

The position of Chamorro and Palauan in the Austronesian family tree: evidence from verb morphosyntax

ERIK ZOBEL

1 Introduction¹

Palauan and Chamorro, spoken in Micronesia on the Palau (Belau) and Mariana Islands, respectively, have long been recognised as outliers in the Pacific region, with stronger ties to the languages of the Philippines and Indonesia than to neighbouring languages. In his ground-breaking monograph, Dempwolff (1934–38) divided the Austronesian language family into an ‘Indonesian’ and a ‘Melanesian’ subgroup. It has become apparent that according to this scheme, Palauan and Chamorro must be included in the ‘Indonesian’ subgroup, since they do not share the innovations characterising the ‘Melanesian’ subgroup (which latter under the label ‘Oceanic’ has remained firmly established as a well-defined subgroup of the Austronesian family).

In Dyen’s lexicostatistical classification of the Austronesian languages (1965), Chamorro and Palauan are isolates of the ‘Malayo-Polynesian’ linkage, coordinate to subgroups of relatively high order. In spite of the problematic nature of lexicostatistics (Dyen’s classification fails to recognise well-established subgroups such as Oceanic), it illustrates the isolated character of Palauan and Chamorro with regard to their lexicon.

Blust (1977) proposed a classification of the Austronesian languages which up to now has gained wide acceptance (Figure 1). It contains two nodes relevant to the discussion here: the Malayo-Polynesian (MP) subgroup, based on phonological, lexical and grammatical innovations (Blust 1995); and the Central-Eastern Malayo-Polynesian (CEMP) subgroup (chiefly based on lexical innovations; Blust 1993), which includes the Central Malayo-Polynesian (CMP) and the Eastern Malayo-Polynesian (EMP) group. The latter contains the

1. I am indebted to Bernd Nothofer, Malcolm Ross and Sander Adelaar for their comments on earlier drafts of this paper.

Oceanic languages. MP languages that are not included in CEMP were grouped together by Blust in a Western MP (WMP) group.

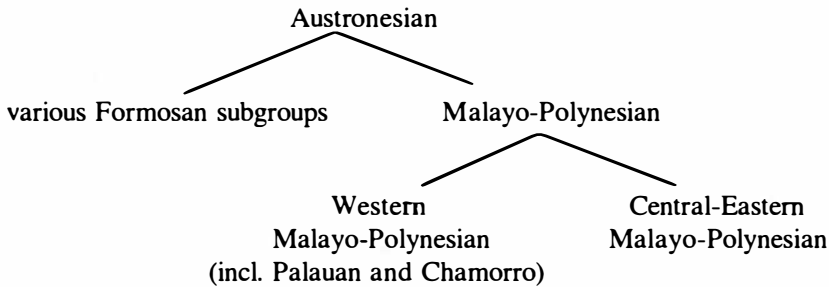


Figure 1: Austronesian family tree (following Blust)

According to this classification, Chamorro and Palauan must be included in the WMP subgroup, since they are clearly MP languages, and do not share the innovations defining CEMP. The next task is to establish closer ties between these languages and other WMP languages.

The phonological histories of the two languages give no clue apart from showing that they are non-Oceanic Malayo-Polynesian languages. The sound changes of both are either found in many other WMP languages (Chamorro: merger of **e* and **u*, Palauan: merger of **D* and **Z*, **ñ* and **n*) or are unique (Chamorro: merger of **D* and **k*, **j* and **q*; Palauan: merger of **j* and **R*, vocalisation of **l* and **p*). There are no phonological innovations common to both languages, apart from trivial ones (loss of **S*, stress on the PMP penultimate). Unlike Chamorro, Palauan shares with neighbouring Yapese and many Nuclear Micronesian languages the loss of final vowels, which is an areal feature in that part of the Pacific.

There have also been attempts to use grammatical evidence to establish the closer affiliations of Palauan and Chamorro. Pätzold (1968) demonstrated that many Palauan affixes (verbal and nominal) are shared with languages of the Philippines and Sulawesi, but this just proves the conservatism of Palauan in this respect.

For Chamorro, Topping (1973) claimed on the basis of its verbal system, that it should be grouped with Philippine languages such as Tagalog or Ilokano. His argument rests mainly on his focus analysis of the Chamorro verbal system, which is not fully appropriate, as I will show later. Starosta and Pagotto (1991) compared the Chamorro verbal system with Formosan and focus-preserving Malayo-Polynesian languages. They note the divergent character of Chamorro, which has led them to state that Chamorro is an early offshoot from PAN. However, if Chamorro is compared with languages further south, it can be seen that this divergence results from innovations which are not peculiar to Chamorro, but are shared by most languages of Indonesia and Oceania.

In this paper, I will use evidence from verbal morphosyntax to propose a modified subgrouping of the Malayo-Polynesian languages based on exclusively shared innovations, and establish the position of Chamorro and Palauan within this subgrouping.² I will discuss the affixation of verbs in the focus system, and in derivations from this system, and the interplay of these verbal affixes with the pronoun sets.

² The subgrouping hypothesis proposed here is elaborated in my dissertation (in progress).

A note on terminology: here, antipassive and passive are defined solely as syntactic surface categories. Both are syntactically intransitive, with the agent (A) or the object (O), respectively, as subject (S), and the other participant absent, incorporated or placed in an oblique relation. This presupposes that transitive and intransitive constructions are clearly distinguishable (e.g. in person marking), which is the case for Palauan and Chamorro.³ The basic transitive construction of both ergative and nominative languages is called 'active'. Focus languages are treated here as neither ergative nor nominative, since it is not clear whether actor focus (AF) or non-actor focus (non-AF) should be the 'basic' transitive form.

2 Reconstruction of the PMP verbal system

Most Formosan and many MP languages have a verbal system usually dubbed a 'Philippine-type' focus system. In line with Blust's subgrouping, it is safe to assume that this verbal system is inherited from PAN. Ross (1995) has reconstructed the focus system of PAN by concentrating on Formosan evidence, supported by additional data from MP languages. His reconstruction is summarised in Table 1.

Table 1: PAN focus system

	Past	Non-past	Atemporal	Projective
AF	<i>um</i>	<i>um</i>	∅	<i>um</i> -a
UF	<i>in</i>	-en	-u	-aw
LF	<i>in</i> -an	-an	-i	-ay

A reduplicated form of the non-past served as a progressive. There was also a 'stative passive' expressed by **ma-/*ka-*. It probably occurred in both non-actor focuses forming the paradigm found in Table 2.

Table 2: PAN stative passive

	Past	Non-past	Atemporal	Projective
UF'	<i>mina-</i>	<i>ma-</i>	<i>ka-</i>	?
LF'	<i>mina-</i> -an	<i>ma-</i> -an	<i>ka-</i> -i	?

The PMP verbal system did not differ much from the PAN verbal system (Table 3). The main innovations are the emergence of a fourth focus and the stem extensions **paN-* and **paR-*. The fourth focus using **Si-* is also found in many Formosan languages (e.g. Paiwan, Atayal, Bunun); the prefix **Si-* probably already existed at the PAN level as a noun forming affix (Ross 1995). The stative passive remained as in PAN.

³ With this definition, it will be seen below that nominative languages can have an antipassive (e.g. Palauan), just as ergative languages can have a passive (e.g. Chamorro).

Table 3: The PMP verb system

	Past	Non-past	Atemporal	Projective
AF	<i>⟨um⟩</i> <i>minaR-</i> <i>minaN-</i>	<i>⟨um⟩</i> <i>maR-</i> <i>maN-</i>	\emptyset <i>paR-</i> <i>paN-</i>	<i>⟨um⟩ -a</i> <i>maR- -a</i> <i>maN- -a</i>
UF	<i>⟨in⟩</i>	<i>-en</i>	\emptyset / <i>-u</i> (?)	<i>-a/aw</i> (?)
LF	<i>⟨in⟩ -an</i>	<i>-an</i>	<i>-i</i>	<i>-ay</i>
IF	<i>ini-/i- ⟨in⟩</i>	<i>i-</i>	<i>-an</i>	?
UF'	<i>mina-</i>	<i>ma-</i>	<i>ka-</i>	?
LF'	<i>mina- -an</i>	<i>ma- -an</i>	<i>ka- i</i>	?
IF'	<i>minai-</i>	<i>mai-</i>	<i>ka- -an</i>	?

The non-atemporal forms **maN-*, **maR-* and **ma-* are portmanteaus of **⟨um⟩* plus **paN-*, **paR-* and **ka-* with **um*. In the following discussion, I will call all affixes containing **⟨um⟩* *M-affixes* (including *⟨um⟩* itself), as opposed to *base affixes* (all non-AF affixes, and AF atemporal).

The stem extensions **paR-* and **paN-* are a characteristic of the Malayo-Polynesian languages: **paR-* is also found in Formosan languages, but restricted to forming reciprocals, while **paN-* with nasal substitution is an innovation particular to the Malayo-Polynesian group. Although it is difficult to establish the original function of the stem extensions, it can be roughly extrapolated from the modern daughter languages that in PMP **paR-* had a durative and reflexive/reciprocal function, while **paN-* had a distributive function, describing an action involving plural agents or objects. Both functions are transitivity-reducing, so it is not surprising that they are found mainly in AF (see below on the pragmatics of AF).

In many Malayo-Polynesian languages, the use of the stem extensions in non-AF has been limited to focussing circumstantial participants, such as location (LF), reason and instrument (IF) and occasionally time and manner, whereas non-AF without stem extensions focuses core roles, such as undergoer and goal. These 'circumstantial' focus forms (Table 4) are used mainly in cleft or equational constructions, which employ only the past and non-past tenses.

