2 *Proto Oceanic phonology and morphology*

MALCOLM ROSS

1 Introduction

This chapter is intended largely for historical linguists who have an interest in the reconstruction of POc phonology and grammar. The non-linguist reader who chooses to skip it will miss little that diminishes understanding of the chapters which follow.

2 Phonology

2.1 Phonemes

Work based on the sound correspondences (Ch. 1, §3.3) of both Oceanic and non-Oceanic languages has resulted in the following reconstructed paradigms of POc consonants and vowels:

*p^w	*p	*t	*c	*k	*q
*b^w	*b	*d	*j	*g	
		*s			
*m^w	*m	*n	*ñ	*ŋ	
		*r			*R
		*dr			
		*l			
*w				*y	
	*i			*u	
	*e			*o	
		*a			

The paradigm reconstructed by Dempwolff (1937) has been modified in various ways by

Haudricourt (1951), Milke (1968), Grace (1969), Wolff (1974), Blust (1978b) and Ross (1988, 1989b, 1996d). The paradigm and orthography here are from Ross (1988), with the addition of *p^w*.

The reconstruction of *p^w* was first proposed (with little commentary) by Blust (1984), and used in reconstructions by Ross (1994a, 1996d). The POc consonant inventory reconstructed by Ross (1988:93–94), as well as its orthographically somewhat different predecessor reconstructed by Grace (1969; see §2.4), included the pairs *p* and *b*, *t* and *d*, and *k* and *g*,[1] but only the single velarised bilabial *b^w* (Grace's *ηp*); so the conclusion that *p^w* is also required by the data is not surprising. However, the reflexes of *p^w* have not been worked out as fully as those of other POc consonants (they are generally missing from the sources listed in Chapter 1, §3.3), and unexplained inconsistencies remain among the small number of widely reflected items in which it is reconstructed. Reflexes which signal its presence are:

a) velarised bilabials (usually p^w) in contexts where they do not reflect *b^w*;

    see *$p^w araq$*, *$p^w araraq$* 'thunder', *$p^w atik$* 'potato yam, aerial yam, *Dioscorea bulbifera*', *$kup(^w)ena$* 'fishing net';

b) apparent fortis reflexes of *p* in Western Oceanic and SES languages in environments where a lenis reflex is usually found;

    see *$p(^w)ilak$* 'lightning', *$p^w araq$*, *$p^w araraq$* 'thunder', *$p^w atik$* 'potato yam, aerial yam, *Dioscorea bulbifera*', *$p^w ita$* 'snare; to snare', *$p(^w)aRaRa$* 'handle';

c) apparent reflexes of *b* or *b^w* in Polynesian or Nuclear Micronesian (and occasionally other) languages;

    see *$lap(^w)a(r,R)$* 'lightning, phosphorescence', *$p^w araq$*, *$p^w araraq$* 'thunder', *$kup(^w)ena$* 'fishing net', *$p(^w)aRaRa$* 'handle'.

In some cases we reconstruct *$p(^w)$* as we are unsure whether the protophoneme was *p^w* or whether we are confronted by one of several other phenomena, including (i) borrowing, (ii) Western Oceanic and SES fortis reflexes of *p*, or (iii) velarisation before a rounded vowel in certain languages. As Blust (1981) remarks with regard to POc *b^w* and *m^w*, velarisation generally does not occur in non-Oceanic languages and sometimes occurs in POc etyma where non-Oceanic cognates give us no reason to expect it.

Although the reconstructed paradigm is fairly secure, questions remain about the phonetics of some segments. The phonemes *p^w*, *b^w* and *m^w* are known in the literature as 'labio-velars'; this orthography reflects their pronunciation in the majority of Oceanic languages in which they remain distinct, but there is evidence to suggest that they may have had the double articulations [k͡p], [g͡b] and [ŋ͡m] that 'labio-velar' suggests, since some languages (e.g. Mwotlap) have these realisations, whilst others (on Malaita and in Fiji) have velar reflexes. Among the apicals, it is possible that *t* was dental, the others alveolar, as in a number of west Indonesian languages (Ozanne-Rivierre 1992) and in Banoni (MM, NW Solomonic). The voiced obstruents in the second row were also prenasalised. Probably the phoneme *r* was an alveolar trill, whilst *dr* was a prenasalised alveolar trill, reflected thus in languages in the Admiralties and Fiji. The phoneme *c* is assumed to have been a voiceless palatal obstruent, because this is the articulation one would predict on the basis of non-Oceanic cognates and of

---

[1] These pairs are derived from PMP pairs of which the first member was an obstruent, the second a nasal + obstruent sequence, and so, viewed diachronically, the POc pairs *r* and *dr*, *s* and *j*, and *c* and *j* (sic) belong here too.

its position in the paradigm. However, it is distinctively reflected only in some Admiralties languages, where its reflexes are mostly alveolar liquids ([l], [r]) or glottals ([ʔ], [h]). Elsewhere it has merged with *s. The phoneme *j is more widely reflected, as [tʃ], [dʒ] or [d], and was more evidently a voiced palatal obstruent. Of the two postvelars (see Ross 1988:31–32), *q was probably a glottal stop, but its uvular stop reflexes in some languages give room for doubt, whilst *R was probably a uvular trill, easily lost or merged with *r or *l in daughter languages.

A noteworthy feature of the reconstructed consonant paradigm is that the only phonemic contrast between stops and fricatives is the one between *t/*d and *s, but, on the basis of widespread reflexes, it is likely that *[ɸ, β] and *[x, ɣ] occurred as allophones of *p and *k (and *[z] of *s). It is also possible that a voiced flap was an intervocalic allophone of *t.

2.2 Phonotactics

POc words were made up of (C)V syllables, with the option of a word-final consonant. These word-final consonants are lost in the majority of Oceanic languages, but retained in a scattering of Western Oceanic languages, in Mussau, and in some cases in South Vanuatu and New Caledonian languages. Quite often, we know that the PMP form had a final consonant, but no reflex occurs in any of the Oceanic languages which reflect final consonants, and so we have no means of knowing whether that consonant occurred in POc or not. In such cases the final consonant is shown in parentheses in the reconstructed POc form: e.g. PMP *kamaliR 'men's house' but POc *kamali(R) (Ch. 3, §3.3). Similarly, where a suffixed form preserves a root-final consonant but the unsuffixed root loses it, the unsuffixed form is reconstructed with a parenthesised final consonant, e.g. POc *kinit-i- (VT) 'pinch (s.t./s.o.)' but *kini(t) (VI) 'pinch'. It appears that PMP word-final consonants were quite consistently retained in POc, but to my knowledge no one has demonstrated that this is so.

PMP permitted CVC syllables both word-finally and word-internally, as in *gapgap 'stammer'. One of the innovations which defines POc is the loss of the final consonant of a word-medial syllable, as in POc *kaka(p). The most common context for this innovation is reduplicated forms like *gapgap (Blust 1977), but it also occurred elsewhere; for example, PMP *beRŋi 'night' became POc *boŋi.

The prenasalised consonants *bʷ, *b, *d, *j, *g, *dr and the glides *w, *y did not occur word-finally. The consonant *d seems to have occurred only intervocalically: if it did occur word-initially, these occurrences were extremely rare.

POc vowel sequences have, to my knowledge, never been systematically investigated, but they seem not to have been particularly common. A check of several geographically and genetically well distributed languages which are otherwise phonologically conservative[2] reveals a consistency which probably reflects the POc pattern, namely that each vowel in a sequence is the nucleus of a separate syllable. Although some Oceanic languages contrast long vowels with short or contrast a sequence of two identical vowels with a single vowel, this kind of contrast is not reconstructed for POc, where only sequences of unlike vowels were permitted. POc *e is derived from PMP word-final *-ay (§2.4), and this historical origin apparently

---

[2] The languages and sources are: Loniu (Admiralties, Hamel 1994:7–8), Manam (WOc, NNG, Lichtenberk 1983:21–32), Tawala (WOc, PT, Ezard 1997), Nakanai (WOc, MM, Johnston 1980:247–248), Kwaio (EOc, SES, Keesing 1985:8–9), Longgu (EOc, SES, Hill 1992), Nguna(EOc, NCV, Schütz 1969:11–12), Boumaa Fijian (EOc, Fij, Dixon 1988:15–16).

precludes its occurrence in POc vowel sequences except in a few probable borrowings and some derived forms. However, it is probable that all sequences of *i, *a, *o* and *u* occurred. Well attested, for example, are *waiR* 'fresh water', *raun* 'leaf', *maosak* 'ready to be eaten (because ripe or cooked)', *bou* 'main bearers supporting raised floor or roof structure, centre post supporting ridgepole', *panua* 'inhabited territory; community together with its land and things on it', *qio(r, R)* 'spear, arrow'. It is probable, incidentally, that the falling sequences *ua* and *ia* were not distinct from *uwa* and *iya*.

