

Benedict's Austro-Tai Hypothesis—An Evaluation

LAWRENCE A. REID

INTRODUCTION

THE POSSIBLE CONNECTIONS among the hundreds of languages of Southeast Asia have been the subject of considerable research and a fair amount of speculation since at least the beginning of this century. Three major language families have generally been recognized—Austroasiatic (including languages such as Vietnamese, Mon, Khmer, Khasi, Nicobarese, and Munda), Sino-Tibetan (including Chinese, Tibetan, Burmese, and scores of other little-known languages), and Austronesian (represented in Mainland Southeast Asia by the Chamic group and Malay). Attempts to relate various of these families into superstocks and to establish the position of Thai within one or another of these families dates back to Schmidt's (1906) attempt to relate Austronesian and Austroasiatic in a family that he called Austric. Thai, primarily because of its monosyllabic word structure and its tonal system which is similar to Chinese, was generally classified as belonging to a Sino-Thai group within Sino-Tibetan.

In 1942, Paul Benedict published a paper in the *American Anthropologist* that proposed what he called "a new alignment." He proposed that Thai was not genetically related to Chinese at all, but to the Austronesian language family, which he at that time, following Dempwolff (1938), called the Indonesian languages. A series of papers followed, culminating in a volume, *Austro-Thai Language and Culture* (Benedict 1975), which reprinted his earlier papers and included a glossary of over 200 pages of Austro-Thai reconstructions with extensive discussion of sound correspondences and lexical evidence from each of the major branches of his proposed family.

Despite a ringing endorsement from Ward Goodenough in the foreword, comparativists in the Austronesian field have been lukewarm at best towards the hypothesis. Two of the better-known linguists in this group, Isidore Dyen and Robert Blust, both reject it. Blust (personal communication) states, "the correspondences simply don't work, unless you fudge every case with special conditions."

The purpose of this paper is to take a close look at what Benedict has done and to provide an evaluation of his proposed hypothesis.

BENEDICT'S HYPOTHESIS

The New Alignment

When Benedict first proposed (1942) that the Thai languages and "Indonesian Languages" were genetically related, he did it on the basis of comparative work that he had been doing on a group of relatively unknown languages, which he concluded were genetically related to each other, and which he labelled Kadai. These languages (Li—the speakers of which, according to Benedict, call themselves Dai—spoken on Hainan Island; and Laqua and Lati spoken in the border area of China and Vietnam) had been noted by early researchers as bearing certain similarities to the languages of Formosa, and (especially in the case of Laqua) also to Cham.¹ It was noted that these languages had similar grammatical structures. They all had Verb-Object word order, and modifiers followed their head words, that is, Adjectives followed Nouns, Possessors followed Possessed, and so on. In addition they had noun prefixes (Lati *m-* and *a-*, Kelao *bu-*, Laqua *ka-*), which are phonetically the same as (or similar to) those commonly found in Austronesian languages.

Benedict examined the lexical material and offered about 36 proposed cognate sets, ten of which were numerals, the pronoun "I," a demonstrative "this," while the remainder were basic lexical items (sun, rain, water, flower, fire, man, father, head, hair, eye, ear, tooth, foot, breast, blood, fat, eat, night, weep, die, alive, black, yellow, and small). He then introduced reconstructed Proto-Tai forms into the cognate sets, adding a number of other basic lexical items such as bird, bone, sour, blind, and grandfather.

Benedict contrasted the kinds of comparisons that he was making with those that suggested a relationship between Chinese and Thai. Of the latter, there were numerals (3-10, 100) and a few body parts (such as arm, leg, and palm of the hand), but the majority were terms for animals or birds and associated cultural items (such as horse, saddle, elephant, tusk, bull, cow, hare, fowl, bee, goose, pigeon, owl) and items of trade (such as silver, indigo, ink, playing cards, salt), all of which suggest something other than a genetic relationship between Thai and Chinese. It was in this paper that Benedict first discussed the problems inherent in comparing basically monosyllabic lexical items carrying tone, which occur in the Thai and Kadai languages, with possible cognates in Austronesian languages which are disyllabic or in some cases even trisyllabic.

It is generally recognized (Matisoff 1973) that Thai and the other tonal languages of Southeast Asia have undergone extensive phonological attrition over the centuries² and that the introduction of tone (as a part of the phonological system) was a way to maintain lexical distinctions, as certain final consonants were being eroded.

It is these phonological developments that are at the heart of the problems in Benedict's methodology. In order to establish cognature, he must sometimes compare a Thai or Kadai monosyllabic form with the initial syllable of an Austronesian disyllabic form (sometimes with the first syllable plus initial consonant of a second syllable in the case of CVCVC forms), and sometimes with the second syllable, for example,³

PAN *walú	= Li du	"eight"
PAN *maCá	= PT *ta	"eye"
PAN *púsuq	= PT *pot	"lungs"
PAN *búta	= PT *bot	"blind"

There can be no question that contrastive stress (which Benedict implies must have been present in the parent language of his proposed Austro-Tai family) has played a significant role in the phonological development of many Philippine languages. It was also

present in the parent language of this group (Zorc 1972, 1978), but the presence of contrastive stress in Proto-Austronesian is not universally accepted, despite attempts by Zorc (1983) and Dahl (1982) to prove that it was. It is perhaps significant, however, that the position of stress in the four forms cited above just happens to agree with its position in cognates in Philippine languages which still retain contrastive stress, for example, Bontok *walú* 'eight', *matá* 'eye', *púsu*, 'heart', Binukid Manobo *búta* 'blind' (Reid 1971). This agreement in stress position is by no means universal, however, as one should expect given the immense period of separation between the languages. That there is *any* agreement is important. What needs to be done is a statistical measure to determine its significance. Nevertheless, to some linguists, Benedict's penchant for choosing either ultima or penultima in his cognate search, depending on phonetic similarity and using stress in the proto-language as an explanation, is not methodologically sound since stress in Proto-Austro-Tai cannot be independently reconstructed. Moreover, the received opinion has been that early stages of Austronesian did not have contrastive stress.

