch, but which occasionally, in isolated words, share in one or another of the special vocalic or tonal features otherwise characteristic of Northern Tai languages exclusively.

The Ning Ming dialect, for example, has already been cited as having sporadic Northern-like forms. In this dialect we also find on some words the tone that otherwise reflects an original voiced initial (see also Ning Ming *maa*^{4} 'to come' and *naa*^{4} 'rice-field'), as in Northern Tai languages, whereas other Southwestern and Central languages always show in the same words the tone that reflects an original voiceless initial.

	NING MING	SIAMESE
right (hand)	kwaa^{4}	khwaa^{5}
son-in-law	kuuy^{4}	khəəy^{5}
needle	hom^{4}	khem^{5}
hole, pit	kom^{4}	khum^{5}

Clearly further study is needed in order to ascertain whether we can define phonological criteria for the Northern branch that will exclude such scattered Northern-like features found in a few Central dialects. If not, then we must be content with regarding these as borderline transitional dialects between the Central and Northern groups.

As regards the division between the Central and Southwestern branches, Li has again proposed lexical and phonological criteria for separating these two groups. With respect to lexicon, it seems fair to say that the Central dialects tend not so much to have distinctive vocabulary items peculiar to this group as they do to share certain words with the Northern languages but not with the Southwestern ones, and in other cases vice-versa. As to phonology, these Central dialects are conspicuous for tending to retain bizarre reflexes of original consonant clusters; for example, the words for 'eye' and 'to die', with initial *t* in all Southwestern and Northern languages except Saek, appear in Central dialects with such initials as *th*, *h*, *ph*, or *phy*.

There is no question that there are basic differences between the Central and Southwestern branches. Clearly the dialects called Tho and Nung in northeastern North Vietnam and adjacent parts of

southern China belong to the Central group, and it is just as clear that in western North Vietnam the White Tai, Black Tai, and Red Tai dialects belong to the Southwestern group, along with Lao, Lue, Siamese, Shan, and so on. But scattered across the intervening part of North Vietnam, between Tho and Nung in the east and White Tai and the others in the west, are a great many local varieties of Tai speech that are still very imperfectly known. Very often we find that the speakers are called Tho by others but Tai by themselves.

In some respects these intermediate dialects are very much like Southwestern languages. In the dialect near Ha Giang cited above, for example, the word *rɨan* 'house' shows an initial consonant and a medial diphthong, which are found also in Siamese *rɨan*^{1} but which intervening dialects have modified in one way or another (see also White Tai *hən*^{4}). On the other hand, this dialect shows such characteristically Central reflexes of original clusters as *phyaa* 'eye', *pyaa* 'fish', *phyom* 'hair of the head', and *suu* 'ear' (all with mid-level tone), or *myɛŋ* 'insect' and *pyaw* 'supper' (both with the high tone found also in *rɨan* 'house').

It seems possible that when we someday have more data on these intermediate dialects we will find that between the Central and Southwestern groups there is no boundary, only gradual transition. An interesting further question will then be whether the gradual transition between Southwestern and Central is on a different scale from the traces of transition discernible between the Northern and Central branches.

These questions seem likely to determine whether Tai has three main branches, as Li has proposed, or only two, the Northern as opposed to all the others. My own prediction, based more on impressions than on solid evidence, is that we will ultimately find that languages of what are now called the Central and Southwestern groups (Haudricourt's "Tai properly so-called") form a continuous dialect area, with only gradual transition throughout and no real language boundary anywhere, while between all these languages on the one hand and the Northern Tai languages on the other we will find that there is a genuine linguistic boundary, giving two rather than three main branches for the Tai linguistic family.

Question 6. To what degree need our reconstruction of Proto-Tai be dependent upon external information?

The earliest students of the history of the more familiar Tai languages such as Siamese took as their starting point various clues provided by the writing system. For example, in Siamese and other Southwestern Tai languages having alphabets of Indian origin, the sounds that we believe reflect original voiced consonants are spelled with letters that in the Indic alphabets represent such sounds as *b*, *d*, *g*, *m*, *n*, *l*, *r*, and *y*. The sounds that we reconstruct as **hm* and the like are still spelled this way in these archaic writing systems. Even **?y* is still so spelled in a few Siamese words such as *yuu*^{2} 'to be in a place'. Archaic alphabets preserving such features are used to write Tai languages in Assam, Burma, Yunnan, Laos, Thailand, and western North Vietnam;

the implication, of course, is that they were adopted in these areas before the great sound shifts occurred.

Our indebtedness to the hints provided by these spelling systems is immense, and one wonders how long it might have taken comparative Tai linguistics to get started without them. Nowadays most scholars tend to work without reference to the writing system, and, of course, for many languages and dialects from which the more important and often crucial data are taken, no such help exists. At most, scholars nowadays are likely to point out from time to time that the Siamese spelling confirms a point, or conversely, that comparative study demonstrates that the modern Siamese spelling is historically incorrect.

Will it one day be possible to adduce inductive proof of the validity of our reconstructions without any reference whatever to such external information as that provided by the writing system? In chart 3 we found that as comparative Tai studies have advanced scholars have been able to add more horizontal divisions representing phonetic distinctions in the earlier consonant system. Some of the material cited here has suggested that even more horizontal divisions may be definable. These divisions differ in their location from one column of the chart to another, so that one can imagine a scholar working without any prior hints finally discovering that the interlocking correspondences from column to column would require precisely the ordering of the horizontal categories that we have displayed. He might end up with the whole chart upside down, but who is to

say he would be wrong? And he would probably never find reason to arrange the three columns in our order A B C (and no doubt speakers of Proto-Tai were quite unaware of any such ordering); there is, as a matter of fact, some reason for regarding column B as being closer to column D than is either A or C, as we shall see below in another connection.

Aside from help from the writing system, it also must be admitted that students of comparative Tai have been aided by the fact that comparative and historical Chinese linguistics also involves correlations of initial and tone reflecting earlier tonal splits conditioned by initial consonants. But, although knowledge of this approach in Chinese linguistics was undoubtedly suggestive to early students of comparative Tai, it seems clear that we are by now at a stage where the conclusions of comparative Tai linguistics are independently valid and would still stand on their own feet even if it were to be demonstrated tomorrow that this approach in Chinese linguistics is wrong.

More serious is the objection raised by R. B. Jones[25] that the consonant system reconstructed for Proto-Tai by most scholars is too similar to that posited for the stage of Old Siamese represented in the earliest written records, dating from the end of the thirteenth century, to be credible, and that modern Tai languages and dialects resemble each other more closely than any of them resembles this reconstructed Proto-Tai. The implication of these criticisms is that scholars have followed the Siamese

writing system too closely in reconstructing Proto-Tai.

There appear to be two answers to these objections. For one thing, it is rather widely believed nowadays that the period of unity of Proto-Tai, the assumed parent language of the family, was no more remote in time than perhaps fifteen hundred, or at most two thousand, years. If this estimate is correct, then the dates of the earliest written records in, for example, Siamese, would be not many centuries removed from Proto-Tai. The implication is a rather rapid dispersion of Tai-speaking peoples in the early centuries, just before their appearance on the threshold of history.

The second answer is that comparative and historical linguistics in more familiar areas is replete with instances of relative stability in some parts of the sound system of a language or group of languages over long periods of time, followed often by drastic shifts during a briefer period. Consonant shifts in the Germanic languages are a familiar example. It is perhaps misleading that the spectacular tonal splits, which Tai languages underwent some time after the date of the earliest Siamese written records, have occupied our attention so fully that we forget that other important changes in other parts of the sound system also took place in each language, in some cases earlier and in others later.

Question 7. How is the relative homogeneity of the Northern Tai languages to be explained?