Table 4: Post-PMP 'circumstantial' focus forms

	Past	Non-past	Atemporal	Projective
LF''	<i>pinaR- -an</i> <i>pinaN- -an</i>	<i>paR- -an</i> <i>paN- -an</i>	<i>(paR- -i)</i> <i>(paN- -i)</i>	<i>(paR- -ay)</i> <i>(paN- -ay)</i>
IF''	<i>ipinaR-</i> <i>ipinaN-</i>	<i>ipaR-</i> <i>ipaN-</i>	<i>(paR- -an)</i> <i>(paN- -an)</i>	(?) (?)

In AF, the choice of *⟨um⟩*, *maR-* and *maN-* has become lexicalised in many Malayo-Polynesian focus languages, although some languages still allow all three forms with one verb.

No Malayo-Polynesian language has retained the system of Table 3 completely. In particular, the atemporal and projective non-AF forms have been conflated into a single category in all Malayo-Polynesian languages. Yet both sets have to be reconstructed 'from the top down', since they can be posited for PAN and reflexes of both are found in modern Malayo-Polynesian languages.

Noun case marking in PMP was as in modern Philippine languages, with nominative (marking the 'focussed' noun phrase), genitive (possessive, unfocused agent) and oblique (unfocused non-agent). The corresponding pronoun sets of PMP can be reconstructed as in Table 5.

Table 5: PMP pronoun sets

	Nominative	Genitive	Oblique
1SG	<i>(i-)aku</i>	<i>-(ng)ku</i>	<i>aken</i>
2SG	<i>(i-)kau</i>	<i>-mu/-nu/-u</i>	<i>iu(n)</i>
3SG	<i>sia</i>	<i>-nia</i>	<i>(ia ?)</i>
1PL.INC	<i>(i-)kita</i>	<i>-(n)ta</i>	<i>aten</i>
1PLEXC	<i>(i-)kami</i>	<i>-mami/-nami</i>	<i>amen</i>
2PL	<i>(i-)kamu(yu)</i>	<i>-muyu</i>	<i>imuyu(n)</i>
3PL	<i>siDa</i>	<i>-niDa</i>	<i>(iDa ?)</i>

In verb-initial sentences, genitive and nominative pronouns immediately follow the verb in that order (ignoring particles that also immediately follow the verb). If the sentence begins with a negative, adverb or any other member of a class of 'preverbs', genitive and nominative pronouns are fronted to immediately follow the preverb. In such constructions, the verb usually is in the atemporal. Examples (1) and (2) from Cebuano illustrate the fronting of pronouns.⁴

- (1) *Gi-tawg-an nako siya.*
 PST-call-LF 1SG.GEN 3SG.NOM
 'I called him.'
- (2) *Wa nako siya tawg-i.*
 NEG 1SG.GEN 3SG.NOM call-ATEMP.LF
 'I didn't call him.'

It can be assumed that the pragmatics of focus choice in PMP functioned as in modern Philippine languages. Focus selection is triggered by syntactic or pragmatic criteria. Syntactic criteria involve cases where the verb is nominalised, as in relative, existential and cleft clauses and in most WH-questions. Here, the verb must take the focus corresponding to the function of the highlighted NP. If the verb is not nominalised, focus is triggered by pragmatic criteria: roughly speaking, if the object NP is definite/referential and totally affected by the action,

⁴ Abbreviations used are: ACT active participle, AF actor focus, ANTI antipassive, APPL applicative, ART article, ASP aspect marker, ATMP atemporal, CONJ conjunction, GEN genitive, GER gerundive, HUM human, IF instrument focus, INTR intransitive, IRR irrealis, LF locative focus, LOC locative, NEG negative, NOM nominative, OBJ object, OBL oblique, PASS passive, PL plural, PERS personal article, POSS possessive, PST past, REAL realis, RED reduplication, RES resultative, SG singular, SUB subjunctive, UF undergoer focus.

this NP triggers the corresponding non-AF as in example (3) from Tagalog (if there is more than one non-agent core NP, case hierarchy determines focus selection). Zero anaphora for focussed non-agents is very common in most focus languages and was certainly a feature of PMP.

- (3) *D-in-alaw ko siya.*
 UF.PST-invite 1SG.GEN 3SG.NOM
 'I visited him.'
 (non-AF, definite object)

If the object NP is indefinite, or definite but partially affected, AF is selected,⁵ as in the Tagalog sentence (4).

- (4) *D-um-alaw ako ng mga kaibigan.*
 AF.PST-invite 1SG.NOM GEN PL friends
 'I visited some friends.'
 (AF, indefinite object)

The syntactic trigger always overrules the pragmatic trigger, as shown in example (5), from Tagalog: on pragmatic grounds, non-AF would be selected, but since the agent is highlighted in a construction that requires nominalisation of the verb, AF is chosen.

- (5) *Sino ang d-um-alaw sa kanya?*
 who NOM AF.PST-invite OBL 3SG.OBL
 'Who visited him?'
 (AF, definite object, but agent is questioned)

3 Grammatical sketch Of Chamorro

This short sketch is largely based on the descriptions by Topping (1973) and Cooreman (1987). Additional material is from Costenoble (1940).

In Topping (1973), Chamorro is described as having a focus system of the Philippine type. I will show below that the notion of focus (in the sense employed here) is not really applicable to Chamorro. Following Cooreman (1987), the Chamorro verbal system is better described as a split-ergative system. The ergativity-split is conditioned by mood: Chamorro distinguishes realis and irrealis mood: in realis there is ergative pronoun marking, while in irrealis there is nominative marking.

3.1 Chamorro pronoun sets and verbal morphosyntax

3.1.1 Pronoun sets

There are four pronoun sets in Chamorro (Table 6) with the following functions:

- Set A has two slightly different subsets depending on the mood of the sentence:
 - the agentive Set A1 marking A in realis mood;
 - the nominative Set A2 marking S and A in irrealis mood;
- Set B (absolute) pronouns mark S in realis mood and O in both moods;

⁵ Note that definiteness of the agent is not decisive for the selection of AF.

- the possessive set functions as possessor and A in certain nominalisations;
- free pronouns occur elsewhere.

Table 6: Chamorro pronoun sets

	Set B absolute	Set A1: agentive	Set A2:IRR nominative	Possessive	Free
1SG	<i>yo'</i>	<i>hu</i>	<i>(bai) hu</i>	<i>-hu/ku</i>	<i>guahu</i>
2SG	<i>hao</i>	<i>ɯn</i>	<i>ɯn</i>	<i>-mu</i>	<i>hago</i>
3SG	<i>gue'</i>	<i>ha</i>	<i>u⁶</i>	<i>-ña</i>	<i>guiya</i>
1PL.IN	<i>hit</i>	<i>ta</i>	<i>(u) ta</i>	<i>-ta</i>	<i>hita</i>
1PLEX	<i>ham</i>	<i>in</i>	<i>(bai) in</i>	<i>-mami</i>	<i>hami</i>
2PL	<i>hamyo</i>	<i>en</i>	<i>en</i>	<i>-miyu</i>	<i>hamyo</i>
3PL	<i>siha</i>	<i>[ma]⁷</i>	<i>uha/w/[uma]</i>	<i>-ñiha</i>	<i>siha</i>

The Sets A1 and A2 only differ in the third person singular and plural, and by the use of optional irrealis markers in some forms of Set A2. Both subsets of Set A are derived from the PMP genitive set (which is also directly continued in the possessive set), while Set B and the free set reflect the PMP nominative set.

The fact that the intransitive subject can be expressed by two pronoun sets makes it possible to distinguish clearly between syntactically transitive and intransitive constructions.

3.1.2 Intransitive verbs

Intransitive verbs can be divided into three classes depending on whether they take *<um>*, *ma(N)-* or \emptyset as singular realis affixes. Plural pronouns with singular verb forms have dual meaning. Intransitive affixes are given in Table 7.

Table 7: Intransitive verbal affixes in Chamorro

Realis singular	Irrealis singular	Realis plural	Irrealis plural
<i><um></i>	\emptyset	<i>maN-</i>	<i>faN-</i>
<i>ma(N)-</i>	<i>fa(N)-</i>	<i>manma(N)-</i>	<i>fanma(N)-</i>
\emptyset	\emptyset	<i>maN-</i>	<i>faN-</i>

The following pair illustrates the use of Set A and B pronouns in irrealis (6) and realis (7) mood:

- (6) Irrealis:
Para ta hanao.
 IRR 1PL.INC.A go
 'We will go.'

⁶ Actually, *u* is an irrealis marker that is obligatory in the third person and optional in the first person plural inclusive; *bai* is the irrealis marker for the first person (exclusive).

⁷ *ma-* is historically — and probably also synchronically — identical to the passive prefix *ma-*.

- (7) Realis:
Humanao hit.
 REAL.SG-go 1PL.INC.B
 'We went.'

Note that the irrealis forms that take Set A pronouns do not use *M*-affixes.

3.1.3 Transitive verbs

In the basic active construction, a transitive verb base is preceded by a Set A pronoun marking the agent, while the object is either a pronoun of Set B, as in (8) and (9), or a definite noun, as in (10). As illustrated in examples (8) and (9), active transitive forms do not change with mood, except for the slight difference in pronoun marking of the agent with Sets A1 (realis) and A2 (irrealis).

- (8) *Hu-li'e' gue'.*
 1SG.A-see 3SG.B
 'I saw him.'
- (9) *Para bai-hu-li'e' gue'.*
 IRR IRR-1SG.A-see 3SG.B
 'I will see him.'
- (10) *Hu-li'e' i lepblo.*
 1SG.A-see ART book
 'I saw the book.'

