Journal of the Linguistic Society of Papua New Guinea

ISSN: 0023-1959

Vol. 35, 2017

The Proto-Oceanic Common Article in Southwestern Malakula Languages

John Lynch
UNIVERSITY OF THE SOUTH PACIFIC
johnlynch123@gmail.com

Abstract

Two forms of the Proto-Oceanic common article marking nonhuman nouns as definite have been reconstructed: *na and *a. In only a couple of Oceanic languages do both occur together, however; usually, if the article is retained—very often in Vanuatu as an accretion to the noun root—it is almost universally a reflex of *na. In a few languages in southwest Malakula, however, the accretion derives from *a, either instead of *na (in some) or along with it (in others). This paper examines that distribution, along with cases where no accreted article occurs, and tries to elucidate (without much success) what might have obtained in Proto-Oceanic.

1. INTRODUCTION

Many languages of Vanuatu have accreted onto noun roots what was once a common article (Crowley 1985, François 2000, Lynch 2001) . In some languages, under some morphosyntactic conditions, this accretion is more separable from the noun root than it is in others, but this point is not at issue here. What is at issue is the shape of the accreted article.

1.1 The Proto-Oceanic common article

Crowley (1985) reconstructed a Proto-Oceanic (POc) common article marking nonhuman nouns as definite, with two forms: *na and *a. He proposed the noun class system for POc as outlined in (1) (1985:184).

(1)	*na/*a	zero
humans	none	all
nonhuman animates	some	most
inanimates	most	some

His survey of Eastern Oceanic languages divided them into four categories based on the way they reflect the article (1985:161):

 I. zero-marking: no trace of any common NP marking system;
 II. residual, nonproductive system, with a morphologically fused article *na/*a;
 III. productive marking on many common nouns by a regularly separable prefix derived from *na/*a; and
 IV. fully productive marking on all common nouns by *na/*a, usually as a free form.
 As to the form(s) of the article, his tentative conclusions are as follows:

> Reflexes of both *na and *a are distributed geographically over such wide areas … that we can hardly deny that both forms are of considerable antiquity.
>
> Since both forms presumably existed in the protolanguage, we should therefore consider the nature of their relationship. There are basically two possibilities. First, they may have been allomorphs of the same morpheme, or second, they may have been morphemes with distinct (but nevertheless closely related) meanings (Crowley 1985:181).

He finds no strong support for either of these hypotheses, and concludes that "this question must therefore remain unanswered. We will continue to simply refer to *na/*a, on the understanding that this is to be ambiguously interpreted" (1985:182).

 Lynch, Ross, and Crowley (2002:71) point out that if *na and *a "were allomorphs, we would expect them to survive together in widely distributed modern languages, but … we find allomorphy [in] only two places": (i) in Label and the north Bougainville linkage, "*a* follows a consonant, *na* a vowel"; and (ii) in Ambae, "*na* occurs in an object noun phrase and *a* elsewhere". They theorise that "*a and *na were once contrasting morphemes, but

the basis for the contrast had been eliminated by changes which had occurred at a stage immediately prior to the break-up of POc, and languages have got rid of the resulting redundancy by eliminating one or the other form".

1.2 The common article in Vanuatu languages

Much of Vanuatu is occupied by Type II languages—that is, languages where many common nouns have a fused article but where the system is not really productive. In these languages, *na is widespread as the form of the accretion; *a occurs as an accretion only in Paamese, and in the languages I am going to discuss in this paper.[1] The examples below are from Mwotlap (Banks Is.), Neve'ei (Malakula), Sye (Erromango), Aɲejoɱ̃ (Aneityum), and Paamese (Paama); the accreted article is marked off from the rest of the noun by a slash.

(2)	POc		MWOTLAP	NEVE'EI	SYE	ANEJOɱ̃	PAAMESE
	*mata-	'eye'	na/mte-	ne/meta-	ni/mtu-	ne/mta-	a/met, mete-
	*lipo-	'tooth'	nɪ/lwo	no/lovu-	ne/lve-	ne/jhe-	a/loh, loho-
	*Rumʷaq	'house'	n/ɪmʷ	ni/yim	n/imo	n/iomʷ	e/im

Terry Crowley's research on Malakula languages, which belong to the Central Vanuatu subgroup, turned up one more language where the accreted article derived from *a rather than *na: Avava. In addition, an examination of the word lists collected by Tryon (1976), Charpentier (1982), and Shimelman (n.d., but collected in 2016) shows that there are other lects in the southwest part of Malakula where *a-accretion is found. In some of these, *a-accretion occurs alongside *na-accretion, with some nouns taking one and some the other. My aim in this paper is (i) to establish where *a-accretion occurs; (ii) in lects where both *na- and *a-accretion occurs, to try to establish conditions for the occurrence of one rather than the other; and (iii) to try to establish conditions for zero-marking in these languages.

2. CROWLEY'S DESCRIPTION OF AVAVA

I will begin with Crowley's description of Avava (Crowley 2006), which is the best-documented and most complete description of all of the language varieties I will be dealing with.

Avava was originally spoken inland from Vinmavis village (where Neve'ei is spoken) , on the central-west coast of Malakula. The language has become geographically discontiguous, and is no longer spoken in its original location: today, speakers of Avava reside in four villages—Tisvel, on the west coast south of Vinmavis; Teremp and Tembimbi, on the east coast just north of Unua; and Khatbol, inland from the east coast near Lingarakh (see map 1). There are 400–500 Avava-speakers in these villages, plus another 300–400 in urban and semi-urban areas elsewhere in Vanuatu (Crowley 2006:3–6).