2.3 Stress

POc stress also remains uninvestigated, but phonologically conservative languages generally agree in displaying primary stress on the penultimate syllable and secondary stress on every second syllable preceding the penultimate, and this was probably the basic POc pattern.

2.4 Phonological innovations and Proto Oceanic orthographies

Oceanic languages reflect a set of shared innovations relative to PMP, and it was on the basis of some of these that Dempwolff (1937) first recognised Oceanic as a major Austronesian subgroup. A number of these innovations occurred among the consonants, as we see when we tabulate the correspondences between the reconstructed consonant paradigms of PMP and POc (for discussion of the PMP consonant paradigm, see Ross 1992 or 1995b). Table 1 kills two birds with one stone, also showing the two current POc orthographies. The first was established by Grace (1969) and has been used with a number of variants (separated by a slash) shown below. The second is the one generally used in this chapter, introduced by Ross (1988).

The terms 'oral grade' and 'nasal grade' were used by Grace (1969) and have become conventional among Oceanic linguists to refer to the outcomes of (c) below. Grace's orthography roughly represents the pre-POc situation. The innovations which occurred over the pre-POc period were mergers and splits, the introduction of new phonemes, and one deletion, as follows:

a)   The PMP voiced/voiceless pairs *p, *b* and *k, *g* merged respectively as early pre-POc *p* and *k*. Ozanne-Rivierre (1992) suggests that the corresponding *t, *d* merger was hindered by their mismatch in point of articulation (dental vs alveolar).

b)   The PMP pairs *s, *Z* and *d, *r* merged respectively as pre-POc *s* and *d* (phonetically probably [r], since Eastern Malayo-Polynesian cognates are liquids).

c)   PMP and a number of its descendants had word-medial homorganic nasal + obstruent sequences (not in Table 1). Some instances of the pre-POc word-initial obstruents *p, *t, *k, *d/r, *s* and *j* also acquired a preceding homorganic nasal (the occurrence of this process is unpredictable and its causes largely unknown; cf. §3.1.3). These sequences became the unitary early pre-POc phonemes *mp, *nt, *ŋk, *nd/nr* and *nj*, subsequently the POc prenasalised voiced obstruents *b, *d, *g, *dr*, and *j* (PMP nasal + *s* and nasal + *j* merged as pre-POc *nj*, POc *j*). It is possible that pre-POc *nt*, POc *d* never occurred word-initially.

Table 1: The phoneme systems of PMP and POc

PMP		*p, b*	—	*t*	*d, r*	*s, z, Z*	*j*	*k, g*
POc	oral grade	*p*	*pw*	*t*	*d/r*	*s*	*j*	*k*
Grace etc	nasal grade	*mp*	*ŋp/mpw*	*nt*	*nd/nr*		*nj*	*ŋk*
POc	oral grade	*p*	*pʷ*	*t*	*r*	*s*	*c*	*k*
Ross	nasal grade	*b*	*bʷ*	*d*	*dr*		*j*	*g*

PMP	*m*	—	*n*	*ñ*	*ŋ*	*w*	*y*	*l*	*q*		*h*	*R*
POc Grace	*m*	*ŋm/mw*	*n*	*ñ*	*ŋ*	*w*	*y*	*l*	*q*		*Ø*	*R*
POc Ross	*m*	*mʷ*	*n*	*ñ*	*ŋ*	*w*	*y*	*l*	*q*		*Ø*	*R*

PMP	*i, -uy(-)*	*e, -aw*	*-ay*	*a*	*u*
POc Grace etc.	*i*	*o*	*e*	*a*	*u*
POc Ross	*i*	*o*	*e*	*a*	*u*

d) The labio-velars *$*pʷ$*, *$*bʷ$* and *$*mʷ$* entered the language. Most of the items containing a labio-velar lack non-Oceanic cognates, and some, at least, must have been borrowed into POc from neighbouring Papuan languages. For example, it can be argued that *$*mʷapo(q)$* 'taro' was borrowed by POc speakers as they acquired more sophisticated taro-growing techniques from Papuan speakers (Ross 1996d). A few of these items were inherited into POc, and the labio-velar was the reflex of a labial occurring next to a round vowel. However, it is not clear in these items that the labio-velar actually occurred in POc (Blust 1981). Thus a number of Oceanic languages reflect *$*tamʷata$* 'man, husband', derived from *$*tau$* 'body, person' + *$*mataq$* 'unripe, immature, young', but we cannot be sure whether this or *$*taumata(q)$* was the POc form.

e) PMP *$*h$* was lost in POc.

f) PMP *$*e$*, phonetically [ə], became POc *$*o$*, and the PMP word-final diphthongs *$*-uy(-)$*,[3] *$*-aw$* and *$*-ay$* were simplified to POc *$*-i$*, *$*-o$* and *$*-e$* respectively, the first two thereby merging with plain vowels.

The combined effect of (a) and (c) is that each of the PMP pairs *$*p$*, *$*b$* and *$*k$*, *$*g$* first merged and then split. As a result, for example, PMP *$*p$* became either POc *$*p$* or POc *$*b$*, and the same was true of PMP *$*b$*, giving the kind of crossover seen in the initial consonants of these examples:

PMP *panas   'hot, warm'   POc *panas
PMP *punay   'wild pigeon'   POc *bune

---

3   The notation *-uy(-) reflects the fact that there is one known case where the change to *i occurred word-medially: PMP *kamuihu (independent 2PL pronoun) > *kamuyu > POc *kamiu.

PMP *baqeRuh*   'new'   POc *paqoRu*
PMP *beRek*   'pig'   POc *boRok*   'domestic pig'.

Similarly, either PMP *k* or PMP *g* could become either POc *k* or POc *g*. For example,

PMP *kuDen*   'cooking pot' POc *kuron*
PMP *kabut*   'mist'   POc *gabu*
PMP *gapgap*   'stammer'   POc *kaka(p)*
PMP *gemgem*   'make a fist' POc *gugu(m)* 'grasp in fist, clench fist'.[4]

3 Word classes

The remarks below on POc word classes and derivational morphology are rather brief, largely because these remain somewhat poorly explored areas. More detail is provided by Lynch, Ross and Crowley (forthcoming), but there is a great deal about POc morphosyntax that we do not know and which is perhaps irretrievable. The major publication on POc grammar is Pawley (1973).[5]

Ideally, the reconstruction of a Proto Oceanic etymon should include not only its form and meaning, but also its word class membership. POc had just two open lexeme classes: nouns and verbs. POc was a head-marking language, and each valent (dependent) noun phrase was cross-referenced on its head noun or verb by a clitic or suffix. Both nouns and verbs fell into two subclasses on the basis of valency.

3.1 Verbs[6]

POc verbs had a valency of either one or two, that is, they were either intransitive or transitive. There were probably no trivalent/ditransitive verbs, i.e., verbs whose role structures required or allowed three noun phrases without case marking, but we cannot be certain about this, as some modern languages do have trivalent verbs (Manam, Hoava).

Verbs apparently took a proclitic cross-referencing their subject and, if transitive, an enclitic cross-referencing their object (in many daughter languages these are a prefix and a suffix), e.g. POc *i-kiniti-au* 'he pinched me' (cf. Manam *i-ʔint-a*).[7] To judge from descriptions

---

[4] PMP etyma with an unambiguous initial *g-* are rare, and this example shows a mismatch between the vowels of PMP and POc (see Ch. 9, §7).