If the agent has to be extracted, as in cleft-, WH-clauses, relative clauses and equi-NP deletion, the agentive Set A pronoun is replaced by the infix $\langle um_2 \rangle$.⁸ The object is still represented by a Set B pronoun, as in (11).

- (11) *Hayi lum₂i'e' gue'?*
 who ACT-see 3SG.B
 'Who saw him?'

Topping has described $\langle um_2 \rangle$ as an actor focus affix. However, the latter sentence shows that the concept of focus — at least in the Philippine-type sense — is inapplicable to Chamorro, since transitive verbs with $\langle um_2 \rangle$ are Janus-faced in an odd way: to the left, the agent is highlighted, which would require AF in Philippine-type focus languages; to the right, the object is marked by an absolutive Set B pronoun, which corresponds to non-AF in focus languages. Avoiding the term focus, I will call forms with $\langle um_2 \rangle$ 'active participles', following Esser (1927) in his description of Mori.⁹ For intransitive verbs the participle is identical to the realis form of the verb.

⁸ This infix $\langle um_2 \rangle$ is homophonous with the intransitive infix $\langle um \rangle$ but not identical, since the latter only occurs with singular subjects, while the former is indifferent to number.

⁹ The term participle is employed, since $\langle um_2 \rangle$ replaces Set A person markers, which in the literature are often referred to as conjugation markers.

The affix *maN-/faN-* is employed to form an antipassive from transitive verbs,¹⁰ i.e. a form that is syntactically intransitive with the underlying agent as subject. The antipassive is mainly used if the object is indefinite; the latter can be left unexpressed or expressed by an unmarked noun, as in (12) and (13). Example (12) is in realis mood, and the agent/subject is represented by a Set B pronoun, while in (13), it is represented by a pronoun of Set A preceding the irrealis variant of the antipassive prefix. The intransitive nature of the antipassive can be seen in the use of either Set A or Set B pronouns for the agent/subject, depending on mood, in the same manner as in examples (6) and (7) above.

- (12) *Man-(t)aitai yo' lepblo.*
 ANTI.REAL-read 1SG.B book
 'I read a book.'
- (13) *Para bai-hu-fan-(t)aitai lepblo.*
 IRR IRR-1SG.A-ANTI.IRR-read book
 'I will read a book.'

With certain transitive verbs, the antipassive can also occur with definite objects, e.g. if the object is partially affected. The object then has an oblique or locative case marker, as in example (14). The active counterpart of (14) is sentence (15).

- (14) *Mam-(p)atek hao gi ga'lagu.*
 ANTI-kick 2SG.B LOC dog
 'You kicked at the dog.'
- (15) *Un-patek i ga'lagu.*
 2SG.A-kick ART dog
 'You kicked the dog.'

There are two passive affixes, *in>* and *ma-*, which are used in both realis and irrealis. Verbs with *in>* and *ma-* are syntactically intransitive, with the patient as subject. The intransitive nature of the passive is apparent in example (16), where the underlying patient is marked by a Set A pronoun, since it is the subject of an irrealis sentence. The agent, if present, is normally marked as oblique case, as in (17), although in cleft and similar constructions the agent can be marked as a possessor, as illustrated in (18).

- (16) *Ti un-hin>engge.*
 NEG 2SG.A-PASS-believe
 'You won't be believed.' (= 'He won't believe you.')
- (17) *Lini'e' si Maria as Pedro.*
 PASS-see PERS Maria OBL.PERS Pedro
 'Maria was seen by Pedro.'
- (18) *Hafa lini'e'-ña si Maria?*
 What PASS-see-3SG.POSS PERS Maria
 'What did Maria see?'

¹⁰ The infix *um₂* instead of *maN-* occurs with at least two verbs, namely *gimen* ('drink') and the suppletive *chocho* ('eat' – the corresponding active base is *kanno*).

- (19) *Ma-li'e' i palao'an.*
 PASS-see ART woman
 'The woman was seen.'

The *ma*-passive is used when the agent is unidentified, as in example (19), or third person plural. Otherwise, the choice between active and *in*-passive is dependent on rather complex discourse factors (Cooreman 1987). The transitive affixes are summarised in Table 8.

Table 8: Transitive verbal affixation in Chamorro

	Realis	Irrealis	Participle
Active	\emptyset	\emptyset	$\langle um_2 \rangle$
Antipassive (SG)	<i>maN-</i>	<i>faN-</i>	
<i>in</i> -passive	$\langle in \rangle$	(in)	
<i>ma</i> -passive	<i>ma-</i>	$(ma-)$	

The applicative affix *-i* (variant form: *-iyi*) has several functions, one of its main functions being the derivation of transitive stems from intransitive verbs and also sometimes from nouns. If suffixed to transitive verbs, it has benefactive function: the beneficiary then becomes the direct object, while the original direct object is put into oblique case. The suffix *-i* is not a focus affix since it can combine with all the above mentioned transitive affixes, giving *maN-i*, $\langle um_2 \rangle -i$, $\langle in \rangle -i$, and *ma-i*. Below I give examples for verbs suffixed with *-i*:

- hanagu-i* 'go to' (< *hanao* 'go')
apas-i 'pay' (< *apas* 'wage')
tugi-i 'write to/for' (< *tuge* 'write something').

Examples (20) and (21) illustrate the use of *-i* in a basic active clause and with the active participle affix $\langle um_2 \rangle$, respectively.

- (20) *Hu-tugi-i si Pedro ni katta.*
 1SG.A-write-APPL PERS Pedro OBL letter
 'I wrote the letter to Pedro.'
- (21) *Hu-konfotme kum₂uentus-i i ma'gas.*
 1SG.A-agree ACT-talk-APPL ART boss
 'I agree to talk to the boss.'

Transitive verbs can take the suffix *-(y)on* to form a stative verb 'capable of being X-ed'. (Occasionally, intransitive verbs can also take this suffix to express 'capable of X'.) Some examples include:

- atan-on* 'nice to look at' (< *atan* 'look at')
honggi-yon 'credible' (< *honggi* 'believe')
falagu-yon 'capable of running' (< *malagu* 'run').

The circumfix *faN-* *-(y)an* forms location nouns from verbs, for example:

- fañochoyan* 'eating place' (< *chocho* 'eat').

3.2 Historical derivation of the Chamorro verbal system

Table 9 gives an overview of how the PMP focus affixes of Table 3 are reflected in Chamorro:

Table 9

PMP	Chamorro		
AF:			
Past	* <i>minaN-</i>	(no reflex)	
Non-past	* <i>maN-</i>	> <i>maN-</i>	antipassive, realis mood
	* <i>um</i>	> <i>um₂</i>	active participle
Atemporal	* <i>paN-</i>	> <i>faN-</i>	antipassive, irrealis mood
Non-AF:			
Past	* <i>in</i>	> <i>in</i>	passive
	* <i>in</i> - <i>an</i>	> (<i>in</i> - <i>i</i>)	passive
Non-past	* <i>en</i>	> -(<i>y</i>) <i>on</i>	(derives stative verbs)
	* <i>paN-</i> - <i>an</i>	> <i>faN-</i> -(<i>y</i>) <i>an</i>	(derives nouns)
Atemporal	*∅	> ∅	active, realis/irrealis mood
	* <i>i</i>	> - <i>i</i>	active, realis/irrealis mood

The following developments from PMP to Chamorro should be noted:

- PMP non-AF has become the Chamorro active, while AF developed into an antipassive, with the notable exception of PMP **um*, which has turned into the active participle, combining features of AF and non-AF.
- In the antipassive, the past/non-past distinction is lost, the non-past becoming the general realis form; the atemporal has become irrealis mood.
- The non-AF atemporal has become the general tense form of the active, while its past form has become a passive; the non-AF non-past is retained residually in lexical derivations.
- Further, there are two innovations that are not seen in Table 9:
- The PMP LF atemporal suffix **i* has become the Chamorro applicative suffix -*i* which can combine with affixes that are derived from the PMP AF, *maN-*, *um₂*, see example (21).
- In the PMP atemporal, a pronominal agent with AF verbs is of the nominative set, while in non-AF it is of the genitive set, and both are fronted to preverbal position if the clause is opened by a negator (or any other preverbal modifier that requires the atemporal form of the verb). This is exemplified in sentences (22) in AF and (23) in non-AF from Waray-Waray (Central Philippines), which has retained the pattern assumed here for PMP.

- (22) *Waray pa ako kaon.*
 NEG yet 1SG.NOM eat
 'I haven't eaten yet.'

- (23) *Waray ko kaun-a.*
 NEG 1SG.GEN eat-ATEMP.UF
 'I haven't eaten [it].'

In contrast, sentences (24) and (25) show that in Chamorro, both in antipassive (< PMP AF) and in active voice (< PMP non-AF), the irrealis is preceded by Set A pronouns, which are derived from the PMP genitive set. Note that the verb forms which combine with these preposed pronouns employ *base affixes* (including \emptyset).

- (24) *Bai-hu-fa[n]-taitai.*
 IRR-1SG.A-ANTI.IRR-read
 'I will read (something).'
- (25) *Bai-hu-taitai i lepblo.*
 IRR-1SG.A-read ART book
 'I will read the book.'

Functionally, selection of active/antipassive in Chamorro corresponds to selection of non-AF vs. AF based on pragmatic criteria in Philippine-type focus languages. Significantly, the syntactic criteria for focus selection in 'focus' languages are *not* relevant for the selection of active/antipassive in Chamorro. This is apparent in the use of the active participle with $\langle um \rangle$. This is used in situations where Philippine-type languages have syntactically conditioned AF, as exemplified in the Tagalog sentence given above in example (5).