The second major problem that Benedict faced, that of reconciling forms that were often only partially phonetically similar, is dealt with only tangentially in his 1942 paper. In this he chose to rely on the presentation of his proposed cognates to establish his thesis, but did not attempt any systematic reconstruction of the parent language.

The Austro-Tai Papers

In 1966 and 1967 Benedict published a series of three articles which developed more fully the thesis introduced in 1942. These are the articles in which the term "Austro-Thai"⁴ first appeared. He stated that the stimulus for these studies came from an examination of a considerable amount of new language material that had become available to him. This material included a number of studies of the so-called Miao-Yao languages,⁵ some publications by Fang-kuei Li on the Kam-Sui languages of south-central China, and some old word lists of the Ong-Be languages spoken on Hainan. These languages are all claimed by Benedict to be part of the Austro-Tai group and to have provided him with large numbers of new cognate sets, many of which are discussed in the papers, and most of which appear in the glossary of his 1975 book.

In addition to the evidence he adduced from mainland languages, he examined material from a number of Formosan languages and claimed to have found a wealth of forms there that are directly relatable to the mainland languages. He also took note of works by Haudricourt (1951) and Goodenough (1962) which supported the reconstruction of labio-velar consonants (*kw, etc.) in Oceanic and which he felt "provided a test case . . . for evaluating the claim that Thai and Kadai are directly related to this ancestral AN language [i.e. Proto-Oceanic]."

It is in these papers that Benedict begins the systematic reconstruction of Proto-Austro-Tai. We also get a clearer picture of his methodology for dealing with the problem, mentioned at the end of the last section, of what to do with possible cognate sets that are only partially phonetically similar. One method is extension of his ultimate or penultimate syllable equation: setting up stress doublets in the proto-language, i.e., pairs of forms differing only in stress, so that a form in one language can be derived from one of the pairs, while a form in another can be derived from the other. Another method he uses is reconstructing complex consonant sequences in the proto-language, which were reduced in different ways in the various daughter languages.

An example of both methods is illustrated by his reconstruction of the Proto-Austro-Tai word meaning 'louse' (1975:22, 333):

IN	*kutú
PT	*thraw
Sui	tu
Mak	tau
Then	tiu
Li	sau~su
Ong-Be	kat

The first six forms Benedict derives from a PAT reconstruction with stress on the final syllable. The Ong-Be form he derives from an alternate form which would have had stress on the penult.⁶ The presence of *s* in the Li form leads him to reconstruct a *-tl-*medial sequence, PAT *kutlu. He was able to find one other Li cognate set which supported this reconstruction, PAT *talú 'three', which, he suggests, became Li sú~su (after loss of the initial vowel). To an Austronesianist accustomed to the extremely limited range of consonant clusters in Proto-Austronesian (even medial nasal-stop clusters in PAN are suspect, Reid 1982), Austro-Thai reconstructions look weird indeed, with syllable initial clusters of stop/nasal + r/l/ɺ combining with prenasalization. This does not mean, of course, that such forms could not have occurred. The assumption that is problematic is that dissimilar correspondences are usually the result of different developments of proto-language clusters. Too little is known of the phonological development of most of the languages being compared to be able to make this assumption.

It is true that many of the non-Austronesian mainland languages do have very complex consonant clusters, some of which are of the type postulated by Benedict for PAT. However, it is just as likely that such clusters are the result of the reduction of disyllabic forms through the loss of unstressed vowels, or the result of prefixation, or other poorly understood processes (e.g., those which produced the initial voiceless nasals of Proto-Tai, Miao-Yao and other languages), than that these clusters were present in PAT. Benedict (1975:233) proposes, for example, a PAT reconstruction *[ma]mlok (from which the well-known PAN *manuk 'bird' would have derived) to account for the following set:

PT	*nok
Dioi	rɔk
Sek	nok
Proto-KS	*()mlok
Sui Mak	nok
Then	nɔk
Kam	mok
Lk	mlok
OB	nok
Lq	nuk
Laha (Tu)	ma/nɔk
Laha (BB)	nok
Proto-MY	**nɔ?
Proto-Miao	**noŋ
Proto-Yao	*nɔ?

Benedict reconstructed the medial *ml cluster to account for two apparently aberrant forms, Kam *mok* and Lk *mlok*. But considering the fact that *ma* appears to be retained in the Laha form, it is just as likely that a reconstruction such as *manuk or *ma[n,l]uk (with the medial consonant ambiguous, or indeterminate) could be made, with the Lk form maintaining the original consonant of the root.

