

# CONSTRUCTION MARKERS AND SUBGROUPING OF FORMOSAN LANGUAGES<sup>1</sup>

RALEIGH FERRELL

1. Topicalisation in Formosan Languages
2. Verb Inflection
3. Marking of Noun Phrases
  - 3.1. 3-CM Systems
  - 3.2. 2-CM Systems
  - 3.3.  $\emptyset$ -CM Systems
  - 3.4. General and Specificational CMs
4. Syntactic, Lexical and Phonological Subgrouping
5. Proto-Formosan Syntax

## 1. TOPICALISATION IN FORMOSAN LANGUAGES

Obligatory sentence focus (topicalisation) is basic to all Formosan aboriginal languages. The verb in all instances is inflected to show the role of the topicalised Noun Phrase (such as agent, object/goal, locus, instrument, cause, beneficiary). The topicalised NP itself is then marked by either (1) an over phrase-introducing particle (Construction Marker), or (2) a fixed-order function slot for the NP, or (3) both of these syntactic devices. These devices, whether overt CMs or constituent NP order, are equational: the same device which identifies a NP as being in primary relationship with the Verb Phrase also equates NPs in non-verbal (equational) sentences.<sup>2</sup>

Non-topicalised NPs are also marked by either CMs or constituent NP order. Typically, all non-topicalised NPs may be marked by identical,

non-equational CMs, except that in languages having a discreet genitive/partitive marker the latter is used to mark the agent/actor in non-Agent Focus sentences.<sup>3</sup>

In certain Formosan languages, the topicalisational CMs perform a dual role by being differentiated to indicate proximity or specificity as well as the focus/non-focus roles outlined above. In at least one language, focus/non-focus roles are marked strictly by fixed NP order; and CMs, which obligatorily introduce each constituent NP, function solely as articles indicating relative proximity or specificity. These unusual developments distinguish these languages from other western Austronesian ones, and heighten the interest of Formosan languages to comparative Austronesian syntactic studies.

## 2. TOPICALISATIONAL VERB INFLECTION

The primary importance of the verb inflections in Formosan focus (topicalisation) system is attested to by the widespread preservation of cognate forms for these inflections. This is in contrast to the variability and relative instability of overt CMs and other NP-marking devices connected with topicalisation.

In Formosan languages generally the various roles which may be filled by topicalised NPs are subsumed structurally under four verbal inflections, herein referred to as AF (Agent Focus), OF (Object Focus), RF (Referent Focus), and IF (Instrument Focus).<sup>4</sup> Figure 1 shows topicalisational verb inflections in various Formosan languages.

	Amis	Atayal	Bunun	Favorlang	Paiwan	Pazeh	Saisiat	Sediq	Siraya	Thao	Yami	Tsou	Saaroa	Rukai
AF	/m/	/m/	m-	/m/	/m/	mV-	/um/	/m/	/m/	/m/	/m/	/m/	/um/	w-
OF	-ən	-un	-un	-en	-en	-un	-Vn	-un	-en	-in	-en	-i(?)	(?)	ki-(?)
RF	/m/+an	-an	-an	-an	-an	-an	-an	-an	-an	-an	-an	-a(?)	-a(na)	ta-t-a(n)
IF	sa-	se-	-is	i-	si-	si-t-i	si-	se-	(?)	sa-	i-	-(n)eni	-(a)ta	sa-t-a(n)

Figure 1. TOPICALISATIONAL VERB FOCUS INFLECTIONS IN FORMOSAN LANGUAGES

	3-CM Systems				2-CM Systems				
	Amis	Atayal	Kuvalan Paiwan	Siraya	Yami	Puyuma	Sediq	Rukai	Saaroa
CM=	o	∅	a	a	ta	u	a	ka	ka
CM#	to	sa	ta	tua	tou	su	da	∅	sa
CMgen	no	na	na	nua	ki	nu			na

Figure 2. GENERAL FOCUS CMs IN SOME FORMOSAN LANGUAGES

## 3. TOPICALISATIONAL MARKING OF NOUN PHRASES

## 3.1. 3-CM SYSTEMS

Amis, Kuvalan, Paiwan, Siraya, Yami, Pazeh, Thao, Favorlang and Ciuli-Atayal each have three topicalisational NP Construction Markers. In these 3-CM systems, the topicalised NP is introduced by an equational CM (CM=), the agent of non-Agent Focus sentences is introduced by the genitive/partitive CM (CMgen), and all other non-topicalised NPs are marked by a non-equational CM (CM≠).

The following examples from Paiwan show how 3-CM systems typically work. Note that in non-Referent Focus sentences, locative NPs may be introduced by secondary, CM-like markers<sup>5</sup> such as the Paiwan *i 'at, in'*, which may either co-occur with the CM≠ or may result in deletion of the latter.<sup>6</sup>

*'(The) man hunts wild-pigs in the mountains with a spear':*

AF	q/m/alup	a	caucau	tua	vavuy	{ tua }	gadu	tua	vu uq
	<i>hunts</i>	CM=	<i>man</i>	CM≠	<i>pig</i>	{ i tua }	<i>mtn</i>	CM≠	<i>spear</i>
						CM≠			
OF	qalup-en	nua	caucau	a	vavuy	{ tua }	gadu	tua	vu uq
		CMgen		CM=		{ i tua }		CM≠	
						CM≠			
RF	qalup-an	nua	caucau	tua	vavuy	a	gadu	tua	vu uq
		CMgen		CM≠		CM=		CM≠	
IF	si-qalup	nua	caucau	tua	vavuy	{ tua }	gadu	a	vu uq
		CMgen		CM≠		{ i tua }		CM=	
						CM≠			

Saaroa must also be included among the languages having 3-CM systems. Although Saaroa has only two overt topicalisational CMs (CM= and CM≠), structurally the equivalent of a 3-CM marking system is maintained by assigning the function of a CMgen to a  $\emptyset$ -marker. The following examples are illustrative:

m-aci?i ka cacil*i*  
*die* CM= *one person*  
*'one person died'*

k/um/ita na a|a|iamu  
*see[AF]* CM≠ *fly*  
*'(they) saw a fly'*