4 Grammatical sketch of Palauan

The following sketch mainly draws from two sources: Josephs (1975), which contains a host of sample sentences, although the analysis of data is inadequate at times; and Lemaréchal (1991), who has reinterpreted a good deal of the former's analysis in a much clearer way (see the appraisal by Josephs 1994). Additional information is taken from Pätzold (1968).

In order to identify PMP morphemes that have been retained in Palauan, note the following sound changes:

- unstressed vowels become ∂ (*e* in Palauan orthography) or \emptyset ;
- pre-stress **pa-* > *o-*, **pina-* > *ul(e)-*;
- Loss of **R* in clusters, as in **maR-* > *me-*, **paR-* > *o-*;
- **n* > *l*, as in **in* > $\langle i \rangle l$.
- The infix $\langle m \rangle$ (< **um*) is often realised as a back semivowel or as backing of the stem vowel.

4.1 Palauan pronoun sets and verbal morphosyntax

Palauan has a nominative pronominal agreement system. There are five sets of pronouns or pronominal affixes, which are given in Table 10.

Table 10: Palauan pronoun sets

	Free	NOM I	NOM II	OBJ	POSS
1SG	<i>ngak</i>	<i>ak</i>	<i>k-</i>	<i>-ak</i>	<i>-k</i>
2SG	<i>kau</i>	<i>ke</i>	<i>òm-</i>	<i>-au</i>	<i>-m</i>
3SG	<i>ngii</i>	<i>ng</i>	<i>l-</i>	<i>-ii</i>	<i>-l</i>
1PL.IN	<i>kid</i>	<i>kede</i>	<i>d-</i>	<i>-id</i>	<i>-d</i>
1PL.EX	<i>kam</i>	<i>aki</i>	<i>kim-</i>	<i>-am</i>	<i>-(m)am</i>
2PL	<i>kemiu</i>	<i>kom</i>	<i>òm-</i>	<i>-emiu</i>	<i>-(m)iu</i>
3PL.HUM	<i>tir</i>	<i>te</i>	<i>l-</i>	<i>-terir</i>	<i>-rir</i>
3PL.NON-HUM				\emptyset	

The free, NOM I and object sets are clearly derived from the PMP nominative set, while the NOM II set is related to Chamorro Set A, and together with the possessive set is derived from the PMP genitive set.

The choice between the two nominative pronoun sets depends largely on syntactic criteria: if the verb is clause-initial, the first set is used. An exception to this are imperative sentences, where the second set is employed. If the verb is preceded by a subject (= S, A) NP (or a part of it), then there is no nominative pronoun; if it is preceded by any other constituent (object, adverbial etc.), the second set is used. The second set is also obligatory after certain conjunctions.

Many verb forms alternate depending on whether they are preceded by a NOM I pronoun (or a nominative NP) or by a NOM II pronoun. For convenience, I will call the first verb form indicative and the latter subjunctive. Subjunctive forms never occur without a preceding NOM II pronoun; they are also never found in sentence initial position, except in imperative sentences.

The following examples illustrate the correlation between nominative pronoun, verb affix and word order. Example (26) is a verb initial sentence with a NOM I pronoun preceding an indicative verb. Examples (27) and (28) are rearranged versions of (26): in (27), the subject precedes the verb, in which case there is no nominative pronoun, while in (28), the object precedes the verb, which therefore has to be in its subjunctive form preceded by a NOM II pronoun.

(26) *Ng-meng-(ch)uiu er a hong a Droteo.*
 3SGI-ANTI-read OBL ART book ART Droteo
 'Droteo is reading the book.'

(27) *A Droteo a meng-(ch)uiu er a hong.*
 ART Droteo ART ANTI-read OBL ART book
 'Droteo is reading the book.'

(28) *A hong l-ong-(ch)uiu er ngii a Droteo.*
 ART book 3SGII-ANTI-read OBL 3SG.FREE ART Droteo
 'As for the book, Droteo is reading it.'

Example (29) illustrates the exceptional sentence-initial position of the subjunctive with a Nom II pronoun in an imperative sentence. The non-imperative counterpart of (29) is (30).

- (29) *D-o-rael!*
1PL.INCII-INTR-go
'Let's go!'
- (30) *Kede-me-rael.*
1PL.INCI-INTR-go
'We go.'

All transitive verbs and many intransitive verbs have two tense forms, past and non-past, and distinct forms for indicative and subjunctive. Transitive verbs can occur in active voice, antipassive voice and three forms of passive voice. Transitive affixes of Palauan are given in Table 11.

Table 11: Overview of basic verbal affixes in Palauan

	Indicative		Subjunctive	
	Non-Past	Past	Non-Past	Past
Intransitive	<i>⟨m⟩</i> <i>me-</i>	<i>⟨il⟩</i> <i>mil-</i>	<i>∅</i> <i>o-</i>	<i>⟨il⟩</i> <i>ul-</i>
Transitive				
Active	<i>⟨m⟩₂</i>	<i>⟨il⟩₂</i>	<i>∅</i>	<i>⟨i⟩l₂</i>
Antipassive	<i>meN-</i>	<i>mileN-</i>	<i>oN-</i>	<i>uleN-</i>
Passive	<i>me-</i>	<i>mil-</i>	<i>(me-)</i>	<i>(mil-)</i>
Resultative	<i>⟨l⟩, ⟨l⟩-el</i>			
Gerundive	<i>-el, -all</i>			

In the active, an object pronoun is obligatory and agrees with an overt object NP, as exemplified in sentence (31). By definition, a verb in the antipassive cannot take an object suffix; if the object is a pronoun or a definite NP, it takes the oblique marker *er*.¹¹ Examples (32) and (33) are antipassive sentences with an indefinite and a definite object respectively. Both active and antipassive have distinct forms for past and non-past, and indicative and subjunctive.

- (31) *Ak-kilisii* *a* *kiokl.*
Ak-⟨il⟩₂-kios-ii *kios-l*
1SGI-ACT-dig-3SG.OBJ ART dig-GER (=hole)
'I (completely) dug the hole.'
- (32) *Ak-milengiis* *a* *kiokl.*
Ak-mileN-kios
1SGI-ANTI.PST-dig ART hole
'I was digging holes.'

¹¹ Josephs describes the difference between what is called active and antipassive here as an aspectual distinction between perfective and imperfective aspect. My interpretation follows the analysis of Lemaréchal, although with a different terminology.

- (33) *Ak-milengiis er a kiokl.*
Ak-mileN-kios
 1SGI-ANTI.PST-dig OBL ART hole
 'I was digging the hole.'

There are three passive forms: a verbal passive (*me-*) with two tense forms, but not distinguishing between indicative and subjunctive (34); and the resultative (35) and gerundive (36) passives, which are better regarded as derivations outside of the transitive voice paradigm, as they are often employed as nouns. All passives generally do not allow the explicit occurrence of the agent.¹²

- (34) *A blai a mil-seseb.*
 ART house ART PASS.PST-burn
 'The house was burnt.'
- (35) *A ulaol a ngɔatech.*
 ART floor ART RES-clean
 'The floor is cleaned.'
- (36) *A ulaol a ngetach-el.*
 ART floor ART clean-GER
 'The floor has to be cleaned.'

4.2 Historical derivation of the Palauan verbal system

Table 12 gives an overview of how the PMP focus affixes are reflected in Palauan:

Table 12: PMP focus affixes and their Palauan reflexes

PMP	Palauan			
AF:				
Past	<i>*minaN-</i>	>	<i>mileN-</i>	antipassive, past indicative
	<i>*amin></i>	>	<i>il₂</i>	active, past indicative
Non-past	<i>*maN-</i>	>	<i>meN-</i>	antipassive, non-past indicative
	<i>*um></i>	>	<i>m₂</i>	active, non-past indicative
Atemporal	<i>*paN-</i>	>	<i>oN-</i>	antipassive, non-past subjunctive
Non-AF:				
Past	<i>*in></i>	>	<i>⟨i⟩l₂</i>	active, past subjunctive
	<i>*in></i>	>	<i>ɔ</i>	(derives resultative)
	<i>*in> -an</i>	>	<i>ɔ -el</i>	(derives resultative)
	<i>*pinaN- (-an),</i> <i>*(i)pinaN-</i>	>	<i>uleN-</i>	antipassive, past subjunctive?
Non-past	<i>*-en, *-an</i>	>	<i>-el</i>	(derives gerundive)
Atemporal	<i>*∅</i>	>	<i>∅</i>	active, non-past subjunctive
	<i>*-i</i>		(no reflex)	

¹² Josephs gives 'awkward' examples of the verbal passive with an agent carrying the oblique marker 'er' which he suspects to be based on an English model.

The following developments from PMP to Palauan should be noted:

- Palauan has nominative agreement, with nominative pronouns always prefixed to the verb. The PMP nominative pronouns in post-verbal position are only preserved with object marking function.
- As in Chamorro, PMP non-AF has become active voice in Palauan, while AF developed into an antipassive, although there is 'cross-over', i.e. some active forms have an AF origin and some antipassive forms are derived from a non-AF source. Again, we find the unusual use of m_2/ai_2 with a following object suffix derived from the nominative set, combining features of AF and non-AF.
- The PMP tense distinction is preserved in Palauan, the only exception being the non-AF non-past, which has become a derivational affix. Its function has been taken over by the non-AF atemporal.
- Among the PMP non-actor focuses, only UF is preserved in the transitive paradigm of Palauan. The past subjunctive forms of antipassive voice are probably derived from LF and IF, since its function corresponds to the function of PMP 'circumstantial' non-AF focus forms, although there is no trace of the characteristic focus affixes.
- The indicative uses affixes derived from PMP *M*-affixes, while subjunctive forms are derived from base affixes. The Nom II pronoun set is equivalent to Set A in Chamorro, and like its Chamorro counterpart, is incompatible with *M*-affixes.