In neighbouring Neve'ei, as in nearly all other languages in Malakula, the accreted article derives from POc *na, with phonological conditioning of the vowel of the article. Cognate forms in Avava, however, are vowel-initial, and I suggest that this vowel derives from the other form of the POc article *a. (I will discuss the actual shape of the vowel below.) Compare (Crowley 2006:23):[2]

(3)	POc	AVAVA	NEVE'EI	
	*baga	a/baŋ	ne/baŋ	'banyan'
	N*maro	a/mar	ne/mar	'hunger'
	*beta	e/bet	ne/bet	'breadfruit'
	*baiga	i/biŋ	ni/biʔiŋ	'giant turban shell'
	E*mʷa(q)ele	i/mwil	ni/mwiyil	'cycad'
	*toqa	o/to	no/to	'chicken'
	*toRas	o/tor	no/tor	'island teak'
	*takuRu-na	u/dru-n	ne/taʔa-n	'his back'

---

[1]   A reflex of *a does occur as a free-standing article in a few Northern Vanuatu languages (for example, Northeast Ambae).

[2]   Reconstructions are Proto-Oceanic unless preceded by a raised letter: E indicates Proto-Eastern Oceanic, N Proto-North-Central Vanuatu (PNCV), R Proto-Remote Oceanic.

MAP 1. MALAKULA LANGUAGES AND SUBGROUPS
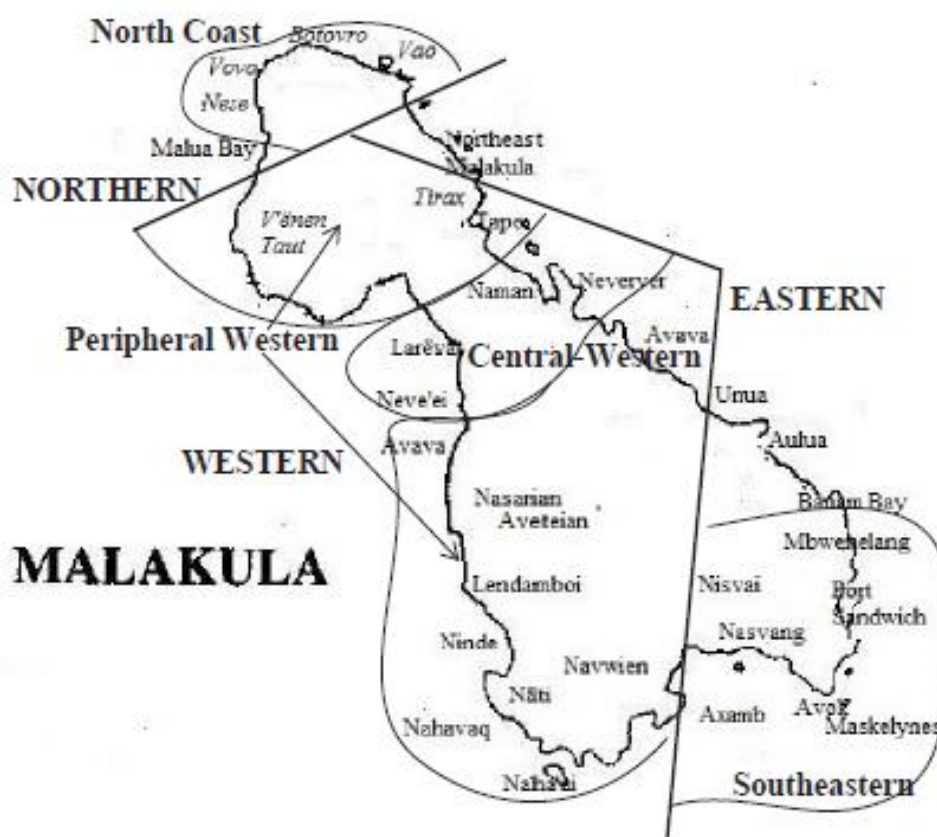
2.1 The form of the accretion

There is some evidence that vowel harmony is involved, as can be seen from the shape of the first two vowels in each noun in (3): the accretion is a copy of the first vowel of the root. A full analysis of vowel-initial nouns in Crowley's lexicon shows the patterns as illustrated in table 1.

TABLE 1. ACCRETED VOWELS IN AVAVA NOUNS

Noun-initial V	Initial segment(s) of noun roots									
	i	(C)Ci	e	(C)Ce	a	(C)Ca	o	(C)Co	u	(C)Cu
i-	**1**	20	**2**	5		16		7		5
e-				**16**						
a-		17			**3**	**54**			2	**18**
o-		2		5		2	**1**	**23**	**6**	1
u-				2		7		1	1	**5**

The following rule accounts for a considerable number of accreted articles (bolded in table 1), but there is also a number of exceptions:

(4)   *a-   → $\begin{cases} \text{i- / \_\_ e, i, many (C)Ci} \\ \text{e- / \_\_ (C)Ce} \\ \text{a- / \_\_ a, (C)Ca} \\ \text{o- / \_\_ o, u, (C)Co, we} \end{cases}$

Rule (4) is basically a rule showing assimilation of the accreted vowel to the next vowel in the noun. There are two significant exceptions in the data shown in table 1:

(i) *a* is the unexpected accretion in 44% of nouns where the first vowel of the root is *i* (17/39) and in 62% of nouns where the first vowel of the root is *u* (18/29); and

(ii) *i* is unexpectedly present noun-initially in 20% of nouns where the first root vowel is *a* (16/79) and 23% of nouns where the first root vowel is *o* (7/31).