[5] Pawley's (1972) reconstruction of PEOc grammar is also relevant to POc reconstruction, particularly as it is not clear that EOc was a discrete interstage, and features reconstructed for PEOc may consequently be attributable to POc.

[6] In order to reconstruct the POc verbal system, I have consulted grammars which (a) are sufficiently detailed and (b) describe languages which seem phonologically and morphologically quite conservative. There are not many of these: I have consulted descriptions of Manam (WOc, NNG) (Lichtenberk 1983), Hoava (WOc, MM) (Davis 1997), Kwaio (Keesing 1985) and Longgu (Hill 1992) (both EOc, SES), Ambae (Catriona Hyslop, pers.comm.) and Malo (Jauncey 1997) (both EOc, NCV) and Boumaa Fijian (EOc, Fij) (Dixon 1988). The criterion of morphological conservatism is the sharing of morphosyntactic features across different Oceanic subgroups. Unfortunately, we have no grammar of an Admiralties language which is sufficiently detailed.

[7] It is not clear how complete the POc clitic sets were. Evidence is strong that an object enclitic occurred only

of verbs in modern Oceanic languages (e.g. Dixon 1988:201–202), a majority of POc verb roots had both intransitive and transitive (or causative) alternants. In at least one language, Hoava (WOc, MM), all verbs in the corpus have transitive and intransitive alternants (Davis 1997).

3.1.1 Verb classes

We can reconstruct three major classes of POc intransitive verb on the basis of semantic and morphological criteria, as shown in Table 2.

Table 2: Classes of intransitive verb in Proto Oceanic

	1 inherently stative or inherently dynamic?	2 intransitive subject	3 forms a transitive?	4 forms a causative?
A verbs	dynamic[8]	A	yes	yes
U verbs	neither	U	yes	yes
U-stative verbs	stative	U	no	yes

As Table 2 shows, the only criterion which distinguishes all three classes from each other is a semantic one: is the intransitive verb inherently dynamic, inherently stative, or inherently neither (column 1)? However, I follow Oceanist convention by using labels which refer to the macrorole of the intransitive subject: A for actor, U for undergoer (column 2), even though this leads to the partially redundant 'U-stative' label where 'stative' would do.[9]

Fijian preserves the classification in Table 2 quite clearly, although individual verb forms in modern Fijian are not always in the same classes as the POc etyma they reflect. Intransitive verbs with an actor subject, i.e. A verbs, are necessarily dynamic (Table 2, columns 1 and 2). The subject of the intransitive is also the subject of the corresponding transitive, as the Boumaa Fijian clauses in (1) illustrate:

(1) a. *Au   rabe.*
    S:1 S kick   'I'm kicking.'

---

if the object was singular or third person non-singular. If it was first or second person non-singular, the object was probably an independent pronoun (Evans 1995). Something similar may have been true of subject proclitics.

[8] I use the term 'dynamic' (rather than 'active') in contrast with 'stative' simply because 'active' also contrasts in a quite different sense with 'passive'.

[9] The terms 'actor' and 'undergoer' are from Foley and van Valin (1984). Dixon (1988) uses A and O. Arms (1974a), Foley (1976) and others use A (actor/agent) and P (patient), but this labelling is infelicitous in today's terms as it confuses macrorole (A, U) and role (agent, patient, experiencer, theme etc.). The distinctions between the three classes were first demonstrated systematically by Pawley (1973:126–140) with data from Motu (PT), Roviana (MM), Kwara'ae (SES), Arosi (SES), and Bauan Fijian. The seminal work on Fijian A and U verbs is Arms (1974a). Biggs (1974) addresses parallel issues in Polynesian, Foley (1976) in Malayo-Polynesian. Boumaa Fijian examples in this section are from Dixon (1988:204, 231).

b. *Au   rabe-t-a   a   polo.*
S:1S kick-TR-O:3S ART   ball   'I'm kicking the ball.'

Intransitive verbs with an undergoer subject, on the other hand, fall into two classes, U and U-stative. With U verbs the subject of the intransitive is the *object* of the corresponding transitive:

(2) a. *E   gagi a   dovu.*
S:3S crush ART   sugarcane   'The sugarcane is being crushed.'

b. *Au gagi-a   a   dovu.*
S:1S crush-O:3S ART   sugarcane 'I'm crushing the sugarcane.'

U verbs and U-statives differ from each other in two respects. First, unlike U verbs, U-statives have no corresponding transitive (but do have a corresponding causative, as described below). Second, U-statives like *loaloa* 'be black' in (3) are inherently stative, whereas U verbs like *gagi* in (4) are inherently neither dynamic nor stative. With appropriate aspect marking and context, the clause in (4) may be given either a dynamic ('the sugarcane is being crushed') or a stative ('the sugarcane is [already] crushed') interpretation.

(3) *E   loaloa a   ʔolii yai.*
S:3S be.black ART   dog   this   'This dog is black.'

(4) *E   gagi a   dovu.*
S:3S crush ART   sugarcane   'The sugarcane is crushed.'

The dividing line between U and U-stative is a thin one. In some, probably many, Oceanic languages, including Fijian, appropriate aspect marking can force a dynamic interpretation of an U-stative (e.g. of 'black' as 'become black'). There does, however, seem to be a semantic difference between the two classes: intransitive U verbs imply some unmentioned agent or instrument, whereas U-statives do not.

There is a tendency both in the reconstruction of POc (Pawley 1973)[10] and in descriptions of modern Oceanic languages to regard all intransitives with an U subject as stative. In some modern languages this is seems to be correct, but in others, and apparently in POc, they are/were distributed between the classes I have labelled U and U-stative.[11]

Although it is not difficult to identify the three verb classes in many modern Oceanic languages, it can often be difficult to determine which class a given POc verb belonged to. The reasons for this are: (i) languages which retain the three classes do not always agree on the class to which the reflexes of a given POc verb belong; (ii) in some languages (e.g. Kwaio) more than others (e.g. Longgu), a verb may belong to more than one class, and indeed some verbs may have belonged to two classes in POc; (iii) particularly in north-west Melanesia

---

[10] Pawley (1973:128) has A-class statives and B-class statives, corresponding respectively to my U-stative and U classes. His Intradirectives are members of my A class, which also includes the intransitive alternants of his Spontaneous Transitives and Deliberate Transitives. The differences among the subclasses of A intransitives are not morphological but lie in the exact semantic roles of their subjects and of the objects of their transitive alternants.

[11] Thus, this example from Hoava is labelled "stative" by Davis (1997) but the verb appears to be an U verb, at least in this usage (*sa leboto* is the subject of intransitive *tuke*):

Tuke   sa   leboto.
be.thrown.away ART:SG bushknife
'The bushknife was thrown away.'

there are languages which have entirely lost the stative class (e.g. Takia) and/or the neutral class (e.g. Takia, Tawala, Mangap-Mbula), replacing all statives and some neutral verbs by adjectives or adjectival nouns (Ross 1998) and transferring other neutral verbs to the dynamic class.

3.1.2 Derivational morphology of verbs

POc transitivising morphology was rather different from that of its Fijian reflexes above. POc verb roots were mostly disyllabic and, in line with POc phonotactics (§2.2), either consonant-final or vowel-final, that is, (C)V(C)VC or (C)V(C)V. In most cases, the canonic shape of the root alone determined its transitivising morphology. The transitive of a consonant-final root was formed by adding the suffix *-i- between the root and the object enclitic. This suffix is known in the Oceanic literature as the '(close) transitive suffix' and was the same regardless of whether the root was A or U:

(5)   intransitive   corresponding transitive

    A verbs   *kinit*[12]   'pinch'   *kinit-i-*   'pinch (s.o/s.t)'   (Ch. 9, §6.3)
              inum   'drink'   *inum-i-*   'drink (s.t.)'