What has been said above about the functional correspondence between non-AF vs. AF in PMP and active vs. antipassive in Chamorro, also holds for Palauan.

4.3 The *-akl* suffix

Unlike in Chamorro, there is no productive applicative suffix. However, as noted by Pätzold (1968), some verbs seem to contain a fossilised affix *-akl*. Combined with the gerundive suffix *-(e)l*, this gives *-ekill* pointing to a synchronic deep form */-akil/* (< **-akin*). There are a few pairs of verbs where the root occurs both with and without */-akil/* (*/-okil/* in one instance):

<i>techolb</i>	'wash, baptize'
<i>techelbaki</i>	'dive into'
<i>toir, tir</i>	'chase'
<i>tiraki</i>	'follow, obey'
<i>iub, ibng</i>	'sneak out, avoid'
<i>ibngokil</i>	'sneak out, avoid'
<i>-renges</i>	'hear'
<i>beko/de/rengesaki</i>	'having sharp hearing' (the prefix <i>beko-</i> , <i>beke-</i> means 'good at doing something'; the additional <i>de-</i> cannot be explained)

It is very hard to extrapolate the original function of */-akil/* from these few examples. There are more examples where however the semantic distance of the pair is too great to exclude mere coincidence:

<i>dibechakl</i>	'cross'
<i>dibech</i>	'invent'

Most verbs that seem to bear a suffix /-akil/ do not occur without it, i.e. there is no way of showing that it is not an integral part of the stem:

<i>bedechakl</i>	'throw down, drop'
<i>ngeriakl</i>	'move forward'
<i>techemakl</i>	'stuff'
<i>techerakl</i>	'pick up with a hook'

5 Chamorro and Palauan innovations and their occurrence in other Malayo-Polynesian languages

If the innovations described in §3.2 and §4.2 are compared, it can be seen that Chamorro and Palauan share the following innovations if compared with PMP:

1. The pronoun set A or NOM II (derived from, but distinct from the possessive set), which occurs before verbs with base affixes;
2. the syntactic and semantic differentiation of $\langle um_2 \rangle / \langle m_2 \rangle$ and *maN-/meN-*, with * $\langle um \rangle$ taking over functions associated with non-AF.

Chamorro also has the following innovations not found in Palauan:

3. the circumfixes *maN-* *-i* and $\langle um_2 \rangle$ *-i*, combining PMP AF and non-AF (atemporal LF) affixes;
4. the loss of the tense (aspect) distinction involving **in*. The infix *in* has become a passive marker.

Apart from being a nominative language, Palauan seems to reflect one important innovation not found in Chamorro:

5. the suffix /-akil/.

5.1 The Set A pronouns

This pronoun Set A is found in many other Malayo-Polynesian languages: Sumatran languages (including languages of the Barrier Islands), Malay, Embaloh, Old Javanese, Sulawesi languages (excluding the focus languages of the North), and CEMP languages. Many of these languages have defective sets, e.g. Batak, Malay, Kaili. Set A is not found in the Philippines, Northern Sulawesi and Borneo (except for Malayic and Tamanic), nor in Sundanese and Balinese.

The most innovative feature of Set A pronouns is that they are placed before the verb, i.e. they are proclitics or — in most languages — prefixes with a fixed position, unlike the nominative or genitive sets in PMP or PAn, which are enclitics that are usually subject to raising.

Three types of languages can be distinguished according to the function of the Set A pronouns:

- (I) Set A pronouns have strictly agentive function, i.e. they are only used as agent markers with transitive verbs, as in Batak, Malay, Lampung, Embaloh, Kaili, Saluan, Totoli, Mandar.
- (II) Set A pronouns occur in a split-ergative system, i.e. generally marking the agent with transitive verbs, but also marking the subject with intransitive verbs¹³ in certain constructions, as in Chamorro, Buginese, Mori, Pamona, Nias.
- (III) Set A pronouns are part of a nominative agreement system, not discriminating between transitive and intransitive verbs, as in Palauan, Muna-Buton languages, and most CEMP languages.

In languages of type I, Set A pronouns co-occur with transitive verb forms derived from the PMP non-AF atemporal. This is also the case in languages of types II and III; here, additionally, intransitive verbs (including derived forms) taking Set A pronouns occur in a form derived from the PMP AF or intransitive atemporal. Usually, Set A pronouns are not compatible with M-affixes.¹⁴

Examples of Set A pronouns in languages of type I:

Karo Batak (Woollams 1996):

- (37) *Ku-guas takal-na.*
 1SG-thump head-3
 'I clobbered him on the head.'

Mandar:¹⁵

- (38) *U-issam-mi.*
 1SG-know-ASP.3
 'I already know.'

For languages of the types II and III, I will restrict myself to giving examples of their occurrence with intransitive/AF verb forms, especially when extended with **paR-/paN-*, to illustrate that these are a continuation of the PMP atemporal. Nias, Bugis and Bungku are western MP languages, while Kambera and Buli represent CMP and EMP, respectively.

Chamorro:

- (39) *Para bai-hu-fa-lagu.*
 IRR IRR-1SG-fa-run.
 'I will run.' (fa- < *paR, cf. ma-lagu)

Palauan:

- (40) *D-o-rael!*
 1PL.INCII-o-go
 'Let's go!' (o- < *paR-, cf. me-rael)

¹³ Including derived intransitive forms of transitive verbs.

¹⁴ Exceptions to this are found in languages which have completely lost the AF atemporal in favour of M-forms, even in imperative function, e.g. in Toraja or Banggai.

¹⁵ Bugis, Mandar, Bungku and Pitu Ulunna Salu data are from my own fieldnotes.

Nias (Sundermann 1913):

- (41) *Mi-o-fanö!*
 2PL-*o-go*
 'Go away (PL)!' (o- < *(p)aR-, cf. *mo-fanö*)

Bugis:

- (42) *Aja' mu-ac-cue:!*
 Don't 2-aC-follow!
 'Don't follow!' (aC- < *(p)aR-, cf. *mac-cue:*)

Bungku:

- (43) *Nahina-po ku-pong-kaa.*
 not-yet 1SG-poN-eat
 'I haven't eaten yet.' (poN- < *paN-, cf. *mongkaa*)

Kambera (Klamer 1994):

- (44) *Nggiki hi u-pa-taru?*
 why CONJ 2SG-pa-watch
 'Why are you watching?' (pa- < *paR-)

Buli (Maan 1951):

- (45) *...fare d-fa-pun-pun.*
 CONJ 3PL-fa-RED-hit
 '...and they hit each other.' (fa- < *paR-)

This use of Set A pronouns with AF/intransitive verb forms is not found in languages of type I. But it has to be noted that in these languages, verb forms derived from the PMP AF or intransitive atemporal are only used as imperatives (Mandar, Totoli, Saluan), or are not reflected at all (Batak, Malay). Many languages of this type have a defective set of preposed pronouns (Totoli, Saluan, Kaili, Embaloh, Batak).

The pronoun Set A could be either taken as a common innovation that occurred in a meso-language from which all above-mentioned languages have derived, or as an independent parallel innovation. The latter view is proposed by Himmelmann (1996) and Wolff (1996), who regard the defective sets as incipient stages to a full paradigm. In contrast, van den Berg (1996) reconstructs a full set for the parent language of at least some languages discussed here (Proto Celebic), but only for transitive verb forms: the extended use of set A pronouns with intransitive or AF verbs in languages of types II and III (both types being represented in his Celebic group) he regards as a later development.

Here I propose that Set A is a common innovation of all the languages in which it occurs, and that it originally was used with *both* non-AF and AF (and intransitive) atemporal verb forms, although the use with AF forms was more limited than with non-AF forms. Languages of types II and III offer ample evidence of this, especially since they include the isolated Chamorro and Palauan languages. Only in a later, and in most cases independent parallel development, have languages belonging to type I restricted the use of AF and intransitive atemporal to functions where there is usually no person marking (e.g. imperative), or have lost this atemporal completely, leading to the restriction of Set A pronouns to transitive forms derived from the PMP non-AF.

5.2 Antipassive *maN*- versus active participle *um₂*

At the end of §2 I discussed the pragmatics of AF in focus languages. Where pragmatic criteria require AF in PMP, Chamorro and Palauan have the antipassive, formed by a reflex of PMP **maN-/paN-*, while in those cases where AF is conditioned by syntactic criteria, a continuation of PMP **um* is used (here marked *um₂*). The objects of verbs carrying this *um₂*, especially pronominal objects, have a case form derived from the PMP nominative.

The syntactic and pragmatic differentiation of **maN-* and **um₂* and the particular object marking after **um₂* are a distinctive innovation that is found in only a few areas, which however have a widely scattered distribution: the South Sulawesi,¹⁶ Bungku-Tolaki and Muna-Buton groups of Sulawesi, older Toba Batak, Nias and Enggano, and probably Old Javanese and Old Balinese.¹⁷ In the following examples, (46) and (47) are equivalent to the Chamorro sentence (11), while the Nias sentences (48) and (49) parallel examples (31) and (32) from Palauan.

Pitu Ulunna Salu (South Sulawesi):

- (46) *Menna mu-hambi-ko?*
 who *mu*-hit-2SG
 'Who hit you?' (cf. *ku-hambi-ko* 'I hit you.')

Bungku:

- (47) *Inai 'umala-o?*
 who *um*-take-3SG
 'Who took it?' (cf. *ku-ala-o* 'I took it.')

