

The Case-marking System of Proto-Formosan

Stanley Starosta

Professor of Linguistics

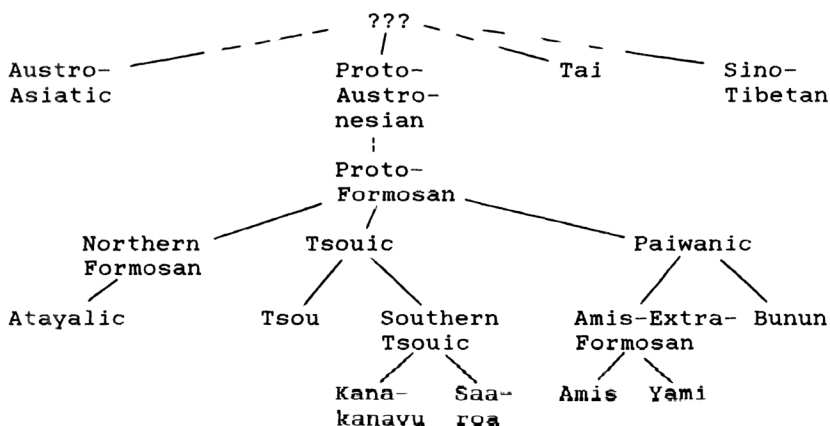
Department of Linguistics

University of Hawaii, Honolulu, Hawaii 96822 USA

I. INTRODUCTION

'Formosan' is the name used by Austronesian linguists to refer to the aboriginal languages of Taiwan. Taiwan is very probably the homeland of all the Austronesian languages of the Pacific and mainland Southeast Asia, and Proto-Formosan (PF), the lowest common ancestor of these languages, is either a first-order subgroup of Proto-Austronesian (PAN), or is identical with PAN itself. If the Austronesian language family is in fact genetically connected with Tai, Sino-Tibetan, and/or Austroasiatic, this is the level at which to look for that connection:

1) Proto-Formosan and its extended family¹



¹The Austronesian part of the family tree shown in this diagram is based on subgroupings by O.C. Dahl (Dahl 1973:124), R.A. Blust (Blust 1977:2), Paul Li (Li 1985) and Laurence Reid. PF had auxiliary verbs which

Formosan languages in general are grammatically quite similar to one another and to Philippine languages, especially in areas such as verb morphology and pronoun and determiner systems. However, there are significant internal differences among them, and the reconstructed proto-system that I will present in this paper in an attempt to account for the modern diversity looks much more isolating and mainlandish than any of its modern descendants.

We have a fairly clear idea of the verbal morphology and the general case-marking typology of this protolanguage thanks to work by Dahl (1973) and Wolff (1973). Starosta, Pawley, and Reid (1982) have reconstructed some of the earlier changes leading to the modern verbal morphology patterns, while Blust (1977) and Reid (1981) have reconstructed several sets of pronouns and determiners. However, many questions about the nature of the original case-marking system and its evolution into the various configurations found in the modern Formosan languages remain open. This paper, which draws on earlier work on Formosan languages by myself and other researchers, is an attempt to partly correct that deficiency. Using the lexibase dependency framework and the comparative method, it concentrates on reconstructing the earliest grammatical devices which overtly marked the grammatical relations between NPs and their regent verbs or nouns, and on showing the historical connections between NP case-markers and verbal focus affixes.

II. THEORETICAL FRAMEWORK

I will maintain at the outset that grammatical reconstruction cannot be done meaningfully in a theoretical vacuum. Within a rigorous and constrained theory, every reconstructed stage must meet the same constraints obeyed by any modern language, and a plausible abductive mechanism must be demonstrated for moving from each stage to the next. If the theoretical framework is sufficiently generative and constrained, then there will be little room for wild divergences in the reconstruction. Two independent reconstructions undertaken within the same generative and constrained framework will ideally reach the same conclusion, and we can have some confidence that this conclusion is in fact the correct one.