The amount of material available on languages of the Northern Tai branch is now considerable. For Wu-ming in the east we have Li's excellent monograph providing texts and glossary.[26] The Russian grammar of Chuang[27] and the Chuang-Chinese dictionary[28] also deal with the language of the Wu-ming area. A Chuang edition of the magazine *China Pictorial* has been published for some years; collation of this with the English edition produces a great deal of lexical material.[29]

In a number of articles Li has published enough data from the Jui dialect of Po-ai to provide fairly ample information on this variety of Northern Tai.[30]

The Chinese have published a voluminous study, using a transcription and an analytical approach in the tradition of Y. R. Chao and F. K. Li, of the Northern Tai speech of the people called Pu-Yi in the southern part of Kwei-chow.[31] Forty geographical points are covered, with an admirable, succinct, phonological description provided for each. In the main part of the book are given the forms in each of the forty dialects for many hundreds of words. For no other part of the Tai-speaking domain except Thailand[32] do we have such accurate phonological information on such a closely spaced network of geographical points, and for no other area at all, even Thailand, has anyone attempted to cope with the typographical problems involved in trying to publish such copious lexical data. We can only dream of the day when we have this kind of coverage for, say, the

Shan dialects of Burma, or the Lao dialects of northeastern Thailand, or of the Tai dialects of North Vietnam. The Pu-Yi book might well serve as a model for future studies in Tai linguistic geography.

For Yay, a Northern Tai language spoken on the North Vietnamese side of the border, a sketch of the phonological structure with some lexical data has been published by Gedney.[33]

For Saek there are early published wordlists.[34] None of the these include marked tones, which are apparently described in the present paper for the first time.

The oldest work on a Northern Tai language, on which earlier students of comparative Tai had to rely exclusively, was the Esquirol and Williatte dictionary of Dioi.[35] This dictionary is exhaustive, and its hundreds of illustrative sentences provide insight into just how each word is used. But, although the transcription is in some ways extremely accurate, some important distinctions, such as that between short *a* and long *aa*, are completely disregarded. Li has identified the location of the Esquirol and Williatte Dioi material as the district of Ts'e-heng,[36] and it turns out that the dialect of this district is represented by point no. 4 among the forty points covered in the Pu-Yi book mentioned above. As a result it will now be possible to reexamine the old Dioi material and rectify many or most of its transcriptional ambiguities.

Many scholars who have worked with Northern Tai linguistic materials have been heard to comment from time to time that languages of this group give the

impression of being relatively homogeneous, despite the fact that they are now scattered over a fairly extensive geographical area. The differences among the various points from which we have data seem on the whole markedly less than the differences among Tai languages in almost any area of comparable extent in the Southwestern and Central branches. Even Saek, though in some respects divergent from all other recorded Northern Tai languages, seems in other ways (vocabulary, tonal system, vowels) remarkably close to them.

The first task, if one were to try to assess this alleged homogeneity, would be to put all the extremely varied Northern Tai data into comparable form, perhaps even to retranscribe all the material into a single system. Once this is done, isoglosses must be drawn for every identifiable lexical and strucural feature. If this task is accomplished someday for both Northern Tai languages and Tai languages of the other branches, then we will be able to judge objectively the degree of homogeneity.

If this impression of relative homogeneity over an extensive area in the Northern Tai branch turns out to be correct, we are faced with the apparent paradox that our earlier discussion of the peculiarities of the Northern branch presumably implies an early date of separation from the other Tai languages, but if the Northern languages did indeed separate from the others at an early date, how did they manage to remain so relatively undifferentiated? One possible hypothesis might be that, although they may have separated from the other Tai languages at a

relatively early date, they remained thereafter undispersed for a fairly long period, during which time they might have undergone in common some of the drastic changes that we have observed, and spread across their present domain in only fairly recent times.

Question 8. What is the source of the final -l of Saek?

Saek is the only known exception among Tai languages to the restriction on permitted finals to zero, to one of the three nasals, *m*, *n*, or *ŋ*, to one of the semivowels, *w*, *y*, or *ɣ*, or to a stop, *p*, *t*, or *k* (or, in some languages, *ʔ*). In addition to these finals, Saek has syllables with final *-l*. At the Saek-speaking village of Ban Atsamat in Nakhon Phanom Province, Thailand, where our data were obtained, speakers over the age of fifty (in the summer of 1966) make a consistent distinction between words having final *-l* and those having final *-n*. Speakers younger than this have only final *-n*, although some younger people are able to imitate this characteristic of their elders' speech at will. The same age boundary also separates a number of initial consonant distinctions made by the older people but lost through coalescence in the speech of the younger generation. Examples are given in chart 19.

Since both the consonantal distinctions and the distinction between final *-l* and *-n* seem to have been lost by all speakers under fifty, but to be consistently preserved by all speakers of fifty or older

	OLDER GENERATION	YOUNGER GENERATION
cliff	phraa^{2}	phraa^{2}
forehead	phraak^{6}	phraak^{6}
vegetable	phrak^{4}	phrak^{4}

(at least throughout this one large village), one infers that a number of sound changes swept across this small area some forty-odd years ago, affecting the speech of all youngsters throughout the village regardless of location, parentage, social status, educational level, or any other factor, suggesting that we have here a miniature laboratory for the study of the little-understood mechanisms of sound change. It might prove interesting to try to find out whether what we are dealing with here is a point in recent history when an entire generation of children became bilingual in Saek and Lao. Nowadays all Saek speakers are fluent in both languages (and many also in standard Thai or Siamese), and unquestionably many of the innovations in the speech of the younger generation are in the direction of the Lao sound system.

Chart 20 gives examples of words that among the older generation are consistently pronounced with final *-l*.

Chart 19.

	OLDER GENERATION	YOUNGER GENERATION
lazy	tliik^{6}	triik^{6}
salt	tlua^{1}	trua^{1}
near	tlɘɘ^{3}	trɘɘ^{3}
smooth	tliaŋ^{3}	triaŋ^{3}
star	traaw^{1}	traaw^{1}
to speak	traaw^{3}	traaw^{3}
to slip and fall	thlɘɘt^{5}	thrɘɘt^{5}
tamarind	thlɛɛw^{2}	thrɛɛw^{2}
to lasso	thlɔɔŋ^{6}	thrɔɔŋ^{6}
Saek	thrɛɛk^{6}	thrɛɛk^{6}
small frog	thrɛɛ^{3}	thrɛɛ^{3}
possessions, things	thrɨaŋ^{5}	thrɨaŋ^{5}
banana blossom	plii^{1}	prii^{1}
fish	plaa^{1}	praa^{1}
leech	pliŋ^{1}	priŋ^{1}
gourd, melon	priaŋ^{1}	priaŋ^{1}
bamboo strip	pruk^{4}	pruk^{4}
to break	preek^{6}	preek^{6}
betel	phluu^{4}	phruu^{4}
to spill	phlaaw^{5}	phraaw^{5}
bracelet	phlam^{4}	phram^{4}

Chart 20.

	SAEK	SIAMESE
stone	riil^{2}	hin^{5}
wasp	thiil^{4}	teen^{1}
to fly	bɨl^{1}	bin^{1}
firewood	vɨl^{1}	fɨin^{1}
slippery	mliɨl^{5}	lɨin^{3}
body hair	pul^{1}	khon^{5}
hot	ruul^{6}	rɔɔn^{4}
earthworm	trual^{1}	dɨan^{1}
forest	thual^{5}	thɨan^{2}
hard	keel^{6}	ken^{2}
to snore	tlel^{1}	kron^{1}
civet cat	ɲel^{2}	hen^{5}
bodylouse	mlel^{4}	len^{1}
classifier for long, thin objects	sel^{3}	sen^{3}
classifier for things	ʔal^{1}	ʔan^{1}
to crow	hal^{2}	khan^{5}
to slash	val^{1}	fan^{1}
dike between fields	ɣal^{4}	khan^{1}
to become	phal^{4}	pen^{1}
sweet	vaal^{2}	waan^{5}
slack	yaal^{4}	yaan^{1}
antelope	vaal^{4}	faan^{1}
to sow	vaal^{6}	waan^{2}
worm	nɔɔl^{2}	nɔɔn^{5}
to teach	sɔɔl^{2}	sɔɔn^{5}
hammer	ɣɔɔl^{6}	khɔɔn^{4}

Other words, which in the older generation are always pronounced with final *-n*, may be suspected of having been borrowed in recent times from Siamese or the local variety of Lao, but the examples given in chart 21 seem certain to represent old inherited forms.

Chart 21.