Nias (southern dialect, Sundermann 1913):

- (48) *Gu-tum>agu(-ya).*
 1SG.IRR-*um*-sew(-3SG)
 'I will sew it.' (definite object)
- (49) *Gu-man-(t)agu.*
 1SG.IRR-*maN*-sew
 'I will sew.' (indefinite object)

Note that the above mentioned languages (with the exception of Old Balinese) also display innovation 1. However, other languages of that group do not have a reflex of **um* in the transitive paradigm, but generally use a reflex of **maN-* (occasionally **maR-*). In some cases, this can be shown to be a later development:

- modern Javanese generally has *N-* (< **maN-*) where Old Javanese used *(m)aN-*, *(m)a-* (< **maR-*) and *um* (Kern 1918–20; Zoetmulder & Poedjawijatna 1961);
- in modern spoken Toba Batak, *um* is replaced by *maN-* (van der Tuuk 1971);
- Embaloh, shown to be closely related to Buginese, generally uses *maN-* (Adelaar 1994, 1995), with *m-* (< **um*) only occurring before vowels.

¹⁶ In Proto South Sulawesi, **um₂* is reflected as a prefix: **um-/mu-*.

¹⁷ In some of these languages, a reflex of **maR-* occurs next to **maN-* in an antipassive function, the choice of which is lexically determined.

Based on these three examples, it is probable that the **ma*[*N/R*]- vs. **um* contrast was also lost in many other languages in favour of **maN*-.

5.3 The affix combinations *maN*- *-i* and *um*₂ *-i*

This innovation can be described as a symmetrisation of the focus system. As described above, the PMP focus system is asymmetrical since two or three non-actor focuses are matched by one actor focus. In Chamorro, the non-AF suffix *-i* can co-occur with the AF affixes *um* and *maN*-. Thus, Chamorro *-i* has become an 'applicative' affix.

Innovation 3 is found in all languages that have the first innovation, provided they have a reflex of the affixes **maN*- and **um* (which is not the case for most CEMP languages) and the atemporal non-AF suffixes (which is not the case in Palauan). It is further found in Balinese, Madurese and Sundanese. The symmetrisation is not restricted to former LF; the AF affixes can also combine with the PMP IF atemporal **-an*, giving **maN*- *-an* and *um*₂ *-an* (e.g. Totoli, South Sulawesi, Selako). Note that in all languages (except Sundanese), this innovation involves the atemporal form of PMP non-AF.

5.4 Loss of past/non-past tense distinction

This innovation is found in almost all languages in which the former three innovations have taken place. An exception to this are Palauan and a few languages in a small stretch of Central Sulawesi, namely Saluan, Balantak, Kaili, and the Tomini-Tolitoli languages. These languages have retained the original PMP tense distinction.

In languages that have lost the past/non-past distinction, it is the non-past form that has been lost in AF. In non-AF, some of these languages retain all tense forms, but the atemporal has become the general form, while past and non-past forms have been relegated to specialised meanings. This is the case for Chamorro (see §3.2), Toba Batak and Buginese.

5.5 Palauan /-akil/

This fossilised suffix is probably related to the widespread applicative suffix **-aken*, although the vowel in the underlying final syllable presents a problem, since Palauan /i/ is not a regular reflex of PAn **e*. However, Sirk (1996) pointed out that there is much variation in reflexes of **-aken*, and the Palauan form falls well within this variation.

Reflexes of **-aken* occur in most languages that have at least one of the above innovations, and functionally, it has taken over the role of the atemporal IF suffix **-an*. It is not found in any language that clearly did not participate in innovation 3, i.e. its introduction must have post-dated the symmetrisation of the focus system. From its distribution, **-aken* can not be reconstructed for PMP as a suffix, but it is likely that it represents the capture of an oblique-case marker *(*a*)*ken* common in the languages of the Philippines, which also survives as the Malay preposition *akan*.

In some language groups on Sulawesi, there is evidence that this capture occurred gradually: here, in some functions, **-aken* already occurs as an inseparable affix, while in others, **aken* still betrays its originally prepositional nature (Mead 1998). The replacement of inherited IF atemporal **-an* by *(*-*)*aken* also must have occurred gradually, with both

morphemes occurring side by side at some stage having related but distinct functions. This is witnessed by the Malayic subgroup, where some languages have a reflex of **-an* (Kendayan, Selako, some Minangkabau dialects), while others have **-aken* (Malay, Serawai, Banjar). In Muna, both reflexes of **-an* and **-aken* are found: **-an* combines with second and third person pronouns, while a continuation of **-aken* is used with nouns.

Functionally, the relation of the Palauan suffix /-akil/ with **-aken* is unclear, since there is little agreement between them. The major functions of **-aken* are: causative (competing with the inherited causative prefixes **pa-* and **paka-*), benefactive, and instrumental, none of which can be assigned to the occurrence of /-akil/ in Palauan. One has to bear in mind however that the Palauan suffix is fossilised, and a closer inspection of its occurrences might reveal a connection with the functions of **-aken* listed above. At the present stage, Palauan is at best considered an imperfect witness for the emergence of the suffix **-aken*. It is even possible that Palauan words with /-akil/ reflect very early loans from an Oceanic language (cf. Proto Oceanic **-akini*), such as Yapese.

6 The Nuclear Malayo-Polynesian subgroup

6.1 Reconstruction of Proto NMP

The innovations discussed in §5 are not independent from each other, especially the innovations proposed in §5.1, §5.2 and §5.3, referred to as innovations 1, 2, 3 respectively in the following discussion:

- All languages displaying innovation 2 also have innovation 1 or 3. In turn, all languages that participate in innovation 1 or 3 and which still make use of the infix ⟨*um*⟩ in the transitive paradigm, also share innovation 2.¹⁸
- No language that shares innovation 1 has evidence that it did not participate in innovation 3, i.e. if a language has Set A pronouns, and has preserved both the AF affixes **maN-* or ⟨*um*⟩ and the non-AF suffixes **-i* or **-an*, it will also make use of the innovative affix combination(s). However, a few languages that display innovation 3 do not share innovation 1, such as Balinese, Sundanese and Madurese.¹⁹
- Innovations 1 and 3 are also functionally interdependent: the use of the atemporal non-AF suffixes in innovation 3 presupposes that in non-AF, atemporal forms are more frequent than the past and non-past forms, because the atemporal has taken over some of the functions of the past and non-past tenses. This is certainly the case in languages which share innovation 1.²⁰

This suggests that innovations 1, 2 and 3 occurred together in a common meso-language which was a daughter language of Proto Malayo-Polynesian and from which all languages

¹⁸ An exception is Acehnese, which has proclitic pronouns that appear to be related to the Set A pronouns, but uses ⟨*eum*⟩ (from **um*) in *de*-transitivising function, which is quite the opposite of innovation 2.

¹⁹ In Sundanese and Madurese, this might be due to the fact that pronouns have largely been relexified with nouns. This argument however does not hold for Balinese, where we still find the original PMP pronouns in Old Balinese and modern Bali Aga dialects, without any evidence for Set A pronouns.

²⁰ This argument however cannot be applied to Balinese and Madurese.

mentioned above derived. I will call this subgroup *Nuclear Malayo-Polynesian* (NMP), as it contains both Malay and the Polynesian family, and the meso-language *Proto Nuclear Malayo-Polynesian* (PNMP). (See Sirk 1978, 1996 for earlier attempts to use innovations 1, 3, and 5 for reconstruction and subgrouping.)

Innovation 4 (§5.4) is found in almost all NMP languages, except for those mentioned in §5.4. In one case, two closely related languages are separated by this innovation, namely Kaili (which has retained the tense distinction) and Pamona (where the distinction is lost). This shows that loss of tense is probably a drift-like phenomenon in the NMP subgroup.

Innovation 5 (§5.5) also postdates PNMP, as it is only found in languages that have innovations 1, 2 and 3, although not in all of them. As illustrated in §5.5, this innovation involved the gradual capture of the preposition **aken*, eventually replacing the atemporal IF suffix **-an*. We can assume that like innovation 4, this capture is the result of drift within the NMP subgroup.

The reconstruction of the PNMP system for transitive verbs given in Table 13 accommodates the evidence given by its daughter languages.

Table 13: The PNMP verb system

	Past	Non-past	Atemporal ²¹
Actor focus	<i>minaR-</i> <i>minaN-</i> (<i>umin</i>)	<i>maR-</i> <i>maN-</i> (<i>um</i>)	<i>paR-</i> <i>paN-</i> (\emptyset)
Actor participle	(<i>umin</i>)	(<i>um</i>)	
Patient focus	(<i>in</i>)		
Gerundive		-en	
Actor focus	<i>minaR- -i</i> <i>minaN- -i</i> (<i>umin</i>) -i	<i>maR- -i</i> <i>maN- -i</i> (<i>um</i>) -i	<i>paR- -i</i> <i>paN- -i</i> (-i)
Actor participle	(<i>umin</i>) -i	(<i>um</i>) -i	
Patient focus	(<i>in</i>) -an	-i	-i
Gerundive		-an	
Actor focus	<i>minaR- -an</i> <i>minaN- -an</i> (<i>umin</i>) -an	<i>maR- -an</i> <i>maN- -an</i> (<i>um</i>) -an	<i>paR- -an</i> <i>paN- -an</i> (-an)
Actor participle	(<i>umin</i>) -an	(<i>um</i>) -an	
Patient focus	(<i>i-</i>) (<i>in</i>)	-an	-an
Gerundive		<i>i-</i>	

This system differs from the PMP system in Table 3 in the following aspects:

- For each patient focus (\emptyset , -i, -an) there is a corresponding actor focus form; this is the result of innovation 3.

²¹ There was still a projective in PNMP, since it is found e.g. in Old Javanese; it is however not relevant for the discussion here.