Exception (i) involves no phonological change to the vowel of the accreted article. This suggests that the article assimilation rule ceased operating before it applied to all relevant nouns, and it also provides some confirmation that the original form of the accretion was indeed *a.

Exception (ii) may involve a dissimilation rule which is widespread in Central and Southern Vanuatu languages and which is not restricted to accreted nominal articles. Basically, *a became a higher vowel—most often *e*, but occasionally *i* or *ə*—when the vowel in the next syllable was also *a (Lynch 2003). In Avava, *a often became *e* (5a) and occasionally *i* (5b) when it occurred before *Ca. (The first three examples in (5a) illustrate the dissimilation in the first vowel of the root, as well as the assimilation of the article.)

(5)   POc   AVAVA
   a. *draRaq 'blood'   e/dre
       ᴿ*maraya 'eel'   e/mer
       *m(ʷ)ata 'point'   e/mwet 'knife'
       *madraR 'fermented breadfruit' + *beta 'breadfruit'   mete/bet
       *taRaq-i 'chop, cut w. axe'   ter
       *saqat 'bad'   se
       *barapu 'long'   ber(ber)
   b. *sapa 'what?'   sive
       ᴺ*ma(t,d)aga 'Kleinhovia hospita, puzzle tree'   midaŋ
       *panako 'steal'   pinok

It is possible then that the general low vowel *dissimilation* rule and the very specific accreted article *assimilation* rule were in conflict in cases where *a was accreted to *Ca-initial nouns, and the conflict was resolved in different ways in different nouns.

2.2 Nouns with no accretion

The accretion seems only to have applied to nonhuman nouns: nouns referring to humans in Vanuatu languages generally did not accrete *na (or *a), as Avava nouns like *vivi-* 'sister (of man)', *bbuah* 'grandmother', *momok* 'woman', *ma-* 'brother (of woman)', *mom* 'father', and *muut* 'man' illustrate. However, not all nonhuman nouns have an accreted article.

There appear to be two, partially overlapping explanations. One of these is word-initial vowel loss. It is worth quoting Crowley at some length:

> My Avava corpus provides numerous examples involving alternation between a longer form of a word and a shorter form, the forms differing in the presence or absence of an initial vowel. We therefore find examples such as the following:

> (6)   **Longer form**   **Shorter form**
>   /avaran/   /varan/   'his/her arm'
>   /atan/   /tan/   'land'
>   /ebet/   /bet/   'breadfruit'
>   /isiːŋ/   /siːŋ/   'wild ginger'
>   /omoːb/   /moːb/   'skink'
>   /uꭰuːn/   /ꭰuːn/   'his/her back'

> In some cases, the longer and shorter forms appear to alternate freely in isolation (including forms found in elicitation), and this process affects even some high-frequency items. In fact, even the language name has been found to alternate between the shapes /avava/ and /vava/.

> In other cases, the longer form is invariably found in isolation, as well as in sentences when the item is not closely linked structurally with any other item within the sentence. The shorter form, by way of contrast, may appear when the item is closely linked to some other form (Crowley 2006:20; my numbering).

Crowley goes on to point out that he has recorded numerous nouns that are consonant-initial—for example, *koit* 'octopus' < *kuRita, *sivir* 'coconut lory' < *sipi(rR)i, or *mari* 'barrel tree' < ᴺ*mariu *'Acacia spirorbis'*— but questions whether these might have had a vowel-initial counterpart that has not as yet been recorded.

A second possible explanation, and one which may well have operated alongside the first, is the effect of syllabic structure. In some Malakula languages, the article is retained whatever the syllabic structure of the noun root. Pearce (2007) first noted for Unua, however, that, when a noun root consists of a single syllable,[3] the accreted article is retained; but where a noun root consists of two or more syllables, the article does not appear. Compare Neve'ei, which represents the first type, with Unua, in (7), in which (7a) shows monosyllabic roots and (7b) multisyllabic roots. It will be seen in (7) that Avava follows the Unua pattern and not the Neve'ei one.[4]

(7) POC   NEVE'EI   UNUA   AVAVA

  a. *kutu   na/ʔat   na/xut   a/ut   'louse'
      *ñamuk   ne/nam   ne/nom   a/nam   'mosquito'
      *manuk   ni/min   ne/men   a/man   'bird'
      *mʷata   ne/mwat   na/mat   a/mwat   'snake'
      *toŋoR   no/doŋ   na/rroŋ   o/doŋ   'mangrove'

  b. *mata-na   ne/meta-n   mete-n   mata-n   'his/her eye'
      *qata-mate   ne/temah   demej   tamat   'devil'
      *bʷatu(k)-na ne/bat-n   bati-n   bat-n   'head'
      ᴺ*buaga   ni/biax   bbuag   bbuax   'swamp taro'
      ᴱ*bakuRa   ne/baʔar   bagur   bakur   'k.o. tree, *Calophyllum* sp.'