	SAEK	SIAMESE
to eat	kɨn^{1}	kin^{1}
tongue	liin^{6}	l̈in^{4}
moon, month	blian^{1}	dɨan^{1}
sky	bɨn^{1}	bon^{1} ('on; top')
rain	vɨn^{2}	fon^{5}
to go up	hɨn^{3}	khɨn^{3}
night	ɣɨn^{4}	khɨɨn^{1}
to carry, transport	khun^{2}	khon^{5}
human being	hun^{4}	khon^{1}
to sleep	nuun^{4}	nɔɔn^{1}
arm	keen^{1}	khɛɛn^{5}
to hang up	veen^{2}	khweɛn^{5}
to see	rɛn^{2}	hen^{5}
silver	ɲɛn^{4}	ŋən^{1}
potato	man^{4}	man^{1}
axe	vaan^{2}	khwaan^{5}
village	baan^{3}	baan^{3}
carrying pole	ɣaan^{4}	khaan^{1}
house	raan^{4}	rɨan^{1}
smoke	ɣɔn^{4}	khwan^{1}
to castrate	tɔɔn^{1}	tɔɔn^{1}
before	kɔɔn^{6}	kɔɔn^{2}

In looking about for an explanation of this final *-l* of Saek, one is reminded that there is an area much farther north, in which are found the Lao dialect of Sam Nuea Province in Laos and the Red Tai dialects spoken just across the border in North Vietnam, where older speakers often have final *-l* occurring as a substitute for final *-n*, apparently in free variation. Available information suggests that the status of final *-l* there is somewhat different from the stable, consistent, *l/n* distinction in Saek. Further study of the extent and circumstances of the replacement of *-n* by *-l* in Sam Nuea and Red Tai is necessary before we can make a judgment as to whether the phenomena are related, for other evidence suggests that Saek may at some time, during its presumed movement southward from its original northern location, have been spoken somewhere in this Sam Nuea or Red Tai area. One of the most striking bits of evidence of this kind is the change of the diphthong *ay* to the monophthong *əə*. This change is found in Sam Nuea and Red Tai, and also in Saek, although in Saek the inventory of words affected is rather different, since Northern Tai languages do not agree with other Tai languages in the list of words having the diphthong *ay*. Saek examples are as follows.

	SAEK	YAY
gizzard	təə^{1}	tay^{1}
leaf	bəə^{1}	bay^{1}
in, inside	rəə^{1}	day^{1}
near	tləə^{3}	cay^{3}

top of the head	-nəə^{4}	nay^{4}
daughter-in-law	khwəə^{6}	pay^{6}
which	nəə^{4}	lay^{4}
to give	həə^{3}	hay^{3}

Obviously more study is needed to determine the possible validity of this suggestion that Saek was formerly in contact with Sam Nuea and Red Tai dialects, that it participated with them in this *ay* > *əə* change, and that it also may have shared with them some sort of innovation, producing a final *-l* pronunciation in certain circumstances that are not yet clear.

If Saek final *-l* cannot be accounted for as an innovation, then it must be old. If it is old, then it is difficult not to conclude that it must be attributed to Proto-Tai, even though the implication would be that all other Tai languages lost the final *l/n* distinction. If the Tai languages are found to belong to a larger group, which includes the Mak-Sui-Kam languages, a question that we will take up next, then this hypothesis becomes even more strained, because all known languages of the latter group have the same limitations on permitted finals as most Tai languages.

If Saek final *-l* is to be attributed to Proto-Tai, and perhaps also to the parent language of any larger grouping that may later be demonstrated, this adds strength to the cause of those who like to believe in an ultimate Malayo-Polynesian relationship for the Tai languages.

Question 9. How are the Tai languages related to the Mak-Sui-Kam group?

The Mak, Sui, and Kam languages are spoken in southeastern Kwei-chow by relatively small numbers of speakers. We owe to Li our information on Mak[37] and Sui,[38] and he has also dealt with the comparison of languages of this group.[39] For Kam there is a locally published Kam-Chinese dictionary[40] providing a large vocabulary of great value for comparative study. This dictionary marks and describes nine tones in contrast on smooth syllables; this surprising assertation must be checked against other descriptions of Kam.[41]

The Mak, Sui, and Kam languages are sufficiently remote from Tai that the sound correspondences are not obvious, but enough evidence is apparent even at first glance to leave no doubt that there is a genetic connection and that scholars will eventually be able to work out the details of the phonological correspondences. What the larger family comprising Tai, on the one hand, and Mak, Sui, and Kam on the other, once it is firmly established, will be called, is an interesting problem in terminology. And it is too early to speculate on the phonological structure of the parent language, which will have to be reconstructed, or on what suggestions and implications the structure of this proto-language may then present for still more remote comparisons and connections.

Even at the present stage of our knowledge, the student who works with Mak, Sui, and Kam material soon gains a certain impression that has important implications for the way in which we view the

relationship of Northern Tai to the other Tai languages. Most scholars, when considering the important differences between the two groups, imply by the wording of their statements that they think of the Northern Tai branch as having diverged from the mainstream, so that the special features characteristic of this group are usually regarded as requiring explanation. Now, whenever one gets a glimpse of a possible systematic correspondence between Mak-Sui-Kam languages and Tai, if there is a difference between Northern Tai and other Tai languages involved it turns out to be the Northern Tai languages rather than the Central or Southwestern, which the Mak-Sui-Kam languages resemble. If it is found that all the Tai languages, including Northern Tai, constitute one branch of a larger family, of which Mak-Sui-Kam constitute the other branch, then these resemblances between Mak-Sui-Kam and Northern Tai suggest a reversal in our habitual ways of viewing the differences between Northern Tai and other Tai languages. That is, a more correct picture may be one in which Mak-Sui-Kam and Tai (in the largest sense) separated first, and within the Tai branch the languages that we now call Southwestern and Central later broke away and subsequently underwent important common innovations and changes resulting in the marked differences now seen between Northern Tai and the other Tai languages. Pursuing this picture further, it may be that the Southwestern languages, which include the best-known and most populous and perhaps historically and politically most important of all Tai languages, such as Siamese, Lao, and Shan, are paradoxically to

be regarded from the point of view of the linguistic prehistorian as the ultimate, farthest-out, most divergent and deviant offshoot of a linguistic mainstream that lies much farther north.

It is perhaps not irrelevant to this question that in cases of discrepancy between Northern Tai and the other branches of Tai, Southwestern and Central Tai languages reflect earlier voiceless consonants versus earlier voiced ones in the Northern Tai languages more frequently than the reverse. If it turns out that this reversal in point of view just suggested requires us to regard the Southwestern and Central languages as divergent, rather than the Northern languages, then one finds some support in the well-known fact that devoicing changes in the history of languages of Southeast Asia and the Far East tend to be much more frequent than processes of voicing.

Question 10. How are the Tai-like languages of Hainan related to the mainland languages?

For the Li language of Hainan we have a glossary published long ago,[42] but in a transcription so inconsistent in its marking of tones as to be scarcely usable for serious comparative study. Fortunately, for Li we now have more reliable information,[43] but it remains to be seen to what extent these more modern descriptions will permit restatement of Savina's older material into more usable form.

Savina also compiled a dictionary of Bê, another Tai-like language spoken on the island of Hainan, and Haudricourt has recently edited and published this

important work,[44] but here again one observes, and indeed the editor regretfully points out, chaotic inconsistency in the marking of tones. It is surprising that Savina did so poorly in his recording of Li and Bê tones, for in his dictionary of Nung[45] he recorded tones with impeccable consistency. In the tonal variations in Savina's transciptions of neither Li nor Bê can one discern any pattern of tone *sandhi* producing systematic alternation. Perhaps the explanation is rather that Savina tended to record tones in terms of the Vietnamese tone system, with which he was most familiar, and when he transcribed a language such as Nung, which he knew well, or where the phonetic similarities were perhaps closer, he did well, but when he attacked a language like Li or Bê, which was more dissimilar, or with which he had briefer contact, his impressionistic habits of recording led to errors and inconsistencies. Any fieldworker who has only a short time to work on a Southeast Asian tone language can understand how this may happen.