- The patient focus non-past is formally identical to the atemporal; the PMP non-past has acquired a gerundive meaning.
- The active participle belongs to patient focus, since it is followed by a patient in nominative case; to the left however, it highlights the agent (innovation 2).
- Pronominal case marking has been reshaped in the atemporal: in AF, a fronted pronoun in pivot function in PNMP is of the genitive set, not of the nominative set as in PMP, while in non-AF, only the genitive pronoun (agent) is fronted, while the nominative pronoun (object) is not fronted, schematically:

	PMP:	PNMP:
AF:	PRV <i>aku</i> V	PRV <i>ku</i> V
UF:	PRV <i>ku sia</i> V	PRV <i>ku</i> V <i>sia</i>
LF:	PRV <i>ku sia</i> V- <i>i</i>	PRV <i>ku</i> V- <i>i</i> <i>sia</i>
IF:	PRV <i>ku sia</i> V- <i>an</i>	PRV <i>ku</i> V- <i>an</i> <i>sia</i>

Thus, in clauses with a verb in atemporal aspect, a pronominal agent is always of the genitive set, whether in AF or non-AF. Before the breakup of PNMP, these pronouns must have shifted their position from enclitic on the preverb to proclitic on the verb; because of their new position, they evolved into a set of their own, distinct from the genitive set, in all NMP daughter languages (innovation 1).

The latter innovation was probably the starting point for the development of the remaining innovations. They probably took place in the following order (see Wolff 1996 and Sirk 1996 explaining the emergence of innovations 1 and 3 in a similar way):

- With the development of pronoun set A, AF and UF atemporal become formally identical (although still differing with regard to the case marking of accompanying NPs). This leads to further symmetrisations of the focus system:
- In AF, **um* (or *M*-affixes) and Set A pronouns are in complementary distribution: since **um* occurs if the agent is in fronted position, this use is extended to non-AF, probably first in UF (*Q* refers to the object NP in nominative case):

	Atemporal	Agent fronted
AF:	<i>ku</i> -V	A <i>um</i> -V
UF:	<i>ku</i> -V <i>Q</i>	A <i>um</i> -V <i>Q</i> .

Thus, in UF the patient is in nominative case even if A is in preverbal focus position and the verb takes **um*, which is a major departure from the original PMP system where patients (especially pronominal patients) of verbs with **um* are in oblique case (innovation 2).

- Later this use is extended to the other non-AFs:

	Atemporal	Agent fronted
LF:	<i>ku</i> -V- <i>i</i> <i>Q</i>	A <i>um</i> -V- <i>i</i> <i>Q</i> .
IF:	<i>ku</i> -V- <i>an</i> <i>Q</i>	A <i>um</i> -V- <i>an</i> <i>Q</i> .

Eventually, **i* and **an* also combine with **maN*- and **maR*-.²² Thus **i* and **an* become applicative suffixes, independent of focus. This co-occurrence of AF and

²² This is a post-PNMP development, as in some NMP languages there is a constraint on the use of **i*-*an* with **maN/RJ*-, e.g. in Buginese and Chamorro.

non-AF affixes leads to symmetrisation of the PMP focus system, where originally one AF contrasts with three non-AFs (innovation 3).

Two widespread phenomena are post-PNMP drifts:

- (d) The loss of the PMP tense distinction: former AF past forms are lost completely, while the former non-AF past tense forms **in* and **in* -*an* acquire passive function without temporal connotation (innovation 4).
- (e) The emergence of the applicative suffix **-aken* (innovation 5).

6.2 From PNMP to Chamorro

Chamorro has retained the PNMP system quite faithfully. However, some of the development described in §3.2 are post-PNMP innovations:

- (i) PMP/PNMP Instrument Focus is lost completely;
- (ii) **in* -*an* has been replaced by *in* -*i*;
- (iii) the tense distinction between past and non-past is lost: with intransitive verbs and in AF, only the non-past form survived as the general form in realis mood, while in non-AF, the non-past form (= the PMP atemporal) becomes the general form for both moods, and the past form becomes a passive;
- (iv) the PMP stative passive **ma* -/*ka* - is generalised as *ma* -.
- (v) the PNMP atemporal is used for irrealis mood.

These developments, especially (ii) and (iii), have also occurred in many other NMP languages, but most probably as a result of drift (loss of tense, see §6.1) or paradigmatic leveling (emergence of *in* -*i*).²³

Incidentally, the morphosyntax of Chamorro as described in §3 is almost identical to that found in Mori (Esser 1927; Barsel 1994).²⁴

6.3 From PNMP to Palauan

The case of Palauan is a little more complicated. The most significant innovation is the development of nominative agreement. Although this probably happened under the influence of Yapese and western Trukic isolects, with which Palauan also shares other areal features, the constructions that are involved already existed in PNMP.

Indicative clauses with Nom I pronouns are derived from PNMP clauses with preposed subject. For that reason, the indicative active contains the PNMP active participle infix **um*₂. Intransitive constructions with a subject pronoun following the verb fell into disuse. Cf. the following examples with the etyma **maR-zalan* 'to walk' and **tanem* 'to plant':

²³ In eastern Central Sulawesi, Saluan has retained *in* -*an*, while the related Balantak has innovative *ni* -*i*.

²⁴ The similarity is a formal one: these two languages differ quite strongly in the pragmatics of the use of the passive and of the atemporal forms of intransitive verbs.

PNMP	Pre-Palauan, with fronting of subject/agent	Palauan
* <i>MaR-zalan akuf.</i> (INTR)	* <i>Aku maR-zalan.</i>	<i>Ak-me-rael.</i> (INTR)
* <i>MaN-(t)anem akuf.</i> (AF)	* <i>Aku maN-(t)anem.</i>	<i>Ak-mel-(d)alem.</i> (ANTI)
* <i>Ku-tanem iaf.</i> (non-AF)	* <i>Aku t_{um}anem ia.</i>	<i>Ak-d_olem-ii.</i> (ACT)

† *aku* and *ia* represent PMP nominative pronouns.

More complex is the origin of the subjunctive in Palauan. As has been illustrated in Table 12, it is partially derived from the original atemporal, partially from non-AF forms. The original atemporal is retained in imperative and negative clauses, and after certain conjunctions; all these only use non-past subjunctive forms, as illustrated in (50) and (51).

(50) *D-o-rael!* (*< *ia-paR-Zalan!*)
1PL.INCII-SUB-go
'Let's go!'

(51) *A le-me a chull...*
if 3SGII-come ART rain
'If it rains...' (lit. 'If the rain comes...')

In the case of fronting of non-subject constituents, the subjunctive is derived from non-AF forms of PMP, which can be seen from the fact that there is a past subjunctive form. Unlike in other NMP languages, *all* genitive pronouns occurring with non-AF verbs have been fronted, not only with atemporal forms, but also with past forms, as in (52).

(52) *A ngikel a le-k_{il}a a bilis.*
ART fish ART 3SGII-ACT.PST-eat ART dog
'The fish were eaten up by the dog'

Here, *le-k_{il}a* is derived from PMP **k_{in}an-nia*, with raising of the genitive pronoun. This is not found in other NMP languages, where the genitive pronoun has become fixed in the position *following* verbs in non-AF past tense (or in the passive that has developed from it).²⁵

As in Chamorro, the PMP stative passive **ma-/ka-* is preserved in Palauan as the verbal passive prefix *me-*, without, however, retaining the atemporal form **ka-*.

6.4 Scope and position of Nuclear Malayo-Polynesian within the Malayo-Polynesian family

If one takes the innovations discussed in §6.1 as diagnostic evidence, the NMP subgroup includes the languages of the CEMP group, Chamorro and Palauan, and most WMP languages of Malaysia and Indonesia. Not included in the NMP group are the following WMP languages: the languages of the Philippines, the three Northern Sulawesi groups

²⁵ Compare the following examples from NMP languages: Kaili *ni-kande-ku*, Saluan *k_{in}aan-ku* 'I ate it', Nias *ni-rongo-mi* 'what you (PL) heard', as opposed to *ku-kande*, *ku-kaan* 'I will eat it' and *mi-rongo* 'you hear'.

(Gorontalo-Mongondic, Minahasan, Sangiric), the Sama-Bajau languages, Malagasy, and all languages of Borneo with the exception of the Malayic and Tamanic groups.

Blust (1999) has pointed out that his Western group of Malayo-Polynesian (WMP) is not to be understood as a subgroup tied together by exclusively shared innovations, but just as an umbrella term for all MP languages not included in the Central-Eastern subgroup. Here, I have shown that indeed this WMP group has to be broken up by adding a node to his tree, as shown in Figure 2.

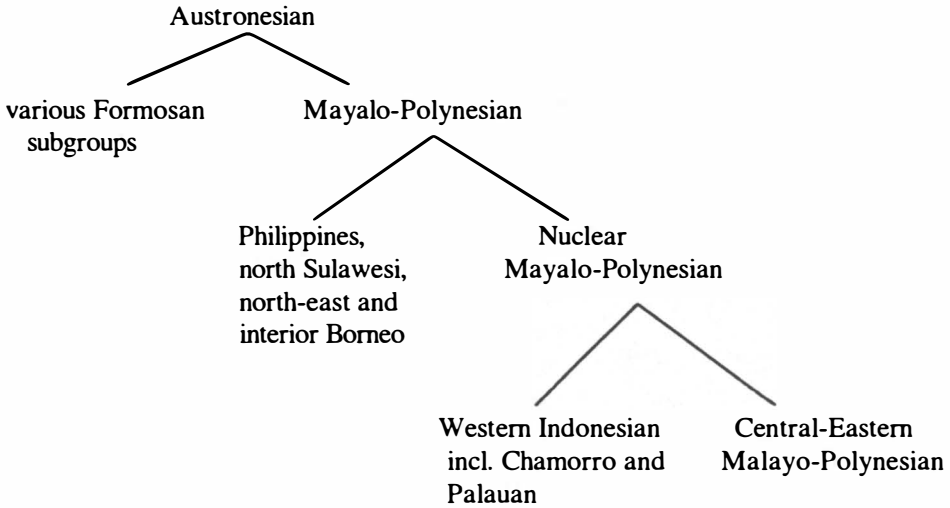


Figure 2: Modified Austronesian family tree

But in turn, it has to be emphasised that the upper branches leaving the MP and NMP nodes are not to be taken as well-defined subgroups. There may be more than one subgroup of MP coordinate to NMP, just as 'Western Indonesian' might contain several subgroups coordinate to CEMP.