Differences in syllabic structure seem to explain the situation reasonably well. As with any statement about Malakula historical phonology, however, there are some exceptions. In particular, there are some noun roots of two or more syllables which unexpectedly retain the accreted article; some examples:

(8) POC   AVAVA

    *b(ou)kas(i)   a/bbuah   'pig'
    *sulati   o/dalah   'worm' (*s > *d* unexpl. but common in Malakula)
    *quluŋ-an   u/kulaŋ   'headrest, pillow'
    ᴺ*laŋa-na   u/laŋa-n   'its scales (of fish)'

Recall from (1) above that Crowley suggested that, in POC, most nonhuman animates were zero-marked. This does not seem to have been inherited in Avava: note, for example, *e/mer* 'eel' in (5), *a/nam* 'mosquito', *a/man* 'bird' and *a/mwat* 'snake' in (7), and *a/bbuah* 'pig' and *o/dalah* 'worm' in (8), all nonhuman animates with *a-marking.

2.3 Residual *na

There is, however, some evidence that *na was present in the language, possibly until relatively recently:

> The situation in Avava is complicated somewhat by the observation that, with at least some very old speakers, the vowel-initial longer forms of nouns are found to alternate very occasionally with forms which begin instead with *nV-*. The following forms have been found to vary in this way in my Avava corpus:

> (9) /emʷet/   /nemʷet/   'knife'
>   /avat/   /navat/   'stone'
>   /ani/   /nani/   'coconut'
>   /adam/   /nadam/   'yam' (Crowley 2006:24: my numbering)

---

3   Whether the original root was multisyllabic or not is not relevant: vowel loss rules have reduced, for example, the two syllables of the root of *na kutu 'louse' to one in Unua *naxut*.

4   I write the velar fricative as *x* irrespective of how it has been written in the original sources; *bb* represents a voiced prenasalised bilabial trill; Unua *rr* is a trill contrasting with the flap *r.*.

Crowley suggests that "*a and *na must both have been retained until relatively recently", although it is also possible that only *a is a retention, and that *na crept in at some stage, in some words, due to contact with neighbouring languages.

2.4 Summary

Unlike nearly all Malakula languages, Avava accreted the article *a rather than *na onto nonhuman nouns. This vowel underwent assimilation, and possibly also dissimilation, such that all five vowels now appear as reflexes. The article was retained if the root was monosyllabic but was lost if the root contained two or more syllables (though there are exceptions).

2.5 A note on Umbbuul

There is a variety of Avava spoken south of the original homeland known as Umbbuul (/uᵐʙuːl/), spoken in an area known as Bangasak or Bangaasak (Crowley 2006:9–11). Shimelman (n.d.) includes a list from Bangaasak, and this corresponds very closely to Crowley's Avava data. In particular, the pattern of *a-marking as opposed to zero marking is very similar, with cognate nouns generally behaving the same way. There are just a few exceptions:

(10)	POc	AVAVA	UMBBUUL	
	*takuRu-	u/dru-	dru-	'back'
	*waiR	o/we	we	'water'
	*kayu	a/ga	na/ɣai	'tree, wood'
	*salan	a/sal	ni/sel	'road'

3. AVETEIAN, LENDAMBOI, AND NASARIAN

The 30+ languages of Malakula belong to three genetic groups: a Northern Malakula subgroup (5 languages), an Eastern Malakula linkage (11 languages), and a Western Malakula linkage (16~ languages). Within Western Malakula, Avava has been classified as belonging to a Southwestern subgroup along with languages that Lynch and Crowley (2001) call Aveteian, Lendamboi, Nasarian, Navwien, Nahavaq, and Naha'ai (Lynch 2016a,b). The first three of these languages show evidence of accretion of *a rather than *na, but the situation is complex in a number of ways.

Data on Aveteian, Lendamboi, and Nasarian come from three sources: (i) Charpentier (1982), a linguistic atlas covering 19 lects in south Malakula, with maps listing the local equivalents of over 1,700 lexical items; (ii) Tryon (1976), a somewhat expanded Swadesh 200-word list; and (iii) Shimelman (n.d), 135 Swadesh-type word lists collected from around the whole island; in many cases, two or more lists come from what is probably the same lect.

In 3.1 through 3.3, I present the facts (as they are available) relating to POc *na and *a in Aveteian, Lendamboi, and Nasarian, respectively, while in 3.4 I examine what might be called zero-accretion—that is, nouns with no trace of the POc articles. I leave any discussion of the implications of these facts, and of the Avava data, till section 4.

3.1 Aveteian

Aveteian was originally spoken just north of Ninde, but may have been extinct by the time of Charpentier's survey, and only its name appears therein, with no data. However, Tryon's (1976) Dixon Reef I wordlist is probably Aveteian (Dixon Reef being a bit of a refugee village).