The framework I am employing in this paper is lexibase, a type of monostratal lexicon-based dependency grammar. This model has been tested and refined in the analysis of parts of almost fifty

languages, and I believe it is generative and constrained enough to use for the purpose of fairly reliable grammatical reconstruction. Nevertheless, although my reconstruction of the Proto-Formosan case-marking system will be stated for the sake of convenience as if it were attested fact, it must be borne in mind that the reconstruction is necessarily conjectural. It approximates factuality to the extent that it is the only possible scenario which is compatible with the facts of the modern languages, with the theory, and with the requirements of abductive grammatical change.

III. NOMINAL ACTANTS

A. Case marking typology

Proto-Formosan case-marking was ergative. NP case marking was somewhat similar to that of modern Atayal, with a small number of prepositions and determiners supplemented by a fairly elaborate system of demonstrative nouns and relator nouns to mark relations of location and possession between a head word (a verb or noun) and an NP dependent.

B. Noun phrases and determiners

There is little doubt that Proto-Formosan, like Thai and Khmer on the mainland, was a head-initial language, with two pre-noun determiners the only exceptions to this rule. No Nominative determiner seems to be reconstructible at this level, but non-subject actants were usually marked by one of two pre-N determiners, **i*[+Det,-Nom,+dfnt] if definite or **a*[+Det,-Nom,-dfnt] if indefinite. Reflexes of both **i* and **a* can be seen vestigially in those modern determiners containing *i-* or *a-*, such as for example the Paiwan determiners *t_i*, *n_i*, and *t_{jai}* (personal) versus *a*, *nu_a*, and *tu_a* (common) (Egali 1990:160).

C. Case functions

The definite non-nominative determiner **i* was used to mark a broad range of non-nominative (non-subject) actants, including preposed topics, locative NPs, definite notional objects of two-argument intransitive verbs (analogous to the function of Tagalog *sa* and Polynesian *i*), and possibly transitive Agents, while **a* was used for indefinite notional objects of two-argument intransitives (comparable to the function of Tagalog *ng* in antipassive sentences).

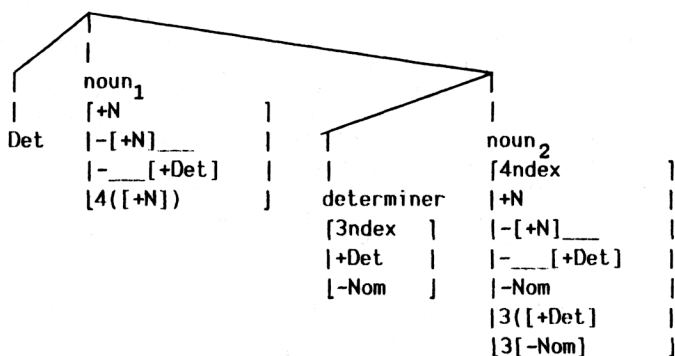
The Paiwan adnominal Genitive determiner *i* (cf. Egli 1989:188) and the Amis (Chen 1982:282,286) and Paiwan (Egli 1989:186) Locative *i* are direct

descendants of **i*, while the Paiwan and Amis ligature *a* reflects **a*, but otherwise neither **i* nor **a* survives in its original clause-level function in any of the daughter languages. Traces of both are however abundant in the determiner and pronoun morphology of the modern languages. The use of **i* to mark topicalized definite NPs can for example be seen in the *i*-initial topicalized Nominative pronoun forms in Kanakanavu (cf. Mei 1982:210-211) as well as the Atayal free Nominative pronouns *isu?* and *ita?*, which contrast with the respective clitic forms *su?* and *ta?* (cf. Egerod 1966:347-48, Huang 1989:117). The *i*-initial alternative forms of the nominative determiners which mark topicalized subjects in the Tapang dialect of Tsou (Tung 1964:64) also reflect this function.