That these minority languages in Hainan are somehow related to the Tai languages there can be no doubt; some sound correspondences, often bizarre but surprisingly consistent, are obvious at first glance. No one has yet been able, however, and it may be that with our present information no one will be able, to determine definitely the exact degree of relationship, or to answer conclusively the question: to which of the major branches of Tai are these Hainan languages most closely related?[46]

Question 11. What are the most serious gaps in our data?

Starting in the west, for Ahom we have a valuable dictionary;[47] this has recently appeared in a revised version,[48] which Western scholars have not yet started using. Since Ahom is extinct, and must be studied entirely from records written in a spelling system that ignores tones and apparently also confuses some vowel sounds (as does the traditional writing system of Shan next door in Burma), there will always be serious limitations on the usability of Ahom data. There is a need for someone who knows Shan well, and who has had sufficient experience in comparative Tai, to learn what sorts of structural features to expect, to reexamine the Ahom writing system and try to determine more exactly which distinctions of the language are accurately represented in the writing, and, conversely, which of the written symbols may be suspected of ambiguously representing more than one sound.

Moreover, there exists at Gauhati, Assam, a collecton of many hundreds of Ahom manuscripts on a rich variety of subjects, which constitute an untouched gold mine for the lexicographer and linguistic scholar. The published dictionaries are based on manuscript glossaries and have not attempted to exploit this wealth of textual material. The extant spoken Tai dialects of Assam have been described in the Linguistic Survey of India and in other early publications but for none of these do we have accurate phonological information, including an analysis of the tones.

Lue, in Sipsongpanna, is so well known to travellers and to scholars in other fields such as anthropology that it is surprising that we do not have a Lue dictionary or published collections of Lue texts.

For Lao there are now two very fine dictionaries,[49] superior in exhaustiveness and authenticity to the older Guignard dictionary,[50] which has been so generally used in the past by comparative Tai scholars, but neither indicates pronunciation so that we do not yet have a full wordlist recorded in the phonological system of any one of the many Lao dialects.

Western Nung, a group of dialects spoken in the area of Lao Kay, about midway across the northern border of North Vietnam, constitutes one of the most serious gaps in available knowledge of Tai languages and dialects. On these Western Nung dialects, apparently the result of relatively recent migrations from the main area of Nung speech farther east, there appears to be no published information whatsoever.

In general, of course, we need more data in two dimensions, first in the density of geographical coverage, and second in the depth of our knowledge of particular dialects. With respect to the first of these dimensions, most major areas in the Tai-speaking domain still wait to be investigated on a fine geographical grid to give us the kind of dense coverage of closely spaced points provided by Brown for Thailand and by the Chinese book on Pu-Yi for southern Kwei-chow. With respect to the second dimension, for only half a dozen languages do we have anything approaching complete lexicographical collections. Recent fieldwork on the Nung dialect of

Lungming, in which the informant proved so adept at providing illustrative sentences that a corpus of over sixteen thousand utterances was collected within a few weeks, turned up a great many rather out-of-the-way words, which could be cognates of words known to occur in Siamese and Lao, but which often seemed too good to be true and cannot certainly be regarded as genuine cognates until we know whether intervening dialects have the words as well. To cite just two examples, Lungming *cooy*^{6} 'drooping gracefully' matches perfectly, in both phonological form and meaning, the Siamese literary word and personal name *chɔɔy*^{4}, but the two languages are so very remote from one another geographically that one cannot rule out coincidence without some supporting evidence, for example, knowledge that this word also occurs in at least some of the Tai languages that stretch between Lungming and Bangkok. White Tai *pɛʔ*^{2} 'like, similar' could phonologically be cognate with Siamese *plɛɛk*^{2} 'strange'. One might imagine that the different, indeed almost opposite, meanings in the two languages could have arisen out of divergent semantic developments from an earlier common meaning 'somewhat similar but not quite alike'; an enigma like this can hardly be solved without copious information on how this word is used in intervening dialects.

We are hampered in these matters by current fashions and trends. It is scarcely fashionable or profitable nowadays for a young scholar to devote his research to fieldwork whose aim is the collection of mere data or the compilation of a dialect dictionary, nor are publications of material of this sort, in

some ways so much more difficult to prepare and publish than other types of scholarly production, generally regarded as really prestigious and elegant. Perhaps we may hope that as time passes it will come to be recognized once again that for at least some first-class young scholars dialect study and the collection of lexicon and texts are perfectly respectable and worthy pursuits.

Question 12. How much linguistic validity is there for the well-known language names in the Tai family?

Familiar names such as Shan, Lue, Lao, White Tai, Red Tai, Black Tai, Tho, and Nung are useful to the linguist only in making rough-and-ready references to the general geographical location and genetic classification of dialects that he is studying. For scarcely any such "languages" are definite geographical boundaries definable. The linguistic fieldworker cannot, of course, fail to make note of the names that the speakers themselves and their neighbors use for the dialect that he is studying, but it is usually more important in the long run to note the precise geographical location of it.

To scholars in other disciplines and to laymen, these well-known names are, of course, irresistibly attractive. Ethnographers and geographers seek constantly to obtain accurate estimates of the number of speakers of each of these "languages," and to define reliable geographical boundaries for them. If, as seems likely to prove true, there are no real language boundaries, but only gradual transition throughout much of the Tai-speaking domain, except

perhaps for the boundary between Northern Tai and the rest, then such efforts are doomed to futility.

On the other hand, some of these names are associated so closely with identifiable cultural traits, and in some cases speakers are so firm in insisting on distinguishing linguistic features (for example, "The word *lɛw*^{6} ['finished'] is Black Tai; we White Tai never use it") that fieldworkers can hardly avoid attempting to establish identifying linguistic criteria for these ethnic names where possible. In some cases one suspects that linguistic criteria corresponding exactly in geographical extent to the ethnic terms will never be found. For example, the speech of those villages and towns in the extreme northeast of Burma, where the people refer to themselves and their languages as "Lue," seems to be closer linguistically to the Shan and Khuen spoken to the west and the dialects of northern Thailand to the south than it is to the dialect of the Lue capital city of Chieng Rung in Sipsongpanna. And, although the Nung people in the extreme northeast of North Vietnam are readily distinguished culturally, and even by their surnames, from their Tho neighbors, no one seems to have tried to identify linguistic features that distinguish Nung from Tho, and one is surprised to find that in the text of Savina's Nung dictionary the name Nung never appears; the language is called in many illustrative sentences not Nung but Tho.

Question 13. Can the three tones assumed for Proto-Tai be reduced to fewer?

Some scholars have attempted to reduce the complexity of the tonal reconstructions necessary in comparative Tai by suggesting that some tonal distinctions may be attributed to lost final consonants. R. B. Jones, for example, has worked toward eliminating the glottalized tones found in so many Tai languages by identifying them somehow with nonglottalized tones plus glottal stop.[51] It is true that these glottalized tones turn up with greatest frequency in the C column of our comparative tonal chart, but often they also occur in the B column, and in some languages even in the A column; for example, the tone of the bottom box of the A column in White Tai is glottalized. The most serious flaw in this approach seems to be the fact, apparently not yet widely noted, that no Tai language of the Northern branch has any of these glottalized tones, with the single exception of Saek, and it seems reasonable to suppose that Saek acquired this feature some time after it moved southward into the area where not only Tai languages but also other tonal languages of Southeast Asia characteristically have tones of this type.

There are other approaches to this question, which seem to call for further study. For one thing, frequency of the Proto-Tai tones, A, B, and C, seems to have been very unequal. No matter what Tai language or dialect one is studying, he finds that the morphemes in column A for which there is evidence of genuine native origin turn out to be very much more

numerous than those in column B or C, perhaps more than both combined. Is it possible that in the parent language tone A was somehow normal, and tones B and C were modifications of it? This notion is supported, perhaps, by evidence from the structure of old Siamese poetry, which had a metrical pattern known as *khlooŋ*[1] in which syllables occurring in certain positions in the stanza were required to have tone B or tone C, as if these were special in some way. The word *khlooŋ*[1] in Siamese is also a verb meaning 'to rock (from side to side, as a boat)'. Since boating songs were so popular in traditional Southeast Asia, could it be that this name for a type of poetic meter originated in playful rowing songs, in which at certain intervals one was to rock to one side (B) and then the other (C)?