Benedict was delighted to discover that labio-velars were present in some Oceanic languages, because he then felt justified in positing them also for PAT. Whether or not it is

possible to reconstruct them at this time depth (?6000 B.C.) for PAT, no evidence has ever been produced to suggest that complex consonants of this sort were present in PAN. To the contrary, the Oceanic evidence indicates fairly clearly that such clusters were innovations in Proto-Oceanic. Furthermore, with our present subgrouping assumptions (see section on other explanations below), in which Oceanic is fairly far down on the subgrouping tree, one would need to assume that the Oceanic labio-velar consonants were independent innovations in that branch rather than retentions from Proto-Austronesian that were lost everywhere else in the family.

EVALUATING THE EVIDENCE

Recognizing that there are serious problems in the way Benedict has chosen to reconstruct PAT⁷ does not mean that the languages under discussion are not genetically related. Proving a genetic relationship is a matter of degree. Usually required are sets of sound correspondences supported by convincingly large bodies of lexical forms. The more recent the linguistic split, the easier it is to prove genetic relationship. Conversely, the more remote the split, the less likely it is that such evidence can be produced. The greater the time depth the greater the number of phonological changes that can take place obscuring forms that are cognate, and the greater the chance that cognate forms are replaced. The great number of items that Benedict reconstructs, and often with meanings of a highly specific nature, casts doubt on the validity of the work, given the great time depth that must be involved.

Having said this, I would like to take a look at some of the items that he cites, and a few others as well, and to suggest that the similarities we find are of such kinds and in such quantities that they are highly unlikely to be accidental, and probably point to a genetic relationship.

The forms I wish to discuss are those that are part of the basic lexicon (sometimes called the core vocabulary) of a language. Such forms are generally believed to be more retentive (that is, less likely to be replaced by competing forms, and less likely to be replaced by borrowing than other forms), and therefore more likely to represent true cognate sets. In addition I will discuss some morphological items, that is, forms having grammatical function or forming part of closed classes such as pronouns and demonstratives. Such items are highly retentive, lingering on in the resources of a language, sometimes with altered functions, but often relatable to similar forms in distantly related languages.

The Tai-Kadai, Miao and Yao forms are mostly taken from Benedict (1975), for which see abbreviations. The Bontok (Reid 1976), Ilokano (Carro 1957), Tagalog (Panganiban 1972) and other Philippine language forms (Isneg, Batak, and Tagbanwa) (Reid 1971) are included in order to exemplify reflexes of the Proto-Austronesian reconstructions and to demonstrate the position of stress on these forms.

Possible Cognate Sets in the Basic Lexicon

1.	AN:		*bulan	month, moon
		Bon	búlan	
	TK:		*bían	
2.	AN:		*ʔəmbun	mist
	TK:		*fon	rain (esp. fine rain)

- | | | | | |
|-----|------------------|-----------------|--------------|----------------------|
| 3. | AN: | | *xapuy | fire |
| | | Bon | ?apúy | |
| | TK: | | *vay, *vi | |
| | | Kam | pwi | |
| | | Sui | wi, vi, vui | |
| | | Lati | pu | |
| 4. | AN: | | *laki | man, male |
| | | Bon | láki | male |
| | TK: | | *la:ŋ | (gr.) child, yg. man |
| | | Dioi | laŋ | male |
| | | Lakkia | lak | man (male) |
| 5. | AN: | | *qulu | head |
| | | Bon | ?úlu | |
| | TK: | SW Thai: Ahom | ru | |
| | | C Thai: Nung | hu, hua | |
| | | Laqua | ru | |
| 6. | AN: | | *maCa | eye |
| | | Bon | matá | |
| | TK: | Si, Po-ai | ta: | |
| | | Sui Mak | da | |
| | | Then | ?da | |
| | | Kam | ta | |
| | | Ong-Be | da | |
| | | Li | sa, sa | |
| | | Dakili | tša | |
| 7. | AN: | | *ma-Cay | die |
| | | Bon | ma-táy | |
| | TK: | Si, Po-ai | tay | |
| | | Sui Mak | tai | |
| | | Then | tai | |
| | | Kam | tei | |
| | | OB | dai | |
| | MY: | | *day | |
| 8. | AN: | | *(ŋ)ipan | tooth |
| | | Ilk | ŋípen | |
| | TK: | | *van | |
| | | Lao | ven | |
| | | Kam | phyan | |
| | | Sui | wyan, vyan | |
| | | S Li | phen, fen | |
| 9. | AN: | Hova | nunu | nipple |
| | | SEP | *nunu | breast (Benedict) |
| | TK: | Lao | hnu | |
| | | OB | nu | breast, milk |
| | Y: | YHN | nu | |
| 10. | AN: | | *kan, *ka?əŋ | eat |
| | TK: ⁸ | SW Thai | kin | |
| | | N Thai: Wu-Ming | kin, kī | |
| | | Dioi | ken | |
| | | Po-ai | kīn | |
| | | OB | kon | |
| | | N Li | kha:n | |
| | | Lq | kə:n | |
| | | N Kl | ka | |
| | | Lt | kho | |