D. The dependency structure of Noun Phrases

Proto-Formosan head nouns allowed NP or S dependents. Except for determiners, all dependents of nouns branched off to the right. NP dependents of the head noun attached either directly to their regent (e.g. inalienable possession and location) or else occurred as dependents of intervening relator nouns or the demonstrative noun **na*:

2) Proto-Formosan NPs



Noun-headed dependents of noun regents bore one of three functions to the head noun: equational, possessive, or locative, the same dependent types which are found in mainland languages such as Thai (Savetamalya 1989) and Old Khmer (Sak to appear).

1. Equational dependents

The equative dependent construction is very common in the languages of mainland and insular Southeast Asia. It is essentially a relative clause construction

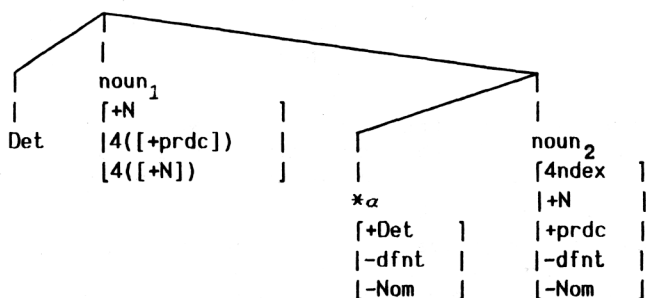
in which the relative clause has a predicate noun as its head, and in which the missing subject of the relative clause is coreferentially chained to the regent noun of the relative clause. As a non-headed construction, an equational attribute has the internal structure of an NP, allowing the same dependent types as other NPs, while its relative clause function is manifested in the internal gap, the missing subject which is coreferential with the regent N (cf. Savetamalya 1989, Sak in progress).

Equational predicates can be divided into two subtypes based on whether the head noun of the dependent NP is indefinite (descriptive predicates) or definite (identificational predicates), and equational relative clauses can be divided into two types accordingly:

a. Descriptive dependents

In the prototypical equational relative clause, the indefinite predicate noun dependent was marked with the indefinite determiner **a*:

3) Descriptive equative dependent



Such descriptive NP predicates are indefinite; they do not refer to a particular individual known from context, but rather state a property, such as an occupational role, that applies to the regent noun. Thus the diagram above can be read as 'noun₁ who/which is a noun₂'. This construction differs from possessive and locative dependent constructions in that the following modifying noun is a predicate [+prdc], rather than a case role-bearing noun.

b. Identificational attributes

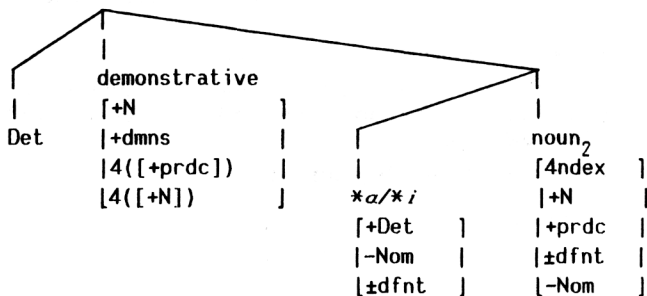
There is some evidence that identificational predicate NPs were also used as noun attributes in Proto-Formosan. In Tsou, equational attributes may be marked by either *no* or *ci*, and Tung's examples (Tung

1964:156) suggest that *no* might be intrinsically indefinite and *ci* definite. This could be the result of derivation from an earlier relator noun (**nV* or **iV*) followed by an indefinite determiner **a* or definite determiner **i* respectively.

c. Demonstrative nouns

Words glossed as demonstrative determiners and demonstrative pronouns in English were demonstrative nouns in Proto-Formosan. As in the modern languages they could occur alone as the sole constituent of a NP, but frequently occurred as the head of an equative construction:

4) Demonstrative nouns

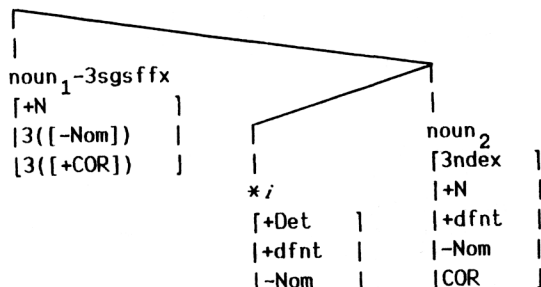


The determiner on the dependent NP could be the indefinite **a* ('this/that one which is a noun₂') or the definite **i* ('this/that one which is the noun₂').