It is also interesting to note that in old Siamese all the loanwords from Sanskrit and Pali acquired the tone of column A on all syllables except checked ones. This again suggests that tone A was the normal level tone, with tones B and C so markedly different from it as never to be used in pronouncing the syllables of words borrowed from a toneless language.[52]

Another possible clue may lie in the tendency of the various earlier tones, A, B, and C, to vary in the degree to which they underwent splitting. In general, column A is found to undergo more splitting than the others, and column B the least; in some Lao dialects column B shows no tonal differentiation whatever. Lack of differentiation in column C is not unknown, for example, in the Lei Ping dialect, but it

is rarer than in column B. The implications of these varying tendencies are not clear.

Moreover, column B is remarkable in that in many dialects it shows the same tonal split, from the point of view of structure and often even with close phonetic similarity in the resultant tones, as column D-long. Can this mean that in column B a lost final consonant is lurking somewhere in prehistory? One is reminded of the curious behavior of such Sanskrit loanwords in Siamese as Skt. *loha* > Si. *loo*^{3} 'shield' (column B), parallel with Skt. *loka* > Si. *look*^{3} 'world' (column D-long),[53] but surely these Siamese phenomena are too recent in history to be relevant to any possibility of a lost final consonant in the B column in the period before Proto-Tai.

During the period since the time of Proto-Tai, the tendency in all Tai languages has undeniably been to increase the number of tones. It seems not incredible that a tendency in the same direction may have been at work even farther back in prehistory.

But for Proto-Tai there seems to be no possibility of positing fewer tones than the three tones, A, B, And C, plus the undifferentiated tone D on checked syllables; nothing else seems to account so completely for all the modern tonal systems. If further study of the Mak-Sui-Kam languages eventually permits us to reconstruct an earlier proto-language representing the speech of the period of Mak-Sui-Kam and Tai unity, it will be interesting to learn how many tones this earlier proto-language turns out to have. One feels that it is too much to hope for that we might ever get so far back, by strict comparative

linguistics, as to reach the stage of a remote proto-language having no tonal distinctions at all, because we would probably then be dealing with languages so distantly related, and expanses of elapsed time so vast, that successive sound changes would have wrought confusion so extreme as to render irretrievable the phonological evidence. On the other hand, what we have seen as to how tonal splits arise through phonetic conditioning by initial consonants makes it easy enough to imagine that tones might once have appeared in a previously nontonal language through an analogous process, even though we may never be in a position to prove that this actually happened in the family of languages that we are studying.

Question 14. What wider relationships are likely to be proven eventually?

Formerly the view was widely held that the Tai languages are ultimately related to Chinese; nowadays many scholars still lean in this direction, while recognizing that we lack the kind of proof now regarded as necessary. The view of an ultimate Chinese-Tai relationship is supported by the general typological similarity between the two groups (tones, monosyllabic morphemes, and so on). The objection that, in the syntax of Tai, the attribute follows the head, while in Chinese it precedes it, is easily refuted by citing the parallel difference between such groups as Germanic and Romance, in spite of their known ultimate genetic relationship; apparently

such syntactic features can change secondarily through time.

Certainly Tai languages everywhere contain many words that are strikingly similar to Chinese. Some of these, for example, the ones represented in Siamese by *mɨk*^{2} 'ink' and *phay*^{3} 'playing card' show such irregularity in form from language to language that no Proto-Tai reconstruction is possible, and we must conclude that they are later borrowings. But for many others the sound correspondences among various Tai languages are so regular as to permit reconstruction of the Proto-Tai forms as readily as for other genuine Tai words for which there are no Chinese parallels to arouse suspicion.

It is surely relevant to this question that many scholars now assume a relatively late date for Proto-Tai unity, perhaps between fifteen hundred and two thousand years ago. This seems to be so late as to preclude the possibility of regarding Proto-Tai and Proto-Chinese as sister languages having a more remote common ancestor. On the other hand, the possibility of connecting Mak-Sui-Kam with Tai, with perhaps even more remote connections still to be established, suggests that we may someday be able to push the date of the period of unity on the Tai side back far enough to place it on a chronological level comparable with that of Proto-Chinese.

In any case, we must reserve judgment until we have firmer reconstructions on both the Tai and Chinese sides. In the meantime our safest course appears to be to regard the Tai words that suspiciously resemble Chinese as part of the common Tai

vocabulary, to work out correspondences and reconstructions for them just as for genuine native Tai words, and to defer until later any decision on the question as to whether they represent common genetic inheritance in two ultimately related families or borrowings from Chinese into Proto-Tai or Pre-Tai.

Another view, first proposed by Paul Benedict in 1942[54] and now widely accepted among linguists, anthropologists, and others, is that the Tai languages are somehow related to Malayo-Polynesian, with four minor languages, called Kadai by Benedict, forming an intermediate group; Kadai includes the Li language of Hainan, mentioned earlier. The author of this bold hypothesis has recently taken up the subject again.[55]

The objections to this hypothesis, and particularly to the evidence and the methodology on which it depends, are so apparent and so strong that one is surprised at its wide acceptance, not among anthropologists, who are always eager for anything that simplifies the picture of language families of the world and eliminates isolated groups, but among reputable linguists who ought to know better. It disregards tones entirely, using much old, badly recorded material, wherein tones were not recorded at all, but also ignoring tonal features in languages where tones are known. The comparisons always involve sporadic partial orthographic similarity between isolated words chosen indiscriminately from one language or another, all in the manner of amateur eighteenth-century efforts to demonstrate the relationship of Hebrew to this language or that, without

regard for systematic comparison of phonological systems, to say nothing of regular correspondences in individual sounds. The suggestion is made that the Tai languages were once nontonal, like Malayo-Polynesian, but acquired tones under Chinese influence. Now, it is undeniable that this kind of influence can occur, though hardly in the manner and on the scale suggested, and all our work in comparative Tai shows that tones are basically important within the Tai family as far back as we are able to carry our reconstructions; it appears that if we ever get back far enough in prehistory to eliminate tones, this will be so far back as to render irrelevant the sporadic lexical similarities that Benedict believes he has found among some of the modern languages.

Moreover, careful study of the Benedict articles on this hypothesis reveal an underlying attitude toward linguistic relationships--for example, in his use of the term *strata*--that is clearly quite incompatible with the now generally accepted views on genetic relationships among languages.

Question 15. To what degree have linguistic changes spread across linguistic boundaries?

When someday we have sufficiently detailed information to plot isoglosses for all linguistic features throughout the Tai-speaking domain, it seems clear that we will have difficulty in some cases in distinguishing between those changes that represent migrations of people and those that represent the spread of innovations from one settlement to another without involving the physical movement of people.

It seems likely that in some cases it will never be possible to decide for certain which of these factors has been at work, or whether both have been involved.

A similar problem arises in studying the origins of the Siamese dialect of Bangkok. This city is known to have been founded less than two centuries ago, and was presumably settled by people of varied geographical origin. Is this, for example, the explanation for the sporadic vowel-lengthening in the Bangkok pronunciation of some words that show short vowels elsewhere in Tai languages, for example, *naam*^{4} 'water' (but not *cham*^{4} 'bruised'), *daay*^{3} 'to obtain, to be able' (but not *may*^{3} 'to burn'), *plaaw*^{2} 'empty' (but not *paw*^{2} 'to blow'), *khaaw*^{3} 'rice' (but not *khaw*^{3} 'to enter')? All these are reminiscent of the processes of vowel lengthening that have occurred in many of the dialects of the southern peninsula of Thailand, where, however, these changes are regular, whereas in Bangkok they have affected only certain words. One is reminded that the early history of Bangkok is filled with the names of prominent persons of southern origin.

Even more curious are cases of the spread of linguistic features and changes across language, and even language-family, boundaries. This may prove to be, for general linguistic theory, the most interesting and important of all our problems and puzzles, and so it has been left until last in our discussion.