    U verbs   *pʷosa(k)* 'crack open'   *pʷosak-i-* 'crack (s.t.) open'   (Ch. 9, §5.2)
              loŋoR   'be audible'   *loŋoR-i-* 'hear, listen to'

With a vowel-final root like *wase-* 'share (s.t.) out' or *kati-* 'husk (s.t.) with teeth', no transitive suffix occurred and the object enclitic was added directly to the root (Evans 1997). The one possible exception to this are roots ending in *-a, where the suffix *-i- probably occurred between the root and the object enclitic, at least when the enclitic itself began with *a (*au O:1S, *a O:3S). Note that the final *-i- of a disyllabic base like *kati was also present when the verb was used intransitively. Hence it was not a transitive suffix in POc, although in some cases it was derived from an earlier suffix. With vowel-final roots, as with consonant-final, there was no formal difference between A and U roots:

(6)   intransitive   corresponding transitive

    A verbs   *kati*   'husk with teeth' *kati-*   'husk (s.t.) with teeth' (Ch. 9, §3.7)
              muri   'follow'   *muri-*   'follow (s.t./s.o.)'
              soka   'pierce, stab'   *soka-i-*   'pierce, stab (s.t./s.o.)' (Ch. 9, §4.1)

    U verbs   *wase*   'be shared out'   *wase-*   'share (s.t.) out'
              poli   'be bought'   *poli-*   'buy (s.t.)'

Below are some examples of POc intransitive/transitive verb pairs with their Boumaa Fijian reflexes. The POc transitive includes the third person object enclitic *a:

(7) POc   Boumaa Fijian

    inum   'drink'   *unu*
    inum-i-a   'drink it'   *unu-m-a*

---

[12] Note that *kinit is reconstructed with final *-t but *pʷosa(k) with parenthesised *-k simply because we have a reflex of *kinit in a language which retains POc final consonants, but none for *pʷosa(k). Following the convention outlined in §2.2, final *(-k) is inferred from reflexes of transitive *pʷosak-i-.

loŋoR	'be audible'	*roŋo*	
loŋoR-i-a	'hear it'	*roŋo-ð-a*	
taŋis	'weep'	*taŋi*	
taŋis-i-a	'weep for it'	*taŋi-ð-a*	
soka	'stab'	*ðoʔa*	'throw a spear'
soka-i-a	'stab it'	*ðoʔa-a*	'throw a spear at'
kati	'husk with teeth'	*ʔati*	'bite'
kati-a	'husk it with teeth'	*ʔati-a*	'bite it'
wase	'be shared out'	*wase*	
wase-a	'share it out'	*wase-a*	

The reader will notice that formal restructuring has occurred in Fijian. The POc final consonant is lost from consonant-final intransitives. This leads to a resegmentation of the transitive, such that the intransitive form is treated as the root in Fijian, the POc final consonant becomes the transitive marker, and the POc transitive suffix *-i- is lost. Because the POc consonant is no longer part of the Fijian root but an allomorph of the transitive suffix, the etymological consonant is sometimes replaced by another or a consonant has been inserted where none is expected (Arms 1974b).[13] This is the case with *roŋo-ð-a* 'hear it' above, where the normal Fijian reflex of POc *R is zero or occasionally *r*, but never *ð*.

Fijian-like restructuring has occurred in many Oceanic languages because of the loss of the POc final consonant. In most of these languages (and in some Fijian dialects), *-i- has not been lost, with the result that a language has a set of transitive suffixes with the form -Ci-, as in Longgu.

(8) POc   Longgu

inum	'drink'	*inu*	
inum-i-a	'drink it'	*inu-vi-a*	
taŋis	'weep'	*aŋi*	
taŋis-i-a	'weep for it'	*aŋi-si-a*	

The consonant of the -Ci- suffix is known as the 'thematic consonant'. Transitive verbs in our data corpus are often cited with a reflex of POc *a* O:3S. Where this can readily be omitted because it permutes with other object enclitics, its absence is marked with a hyphen. Where it is not synchronically separable, it is parenthesised in accord with the conventions of Chapter 1, §4.2.

Above, I wrote that transitivising morphology is determined in most cases by the canonic shape of the root. There are some exceptions to this generalisation, and one set of these is discussed by Blust (1977). These consist of forms like those in (9), descended from a PAn or PMP reduplicated monosyllabic root. The intransitive POc form in each case reflects the earlier form with regular loss of the final consonant of the first syllable (§2.2). The transitive form reflects a single monosyllabic root plus an *-i- which by POc times was no longer

---

[13] A special case of non-etymological consonant insertion occurs with (A) verbs of motion and posture, where the transitive form takes a location as its object. For example, Longgu *eno* 'lie down' (< POc *qeno*) vs *eno-vi-* 'lie on', *dio* 'fall' (< POc *sipo* 'descend') vs *dio-ŋi-* 'fall on'; Boumaa *laʔo* 'go' (< POc *lako*) vs *laʔo-vi-* 'go for'. To date I have found such cases only in EOc languages and do not reconstruct this feature for POc.

separable, so that the transitive forms were similar in their behaviour to the transitive forms of vowel-final roots like *kati-* 'husk with teeth'.

(9)   Proto Oceanic
                                  intransitive   transitive   Ch. 9[14]

		intransitive	transitive	Ch. 9[14]
PAn	*tektek* 'hack off'	*toto(k)*	*toki-*	3.2
PMP	*(c,s)uk(c,s)uk* 'skewer'	*(su)su(k)*	*suki-*	4.1
PAn	*tuqtuq* 'hammer, pound'	*tutuk*	*tuki-*	5.1
PMP	*pakpak* 'clap, slap'	*baba(k)*	*baki-*	5.1
PAn	*buCbuC* 'pluck out'	*pupu(t)*	*puti-*	6.1
PMP	*pespes* 'squeeze, press out'	*popo(s)*	*posi-*	7

We also find 'deponent' cases in POc where a PAn/PMP reduplicated monosyllable and a POc transitive monosyllable + *-i-* form are reconstructable, but no intransitive form. This is perhaps a reflection of the fact that in many cases the formal relationship between reflexes of the intransitive and transitive forms, especially after loss of the final consonant from the intransitive, has become so opaque that the pairing of all but the most frequently used forms has been lost.

We have seen that a majority of non-stative POc verb roots had both intransitive and transitive (or causative) alternants, and that the transitive is usually derived from the intransitive. There seem to have been a few POc verb roots, however, which were intrinsically transitive and from which either an A or an U intransitive could be formed.

An A intransitive was formed by reduplicating the disyllabic (transitive) root. One such verb was apparently POc *kani-* 'eat (s.t.)', whose corresponding intransitive was *kani-kani*. Hence we find Malo (NCV) *hani-* 'eat (s.t.)' vs *han-hani* 'eat' and Motu (PT) *ani-* vs *ani-ani*. A Boumaa Fijian example is *ʔusi-*'wipe (s.t.) with a cloth' vs *ʔusi-ʔusi* 'wipe hands after washing them at the end of a meal'.[15]

An U or U-stative intransitive was apparently formed by prefixing the anti-causative prefix *ma-* to the transitive root, reflected, e.g., in Malo *ma-duru* 'be split' from *duru* 'split (s.t.)' and *ma-mbila* 'be shattered' from *bila* 'shatter (s.t.)', and in Arosi *ma-hita* 'be split, broken' from *hita* 'split, hit, strike (s.t.)'. Like these examples, most modern reflexes of *ma-* derive U-statives from transitives, but we have already noted a tendency for U intransitives to become statives. There is, however, a scattering of languages in the North New Guinea and Papuan Tip clusters where *ma-* derives A verbs. Since these are languages in which U intransitives have been reinterpreted as A intransitives, I take it that these A verbs reflect earlier U verbs. For example, from Sio *liŋi*, Misima *liŋi-n* 'pour (s.t.) out' we can reconstruct transitive POc *liŋi* 'pour (s.t.) out', and from Sio *ma-liŋi*, Misima *ma-liŋi-n* '(liquid) run away' we can reconstruct the POc U verb *ma-liŋi* 'be poured out'.

Thus there were five morphological relationships between intransitive and transitive forms in POc, illustrated in (10), the first two being the most widely represented:

---

[14] These reconstructions are drawn from Ch. 9. Glosses are abbreviated or omitted here for the sake of clarity, and cross-references are to the full presentations in Ch. 9.