Of the 77 nouns in Tryon's list that have POc or PNCV antecedents, 46 occur prefixed with what is almost certainly a reflex of *a,[5] 12 with *na, while 15 have no accreted article. (The remaining four are kin terms that

---

[5]   The reflex is most commonly *a* but, in decreasing order of frequency, *e*, *ə*, and *o* also occur. Conditioning patterns—like those in Avava, for example—are not readily apparent: *e* often occurs when the first vowel of the root is *i* or *e* (*ikan > e/ika* 'fish', *lipo- > e/lvi* 'tooth', *mata- > e/meti* 'eye'), but *a* also occurs in this environment (ᴺ*yalo > a/iale* 'sun', *butoŋ- > a/bti* navel', *salan > a/sele* 'road').

have an accreted *kV-, which is not of relevance here.) However, there seems to be no phonological conditioning determining whether *a- or *na- appears; compare:

(11) POc   AVETEIAN   **but**   POc   AVETEIAN
     *manuk   a/man   'bird'   N*matuqu   na/mat   'coconut'
     *tali   e/tal   'rope'   *tasik   ne/tas   'sea'
     *pulu- 'body hair' e/vli 'head hair'   *para-   na/vari   'arm, hand'

I will return to the question of whether a noun did or did not have an accreted article in §3.4.

3.2 Lendamboi

What I refer to here as Lendamboi is the most internally diverse of all the languages discussed in this paper. Lynch and Crowley (2001:71–72) used the name Lendamboi to cover four mutually intelligible varieties recorded by Charpentier: Ayiaulien, Mbotkot, Natanggan, and Nioleien. These varieties "were spoken traditionally over a wide area of the interior of south Malakula", but "the language currently has a geographically discontiguous distribution as a result of post-contact movements of people from the interior to a variety of coastal locations" on the west, south, and east coasts (Lynch and Crowley 2001:71–72; see map 1).

3.2.1 Charpentier's data.
In three of these four lects, around 90 percent of nonhuman nouns take *na, and the remainder are unmarked. Only the westernmost of these four lects, Ayiaulien, shows significant accretion of *a: the marker is invariably *a*, irrespective of the nature of the vowel in the next syllable:

(12) POc   AYIAULIEN
     *tanoq   a/tan   'earth'
     *mata-   a/met-   'eye'
     *ikan   a/ik   'fish'
     *qusan   a/us   'rain'
     *pulu-   a/vəl-   'hair'

Almost no Ayiaulien nouns with *na, one of the few I have found being *salan > *na/sal* 'road'.

3.2.2 Shimelman's Newotenyene and Nevatanyene
Shimelman has a number of lists that relate to what I refer to as the Lendamboi language. In most of those labelled Letemboi and Nombotkote, the large majority of nonhuman nouns take *na, and nearly all the rest are unmarked. However, her Letemboi Newotenyene list shows a fifty-fifty distribution between *na (> *na, ne, ni* with no clear conditioning) and *a (> *a, e, i* also with no clear conditioning). (The figures are similar for her Nevatanyene list.)

(13) POc   NEWOTENYENE   **but**   POc   NEWOTENYENE
     *tanoq   ne/tene   'earth'   *tali   e/tal   'rope'
     N*kabati(ao)   n/ebasi   'moon'   *bʷatu (k)   i/ba:ti- 'head'
     *salan   ni/sele   'road'   *tinaqe   e/səne   'intestines'

There appears to be no phonological or semantic explanation for which nouns take *na and which *a.

3.2.3 Shimelman's Novol
Shimelman indentified one other "language" which I believe belongs under Lendamboi. This is labelled Novol, with two lists from the Dixon Reef area, and one from Mbangir on the east coast; we know from one of Charpentier's maps (1982:25) that Mbangir almost certainly represents a resettlement of Mbotkot. In these three lists, *a predominates, but around one-quarter of nouns take *na. Again, no conditioning is apparent. Here are some examples from the Tavendrua Novol list.

(14) POc   TAVENDRUA   **but**   POc   TAVENDRUA
     *takuRu-   e/tak-   'back'   *tasik   ne/taz   'sea'
     *mata-   e/met-   'eye'   *mʷata   nas/mʷote 'snake'
     *pulu-   e/vəl-   'hair'   *patu   ne/vat   'stone'
     *kutu   a/xut   'louse'   *(q)aca-   næ/ks-   'name'

3.3 Nasarian

Nasarian—which Charpentier suggests is /na-sar-ián_a/ (ART-speak-NMLZ) 'language'—was spoken, before 1949, about three hours walk inland from Dixon Reef on the west coast. The community fractured, with a migration to Dixon Reef—Tryon's Dixon Reef II list (1976)—and another first to Tembimbi on the east coast and then back to the west coast, to Tisvel village. Charpentier (1982:25) found only 5 speakers, all bilingual in some other nearby language. Charpentier's data suggest that nearly all nonhuman nouns have accreted *na, and nearly all of those have invariant *na-* irrespective of what follows. Most of the roots which were unmarked in Avava and Aveteian take *na-* in his data.

Shimelman (n.d.), however, has two lists labelled Nasarian: one spoken in Ngava and the other in Wileven. both southwest of but very close to Dixon Reef. It is not known how the speakers of these two lists connect with those interviewed by Charpentier 35 years earlier, but there is no reason to believe that we are dealing with a different language that has been given the same name: geographically, they "belong". (And even if it turns out that they are not Charpentier's Nasarian but some other similarly named language, they present an interesting picture of *a-accretion.)

Whereas Charpentier's data showed that almost all nonhuman nouns took *na-, these two lists are very different. The data are as follows:

(15)		Total nouns	nV-	V-	hV-	zero	
Nasarian: Wileven	69	3	20	40	6	
Nasarian: Ngava	66	4	41	—	20	(+ one *la-* ?)