11.	AN:		*[dD]anum	water
	TK:	Bon	danúm	
		LK	*nam	
		OB	num	
		S Li	nam, nòm	
			nom	
12.	AN:		*DəDəm	dark
	TK:	Bon	sedém	evening
		SW Thai	dam	dark, black
		C Thai	dam	
		Mak	?dam	shade; dark
		S Li	dəm	black
		N Li	dam	
13.	AN:		*qa()jaw	sun
	TK:	Bon	?algew	
		Sek	*?da:w	star
		Li	tra:w	
			ra:w	
14.	AN:		*laŋuy	swim
	TK:	Ilk	láŋuy	
		KS	*lo:y	swim, float
		Li	lui	float
			lei	swim
15.	AN:		*bəna	low-lying
	TK:		*na:	(flooded) land
		OB	nea	ricefield
		S Li	na	
16.	AN:		*manuk	bird
	TK:	Bon	manúk	chicken
		Si	nok	bird
		Lungchow	nuk	
		Sui Mak	nok	
		Then	nɔk	
		OB	nok	
17.	AN:		*aLak	child
	TK:	Bon	?anák	
		SW Thai	lu:k	
		Dioi	lək	
		Sek	lik	
		Sui	lak	
		Then	la:k	
18.	AN:		*qa-lsəm	sour
	TK:	Ilk	?alsém	
		SW Thai	som	
		C Thai	Lum	
		N Thai	Ləm	
		Sui	hum	
		Mak	sum	
		Kam	sem	
		Then	them	

19.	AN:	Bon	*bayi bá?i	female, mother female
	TK:	Thai	*me: mai	mother pref. for young girls or young women
		OB	mai	female
		WS Li	pa:i	woman, wife
		WS Loi	bai	older sister
20.	AN:	Bon	*inay ?iná	mother mother (relationship term)
	TK:	Thai	na:y	
		Sui	ni	
		Mak	nay	
		Then	nei	
21.	AN:	Bon	*ətut ?utút	fart
	TK:	Sui Kam	*tot tət	
		Mak	tut	
		OB	*tut	
		WS Li	thut	
22.	AN:	Bon	*m-aRi ?-um-áli	come
	TK:	Thai	*hma ma	
		Dioi	ma	
		Sek	ma	
23.	AN:	Bon	*dataŋ datáŋ	reach, arrive
	TK:	SW Thai	thiŋ	
		C Thai	thiŋ	
		N Thai	taŋ	
		OB	*taŋ	
24.	AN:	Isneg	*qabaRa? ?abága	shoulder
	TK:	Si	*?ba bá	
		Sek	va	
		Mak	ha	
		Li	*va	

Possible Cognate Sets in the Functional Morphology

25.	AN:	Bon	*-ku -ku	1sg pronoun
	TK:	Si	ku	1sg pron (superior to inferior)
		Lao	ku	
		Dioi	ku	
		Khamti	kaw	
		OB	hao	
		S Li	həu, du	
		N Li	ho	
		Laqua	khəu	
		Lati	ku	
MY:	Miao		ku	

26.	AN:		*-Su	2sg pron
	TK:	White Thai	su, miŋ	
		Khamti	su	you
		Ahom	su, si	
		Si	su:	2sg pron (superior to inferior)
27.	AN:		*-mu	2pl pron
	TK:	W Li	mo	
		N Kelao	mu	thou, you
28.	AN:		*ni	this
		Tag	?ini	
	TK:	Si, Lao, BT, WT, Lk	ni	
29.	AN:	Bon, Ilk	nay	here, this
	TK:	Nung, Lk, KS	nai	this
		S Li	nei	
		Li	nay	
	MY:	Yao	nai	
30.	AN:		*na	that; ligature; determiner
		Bon	na	this; -na 3s pr;
			nan	determiner
	TK:	Li	na	3s
		Loi	na	
		WS Li	na	that, there
		Si, Lao	nan	that (near addressee)
		Lk	nan	that
	MY:	Yao	nan	3s pron
		Miao	ni	
31.	AN:		*tu	that
		Tag	?itu	
	TK:	Lk	tu	they
		Lq	to	
		Lt	ato	that
	MY:	Miao	to	
32.	AN:	Tag	no?on	that far (genitive)
	TK:	Si	no:n	that far
33.	AN:		*di	that (far); locative
		Bon	di	that (far)
	TK:		*di:	place; locative
		SW Thai	thi:	
		C, N Thai	ti:	
34.	AN:		*ti	that
		Ilk	?iti	determiner
		Batak	?iti	that (far)
		Tagbanwa	?iti	
	MY:	Miao	ti	
35.	AN:		*Di	negative
	MY:	Miao	tsi	

A third set of possible cognates is found in the numerals, a number of which are very similar in form to those found in Austronesian languages. Benedict (1975:444) provides a chart of numerals from a number of languages, and discusses those that he believes are cognate.

OTHER EXPLANATIONS

There are a variety of explanations, other than postulating a genetic relationship, for the similarities that are discussed above and that have been noted by so many other linguists. One class of explanations may be categorized as language contact explanations. These include borrowing, substrata (and other kinds of stratal influence), and areal influence or diffusion. These are the kinds of explanations that are paramount in accounting for many of the similarities found between Thai, Vietnamese, and Chinese, including their monosyllabic syllable structure and tone. They all assume a period of linguistic contact, varying from intermittent and casual trading relationships to extended periods of geographical contiguity resulting in "diffusion" of features between languages, and periods of contemporary occupation of the same geographical territory resulting in "stratal" influence. Both of the latter types of contact imply extensive bi- or multilingualism.