[15] Possible counterevidence is provided by Longgu, where reduplication forms both A and U intransitives: *ale-a* 'bite him' vs *ale-ale* 'bite', but *ʔave-a* 'bend it' vs *ʔave-ʔave* 'be bent'. But comparative evidence suggests that *ʔave* was originally an U verb ('be bent') from which *ʔave-a* was derived, and that *ʔave-ʔave* is the result of pattern extension.

(10) Class   intransitive   transitive

both	*kinit*	'pinch'	*kinit-i-*	'pinch (s.o/s.t)'
both	*kati*	'husk with teeth'	*kati-*	'husk (s.t.) with teeth'
both	*toto(k)*	'cut, chop'	*toki-*	'cut, chop (s.t.)'
A verb	*kani-kani*	'eat'	*kani-*	'eat (s.t.)'
U verb	*ma-liŋi*	'be poured out'	*liŋi*	'pour (s.t.) out'.

As the foregoing examples show, in the reconstructions a transitive verb is marked with a final hyphen. If an intransitive/transitive pair is reconstructed, it will be shown as, e.g., *kinit, *kinit-i- 'pinch'. Where a vowel-final verb like *wase 'distribute' is reconstructed, consistency would require us to show this as *wase[-], that is, as reconstructable with and without a following object enclitic. This convention is followed in Chapter 9, but not in other chapters; it is in any case often very difficult to determine whether a vowel-final root was used both transitively and intransitively in POc.[16] Where a pair of verbs is reconstructed without and with *-i-, the supporting cognate sets are usually combined. Occasionally, where the cognate sets supporting the intransitive and transitive forms are of considerable size, they are given as separate lists.

One final point must be made with regard to POc *-i-. I have adopted the usual convention of calling it a transitive suffix. However, in a number of modern languages (e.g. Hoava, Davis 1997), in circumstances where the verb is immediately followed by a modifier, the 'suffix' (if any) and the object enclitic follow the whole verb-modifier complex, suggesting that *-i- may have been a transitive enclitic rather than a suffix. However, this has no bearing on lexical reconstruction.

Implicit in the discussion above are two slightly unusual features of transitivity in POc and many daughter languages. One is the division of non-stative intransitives into A and U verbs. The other is the use of the A intransitive alternant of verbs which are semantically transitive. If the object of, say, POc *inum 'drink' or *kati 'husk with teeth' was not mentioned, the intransitive form was evidently used.

Two other morphemes were productive in the derivation of POc verbs: the causativiser *pa- or *paka- and the applicative *-aki or *-akini. As these occur less often in our reconstructions than the transitivising and detransitivising morphemes discussed above, they are described only briefly here.

Although transitives could not be formed from U-stative roots, causatives could be formed from roots of all three classes. The POc prefix deriving causatives was *pa- or *paka- (both forms are reconstructable, and the difference between them needs more research). A causative formed with this prefix was a transitive verb whose subject was always the causer and whose object was the same as the subject of the corresponding intransitive verb (Table 2, column 2). Thus an U-stative verb like POc *ponuq 'be full' could be causativised (= transitivised) with *pa[ka]- to give *pa[ka]-ponuq-i- 'cause (s.t.) to be full, make (s.t.) full' (object=U).[17] An U verb like *wase 'be shared out' gave *pa[ka]-wase- 'cause (s.t.) to be shared out' (object=U), and an A verb like *inum 'drink' gave *pa[ka]-inum-i- 'cause (s.o.) to drink' (object=A).

---

[16] It could be said that consistency requires us to put a hyphen before every verb to mark the presence of the subject proclitic, but we find this redundant since in this regard every verb behaves in the same way.

[17] The fact that U-stative roots could only be transitivised with *pa[ka]- was one of the features of this verb class recognised by Pawley (Pawley 1973:128–129) and used to distinguish it from what I have here called the U class.

Presumably the liquid which was drunk was expressed by an oblique phrase, but this needs more research. The presence or absence of *-i-* in causative verbs was again determined by whether the root ended in a consonant or a vowel.

The applicative *-aki* or *-akini*, reconstructed by Pawley (1973) as the 'remote transitive' suffix, is attached to an A or U root to form a transitive verb. In its canonic usage, the object of this verb is a referent which would appear as an oblique with the corresponding direct transitive. For example:

(11) POc   Boumaa Fijian

taŋis	'weep'	*taŋi*
taŋis-i-a	'weep for it'	*taŋi-ð-a*
taŋis-akini-a	'weep about it'	*taŋi-ðaʔin-a*
soka	'stab, spear'	*ðoʔa*
soka-i-a	'stab, spear it'	*ðoʔa-*
soka-(C)-akini-a	'stab, spear with it'	*ðoʔa-taʔin-a*
puni	'hide'	*vuni*
puni-a	'hide it'	*vuni-a*
puni-(C)-akini-a	'hide (s.t.) for s.o.'	*vuni-taʔin-a*

Again there are two reconstructable forms, POc *-aki* and *-akini*. The difference between them needs more research, as does their status, their POc function, and the history of their reflexes (Harrison 1978, 1982). It was noted above that the POc transitiviser *-i-* occurred only with consonant-final or *-a-* final roots. The applicative *-aki[ni]* was not subject to this limitation. In Boumaa Fijian, like other languages, an apparently non-etymological thematic consonant is in many cases inserted between a vowel-final root (including one in *-a-*) and the reflex of *-aki[ni]* to form *ðoʔa-taʔin-a, vuni-taʔin-a* and so on. The inserted thematic consonant is apparently a lexically determined choice between *-t-* and *-v-*.[18] This gives rise to the reconstructive difficulty seen in (11): we do not have enough information to know what happened in POc when *-aki[ni]* followed a root ending in a vowel.

3.1.3 Fossilised verbal morphology

The morphology described in §3.1.2 was largely productive when POc diversified into daughter languages. There are a few patterns in our reconstructions of verbs, however, which reflect morphology that was already dead by this stage.

One of these is illustrated in (9), where intransitive forms descended from PAn or PMP reduplicated monosyllabic roots correspond with transitive forms reflecting the unreduplicated root plus inseparable *-i-*. PAn monosyllabic roots have been investigated in some detail by Blust (1988) and are also often reflected as the second syllable of POc CVCVC intransitive roots.[19] As a result, consonant-final disyllables with related meanings often share their second syllable. Thus in Chapter 9, §6.1 we find the following reconstructions forming the following set:

---

[18] Cf. *dusi-a* 'point it out (nearby)' vs *dusi-vaʔin-a* 'point it out (far off)'; *te-a* 'plant (crop)' vs *tee-vaʔin-a* 'plant (land)'; *tala-a* 'send her/him' vs *tala-vaʔin-a* 'send for s.o.'.

[19] A tentative explanation of the origin of PAn monosyllabic roots is offered by Ross (1995b:95–96).

(12)   PAn root *-buC* 'weed, pull, pluck out' (Blust 1988:86–87)

POc	*pupu(t), *puti-*	'pick (fruit +), pluck (feathers +)'
POc	*(s, j)apu(t), *(s, j)aput-i-*	'pull out, pull up, pluck (fruit, nuts)'
POc	*tapu(t), *taput-i*	'strip (crops), pull off'

The pair *pupu(t), *puti-* reflects PAn *buC-buC* in accordance with the paradigm in (9), whilst the intransitive roots *(s, j)apu(t)* and *taput* apparently reflect PAn forms **sa-buC* and **ta-buC*.[20] Other such sets are:

(13)   PAn root *-pak* 'break, crack, split' (Blust 1988:135–136)

POc *sapaki* 'pluck off, break off (leaves) with the hand' (Ch. 9, §6.1)
POc *paki* 'pluck, break off (leaves) with the hand' (Ch. 9, §6.1)
POc *lopa(k)* 'break' (Ch. 9, §6.2)

(14)   PAn root *-Tuk* 'knock, pound, beat' (Blust 1988:160–161) (Ch. 9, §5.1)

POc *tutuk, *tuki-* 'pound, mash by pounding, hammer, crack by hammering'
POc *putu(k)* 'repeatedly knock, pound, beat'
POc *butu(k), *butuk-i-* 'repeatedly knock, pound, beat'

(15)   PAn root *-Tak* 'sound of cracking, splitting, knocking' (Blust 1988:157–158)

POc *potak, *potak-i-* 'crack open, split open, make incision' (Ch. 9, §3.8)
POc *botak, *botak-i-* 'crack open, split open, make incision' (Ch. 9, §3.8)
POc *pita(k), *pitak-i-* 'break, split' (Ch. 9, §5.2)

An additional complication consists in the fact that Blust finds PAn/PMP roots which differ only in the voicing of the root-initial consonant and which have similar meanings. Since PAn/PMP voicing distinctions were not retained in POc (§2.4), their reflexes are indistinguishable in Oceanic languages. Thus POc *tupu(k), *tupu(k)-i-* 'knock against', apparently reflecting PMP *tu(m)buk* 'pound', is attributed by Blust (ACD) to the PAn root *-buk* 'pound, thud, heavy splash' (Blust 1988:87–88). POc *sapu(k), *sapu(k)-i-* 'hit', on the other hand, apparently reflecting PMP *sa(m)puk* 'collide, bump into', is attributable to the PAn root *-puk* 'throb, thud, clap, break' (Blust 1988:87–88) (cognate sets in Ch. 9, §5.1).