2. Possessive dependents

Possessive dependents of nouns were definite noun phrases bearing the Correspondent case relation to their regent nouns and designating a 'possessor' in very general sense. First or second person possessor in the modern languages are usually encoded as suffixes on the head noun, and this may have been the situation already in PF or before. Third person possessors were encoded as right dependent non-nominative noun phrases, third person possessive suffixes, or possibly both, as is the case in Saaroa (Tsuchida 1976:67), Chamorro, and various Micronesian languages. Possessors are typically definite, and so were marked with the definite non-nominative determiner **i*:

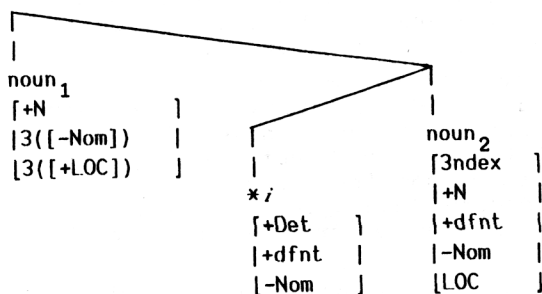
5) Possessive attributes



3. Locative dependents

The locative dependent construction was identical to the possessive construction except for its Locus case relation and the absence of a possessive suffix on the regent noun:

6) Locative attributes



E. Noun classes

1. Common and personal nouns

PF non-pronominal nouns (all nouns other than pronouns, demonstrative nouns, and relator nouns) were invariant in form, and able to cooccur with a full range of attributes. The grammatical distinction between common nouns and personal nouns (proper nouns, kinship terms, and personal pronouns) is a fairly early one, and occurs in several Paiwanic languages, including Rukai, Paiwan, and Amis. It does not however occur in Tsouic or Atayalic, and so cannot be reconstructed for Proto-Formosan.

The most salient characteristic of personal nouns is that they are lexically definite. Thus the presence of reflexes of the definite **i* in personal determiners

is striking in Rukai (Li 1973:86-87) and Amis (Chen 1982:286), in Tagalog personal nominative *si*, genitive *ni*, and locative *kay*, and in Paiwan, where the personal forms reflecting definite **i* contrast neatly with non-personal forms reflecting indefinite **a* (cf. Egli 1990:160).

2. Pronouns

PF was an ergative language, and had two basic sets of pronouns, Nominative and non-Nominative, each with clitic and non-clitic subtypes. In the original PF system, the clitic and free forms were formally distinct, with the free forms being longer.

First and second person clitic nominative pronouns immediately followed the root verb.² Transitive clauses in modern Formosan languages carry either two clitic pronouns, Nominative and either Genitive or Locative, as in Atayal (Huang 1989:122-28) or just one, Nominative in intransitive clauses and Genitive in transitive clauses, as in Tsou. Since the two-clitic pattern is found at least vestigially also in Paiwan (Egli 1989:156,296), in a different subgroup and located at the other end of the island from Atayal, it can probably be reconstructed for PF.

As in Tsou, there was a third person non-Nominative clitic in PF, identical to the third person possessive suffix on nouns, while the third person nominative clitic form did not exist. There were probably no third person free pronouns at all in Proto-Formosan. In modern languages, all third person free pronouns seem to derive from earlier demonstrative nouns.