In the context of the present discussion we need to ascertain the degree of likelihood that one or more of the above relationships existed between the ancestors of the Austronesian languages and the ancestors of the Thai languages. At the present time there are only two Austronesian groups on the mainland: the Chamic languages, and Malay and its related languages (such as Urak Lawoi' which is spoken on Phuket in Thailand). Neither of these groups gives any evidence that they are residual enclaves of some Austronesian homeland. Both Malay and the Chamic group can be unambiguously assigned to a Western Malayo-Polynesian subgroup. Virtually all Austronesianists believe that the ancestors of the Chamic group moved back to the mainland from some area within the Malay-Indonesian language area. We have no evidence moreover that the Chamic group on the mainland has ever been in contact with other than the Austroasiatic languages (Khmer, Bahnaric and Viet-Muong) that presently surround it. And although the Malay group presently is contiguous to the languages of southern Thailand, we know that this contact is relatively recent—within the last 800 years or so—as Thai speakers moved from the South China area and split the Mons on the Burmese side from their Austroasiatic Khmer cousins on the Kampuchean side.

To establish areal diffusion (or stratal influence) as a likely explanation, it would be desirable to have evidence that there was an Austronesian homeland on the mainland, and preferably in the South China area. But we know of no languages spoken in this area that are unambiguously Austronesian, let alone one which would show evidence of representing an Austronesian homeland. If there had been one in the past which has now been Sinitized or for some other reason has disappeared, or if the Pre-Austronesian ancestors of the Proto-Austronesians lived in this area on the mainland (as indeed is probable), the possibility of language contact as an explanation for the similarities we are discussing would exist.

Borrowing as a result of some kind of trading relationship is unlikely, for at least two reasons. First, the kinds of terms we discussed above are not the kind that are likely to be borrowed in such a contact situation. Second, we have no other evidence that early Austronesians and early Thais carried on such trading, although they may have. If Pre-Austronesians moved from Mainland Southeast Asia to what is now Taiwan they could also have returned. But at the time depth we are talking about, which must have been prior to the dissolution of the Proto-Tai community, there is no evidence that such trade occurred.

The other kinds of possible explanations are of two types. The first accounts for similarities which are the result of the inherent character of language. Such an explanation is

often given for agreement in syntactic patterning which is not infrequently found between genetically unrelated languages. Such agreement could be considered supportive of a genetic relationship based on other criteria but may not of itself establish the relationship. The fact that Thai and Indonesian are both SVO tells us nothing about their genetic relationship. Chinese is also SVO.⁹ Thai and its related languages do have a head-modifier word order, like Proto-Austronesian and most Austronesian languages, with adjectives following their head nouns, and possessor nouns following possessed nouns, unlike Chinese. But explanations of this sort do not account for the similarity in lexical forms that has been discussed above.

The other explanation that has been used to account for these similarities is coincidence. It is proposed that because of the limited number of phonemes in language and their limited combinatorial possibilities, accidental similarities are bound to occur, and therefore coincidences of the type Thai *fai*—English *fire*, Thai *tai*—English *die*, Thai *rim*—English *rim*, must be expected and do not suggest genetic relationship. This is true, but it would be extremely improbable for coincidence to bring about such striking similarities, not only in the core vocabulary but in the morphology as well.¹⁰

AUSTRONESIAN SUBGROUPING AND THE HOMELAND HYPOTHESES

Postulating an Austro-Thai genetic relationship, or even a close contact relationship, has implications for an Austronesian homeland. It can probably be assumed that homelands were located in the geographical vicinity of the "seams" between first order subgroups. If Austronesian and Thai are genetically related, then postulating Formosa as the Austronesian homeland as Benedict and others have done is reasonable, because of its geographical contiguity to Southeast China, the presumed homeland of Proto-Tai. Even if they are not genetically related, and if the similarities between the groups are the result of extensive linguistic contact, we would still need to place pre-Austronesians in Southeast China.

It may be useful at this point to review the various hypotheses regarding the homeland of the Proto-Austronesians, considering the degree to which they are supportive of a Formosan homeland, or whether other possibilities exist. Crucial to this discussion is the position of Oceanic languages within Austronesian. Early subgrouping hypotheses divided Austronesian into two major families, a Western group—Hesperonesian, and an Eastern group—Oceanic. Such a subgrouping is explicit in the work of Tsuchida (1976), who makes Formosan languages a branch of the Hesperonesian group. Haudricourt's (1965) subgrouping is similar to Tsuchida's, but with Formosan made a third primary branch. Dyen's (1965) lexicostatistical subgrouping had a large number of "highest order subgroupings" (24 of them simple languages, 12 of them small groups), the majority of which were clustered in three areas; Melanesian-East New Guinea, West New Guinea, and Formosa. Dyen suggested each of these areas as a possible homeland, based on the assumption that areas of greatest linguistic diversity implied longest periods of settlement. Dyen (1964) subsequently removed Formosa from the list of possible homelands because he claimed that the lexicostatistical percentages in this area were deflated. His most recent statements (Dyen 1982) continue to keep Oceanic as a primary subgroup, and he still believes (Dyen, personal communication) that the homeland was probably in one of the two areas he formerly cited.