The Shan dialects of Burma have, like the unrelated but geographically contiguous Burmese language, two sibilants, one described as a plain *s*, and the other as aspirated. To the east in many

dialects of Tho and Nung the sound that is *s* in most other Tai languages (whether originally from a voiced or voiceless source) is represented by a voiceless lateral *l̥*, a phenomenon that is said to occur in other non-Tai languages of that area. Some dialects of the Nung group (Central branch of Tai) in southern Kwangsi have diphthongized the high vowels, so that *ii* becomes *ey*, *uu* becomes *ow*, and the like. A similar change has affected the high vowel in the Tai dialect of Wu-ming, farther north, but Wu-ming belongs to the Northern branch of Tai, and it appears that this change spread from one branch to another long after their genetic divergence. Similar diphthongizing changes are known to have occurred in some non-Tai languages in the same area, as noted above.

We have had occasion already to mention the suspicion that Saek may have participated in a change of *ay* to *əə* while in contact with other, only very remotely related, languages that made this change. We also have referred to the geographical distribution of the feature of glottalized tones, which are found in all known Tai languages of the Southwestern and Central branches, and also in Saek, which belongs genetically to the Northern branch but is now located in the Southwestern area. Tones of this type are found also in many non-Tai languages of Southeast Asia such as Burmese and Vietnamese. It seems likely that we may someday be able to prove that glottalized tones in Southeast Asia are an areal feature that has spread across linguistic boundaries without regard to genetic relationships, perhaps in relatively recent times.

For the student of comparative Tai these phenomena suggest that in working on the history and prehistory of any dialect or group of dialects one must take great care in trying to distinguish among inherited features, local changes, and changes that may have spread from one group to another. This is, of course, the familiar question as to how seriously waves can distort and obscure a family tree.

Somehow involved in this question of the spread of sound changes across language and language-family boundaries is the striking typological resemblance in phonological structure among the modern languages in the area. Tai languages today show many basic structural similarities, though often the various similar elements of the sound systems may be shown to have very different historical sources. Many such typological similarities appear even between languages not known to be related, or for which any possible genetic relationship that might someday be proved would be so far back in time as to be irrelevant. Why, for example, do so many Southeast Asian languages, whether related or not, show an obstruent system of the shape *p t c k, ph th ch kh*? Why is voiced *g* so rare nowadays in Southeast Asia, although scholars often find it necessary to posit a voiced velar for earlier stages? Why, in languages of Southern and Southeast Asia, is the short vowel corresponding to long [a:] so often phonetically higher, often similar to the vowel of English *but*? Why, indeed, are languages of the tonal and monosyllabic type so common in Southeast Asia and the Far

East? These areal tendencies toward convergence are still scarcely understood.

Perhaps most curious of all such areal phenomena is the fact that the process of tonal splits conditioned by phonetic features of initial consonants appears to not only have swept across all Tai languages and dialects at some period in the past, but also to have affected many other non-Tai families of tonal languages in the Far East and Southeast Asia.[56] Also, in the nontonal Mon-Khmer languages a parallel set of changes occurred, at roughly the same time so far as we can tell, in which each of the vowels was split into two, depending upon the voiced or voiceless nature of the initial consonant.

If we someday achieve a clearer understanding of how these changes have operated across language and language-family boundaries, we may find that we have at last laid the ghost of what has been called "substratum." For example, whatever explanation we arrive at also may be found to account for such famous problems as the retroflex consonants shared by Dravidian and Indo-Aryan languages in India.

We have discussed only some of the problems and enigmas of comparative Tai linguistics, selected and ordered more or less at random. Any scholar in the field could add others of the same general sort--puzzles that tend to lurk in the back of the mind as one works on field notes or sorts out sound correspondences. Some of these problems are of the kind that could be solved at once if we had more people doing fieldwork. For others it may turn out that the required information is already available, but we

need more investigators able to put in more time studying the material. There are other puzzles, however, for which solutions seem not to be in sight at all. For these we can only hope that someone, someday, while working on this material will happen to be struck by a new insight or inspiration suggesting a solution or a promising approach that no one has thought of before.

Notes

1. He has frequently used the expression "thai proprement dit"; for example, see Haudricourt 1960, 162, 166.
2. Li 1959, 1960.
3. Maspero 1911 is usually regarded as the first scientific work on comparative Tai linguistics.
4. The Lao dialect of Nong Khai, in Northeastern Thailand, has no initial consonant clusters whatsoever. My Nong Khai data were furnished by Mr. Nikom Buddhamatya at Ann Arbor in August and September 1967.
5. The symbol ɨ represents a high-back unrounded vowel; ə is mid-back unrounded; ɛ is low-front [æ]. The Siamese vowel system is described by many authors; see, for example, Gedney 1964, 28, for both Siamese and White Tai vowels. The Yay vowel system is described in Gedney 1965, 182-83 [reprinted in this volume. pp. 415-62 and 401-14].
6. Of the six languages cited, Siamese and White Tai belong to Li's Southwestern branch of Tai, Lei

Ping and Lungming to the Central branch, and Yay and Saek to the Northern branch.

The provenience of the White Tai data is identified in Gedney 1964, 2-3 [reprinted in this volume, pp. 415-62]. Lei Ping and Lungming are located in southern Kwangsi; my data on these two dialects were obtained in Hong Kong in 1966 from Mr. Liang Shao-lu (Lei Ping) and Mr. Tong Tin Sum (Lungming). Data from Ning Ming, cited below, were also obtained in Hong Kong in 1966 from Mr. Wohng Gong. For the exact location of the Yay data, see Gedney 1965, 180 [in this volume, p. 401]. Saek is spoken in a few villages along the Mekhong River in Nakhon Phanom Province in northeastern Thailand and also across the river in Laos. My data were collected in the village of Ban Atsamat, five kilometers north of Nakhon Phanom. Saek is incorrectly classified as Mon-Khmer in many publications, e.g., Lebar 1964, 149-50.

The tones of the six dialects cited are as follows.

Siamese: 1 level, slightly lower than mid, 2 low level, 3 falling, 4 high level, or with a slight rise and fall, 5 rising. Tones 3 and 4 are glottalized.

White Tai: 1 level, slightly lower than mid, 2 high rising, 3 low rising, 4 level, somewhat higher than mid, 5 level with a slight rise and fall, at a pitch somewhat higher than mid, 6 falling. Tones 3, 4, and 6 are glottalized.

Lei Ping: 1 high rising-falling, 2 high level, with a final drop before pause, 3 rising, 4 low falling, 5 low level. Tone 3 is glottal-ized.

Lungming: 1 high level, 2 high rising, 3 mid level, 4 low falling, 5 low level, 6 low falling-rising. Tones 3 and 6 are glottalized.

Yay: 1 level, slightly lower than mid, 2 low level, 3 rising, 4 high, with a slight rise and fall toward the end, 5 falling, 6 higher than mid, with a slight rise toward the end.

Saek: 1 rising, 2 low level, 3 low falling, 4 high rising-falling, 5 high falling, 6 mid level. Tones 3 and 6 are glottalized.

7. Some remarks regarding transcription are in order. Each scholar has his own preferences among symbols, and as soon as one starts comparing data recorded by a number of different authors he finds it necessary to draw a distinction between mere transcriptional variations, where different scholars use different symbols for the same sound, and genuine structural differences that have some significance for comparative study.

Even in material recorded by the same person at different times, variations will occur. For example, the representation of the phonetically long syllable-final vowels (found in all Tai languages) by double letters, as in the Lei Ping, Lungming, and Saek data here, is in accordance with the author's present practice. In Siamese, double-vowel letters are used by almost everyone, as is done here. But in the White Tai forms it will be noted that final *aa* is written double but

other vowels are given single letters, while in Yay all these final vowels are written with single letters. The reason for this variation is simply that the author has transcribed White Tai and Yay vowels in these ways in previous publications dealing with these two languages and felt that it was unwise to introduce a change here.

Serious problems arise when the comparatist faces the need to cite forms from the publications of other scholars whose transcriptional preferences differ from his own and from each other's. Such preferences differ widely, and are sometimes accentuated by varying circumstances and places of publication. If one cites forms from the works of other scholars without alteration, he cannot expect his readers to understand the various conventions involved, as an Indo-Europeanist might reasonably expect his readers to understand transcriptional differences in forms from, say, Greek, Latin, Sanskrit, and Gothic. Moreover, if one cites unaltered forms from the works of a number of other scholars, the resulting typographical variety may soon reach impracticable extremes. As comparative Tai studies develop, and as there are more and more publications dealing with data cited from various sources, it seems likely that we will to have to allow each author to retranscribe the material he cites from the works of others. Such retranscription requires careful and systematic analysis of all the data from each dialect, to ensure that every phonological distinction is maintained in the retranscription, and, of course,

an explanation of the correspondences of symbols must be included.