PAn monosyllabic roots probably ceased to be independent morphemes (if indeed they ever were independent) sometime around the break-up of PAn, although they may well have played a role in the phonaesthetics of daughter languages for some time after the break-up. The other dead patterns I wish to consider, however, probably arose from morphology which remained productive in Eastern Malayo-Polynesian until not long before the break-up of POc. This is morphology which reflects the Austronesian 'focus' system, variants of which occur in many non-Oceanic Austronesian languages. I present the reconstructed systemic changes first, then the instances of them, partly because the systemic changes also have a bearing on POc nominal morphology, discussed in §3.2.1.

The essence of the focus system is that the semantic role of the 'topic' (alias 'subject', 'nominative', 'pivot', 'trigger') of a verbal clause is indicated by an affix or affixes on the verb.[21] The morphology of the PMP focus system was, at least roughly, as in (16).

---

[20] The double asterisks indicate that I do not know of non-Oceanic cognates which would independently support the PAn reconstructions.

[21] The question of how 'focus' systems should best be described has a long and controversial history. A good

(16) The PMP verbal system (based on Wolff 1973 and Ross 1995a)

| | nominalisations *or* indicative independent | | non-indicative |
	imperfective	perfective	or dependent
Actor focus or intransitive	‹um›√	‹um-in›√	√
Actor focus only	[ma]N-√	naN-√	paN-√
Patient focus or intransitive	√-en	‹in›√	√-a
Location focus	√-an	‹in›√-an	√-i
Instrument or beneficiary focus	i-√	i-‹in›√	√-án

The symbol √ represents the verb root and ‹...› an infix after the root-initial consonant. PMP -*N*- represents an underlying velar nasal which combined with a root-initial voiceless obstruent to give the homorganic nasal, and with a root-initial voiced obstruent to give either the homorganic nasal or a nasal + obstruent sequence:

(17) **paN-* + **takaw* 'steal' ⟶ **panakaw* 'steal (actor focus)'
　　*paN- + **deŋeR* 'hear' ⟶ **pandeŋeR* 'hear (actor focus)'

The systemic features relevant to this discussion are that (i) there were two sets of verb forms, the first used in indicative independent clauses and the other in non-indicative independent and dependent clauses; and (ii) those in the first set were formally identical with nominalisations. However, there were probably no nominalisations corresponding with the actor focus forms.

　　The historical relationship between the PMP and simpler POc system has intrigued various scholars (Pawley & Reid 1980, Starosta, Pawley & Reid 1982, Wolff 1980). The POc system in (18) is set out in such a way that it corresponds with (16).

(18) The POc verbal system

| | nominalisation | | verb |
	imperfective	perfective	
Intransitive	—	—	√
(relic transitives)	(√-on)	—	(paN-√, N-√)
Transitive	√-an	‹in›√(-an)	√-i-
Applicative	i-√	—	√-aki[ni]-

The stages by which the PMP system became the POc system lie beyond the scope of this chapter (Ross 1997 provides a hypothesised sequence), but the main changes (not necessarily in diachronic order) were:

a)　Instrument or beneficiary focus **√-án* was replaced by **√-aki[ni]-*. This change is also reflected in many languages in Indonesia.

b)　PMP patient focus forms were lost, and the function of location focus forms was extended to include patient focus. These forms became the POc transitives.

c)　The PMP indicative independent verbal forms lost their verbal functions and remained

---

basic description of such a system is provided by Schachter (1987).

only as nominalisers, leaving the erstwhile non-indicative/dependent verbal forms as the only verbal forms.

Note that the verbal morphology reconstructed in the rightmost column of (18) is what has been described in §3.1.2. The only additions are the parenthesised forms, representing possible fossils. Evidence for these is given in (19).

(19) POc relic verb forms (Ross 1988:41–42 gives supporting data)

PMP		POc	
root	*[pa]N-root		
*takaw	*panakaw	*panako	'steal'
*ka(ʔe)n	*paŋa(ʔe)n	*paŋan	'eat'
*ka(ʔe)n-i		*kani-	'eat'
*sepsep	*ñep	*sopi-, *ñopi-	'suck'
*da(n)daŋ	*naŋ	*raŋi-, *naŋ	'shine' (Ch. 9, §11)
*buni	*muni	*puni-, *muni-	'hide (VT)'

In POc *panako and *paŋan, we have clear cases where *paN- is preserved. The pair *paŋan and *kani is evidence that some relic of the focus system may have continued to exist until shortly before the break-up of POc, *paŋan reflecting the actor focus in this sytem, *kani the patient focus. This in turn allows us to intepret the POc pairs *sopi, *ñopi etc. in (19) as patient/actor pairs. Note, however, that this interpretation is not watertight. First, the expected outcome of *-N-raŋ is **draŋ, not *naŋ. Second, *ñopi 'suck' displays both initial *ñ- and suffixed *-i-, i.e. the morphology of both actor and patient focus simultaneously.[22]

A third set of morphological fossils also seems to date from the pre-POc period and may also be associated with the focus system. We find a number of POc verbs, mainly in Chapter 9, where the data support the reconstruction of a pair of forms differing in the grade (oral or nasal; cf. §2.4) of their initial consonant. For the verbs in (20), the data justify the reconstruction of two forms. For those in (21), the case is not quite so clear, but there are forms which indicate that there may also have been a POc form with a nasal-grade initial.

(20) POc   POc

   *puru(k), *puruk-i- 'pierce, bore (hole)'   *buru
   *putu(k) 'repeatedly knock, pound, beat'   *butu(k), *butuk-i-
   *kiri 'file, rasp, saw'   *giri
   *kora(s), *koras-i- 'scrape out'   *gora(s), *goras-i-
   *rama(R) 'torch; fish with torch'   *drama(R)

(21) POc   Reflexes with voiced initials

   *poka 'to divide, cut up'   Wayan boka(ti-) 'split or cut s.t. in half'
                                                   Carolinian pax 'be cut, split'

   *potak, *potak-i- 'crack open'   Motu bota(i) 'beat, thrash'
                                                   Wayan bote(ki) 'split or crack s.t. open'

---

[22] This doubling up of morphology has at least two possible explanations. One, *-i- was added to **ñop by analogy after the focus system had collapsed. Two, the system reconstructed in (21) is wrong and the pre-POc system was more like that of a number of Indonesian languages, where reflexes of *maN- and *-i co-occur in a single verb form. The evidence to date is too thin to permit a choice.

p(ʷ)ip(ʷ)i(t) 'press, wring, squeeze s.t.'	Gumawana *bibi* 'squeeze (boil +)' Arosi *bibi* 'crush, squeeze, crowd'
kili(s), *kilis-i-* 'twist, bore, rotate'	Bauan *gili-* 'twist or rub in the hands' Nadroga *gili-* 'braid'
kutu 'cut'	Bauan *gutu*, *gutuv(a)* 'cut off, sever'
kinit, *kinit-i-* 'pinch off'	Gumawana *ginisi* 'pinch s.o.' Lau *gini-gini* 'pinch off with the nails'
kawit, kawit-i- 'hook, catch hold of'	Lukep -*gaot* 'pick (breadfruit)' Dobu *geuta* 'hook fruit, fruit hook' Lau *gau* 'pluck fruit with bamboo, crook'
sapu(t), *saput-i-* 'pull out, up, pluck'	Roviana *zapu* 'pull coconuts from a tree'

The pairs *p/*b, *k/*g, *s/*j and *r/*dr are represented in (20) and (21). Differences in token frequency are probably largely attributable to differences in the frequency of these consonants across POc vocabulary as a whole. The other logically possible pair *t/*d is not represented, but this reflects the fact that word-initial occurrences of *d were either rare or zero (§2.2).