Other linguists do not consider Oceanic languages to be a primary subgroup, thus in effect removing the Oceanic-Hesperonesian seam (that is, the area of Eastern Indonesia—

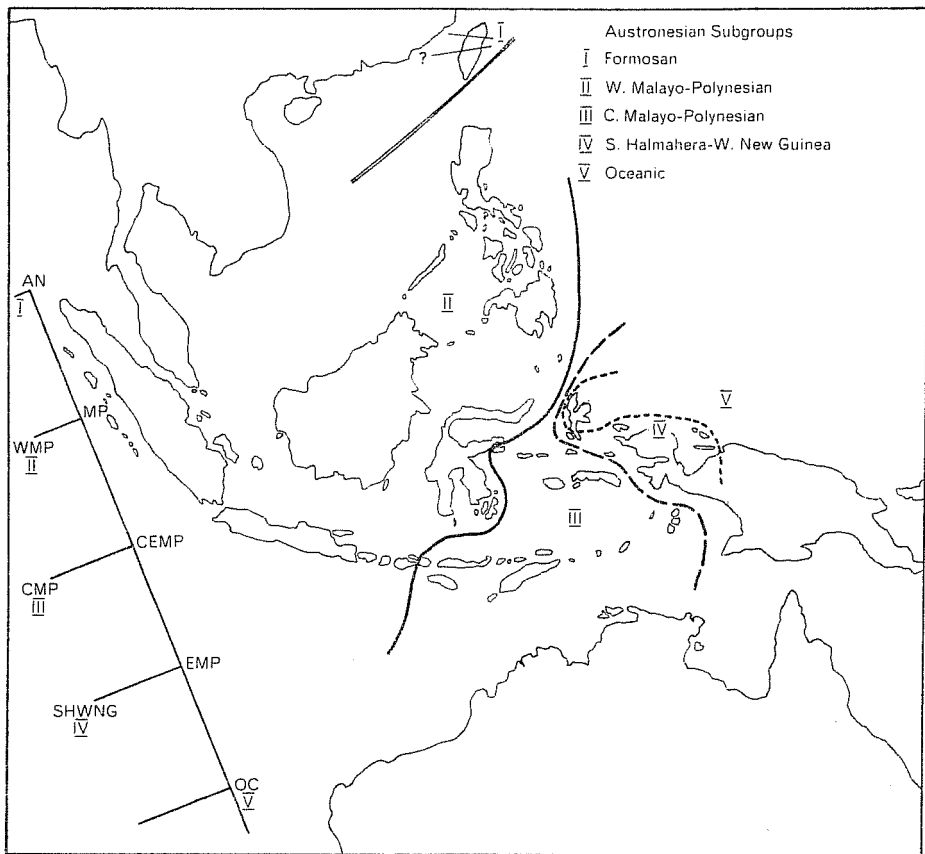


Fig. 1. Blust's (1978) tentative subgrouping of Austronesian (drawn by V. M. Lyon).

Western New Guinea-Melanesia) as a possible homeland. Dahl's (1973) subgrouping makes all non-Formosan languages a single subgroup, a position also taken by Blust (1977), who labels this subgroup Malayo-Polynesian. In Blust's version, Oceanic languages are at the end of the line, a fifth order subgroup (see Fig. 1).

The evidence that has been produced for considering all non-Formosan languages a single subgroup is not extensive, but it is persuasive. It includes several phonological innovations found in all non-Formosan languages, plus a number of morphological innovations. The phonological innovations are as follows:

1. *t and *t' > *t
2. *l and *L- > *l
3. *n and *-L-, -L > *n

Of these, numbers (2) and (3) represent a "split merger" that would not likely have developed independently in more than one group.

The morphological innovations in the non-Formosan languages which can be reconstructed for their parent (but not for Proto-Austronesian) include the following:

- | | |
|-----------|--|
| 1. *paN- | (and associated *maN- and *minaN-) transitivity prefix |
| 2. *siDa | 3 pl. Nominative pronoun |
| 3. *-mu | 2 sg. Genitive pronoun (PAN 2 pl.) |
| 4. *-atən | 1 incl. pl. Predicative pronoun (PAN *-itən) |

In summary, an eastern Indonesian-western Oceanic homeland for Proto-Austronesian, which would contraindicate an Austro-Thai genetic hypothesis, is supportable only on the basis of lexicostatistics. Subgrouping hypotheses that imply a Formosan Austronesian homeland are more generally accepted and are supported by fairly persuasive qualitative evidence.

THE AUSTRIC HYPOTHESIS

First proposed by Schmidt (1906), and both supported and refuted by many linguists since, the Austric hypothesis would link Austronesian and Austroasiatic languages into one superstock. This hypothesis is relevant to the present discussion, because if it is possible to show that Austronesian and Austroasiatic languages are probably genetically related, it would enhance the possibility of an Austronesian-Thai genetic relationship because of the geographic distribution of the languages.

The similarity of some of the Austronesian morphology to Austroasiatic, especially the *pa- 'causative', has been noted in the literature. It appears, however, that the full extent of the similarities has not been recognized. They are in fact so distinctive that only a genetic explanation can adequately account for them.

Nancowry, a language classified by all who have examined it as Austroasiatic, has a morphological apparatus so similar in form and function to what has been reconstructed for PAN that an Austronesianist looking only at the morphology would immediately consider it to be Austronesian. Nancowry, spoken in the Nicobar Islands, is apparently a relic area, sufficiently isolated from the Mainland Southeast Asian linguistic area to have escaped the innovative pressures that resulted in the loss or modification of the original morphological features from many of its sister languages.