In the introduction I mentioned that Chamorro and Palauan have no (non-trivial) common phonological innovations. Also the grammatical aspects discussed here do not point to a closer relation of these two languages to each other, compared with the remaining NMP languages. What they share is a certain morphological conservatism (e.g. the retention of the *maN-/um* distinction), and the retention of the passive prefix **ma-*, with loss of the atemporal form **ka-*. Only a few 'Western Indonesian' languages have preserved **ma-/ka-* as a productive morpheme. In most languages it has been replaced by a generalised **ka-* (Javanese, Toba Batak), or the widespread innovation **taR-*, which is also found in CEMP languages. Another feature that is widespread among other NMP languages but not found in Chamorro and Palauan, is the capture of the preposition **(a)ken* as a benefactive and instrumental applicative suffix. Taken together, these points indicate that Chamorro and Palauan are early offshoots from PNMP.

7 The dispersal of MP and NMP: a scenario for early migrations

The family tree proposed by Blust allows some tentative conclusions about the homeland of Austronesian speakers and the way they expanded into the archipelago and Oceania. Most certainly, Taiwan must have the longest history of Austronesian settlement; from there speakers of PMP moved south to the Philippines and further to Sulawesi, Borneo, and the Sunda islands. Speakers of Proto CEMP then broke away to migrate to eastern Indonesia and on to Oceania. These movements can be partly correlated with archeological findings (Bellwood 1985).

The NMP hypothesis allows us to refine this scenario in the following way: From the Philippines, there were movements to Borneo and Sulawesi. The latter island is a good candidate for the center of NMP dispersal, since there we find a large number of morphologically conservative languages with diverse grammatical systems. From Sulawesi, NMP speakers expanded to the Sunda islands, to parts of Borneo (as Malayic and Tamanic speakers), and to the east and south-east. It was probably also from Sulawesi that the speakers of Chamorro and Palauan (or better: Pre-Chamorro and Pre-Palauan) sailed to the northeast to the distant islands of Palau and the Marianas. The early dates of the first settlement of these islands are in accord with the assumption that Chamorro and Palauan are early breakaways from PNMP.

Of course, these speculations do not take into account the possibility that NMP languages were also spoken in other areas, e.g. in the Philippines, and have later been replaced by focus-retaining languages. Thus, Palauan and Chamorro speakers may have departed from an area outside of the present-day NMP area.

Chamorro must have remained in relative isolation for a long time, although a few loanwords from Western MP languages and Oceanic can be detected (Costenoble 1940). At one point during colonial times, the Chamorro population experienced a catastrophic reduction. This and other factors resulted in massive borrowing from Spanish and, to a lesser degree, from Philippine languages. Yet, its basic verbal morphosyntax seems to have remained quite unchanged.

In the case of Palauan, speakers of it had contacts already in pre-colonial times with speakers of Oceanic languages, first with Yapese, then with Trukic. Ross (1996) has shown that Yapese is an early offshoot of Oceanic that has borrowed heavily from Palauan and Nuclear Micronesian languages. In turn, some of the aspects where Palauan deviates from PNMP might be explained by convergence with Yapese.

References

- Adelaar, K. Alexander, 1994, The classification of the Tamanic languages. In Tom Dutton and Darrell Tryon, eds *Language contact and change in the Austronesian world*, 1-42. Berlin: Mouton de Gruyter.
- 1995, Problems of definiteness and ergativity in Embaloh. *Oceanic Linguistics* 34:375-409.
- Barsel, Linda A., 1994, *The verb morphology of Mori, Sulawesi*. Canberra: Pacific Linguistics.
- Bellwood, Peter, 1985, *Prehistory of the Indo-Malaysian archipelago*. Sydney: Academic Press.

- Berg, René van den, 1996, The demise of focus and the spread of conjugated verbs in Sulawesi. In Steinhauer, ed. 1996: 89-114.
- Blust, Robert A., 1977, The Proto-Austronesian pronouns and Austronesian subgrouping: a preliminary report. *University of Hawaii Working Papers in Linguistics*, 9/2:1-15.
- 1993, Central and Central-Eastern Malayo-Polynesian. *Oceanic Linguistics* 32:241-293.
- 1995, The position of the Formosan languages: method and theory in Austronesian comparative linguistics. In Li et al. eds, 1995:585-650.
- 1999, Subgrouping, circularity and extinction: some issues in Austronesian comparative linguistics. In Elizabeth Zeitoun and Paul Jen-kuei Li, eds *Selected papers from the Eight International Conference on Austronesian Linguistics*. Taipei: Institute of Linguistics, Academia Sinica.
- Cooreman, Ann M., 1987, *Transitivity and discourse continuity in Chamorro narratives*. Berlin: Mouton de Gruyter
- Costenoble, H., 1940, *Die Chamoro-Sprache*. 's-Gravenhage: Nijhoff.
- Dempwolff, Otto, 1934-38, *Vergleichende Lautlehre des austronesischen Wortschatzes*. Berlin: Reimer.
- Dyen, Isidore, 1965, *A lexicostatistical classification of the Austronesian languages*. International Journal of American Linguistics Memoir 19.
- Esser, S.J., 1927-33, *Klank- en vormleer van het Morisch*. Verhandelingen van het Koninklijk Bataviaasch Genootschap van Kunsten en Wetenschappen 67, 3/4. Leiden, Bandung.
- Himmelmann, Nikolaus, 1996, Person marking and grammatical relations in Sulawesi. In Steinhauer, ed. 1996:115-136.
- Josephs, Lewis S., 1975, *Palauan reference grammar*. Honolulu: University Press of Hawaii.
- 1994, Review of Lemaréchal (1991). *Oceanic Linguistics* 33:231-256.
- Kern, H., 1918-20, *Spraakkunst van het Oudjavaansch: Verspreide Geschriften VIII + IX*. 's-Gravenhage: Nijhoff.
- Klamer, Marian, 1994, *Kambara: a language of Eastern Indonesia*. Den Haag: Holland Academic Graphics.
- Lemaréchal, Alain, 1991, *Problèmes de sémantique et de syntaxe en Palau*. Paris: Edition du Centre National de la Recherche Scientifique.
- Li, Paul Jen-kuei, Dah-an Ho, Ying-kuei Huang and Cheng-hwa Tsang, eds, 1995, *Austronesian studies relating to Taiwan*. Symposium Series of the Institute of History and Philology, Academia Sinica 4. Taipei: Institute of History and Philology, Academia Sinica.
- Maan, G., 1951, *Proeve van een Bulische Spraakkunst*. 's-Gravenhage: Martinus Nijhoff.
- Mead, David, 1998, *Proto-Bungku-Tolaki: reconstruction of its phonology and aspects of its morphosyntax*. PhD dissertation, Rice University, Houston.
- Nothofer, Bernd, ed., 1996, *Reconstruction, classification, description: festschrift in honor of Isidore Dyen*. Hamburg: Abera.
- Pätzold, K., 1968, *Die Palau-Sprache und ihre Stellung zu anderen indonesischen Sprachen*. Berlin: Reimer.
- Ross, Malcolm D., 1995, Reconstructing Proto-Austronesian verbal morphology: evidence from Taiwan. In Li et al., eds, 1995:727-791.
- 1996, Is Yapese Oceanic? In Nothofer, ed. 1996:121-166.

- Sirk, Ülo, 1978, Problems of high-level subgrouping in Austronesian. In S.A. Wurm and Lois Carrington, eds *Second International Conference on Austronesian Linguistics: proceedings*, 255-273. Canberra: Pacific Linguistics.
- 1996, On the history of transitive verb suffixes in the languages of western Indonesia. In Steinhauer, ed. 1996:191-205.
- Starosta, Stanley and Louise Pagotto, 1991, The grammatical genealogy of Chamorro. In Ray Harlow, ed. *VICAL 2: Western Austronesian and contact languages. Papers from the Fifth International Conference on Austronesian Linguistics*. Auckland: Linguistic Society of New Zealand.
- Steinhauer, Hein, ed., 1996, *Papers in Austronesian linguistics* No.3. Canberra: Pacific Linguistics.
- Sundermann, D. Heinr., 1913, *Niassische Sprachlehre*. 's-Gravenhage: Martinus Nijhoff.
- Topping, Donald, 1973, *Chamorro reference grammar*. Honolulu: University Press of Hawaii.
- Tuuk, H.N. van der, 1971, *A grammar of Toba Batak*. The Hague: Nijhoff.
- Wolff, John U., 1996, The development of the passive verb with pronominal prefix in Western Austronesian languages. In Nothofer, ed. 1996:15-40.
- Woollams, Geoff, 1996, *A grammar of Karo Batak, Sumatra*. Canberra: Pacific Linguistics.
- Zoetmulder, P.J. and I.R. Poedjawijatna, 1961, *Bahasa Parwa*. Jakarta: Obor.

— PART V —

Discussion notes