Nominative clitic pronouns marked the non-third person subject (Nominative Patient) constituent of all clauses headed by an auxiliary verb, and this category corresponds with the actor in all intransitive clauses, including anti-passives. The non-Nominative clitics on the other hand appeared only in transitive clauses marking an Agent actor:

²PF had auxiliary verbs which functioned as root verbs and 'attracted' clitic pronouns.

7) Proto-Formosan pronouns

	Clauses:		NPs:
	Nominative	Non-Nominative	Genitive
Clitic	<i>PAT</i>	<i>AGT</i>	<i>Possessor</i>
(non-3rd)	(all clauses)	(transitive clauses)	
	<i>actr</i>	<i>actr</i>	
	(intransitive clauses)	(transitive clauses)	
Free	<i>PAT</i>	<i>AGT, LOC, etc.</i>	--

The **i* definite non-nominative determiner cooccurred with fused with non-nominative personal pronouns, and fused with them in the formation of *i*-initial Locative personal pronouns or pronoun paradigms in Amis (Chen 1982:306) and Bunun (cf. Jeng 1977:130-31).

Free nominative pronouns could optionally appear in normal clause-final subject position for emphasis. The more common position for free nominatives however was in 'exposure', functioning as preposed topics or noun predicates. The free [+Nom] pronouns occurring the topic position in the Southern Tsouic group and as well as some in Atayal acquired a reflex of the [-Nom] **i*, possibly by contamination from definite NP predicates. This feature is amply attested for example by *i*-initial free topicalized nominative pronouns in Kanakanavu (Tsuchida 1976:40-41, Mei 1982:210-211) and Atayal (Huang 1989:117).

IV. PROTO-FORMOSAN CLAUSE STRUCTURE AND VERBAL MORPHOLOGY

A. Transitivity and case marking

1. Ergative, antipassive, and the evolution of verbal focus

At the earliest stage reconstructible internally and comparatively, PF verbal morphology was relatively isolating. Clause structure, like NP structure, was right-branching except for optional sentence-initial topics. As stated earlier, Proto-Formosan, like most or all of its Formosan and Philippine descendants, was an ergative language, which means in lexibase terms that the Patient constituent was always marked by the nominative case form and vice versa. The non-

nominative case form marked transitive agents, locatives, antipassive 'objects', and adnominal adjuncts. In these respects, PF was quite similar to modern Rukai and Tsou.

The simple intransitive, antipassive intransitive, and transitive clause patterns of Proto-Formosan can be represented schematically as follows:

8) Simple Intransitive clauses: Atayal (Huang 1986:7)

pima	saku?	'I am going to wash.'
wash	I	
[-trns]	[<u>Nom</u>]	
[-apsv]	[<u>PAT</u>]	
	[actr]	

9) Antipassive intransitive clauses: Atayal (Huang)

pəpima	saku?	sunan	'I am going to wash
wash	I	to you	you.'
[-trns]	[<u>Nom</u>]	[-Nom]	
[+apsv]	[<u>PAT</u>]	[Lcv]	
	[actr]	[LOC]	

10) Transitive clauses: Atayal (cf. Huang 1986:7)

pəman	saku?	nya?	'He is going to wash
wash	I	by him	me.'
[+trns]	[<u>Nom</u>]	[-Nom]	
	[<u>PAT</u>]	[Gen]	
		[AGT]	
		[actr]	

As discussed above, PF had a [\pm dfnt] definiteness distinction in non-Nominative NPs, with a determiner * α marking indefinite noun phrases and * i marking definite NPs. Since antipassive is a construction which puts the performer into the center of the action and downgrades the undergoer, the non-nominative actant in the antipassive construction tended to be indefinite, marked by indefinite * α or by no determiner at all.