The following comparisons, based on a brief article by Radhakrishnan (1976), should be sufficient to illustrate the point.

	<i>Nancowry</i>		<i>Austronesian</i>
ha-	causative	*pa-	causative
h-an-	caus + instrumental	*paN-	instrumental
ma-/		*maR-	agentive
-am-	agentive		
-in-	instr. nominal	*-in-	completive
-a	objective	*-a	objective
-um-	causative	*-um-	agentive (also causative in some forms)

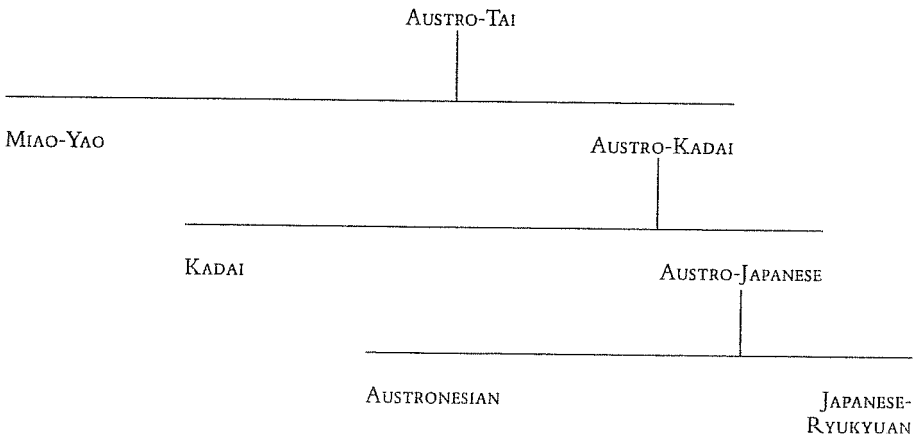
The similarities both in form and function are striking. But perhaps what is even more striking is that the process that produces infixation in the Austronesian languages is apparently still operating in Nancowry. Infixes in all of these languages appear after the

initial consonant of the word. In many Austronesian languages, infix *-um-* alternates with prefix *mu-*, and infix *-in-* alternates with prefix *ni-*. In Nancowry, infix *-am-* alternates with prefix *ma-*. For both Austronesian and Nancowry, a unique phonological process (metathesis of the first two consonants of the affixed word) must be operating to produce these alternations.

If Nancowry really is an Austroasiatic language its morphology clearly shows it has Austronesian connections. These could not be the result of language contact with perhaps Malay to the south, since it is highly unlikely that a language would borrow morphology without borrowing the forms to which they are attached. Neither could Nancowry be an Austroasiatic language with a Malay substratum—the affixes are too archaic to be considered to be Malay. The only alternative is to consider Austronesian and Austroasiatic to be genetically related, and if we accept that, Austro-Tai is only a step away.

CONCLUSION

The accumulated evidence presented by Benedict to support a genetic relationship between Austronesian, Thai, and its linguistic relatives on the mainland provides us with a foothold for further research. Benedict (personal communication) admits that the presentation of these data was not satisfactory and that this is partly the reason that linguists in general continue to be skeptical. His most recent work however is likely to open up another area of controversy. Benedict has become convinced that Japanese is also genetically related to Austronesian, and at a shallower time depth than that of Austro-Tai. In a soon-to-be published monograph, he charts the relationships between the families as follows (Benedict N.D.):



NOTES

1. More recently, Benedict has included the Thai languages within the Kadai family.
2. The Thai writing system, less than 700 years old, makes a number of phonological distinctions, and represents various consonant clusters no longer used in modern Thai.

3. The asterisks before the forms in this list, and elsewhere in the paper, mark lexical items which are reconstructions for some proto-language, based on comparisons between apparently cognate forms in its supposed daughter languages. The equals sign (=) denotes a proposed genetic relationship between the protoform on the left (in this case Proto-Austronesian) and the form on the right of the arrow. This language may itself be a proto-language, as in the case of the Proto-Tai forms for "eye," "lungs," and "blind."
4. Although Benedict used the term "Austro-Thai," the term more commonly used today is "Austro-Tai," in which "Tai" refers to the family of languages and dialects to which "Thai," the standard form of the language of Bangkok, also known as Siamese, belongs.
5. Miao is perhaps better recognized today as Hmong, or Mong, the language of large numbers of refugees in the United States and elsewhere.
6. Philippine languages show penultimate stress on this root.
7. It should be noted that Benedict himself calls his cognate sets LCG's, i.e., Likely Cognate Groups having a better than 50 percent chance of being cognate (1975:139). He furthermore labels his reconstructions "provisional . . . of a kind that might be labeled simply 'work in progress'" (1975:146).
8. Li (1977:262) suggests the vocalic alternation in this set "may be due to an original diphthong *i i, but the reconstruction is doubtful."
9. Proto-Austronesian was probably VOS, or VSO, and Indonesian has only relatively recently undergone a syntactic innovation which reordered the basic sentence constituents.
10. For an excellent discussion of the various possibilities discussed in this section, but with reference to the languages of Thailand, see Matisoff (1973).