Wileven shows a predominance of nouns with initial *h*V- where Ngava has just V-: Wileven *he/dele-n* vs. Ngava *e/ⁿtele-n* 'his/her neck', *ha/gut* vs. *a/xut* 'louse', *ha/tas* vs. *a/tas* 'sea'. This *h* also occurs initially on most verbs in Shimelman's Wileven list: verbs in her lists are generally given in the 3SG form, and are thus usually prefixed with *i-* or *e-*, but most of her Wileven verbs are prefixed with *hubu-* or a similar form, whose meaning I do not know. These occurrences of *h* seem to be the only ones in the language: that is, none of the noun or verb *roots* contain *h*. I will therefore treat this *h* as a phonetic accretion—rather along the lines of the way some speakers of Bislama pronounce vowel-initial words with an initial *h*: *hae* for *ae* 'eye', *heg* for *eg* 'egg', etc.

The article deriving from *a has two main reflexes: *(h)a-* and *(h)e-*: *(h)e-* seems to only occur when the next vowel is *e*, while *(h)a-* occurs elsewhere, but also occurs in some cases where the next vowel is *e*.[6] Examples:

(16) | POC | WILEVEN | NGAVA | |
|---|---|---|---|
| *pulu- | e/vle-ne | e/velel | 'hair' |
| *draRaq | he/dre | e/dre | 'blood' |
| *tanoq | ha/tene | a/tan | 'earth' |
| *salan | a/iale | a/sal | 'road' |
| — | a/vwigai | a/vikal | 'bone' |
| — | a/movəne | a/moʙi | 'egg' |
| *kutu | ha/gut | a/xut | 'louse' |

The accretion of *na is rare, and does not operate consistently across the two lects, suggesting that no conditioning can be established:

(17) | POC | WILEVEN | NGAVA | |
|---|---|---|---|
| *(q)aca- | na/kia-ne | a/kse-na | 'name' |
| *waiR | na/wo | o/wo | 'water' |
| —[7] | na/tavse-ne | a/tevsi | 'flower' |
| *lipo- | he/levə-m | ne/luvʷo-n | 'tooth' |
| *tinaqe- | a/sna | ni/sine | 'intestines' |
| *susu- | a/u | na/sis | 'breast' |
| *kamaliR | [ha/matal] | na/xamal | 'house' |

---

6   Ngava also shows *o*, as in *waiR 'water' > *o/wo*.

7   Available data suggest a Proto-Southern Oceanic reconstruction *tawaji: cf. Mwotlap *na/tawhi*, Mota *tawas*, Avava *a/tah*, Nāti *ne/tüvüh*, Sye *tasisi*, Kwamera *tihi-*, Anejo̱m̃ *n/tesia-*.

Finally, Shimelman has a list labelled Angavae, collected in Dixon Reef (which as mentioned before is linguistically quite a mixed village as a result of various movements of people). This may be either a variety of Nasarian, or else a previously unrecognised language closely related to it. As with Wileven, most nouns take *hV-* (and many verbs are also *hV-*initial in the, presumably, 3SG form listed there), and I suggest that the presence of *h* has the same explanation as in Wileven. There are very few occurrences of *nV-*, though a few nouns are unmarked. The examples in (18) are of *a-marking (18a), *hV-*marking (18b), and *na- marking (18c).

(18) POC   ANGAVAE   POC   ANGAVAE
  a. *para-   a/vara-   'hand'   b.   *kutu   ha/agt   'louse'
      *kayu   a/gai   'wood, tree'   *pulu-   he/vele-   'hair'
      *tali   a/tal   'rope'   *taliŋa-   hi/deliŋa-   'ear'
      *draRaq   e/drei   'blood'   *bʷatu(k) hu/baði-   'head'

  c. *(q)aca-   na/kia-   'name'
      *waiR   na/wo   'water'

3.4 Zero accretion

In this section, I examine cases in the languages discussed in §3.1–3.3 in which neither *na nor *a has been accreted to the noun root: in other words, cases like the Avava nouns discussed in §2.2 in which just the bare noun root appears, and which I will refer to by the expression "zero accretion". We can divide these into three groups.

First, there are lects in which there is almost no zero accretion. These all occur in Charpentier's lists: almost all nonhuman nouns occur with either fused *na (in the Mbotkot, Nioleien, and Natanggan dialects of Lendamboi, and Charpentier's Nasarian), or else with fused *a (in the Ayiaulien dialect of Lendamboi).

Second, there are lects where zero accretion occurs, but where no conditioning can be easily demonstrated.

(a) In Wileven and Angavae (Nasarian), zero accretion is relatively uncommon, and no pattern is discernible.

(b) In Nevatanyene (Lendamboi), it occurs quite frequently, but appears to be almost random. I can find no phonological conditioning, nor does semantics help: some nonhuman animate nouns have zero accretion, along the lines suggested in (1)—for example, *lipax* 'dog', *ləpo* 'rat', *tilao* 'fish', *siləs* 'worm'—but others have an accreted article, like *ni/mwat* 'snake', *na/nəx* 'bird', or *ni/lao* 'spider'. In addition, some human nouns do take the accreted article: *e/mægut* 'man', *e/tuvwate* 'woman', *a/teiborai* 'child'.