11) PF antipassive clauses, indefinite object

V	* α NP	NP	e.g. 'The man will look
look at	dog	man	at a dog.'
[-trns]	[-dfnt]	[+Nom]	
[+apsv]	[-Nom]	[actr]	
	[MNS]	[PAT]	

When it was definite, it was marked by the definite **i*:

12) PF antipassive clauses, definite object

V	<i>*i</i> NP	NP	e.g.	'The man will
look at	dog	man		look at the
[-trns]	[+dfnt]	[+Nom]		dog.'
[+apsv]	[-Nom]	[actr]		
	[LOC]	[PAT]		

The intransitive two-argument **i* pattern included not only antipassives derived from transitives, but also underived two-argument intransitive constructions such as locative intransitives, e.g.

13) PF Locative intransitive clauses

V	<i>*i</i> NP	NP	e.g.	'The man will walk
walk to	house	man		to the house.'
[+apsv]	[+dfnt]	[+Nom]		
[-trns]	[-Nom]	[actr]		
	[LOC]	[PAT]		

Essentially this system is preserved in Rukai, where traces of the definite **i* and indefinite **a* remain in the demonstrative determiners and articles, though the original definiteness dimension has been specialized to a distinction between personal and common nouns.

We have now almost arrived at the point of departure for Starosta, Pawley, and Reid's *The evolution of focus in Austronesian* (Starosta, Pawley, and Reid 1982, hereafter *SPR*). The verbal focus system which they reconstructed as the earliest ancestor of the modern Philippine focus systems can be derived from the configuration above in one simple step: 'recentralization', a transitivizing verbal derivation process of a common type (cf. Starosta 1988:171-174) reinterprets the non-Patient arguments of these antipassive constructions as definite Patients ('recentralization'; cf. Starosta 1986), in the process capturing the determiners of the immediately following reinterpreted NPs as suffixes and upgrading the original Patient to Agent:

14) 'Goal focus'/'Object focus'

V- <i>a</i>	NP	NP	e.g.	'The man will view
look at	dog	man		the dog.'
[+afct]	[+dfnt]	[-Nom]		
[+trns]	[+Nom]	[actr]		afct = 'affect'
	[PAT]	[AGT]		

V- <i>i</i>	NP	NP	e.g. 'The man will view
look at	dog	man	the dog.'
[+afct]	[+dfnt]	[-Nom]	
[+trns]	[+Nom]	[actr]	
	[PAT]	[AGT]	

15) 'Locative focus'

V- <i>i</i>	NP	NP	e.g. 'The man will visit
walk to	house	man	the house.'
[+lfct]	[+dfnt]	[-Nom]	
[+trns]	[+Nom]	[actr]	lfct = 'local affect'
	[PAT]	[AGT]	

This stage is exactly of the form which SPR posited as the input to the changes which account for the evolution of at least the transitive part of the modern focus systems. SPR reconstructed two systems of verbal focus for PAN, a newer Philippine-looking system (which I will refer to here as F2), which they postulated to have arisen by the reinterpretation of nominalized cleft equational sentences, and an earlier system (referred to here as F1) preserved in languages such as Tsou and fragmentarily in various languages in subordinate clause contexts, whose origin they did not attempt to account for:

Original (F1) and noun-derived (F2) PAN verbal focus affixes

	AF	OF	LF	IF
	[-trns]	[+trns]	[+trns]	
F1	mu-/-um-	O/-a O/-i	-i	-
F2	mu-/-um-	O/-ən	-an	sa-, si-

This paper now pushes the syntactic prehistory of the Austronesian language family one step farther back than SPR did, to propose a source for the Object Focus and Locative Focus affixes *-a* and *-i* of the earlier F1 system in the PF non-nominative determiners **a* and **i*.

V. REFERENCES

- Blust, R.A. (1977) 'The Proto-Austronesian pronouns and Austronesian subgrouping: a preliminary report.' University of Hawaii Working papers in linguistics 9:2.1-15.