A salient fact about these pairs is that there are generally more reflexes with the oral-grade alternant than the nasal-grade. This suggests that the nasal-grade alternants reflect morphologically marked forms of the unmarked oral-grade forms. If this is correct, then we have to ask whether all POc verb roots had an unmarked form with an initial oral-grade consonant. The answer seems to be a qualified 'yes'. Exceptions occurred when a noun with an initial nasal-grade consonant was used as a verb, e.g. POc *buku* 'node, knot, protuberance' was used as the verb *buku* (VI), *bukuti* (VT) 'tie (a knot); fasten'. Otherwise we find, in Chapter 9 for example, that the large majority of reconstructed POc verbs begin with an oral-grade consonant. A few, like POc *baba(k)*, *baki-* 'strike one against another, knock', begin with a nasal-grade initial, and I have no explanation for these other than to speculate that an oral grade-initial root occurred but is not reflected in our data.

If the suggestion of the previous paragraph is correct, then we are looking for a morphological alternation that ceased to be productive shortly before the break-up of POc. Initial nasal-grade consonants (which did not occur in this position in PMP) arose from a sequence of nasal + obstruent. That is, *b, *g, *dr, *j developed from pre-POc *mp, *ŋk, *nr, *ns (and *nc) (§2.4). It is only a short step to infer that this nasal feature was the actor focus morpheme *-N- seen in (16) and (18) and that these pairs are systemically parallel to those in (19). But we should be wary of making this inference too quickly. PMP nasal + obstruent sequences occurred in actor focus verbs only when the root-initial obstruent was voiced, as illustrated in (17). If the nasal feature in the voiced-initial members of the pairs in (16) and (18) really was *-N-, then we would expect to find that they were all descended from PMP voiced-initial roots. But when the known PMP ancestors of these pairs are listed against them, we find that they often have a voiceless initial obstruent:

(22) POc   PMP

putu(k) 'repeatedly knock, pound, beat'   *buTuk*
kiri 'file, rasp, saw'   *kirkir*
poka 'to divide, cut up'   *peka*
potak, *potak-i-* 'crack open'   *beTak*
p(ʷ)ip(ʷ)i(t) 'press, wring, squeeze s.t.'   *pitpit*

*kinit, *kinit-i- 'pinch off'	**kinit*
*kawit, kawit-i- 'hook, catch hold of'	**kawit*

This means that instead of the pair **kiri/*giri*, for example, descended from PMP voiceless-initial **kirkir*, we would expect a pair **kiri/**ŋiri,* corresponding to the pairs in the lower half of (19). But the only such pairs I have found are those in (19). Thus, although I cannot exclude the possibility that the pairs in (16) and (18) reflect a feature of the focus system, I am unable to give a principled account of them.[23]

3.2 Nouns

Nouns had a valency of either one or zero. A monovalent noun normally took a suffix which cross-referenced the person and number of the dependent noun phrase, usually its possessor, e.g. POc **tama-gu* 'my father', **tama-ña* 'her/his father' (cf. Bali-Vitu, Tolai, Fijian *tama-gu, tama-na*), **tama-ña tam**ata* 'the man's father'. I assume that, as in many modern languages, the dependent noun phrase usually occurred only if it was third person, and was optionally omissible even there.[24] A zero-valency noun had no affixation, e.g. POc **Rumaq* 'house' (Ch. 3, §3.3). Monovalent nouns were nouns that are possessed by default: they included kin terms (like **tama-* 'father'), body parts (e.g. POc **qaqe-* 'leg') or parts of wholes (e.g. POc **gab**ari-* 'area underneath a raised house' (Ch. 3, §3.4)), and are shown in reconstructions and supporting data with a final hyphen. Zero-valency nouns were all other nouns. These two subclasses are reflected in numerous Oceanic languages, and are referred to in grammars either by the semantic labels 'inalienable noun' and 'alienable noun' or, since Lichtenberk (1985), by the syntactic labels 'directly possessed noun' and 'indirectly possessed noun'.

3.2.1 Derivational morphology of nouns

Morphemes which derive nouns have rather a complex history in Oceanic. As noted above (§3.1.3), PMP nominalisations were identical to certain independent indicative forms of the verb. These are repeated in (23), with examples based on the root **kaʔen* 'eat' to illustrate how the system apparently worked. The meanings based on the focus labels are the probable central meanings of these nominalisations: as in the modern languages, each affix must have had other senses too.

---

[23] Explanations might include (i) errors in the association of POc and PMP reconstructions, i.e. more items in (23) and (24) are descended from voiced-initial PMP roots than is apparent (unlikely, as PMP **g-* had a very low functional load); (ii) some unrecognised developments had occurred in the pre-POc focus system; (iii) the nasal feature came not from PMP **-N-* but from PMP **‹um›* in (19), i.e. the POc system in (21) is incorrectly reconstructed.

[24] When a monovalent noun had a non-specific possessor, it was evidently linked to the latter by the preposition **qi,* e.g. **natu qi boRok* 'piglet' (lit. 'child of pig'), but there is considerable evidence that **qi* was bound to the preceding monovalent noun (cf, for example, Seimat [Admiralties] *nat-i pou* 'piglet'; see Hooper 1985 for further evidence).

A description of the (quite complex) POc possessive system is beyond our scope. The interested reader is referred to Pawley (1973:153–169), Lichtenberk (1985), Hooper (1985), Lynch (1996a).

(23) PMP	nominalisation		
	imperfective	perfective	*ka?en* 'eat'
Patient	√-en	‹in›√	*ka?en-en* 'thing to be eaten, food'
			k‹in›a?en 'thing eaten'
Location	√-an	‹in›√-an	*ka?en-an* 'place where one eats'
			k‹in›a?en-an 'place where one has eaten'
Instrument	i-√	(i-‹in›√)	*i-ka?en* 'thing one eats with'

The affix combination **i-‹in›√* is parenthesised because I know of no reflexes of it as a nominaliser. Reflexes of **‹in›√-an* are known only from Oceanic, and may reflect a local innovation.

The corresponding POc nominalisations are tabulated in (24). Although the verbal focus system had disappeared in POc, it is reasonably clear that the nominalising morphology continued to be associated with semantic roles:

(24) POc	nominalisation		
	imperfective	perfective	*kani* 'eat'
Patient	(√-on)	‹in›√	*kanoŋ* 'thing to be eaten, food'
			k‹in›ani 'thing eaten'
Location	√-an	‹in›√-an	*kan-an, *kanaŋ* 'place where one eats'
			k‹in›an-an 'place where one has eaten'
Instrument	i-√	—	*i-kani* 'thing one eats with'

The affix combination **√-on* is parenthesised because it survived only in fossilised forms like POc **kanoŋ(a)* 'flesh, meat, coconut flesh' (Ross 1996d:174). However, it is clear that the rest of the system remained productive in POc (and much of it remains productive in various modern languages), as **‹in›* and **i-*, at least, were evidently added to the productive POc root, e.g. **kani*, not the stem **kan* reflecting PMP **ka?en*. However, it is less clear that this is true of **-an*.

Instrumental **i-√* has kept its PMP function, e.g. **asa(q)* 'grate', **asaq-i-* 'grate (s.t.)'; **i-asa(q)* 'grater' (Ch. 9, §2.1). It has been lost in a number of languages, however, sometimes where it was in competition with initial-syllable reduplication, which also formed instruments.