Third, there appear to be four cases where the absence of an article is conditioned by the nature of the initial consonant of the root: Newotenyene and Novol (Lendamboi), Aveteian, and Ngava (Nasarian) show no article when the root begins with a voiced prenasalised stop or trill. Some examples are given in (19): nouns whose roots fit this structural description for zero accretion but which nevertheless take an article are enclosed within square brackets.[8]

(19) POC   NEWOTENYENE   NOVOL (BANGIR) AVETEIAN   NGAVA
      *bʷatu(k)-   [i/bʷa:ti-]   bʷa:ti-   bat, bat-   batu-   'head'
      N*ganisu-   garse-   ga:rse-   gars-i   garse-[9]   'nose'
      N*daliŋa-[10] deliŋa-   deliŋa-   deliŋi-   deliŋa-   'ear'
      N*bʷaŋo-   [i/bʷə-]   [a/bʷoŋ]   boŋ-   boŋo-   'mouth'
      —   [ə/burman]   buruman   burman   —   'dog'
      —   deŋə   del   [i/del]   —   'smoke'
      *pʷilak   belebele   belebele   [a/ble]   belbe   'lightning'

---

[8] There are a few cases that are exceptions to this generalisation, where an article precedes a voiced stop (i.a), or where a noun with some other initial consonant takes no article (i.b):

(i)   POC   NEWOTENYENE NOVOL (BANGIR) AVETEIAN NGAVA
  a.   *sulati   e/dela:s   e/dela:s   —   a/dalas   'worm' (*s > d irregular but widespread)
  b.   —   melase   milazë   melase   melaas   'wind'

[9] The *r* reflex of *n in all four of these items is unexplained, but occurs in a number of other Western Malakula languages, including Navwien, Nasvang, and Nisvai.

[10] PNCV reflects POC *taliŋa- 'ear' as *daliŋa-.

My initial reaction to these forms is that a kind of dissimilation had taken place: *na + N > N (where N is a prenasalised consonant). Whether this explanation is valid for languages that have *a rather than *na, though, is problematical.

4. DISCUSSION

Table 1 outlines the relevant facts concerning the major speech varieties that I discussed in sections 2 and 3.

TABLE 1. ARTICLE ACCRETION IN THE LANGUAGES UNDER DISCUSSION

| LANGUAGE	"DIALECT"	*a OR *na?	*a > *ha?	ZERO
Avava	[Crowley's data]	*a with about 25% of nouns (*na rare)	no	widespread with roots of 2+ syllables
	Umbbuul			
Aveteian	[Tryon's data]	*a predominates, *na not rare	no	about 10%, before voiced stops?
Lendamboi	Mbotkot, Nioleien, and Natanggan	*na	no	very rare
	Ayiaulien	*a on almost all nouns, *na very rare		
	Newotenyene	*a and *na 50-50		about 25%, before voiced stops
	Nevatanyene	*na 45%, *a 23%		about 30%, conditioning not known
	Novol	*a predominates, *na not rare		about 20%, before voiced stops
Nasarian	[Charpentier's data]	*na	no	very rare
	Ngava	*a (*na rare)		about 30%, before voiced stops
	Wileven	*a (*na rare)	about 60%	about 10%, conditioning not known
	Angavae	*a (*na rare)	about 60%	about 15%, conditioning not known

Legend

▨	only or predominantly *a	▨	zero accretion with roots of 2+ syllables
▨	only or predominantly *na	▨	zero accretion before voiced stops
▨	neither *na nor *a predominates	▨	zero accretion apparently unconditioned
▨	*ha instead of *a		

Figure 1 shows a partial family tree of Malakula languages, featuring the languages that show *a-accretion—in bold capitals—along with their closest relatives (Lynch 2016a,b). An asterisk following the name of the language indicates that only some (dia)lects of this language show *a.

In all branches of the Malakula subgroup except for Southwestern Malakula, the accreted article is *na: there is no evidence of *a at all. This might suggest that, relatively early, the "competition" between *a and *na was "won" by *na. But it is clear from looking at this family tree that both *na and *a must have been present in Proto-Southwestern Malakula (PSWM), to account for the presence of (i) *a in Aveteian and Avava, (ii) *a and *na in Lendamboi and Nasarian, and (iii) *na in Navwien, Nahavaq, and Naha'ai. But if both were present in POC *and* in PSWM, both also must have been present in every intervening protolanguage—that is, on the tree in figure 1, Proto-Southern Oceanic, Proto-Malakula (if such existed), Proto-Western Malakula, and Proto-Peripheral Western Malakula.

This, plus the fact that *a (and not *na) is reflected in Paamese, another member of the Central Vanuatu subgroup, implies that both *na and *a were present together in a number of intermediate languages spoken long after the break-up of POC. The Malakula data, unfortunately, do not cast any light on what differences there might have been between *na and *a. But they suggest that both coexisted until relatively recently, and that although *a was usually the one to drop out, this did not always happen. Indeed, in a number of lects that I have examined in this paper, that "competition" seems to be still going on.

I also examined what I referred to as zero-accretion: nouns that do not take the accreted article. Crowley suggested that the presence or absence of the article in Proto-Oceanic was rather fuzzily semantically conditioned: as shown in (1), human nouns did not take the article; some nonhuman animate nouns took the article, but most

did not; and most inanimate nouns took the article, though some did not. This does not seem to be the case in the languages I examined. I found no semantic conditioning in relation to zero accretion; and indeed there are a number of cases in many of these languages where nouns referring to humans take the article; as exemplified in (20).