- Chen, T.M. (1982). **Verbal constructions and verbal classification in Nataoran-Amis.** Ph.D. dissertation, University of Hawaii. Pacific Linguistics.
- Dahl, O.C. (1973) **Proto-Austronesian.** Scandanavian Institute of Asian Studies monograph series no. 15. Lund: Studentlitteratur.
- Egerod, S. (1966) 'Word order and word classes in Atayal.' **Language** 42.346-369.
- Egli, H. (1990) **Paiwanggrammatik.** Wiesbaden: Otto Harrassowitz.
- Ho, A.Y.L. (1990) **Yami structure: a descriptive study of the Yami language (English version).** M.A. thesis, National Tsing Hua University, Taiwan.
- Huang, L.M-J. (MS, 1986) **Some remarks on Atayal.** Taipei: Department of English, National Taiwan Normal University.
- (1989) 'The pronominal system in Atayal.' **Studies in English Literature and Linguistics** 15:115-33. Taipei: National Taiwan Normal University.
- Jeng, H. (1977) **Topic and focus in Bunun.** Special publication no. 72; Taiwan: Institute of History and Philology, Academia Sinica.
- Li, P.J-K. (1973) **Rukai structure.** Academia Sinica special publications no. 64. Taipei: Academia Sinica.
- (1985) 'The position of Atayal in the Austronesian family.' In Andrew Pawley and Lois Carrington (eds), **Austronesian linguistics at the 15th Pacific Science Congress; Pacific Linguistics** C-88, 257-280.
- Mei, K. (1982) 'Pronouns and verb inflection in Kanakanavu.' **Tsing Hua Journal of Chinese Studies, New series** 14, 1-2:207-231.

- Reid, L.A. (1981) 'Proto-Austronesian genitive determiners.' In Andrew Gonzalez and David Thomas (ed), *Linguistics across continents : studies in honor of Richard S. Pittman*. Philippine Journal of Linguistics monograph series. Manila: Summer Institute of Linguistics (Philippines) : Linguistic Society of the Philippines monograph series, no 2.
- Sak, Chhany (In progress) *The syntax of dated Pre-Angkorian inscriptions*. M.A. thesis, Department of Linguistics, University of Hawaii.
- Savetamalya, S. (1989) *Thai nouns and noun phrases : a lexicase analysis*. Ph.D. dissertation, University of Hawaii at Manoa
- Starosta, S. (1988) *The case for lexicase: an outline of lexicase grammatical theory*. London: Pinter Publishers.
- (1986) 'Focus as recentralization.' In Paul Geraghty, Lois Carrington, and S.A. Wurm (eds), *FOCAL I: papers from the Fourth International Conference on Austronesian Linguistics*, Pacific Linguistics, C-93.
- Starosta, S., A. Pawley, and L.A. Reid (1982) 'The evolution of focus in Austronesian.' In S.A. Wurm and Lois Carrington (eds), *Papers from the Third International Conference on Austronesian Linguistics*, vol. 2: *Tracking the travelers*; Pacific Linguistics C-75.
- Thompson, L.C. (1965) *A Vietnamese grammar*. Seattle: University of Washington Press.
- Tsuchida, S. (1976) *Reconstruction of Proto-Tsouic phonology*. Monograph series no.5, Study of languages and cultures of Asia and Africa. Institute for study of languages and cultures of Asia and Africa. Tokyo Gaikikugo Daigaku.
- Tung, T-H. (1964) *A descriptive study of the Tsou language, Formosa*. Institute of history and philology special publication no. 48. Taipei: Academia Sinica.

Wolff, J.U. (1973) 'Verbal inflection in Proto-Austronesian.' In Andrew B. Gonzalez (ed), Parangal Kay Cecilio Lopez (Essays in honors of Cecilio Lopez on his seventy-fifth birthday). Philippine journal of linguistics special monograph issue no.4. Quezon City: Linguistic Society of the Philippines, 71-94.

(1979) 'Verbal morphology and verbal sentences in Proto-Austronesian.' In Paz Buenaventura Naylor (ed), Austronesian studies: papers from the second eastern conference on Austronesian languages. Michigan Press on South and Southeast Asia no. 15. Ann Arbor: Center for South and Southeast Asian Studies, University of Michigan, 153-168.