ABBREVIATIONS

AT	Austro-Thai	PAN	Proto-Austronesian
BON	Bontok	PAT	Proto-Austro-Thai
BT	Black Tai	PT	Proto-Tai
C	Central	S	South
ILK	Ilokano	SEP	Southeast Papua
IN	Indonesian	Si	Siamese
KS	Kam-Sui	SW	South-West
LK	Lakkia	Tag	Tagalog
Lq	Laqua	TK	Thai-Kadai
Lt	Lati	W	West
My	Miao-Yao	WS	White Sand
N	North	WT	White Tai
OB	Ong-be	YHA	Haininh Yao

REFERENCES

- BENEDICT, PAUL K.
 1942 Thai, Kadai and Indonesian: a new alignment in Southeastern Asia. *AA* 44:576-601.
 1966 Austro-Thai. *Behavior Science Notes* 1:227-261.
 1967a Austro-Thai studies. *Behavior Science Notes* 2:203-244.
 1967b Austro-Thai studies: 3. Austro-Thai and Chinese. *Behavior Science Notes* 2:275-366.
 1975 *Austro-Thai: Language and Culture, with a Glossary of Roots*. New Haven: HRAF Press.
 N.D. *Japanese/Austro-Tai*.
- BLUST, ROBERT
 1977 The Proto-Austronesian pronouns and Austronesian subgroupings: a preliminary report. *Working Papers in Linguistics* 9 (2): 1-15. Honolulu: Department of Linguistics, University of Hawaii.
- CARRO, ANDRES
 1957 *Iloko-English Dictionary*. Translated, augmented, and revised by M. Vanoverbergh. Baguio: Catholic School Press.
- DAHL, OTTO C.
 1973 *Proto-Austronesian*. Scandinavian Institute of Asian Studies Monograph Series, No. 15. Lund: Studentlitteratur.
 1982 *Early Phonetic and Phonemic Changes in Austronesian*. Instituttet for Sammelningende Kulturforskning, B-63. Oslo: Universitetsforlaget.

- DEMPWOLFF, OTTO
1938 *Vergleichende Lautlehre des austronesischen Wortschatzes*, vol. 3, *Austronesisches Wörterverzeichnis*. Berlin: Dietrich Reimer.
- DYEN, ISIDORE
1964 The position of the Malayopolynesian languages of Formosa. *AP* 7:261-271.
1965 *A Lexicostatistical Classification of the Austronesian Languages*. International Journal of American Linguistics, Memoir 19. Baltimore: Waverly Press.
1982 The present status of some Austronesian subgrouping hypotheses, in *Papers from the Third International Conference on Austronesian Linguistics*, vol. 2: *Tracking the Travellers*: 31-35, eds. Amran Halim, Lois Carrington, and S. A. Wurm. Pacific Linguistics C75. Canberra: Australian National University.
- GOODENOUGH, WARD
1962 Comment in Arthur Capell, *Oceanic linguistics today*. *CA* 3:371-396.
- HAUDRICOURT, ANDRE
1951 Variations parallèles en Mélanésien. *Bulletin de la Société Linguistique de Paris* 47:140-153.
1965 Problems of Austronesian comparative philology. *Lingua* 14:315-329.
- LI, FANG KUEI
1977 *A Handbook of Comparative Tai*. Oceanic Linguistics Special Publication No. 15. Honolulu: University of Hawaii Press.
- MATISOFF, JAMES A.
1973 Tonogenesis in Southeast Asia, in *Consonant Types and Tone*: 73-95, ed. Larry M. Hyman. Southern California Occasional Papers in Linguistics No. 1. Los Angeles: University of Southern California.
- PANGANIBAN, JOSE VILLA
1972 *Diksiyunaryo tesauro Pilipino-Ingles*. Quezon: Manlapaz Publishing Co.
- RADHAKRISHNAN, R.
1976 A note on the morphology of the causative in Nancowry, in *Austroasiatic Studies Part II*: 1035-1040, eds. Philip N. Jenner, Laurence C. Thompson, and Stanley Starosta. Oceanic Linguistics Special Publication No. 13. Honolulu: University of Hawaii Press.
- REID, LAWRENCE A.
1971 *Philippine Minor Languages: Word Lists and Phonologies*. Oceanic Linguistics Special Publication No. 8. Honolulu: University of Hawaii Press.
1976 *Bontok-English Dictionary*. Pacific Linguistics C36. Canberra: Australian National University.
1982 The demise of proto-Philippines, in *Papers from the Third International Conference on Austronesian Linguistics*, vol. 2, *Tracking the Travellers*: 201-216. Pacific Linguistics C75. Canberra: Australian National University.
- SCHMIDT, WILHELM
1906 Die Mon-Khmer-Völker, ein Bindeglied zwischen Völkern Zentralasiens und Austronesiens. *Archiv für Anthropologie* 33:59-109.
- TSUCHIDA, SHIGERU
1976 *Reconstruction of Proto-Tsoui Phonology*. Monograph Series No. 5. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
- ZORC, R. DAVID
1972 Current and proto-Tagalic stress. *Philippine Journal of Linguistics* 3 (1): 43-57.
1978 Proto-Philippine word accent: Innovation or Proto-Hesperonesian retention?, in *Second International Conference on Austronesian Linguistics: Proceedings*. Pacific Linguistics C61 (Fascicle 1): 67-119. Canberra: Australian National University.
1983 Proto Austronesian accent revisited. *Philippine Journal of Linguistics* 14 (1): 1-24.