I noted in §3.1.3 that the function of location focus forms was extended to include patient focus. This also happened to a degree with nominalisations, in that **√-on* survives only in fossils. However, its perfective counterpart **‹in›√* survives as patient and general nominaliser in Mussau and the Meso-Melanesian cluster.[25] In Roviana and Hoava (MM) the general nominaliser is ‹in›, e.g. Roviana *k‹in›era* 'song' from *kera* 'sing', whilst **-an* retains its local meaning, e.g. Roviana *huhuve-ana* 'bathing place, bath' from *huhuve* 'bathe' (Roviana preserves POc final consonants with a following echo vowel, so *-ana* is the regular reflex of **-an*). In other languages the reflex of the ertswhile locational **√-an* has taken over the function of general nominaliser, so that some reflexes of POc **mate-an*, e.g. Vitu (MM) *mate-a*, Longgu (SES) *mae-a-*, mean 'death', rather than 'deathbed' or 'cemetery'. The affix combination **‹in›√-an* is reflected in fossilised reflexes of POc **k<in>ani-ana* 'food' in the languages of Epi (central Vanuatu) (Tryon 1976:289).

The history of **√-an* as a nominaliser in Oceanic has several complications. First, as **kanaŋ* and **kanoŋ* above indicate, forms in final **-ŋ* occurred in POc alongside those in

---

[25] Zero-derivation or reduplication is also used to form deverbal nouns in a number of languages.

final *-n*, so *√-an* and *√-aŋ* apparently coexisted, a fact for which I have no explanation. Secondly, there is good evidence that both also occurred with an additional *-a*, i.e. as *√-ana* and *√-aŋa*. The only known language where a contrast has been found between forms without and with *-a* is Mangseng (NNG), where *-ŋ* is the general nominaliser, *-ŋ-a* the instrumental formative (e.g. *puno-ŋ* (N) 'fight', *puno-ŋ-a* 'weapon', both from *pun* (V) 'fight'). This suggests that the forms with *-a* may have had a separate function in POc.

Forms reflecting *√-an* and *√-aŋ* can be disambiguated only in languages which preserve POc final consonants faithfully, and in fact only two such languages are known to reflect *√-an*. These are Roviana and Hoava, as illustrated above. However, if reflexes of *√-ana* are also taken as evidence of POc *√-an*, then there are a number of these in the Admiralties, e.g. Seimat *paku-an*[26](N) 'dance', from *pak* (V) 'dance', and in Central and South Vanuatu, e.g. Paamese *sau-ene* 'singing' from *sau* 'sing', Lenakel *akar* 'talk' from *n-akar-aan* 'talking' (where *n-* reflects a POc article *na*).

Forms reflecting *√-aŋ* are found in languages of the NNG cluster. Here, however, there is a different complication, namely that *√-aŋ* and *√-aŋa* are reflected as *√-ŋ* and *√-ŋa*, e.g. *mate-aŋa* 'death' is replaced by *mate-ŋa*, getting rid of the vowel sequence in favour of the strongly CV canonic shape of early Oceanic (§2.2). This is evidently a local innovation. Thus we find, for example, Lukep-Pono *kani-ŋ* 'yam' (from *kani* 'eat'), Gitua *gururu-ŋ* (N) 'thunder' (from *guru* (V) 'thunder').

Forms reflecting *√-aŋa* and *√-ŋa* are well scattered, for example, Poeng (NNG) *mate-ŋa* 'death' (from *mate* 'die'), Mussau *palapala-ŋa* (N) 'thunder' (from *pʷaraq* (V) 'thunder', Samoan (Pn) *inum-aŋa* 'draught, dose' (from *inum* 'drink'), *tafi-ŋa* 'removal' (from *tafi* 'remove'), and throughout Polynesia.

There is a wealth of languages in which final consonants are lost, so that both *√-an* and *√-aŋ* are regularly reflected as *-a*, leaving us with no way of knowing which form was ancestral (e.g. Loniu *hɛ-ya* 'washing' from *hɛ* 'wash'; Malo *dule-a* 'clearing bush' from *dule* 'clear bush'). The Bali dialect of Bali-Vitu retains final consonants with a following echo vowel, but neutralises POc *-n* and *-ŋ* as *-ŋ-*, so that the ancestral consonant is again ambiguous (e.g. *moŋe-aŋa* 'sleeping' from *moŋe* 'sleep').

In the light of these complications—and because we do not understand them well—we reconstruct each nominalisation as the data require, with *√-an*, *√-aŋ*, *√-ana* or *√-aŋa* (or on occasion *√-ŋ* or *√-ŋa*), but recognising that this is probably not an accurate rendering of the POc form. As I noted above, however, we find cognate sets which allow the reconstruction of, say, both *mate-an* (or *mate-aŋ* or whatever) and *m‹in›ate* 'death'. These pairs occur almost certainly because the reflex of one of *√-an* (etc.) or *‹in›√* has extended its productivity at the expense of the other in various languages, creating the appearance that POc had both forms. In such cases we reconstruct, e.g., POc *mate* ('die') + NOMINALISER 'death', as we cannot tell which form was in fact lexicalised in POc.

3.3 Adjectival classes

POc had no separate adjective class. Instead, it had a large class of adjectival verbs and a small class of adjectival nouns (Ross 1998). The class of adjectival verbs appears to have included all U-statives, and at least some U intransitives (§3.1.1). Many reconstructed adjectival

---

[26] The presence of *-n* attests the earlier presence of a following vowel.

verbs include the U intransitive prefix **ma-* (§3.1.2). Both adjectival verbs and adjectival nouns could apparently follow the noun they modified without any morphological marking, but their behaviour differed in the predicate. There, an adjectival verb behaved like any intransitive verb, whilst an adjectival noun behaved as a (zero-valency) noun whose property is attributed to a referent or referents. Thus POc **saqat* 'bad' was an adjectival verb, **paqoRu* 'new' an adjectival noun. Either could modify a noun: POc **a Rumaq saqat* 'a bad house', POc **a Rumaq paqoRu* 'a new house' (where **a* was a common article). Compare Bali-Vitu *a rumaka zayata* and *a rumaka vayoru* (Bali-Vitu is a conservative Oceanic language of the Meso-Melanesian cluster). However, as predicates they behaved differently, probably as in the Bali-Vitu examples below, where *zayata* 'bad' is a verb preceded by the proclitic *ti*, a portmanteau marker of third person subject and perfective aspect, whilst *vayoru* 'new' is preceded by the article *a* and means 'a new one'.

(25) Bali-Vitu:

a. *A   vaga beini ti   zayata.*
    ARTICLE canoe that   PERFECTIVE:3 bad
    'That canoe is broken.'

b. *A   rumaka beini a   vayoru.*
    ARTICLE house   that   ARTICLE new
    'That house is new.'

4 Assigning reconstructions to word classes

Although we know with reasonable certainty what the open word classes were in POc and what their major subclasses were,[27] we cannot always assign a reconstructed etymon to a single word class or subclass. Some items, of course, like **tama-* 'father', **Rumaq* 'house' and **kinit-i-*, are easily assigned: they are, respectively, a monovalent noun, a zero-valency noun and a transitive verb. Quite a number of the items we reconstruct are derived items, and their morphology allows us to assign them to a class. Relevant morphemes are described above in §3.1.2 and §3.2.1, Other items, especially zero-valency nouns and intransitive verbs, the comparative evidence suggests, could readily serve in more than one word class without any morphological change. Thus in cases like POc **p(ʷ)anaq* (N) 'bow', (VI) 'shoot' (Ch. 8, §9) we assume that the etymon served as both a noun and verb in POc.

In other cases our data sources simply give insufficient or inaccurate information about word class and subclass memberships, so that we often do not know whether a disyllabic vowel-final verb base in a given language is transitive, intransitive or both. In the case of adjectival nouns and adjectival verbs, we cannot always be sure which of the two subclasses an etymon belonged to. Confronted with situations of this kind, we have not always attempted formally to assign POc etyma to their word classes, but allow our glosses and the hyphenation conventions referred to in the discussion above to speak for themselves.

---

[27] For more detailed information the reader is referred to Pawley (1973:126–140), Crowley (1985), Ross (1998).