The transcription of tones presents special problems, as a great variety of devices are available. Each language or dialect has its own tone system, unique in the number of tonal distinctions, in the phonetic characteristics of the tones, and in the assignment of tones to particular morphemes. Diacritics over the vowels are useful in presenting a great deal of material in a single dialect, as in a glossary or collection of texts. In handling more than one dialect, however, diacritics tend more than any other device, perhaps, to mislead the reader by suggesting closer phonetic resemblances among tones of different dialects than is intended. And, if one uses a particular diacritic for rising or falling tones, what does one do with a dialect that has two rising or two falling tones? More satisfactory are the various devices that provide for each syllable a marker that somehow reminds the reader of the phonetics of the tones. One such device is a pictorial symbol at the end of the syllable depicting the tone height and contour. Another, more easily printed, consists of a combination of two or three numerals, so that, for example, 11 represents low level, 55 represents high level, and 31 and 53 represent various types of falling tones. Letter abbreviations, for example, *h* for high or *lr* for low rising, achieve much the same effect. It can be argued that such devices are really similar to vowel and consonant symbols,

since they suggest at least roughly the phonetic facts. My own feeling is that these devices, despite their clear advantages, are unnecessarily cumbersome. After only a little work with a particular dialect, one has difficulty remembering which tone is high, which low, and so on. Moreover, in comparative work the important point is not usually the phonetics of the tones, but rather the particular list of morphemes in a given dialect that have the same tone and contrast in tone with all other morphemes. Therefore, I have come to prefer raised numerals, as in the data cited in this study. In assigning numbers to the tones of a particular dialect, one sometimes finds that there is an established native order, as in Siamese, or an order already used by other scholars, as in White Tai. In other cases one must decide on an arbitrary order. Some of those who use raised numerals have attemped to base the numerical order on historical principles. For example, in much of the dialect work now being published in China, odd-numbered tones are those believed to derive from originally voiceless-initial syllables, while even numbers reflect originally voiced initials. This occasionally results in the necessity to omit a particular number for some dialects, if syncretism has obliterated a category. (Apparently those who try to follow this principle have not yet had to deal with a dialect having two or more tones in one of the historical categories.) My own practice, when it becomes necessary to establish a numerical

order for the tones of a new dialect, is to follow the order of some closely related dialect for which I have already decided on an order; when this is not possible I make a purely arbitrary decision.

Still another device much used nowadays by some of the best writers on comparative Tai is to remove entirely from the individual forms the notations of tones, in terms of the tonal system of that dialect, and to assign to each morpheme a symbol, such as C1 or H2, which indicates the historical source of its tone. Authors who do this must, just as those who use other devices must, give elsewhere a description of the tones of each dialect. Although this procedure has clear advantages, it has always seemed to me methodologically improper. We would scarcely allow an Indo-Europeanist to substitute a reconstructed Proto-Indo-European *bh* for the actual reflexes of this sound in, say, the Sanskrit or Greek or Old English forms that he cites as evidence.

8. In discussing realism as a general requirement of the comparative method, Hockett (1958, 506) says, "the parent language should be expected to be somewhat more like each of its descendants than they are like each other." We are forced by the evidence to defy this principle in reconstructing Proto-Tai, at least as regards some parts of the sound system. In addition to the examples of initial consonants above, where we reconstruct original consonants different from those found in any modern dialect, Hockett's principle is again

contradicted by the fact that almost all Tai scholars reconstruct for the parent language a tonal system having fewer tones than any modern dialect has. No doubt counterevidence for Hockett's dictum also could be found in more familiar language families. For example, the so-called laryngeals reconstructed nowadays for Proto-Indo-European would seem to render the reconstructed parent language more different from any of the daughter languages than these are from one another.

9. Li 1947

10. For example, Haas 1958, Egerod 1961, Brown 1965.

11. Data were supplied by Tran Phuc Ky, on 19-20 July 1964, at Tung Nghia, South Vietnam.

12. It was Miss Saul and her colleague, Miss Freiberger, who found my Bac Va informant for me, in a large Nung-speaking settlement of refugees from North Vietnam who were living and working at Nam Sởn in South Vietnam. Apparently my informant's dialect is identical with, or closely related to, that of the informants with whom these women worked.

13. When a resumé of this paper was presented at the Linguistics Club of the University of Washington in May 1967, Professor F. K. Li pointed out in the discussion period that other cases are known of tonal languages in which an initial *h* or an aspirated stop like *ph*, *th*, or *kh* has conditioned a special tonal alteration, probably at some relatively late period in time.

14. Li 1957a, 1957b, 1959, 1960; Haudricourt 1960; Gedney 1965.
15. Li 1965, 153.
16. Li 1947.
17. Li 1954.
18. The Siamese form *yet*^{4} is now the taboo verb referring to sexual intercourse; Siamese speakers when they first hear White Tai spoken are mortified by the casual and incessant use of this verb in the ordinary meaning 'to do or make'. For the meaning 'to do, make' such a variety of morphemes occurs across the Tai-speaking domain that one suspects that a similar semantic specialization leading to taboo and replacement has occurred in a number of areas.
19. Li 1965, 165. The Sui forms have low rising tone.
20. Li 1954, 375.
21. Data were supplied by Hoàng van Hiên in July 1964, at Tung Nghia, South Vietnam.
22. For example, Brown 1965.
23. Published as Brown 1965.
24. Li 1965, 157.
25. Jones 1966.
26. Li 1956.
27. Serdiuchenko 1961.
28. *Kwangsi People's Press* 1960. M. Haudricourt kindly allowed me to borrow and microfilm his copies of this Chuang dictionary and of the Kam dictionary cited below, in Paris, June 1965.
29. Twelve issues of the Chuang edition, of forty-odd pages each, appear annually, identical in

illustrations, captions, and text with the English and other editions; these are published in Peking.

30. Li 1957a, 1957b. Wang 1966, a dissertation by one of Li's students, includes Po-ai data from Li's field notes not available in other publications.
31. Chinese Academy of Sciences 1959a.
32. In Brown 1965.
33. Gedney 1965.
34. Rivière 1902, Macey 1906, Fraisse 1950.
35. Esquirol and Williatte 1908.
36. Li 1960, 959, n. 8.
37. Li 1943, 1948a.
38. Li 1948b, 1965.
39. Li 1965. Li has recently published an article on still another language of this group (Li 1966).
40. Chinese Academy of Sciences 1959b.
41. For example, Liang 1965.
42. Savina 1931.
43. For example, Wang and Ch'ien 1951, Ch'en et al. 1958, and Ou-yang and Cheng 1963.
44. Haudricourt 1965.
45. Savina 1924.
46. Lakkia, a mainland language now accessible to us through Haudricourt 1967, appears to be just as certainly related to Tai as the Hainan languages Li and Bê, but just as puzzling as to the exact nature and degree of the relationship.
47. Barua 1920.
48. Barua 1964.
49. Laos, Ministry of Education 1962; Reinhorn 1955.

50. Guignard 1912.

51. Jones 1965.

52. Benedict (1942, 598) noted a similar phenomenon with regard to Siamese words for which he believed he had found Indonesian cognates: "It is probably significant that almost all the Thai roots having IN [Indonesian] correspondences are associated with a single toneme, represented in Siamese by the mid-level tone (with sonant and unaspirated surd stop initials) or the high-rising tone (with other surd initials) [i.e., words belonging in column A of our charts]." He is right in regarding this as probably significant but from our point of view it would not support his view of a genetic connection between Thai and Indonesian. Rather, it would, insofar as one may accept his alleged cognates as having any significance at all, arouse suspicion of borrowing, as in the case of the Sanskrit and Pali words.

53. Gedney 1947, 66, 79.

54. Benedict 1942.

55. Benedict 1966.

56. For important discussions of this subject, see Downer 1963 and Brown 1965, 62.

References

Barua [spelled Borua on title page], Golap Chandra. *Ahom-Assamese-English Dictionary*. Calcutta, 1920.

Barua, B., and N. N. Deodhai Phukan. *Ahom Lexicons (Based on Original Tai Manuscripts)*. Gauhati, 1964.

Benedict, Paul K. "Thai, Kadai, and Indonesian: A New Alignment in Southeastern Asia." *American Anthropologist* 44 (1942):576-601.

Benedict, Paul K. "Austro-Thai." *Behavior Science Notes* 1 (1966):227-61.

Brown, J. Marvin. *From Ancient Thai to Modern Dialects*. Bangkok, [1965].

Ch'en Shih-min, Huang Chia-chiao, and Chuang Yung-yen. "The phonemic system of the Li language of Pao-t'ing, Hainan Island." *Hsüeh-shu Lun-t'an* (Amoy) 1 (1958):1-7.

China Pictorial. Chuang and English editions. Peking.

Chinese Academy of Sciences. *Report on the Investigation of the Pu-yi Language*. Peking: Minority Languages Research Institute, 1959a.

Chinese Academy of Sciences. *Concise Kam-English Dictionary*. Kuei-yang: Second Working Group of the Minority Languages Survey, Kweichou People's Press, 1959b. (preliminary draft)

Downer, G. B. "Chinese, Thai, and Miao-Yao." In H. L. Shorto, ed., *Linguistic Comparison in South East Asia and the Pacific*, pp. 133-39. London, 1963.

Egerod, Søren. "Studies in Thai Dialectology." *Acta Orientalia* 26 (1961):43-91.

Esquirol, J., and G. Williatte. *Essai de dictionnaire dioi-français reproduisant la langue parlée par les tribus thai de la haute Rivière de l'Ouest suivi d'un vocabulaire français-dioi.* Hong Kong, 1908.

Fraisse, André. "Les tribus Sek et Kha de la province de Cammon (Laos)." *Bull. Société des Etudes Indochinoises* 25 (1950):333-48.

Gedney, William J. "Indic Loanwords in Spoken Thai." Ph.D. diss., Yale University, 1947.

Gedney, William J. "A Comparative Sketch of White, Black, and Red Tai." *Social Science Review* (Bangkok), special number (14 Dec. 1964):1-47. [Reprinted in this volume]

Gedney, William J. "Yay, a Northern Thai Language in North Vietnam." In Milner and Henderson 1965, pp. 108-93. [Reprinted in this volume]

Guignard, Théodore. *Dictionnaire laotien-français.* Hong Kong, 1912.

Haas, Mary R. "The Tones of Four Tai Dialects." *Bulletin of the Institute of History and Philology*, Academia Sinica, 29 (1958):817-26.

Haudricourt, André G. "Note sur les dialectes de la région de Moncay." *Bulletin de l'École Française d'Extrême Orient* 50 (1960):161-77.

Haudricourt, André G., ed. *Le vocabulaire Bê de F. M. Savina.* Paris, 1965.

Haudricourt, André G. "La langue lakkia." *Bulletin de la Société de Linguistique de Paris* 62 (1967):165-82.

Hockett, Charles F. *A Course in Modern Linguistics.* New York, 1958.

Jones, Robert B. "On the Reconstruction of Proto-Thai." In Milner and Henderson 1965, pp. 194-229.

Jones, Robert B. "Comparative Thai Studies: A Critique." In Ba Shin, Jean Boisselier, and A. B. Griswold, eds. *Essays Offered to G. H. Luce by his Colleagues and Friends in Honour of his Seventy-fifth Birthday,* vol. 1, pp. 160-63. Ascona, Switzerland, 1966.

Kwangsi People's Press. *Chuang-Chinese Lexicon.* Rev. ed. Nanking, 1960. Revision of first edition, dated 1958.

Laos, Ministry of Education. *Dictionary of the Lao Language.* Vientiane, 1962. In Lao.

Lebar, Frank M., Gerald C. Hickey, and John K. Musgrave. *Ethnic Groups of Mainland Southeast Asia.* New Haven, 1964.

Li, Fang-kuei. *Notes on the Mak Language.* Monographs of the Institute of History and Philology, Academia Sinica, series A., no. 20. Peking, 1943. Identical with Li 1948a.

Li, Fang-kuei. "The Hypothesis of a Pre-glottalized Series of Consonants in Primitive Tai." *Bulletin of the Institute of History and Philology*, Academia Sinica, 11 (1947):177-88.

Li, Fang-kuei. "Notes on the Mak Language." *Bulletin of the Institute of History and Philology*, Academia Sinica, 19 (1948a):1-80.

Li, Fang-kuei. "The Distribution of Initials and Tones in the Sui Language." *Language* 24 (1948b):160-67.

Li, Fang-kuei. "Consonant Clusters in Tai." *Language* 30 (1954):368-79.

Li, Fang-kuei. *The Tai Dialect of Wu-ming (Texts, Translations, and Glossary)*. Academia Sinica, Institute of History and Philology Monographs Series A, no. 19. Taipei, 1956.

Li, Fang-kuei. "The Jui Dialect of Poai: Phonology." *Bulletin of the Institute of History and Philology*, Academia Sinica, 28 (1957a):551-66.

Li, Fang-kuei. "The Jui Dialect of Po-ai and the Northern Tai." *Bulletin of the Institute of History and Philology*, Academia Sinica, 29 (1957b):315-22.

Li, Fang-kuei. "Classification by Vocabulary: Tai Dialects." *Anthropological Linguistics* 1-2 (1959):15-21.

Li, Fang-kuei. "A Tentative Classification of Tai Dialects." In S. Diamond, ed., *Culture in History: Essays in Honor of Paul Radin*, pp. 951-99. New York, 1960.

Li, Fang-kuei. "The Tai and the Kam-Sui Languages." In Milner and Henderson 1965, pp. 148-79.

Li, Fang-kuei. "Notes on the T'en Language." *Bulletin of the Institute of History and Philology*, Academia Sinica, 36 (1966):419-26.

Liang Min. "A General Description of the Kam Language." *Chung-kuo Yü-wen* (Peking), no. 3 (1965): 232-51.

Macey, Paul. "Etude ethnographique sur diverses tribus, aborigenes ou autochtones habitant les provinces des Hua-Phano-Ha-Tang-Hoc et du Cammon, au Laos." *Actes, IV^e Congrès International des Orientalistes, Alger 1905*. Paris, 1906. 1^re partie, 5^e sect., pp. 36-43, 52-57.

Maspero, H. "Contribution à l'étude du système phonétique des langues thai." *Bulletin de l'École Française d'Extrême Orient* 11 (1911):153-69.

Milner, G. B., and Eugénie J. A. Henderson, eds. *Indo-Pacific Linguistic Studies*. Part I: *Historical Linguistics*. Amsterdam, 1965. Also published as *Lingua*, vol. 14.

Ou-yang Chüeh-ya and Cheng Yi-ch'ing. "A General Description of the Li Language." *Chung-kuo Yü-wen* (Peking), no. 5 (1963):432-41.

Reinhorn, Marc. *Dictionnaire laotien-français.* 5 vols. Paris, 1955. Mimeographed.

Rivière, Armand. *Mission pavie, géographie et voyages.* Vol. 4. Paris, 1902.

Saul, Janice E. "Classifiers in Nùng." *Lingua* 13 (1965):278-90.

Savina, F. M. *Dictionnaire étymologique français-nùng-chinois.* Hong Kong, 1924.

Savina, F. M. "Lexique dày-français accompagné d'un petit lexique français-dày et d'un tableau des différences dialectales." *Bulletin de l'École française d'Extrême Orient* 31 (1931):103-99.

Serdiuchenko, G. P. *Chuanskii iazyk.* Moscow, 1961.

Wang Li and Ch'ien Sun. "Preliminary Investigation of the Li Language of Pai-sha on Hainan Island." *Ling-nan Hsüeh-pao* (Canton) 11, no. 2 (1951): 253-300.

Wang, Stephen S. *Phonology of Chinese Loanwords in a Northern Tai Dialect.* Ann Arbor, 1966. [University Microfilms]