FIGURE 1. PARTIAL MALAKULA FAMILY TREE
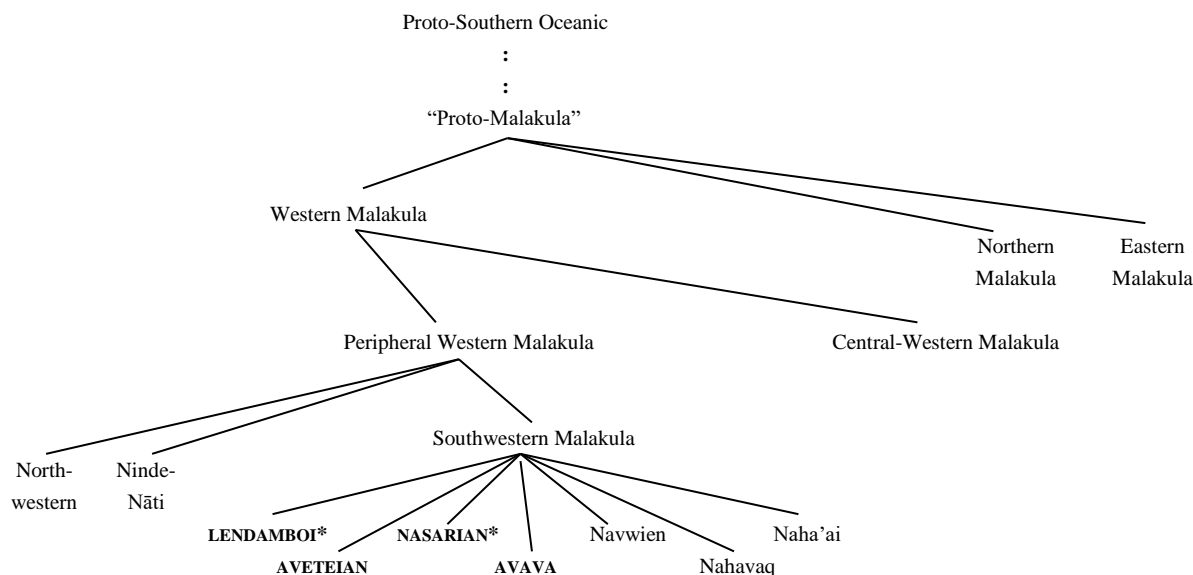
(20)	NOVOL (BANGIR)		but cf.	AVAVA
	na/muɣut	'man'		muut
	ne-tuvwote	'woman'		momok
	a-teiʙorai	'child'		at- (< *natu; initial *a* is part of the root, not the article)

In some of the languages I examined, the presence or absence of the accreted article seems to have been conditioned phonologically rather than semantically (although in other cases I could not establish any conditioning at all). There seem to be two different types of phonological conditioning occurring.

1. In Avava, the syllabic structure of the root is the important factor: a one-syllable root will have the article, a root of two or more syllables will not—see §2.2 above. This is not unique to Avava. As pointed out in §2.2, this also occurs in Unua. In fact, it occurs in quite a number of languages of the Eastern Malakula linkage (Unua, Uripiv, and probably Port Sandwich, Banam Bay, Aulua, and Vao), as well as some Western Malakula linkage languages (Naman, Tirax, and Nahavaq), and probably also Paamese. Further investigation may turn up other languages that follow this pattern.
2. In Aveteian, Newotenyene and Novol (Lendamboi), and Ngava (Nasarian), there is zero accretion when the root begins with a voiced prenasalised stop or trill. Something at least partly similar appears to occur in Uripiv. Here, *b is usually reflected as *p*, *bʷ as *pʷ*, and *t and *t* (21a). When preceded by *na, however, *b > *b* and *bʷ > *bʷ* (though *t remains *t*): see (21b). And when the article is omitted, following the Unua pattern described above and illustrated in (6), the initial consonant is voiced and prenasalised (*b*, *bʷ*, *d*), suggesting assimilation to the article before it was lost: see (21c).

(21)	POc		URIPIV	
a.	*barapu		periv	'long'
	ᴺ*bʷaro		o-pʷur	'unripe'
	*tanum		e-tanu	'bury'
b.	*boŋi 'night'		na/boŋ	'day'
	ᴺ*bʷea		na/bʷe	'slitgong'
	*toqa		na/to	'fowl'
c.	*bati-		bati-	'canine tooth'
	ᴺ*bʷero-		boro-	'ear'
	*taRutu(m,ŋ)		daut	'porcupine fish'

Again, further research is needed to see whether this pattern is more widespread.

5. CONCLUSION

Both forms of the POC common article, *na and *a, were inherited into Proto-Malakula, and both are still reflected in the Southwestern Malakula linkage; indeed, in a couple of cases, both are still reflected in the same language. This double retention is unusual within the wider Oceanic context, and might raise hopes that we should be able to discern any differences in function between *a and *na.

This situation implies that, in Vanuatu languages where one form was lost, that loss must—at least in many cases—have been quite recent. Unfortunately, the languages in which both forms of the article are retained shed no further light on what might have been the difference(s) between them. We are left then with more cases of *a, an almost unique situation in which both *a and *na occur in the same languages, but no further advances in explaining any differences between them.

The absence of the article before inanimate nouns—so-called zero accretion—appears not to be due to semantic factors in these languages, but rather to phonological factors, though there is no single common phonological conditioning running through all the Southwestern Malakula languages.

References

Charpentier, Jean-Michel. 1982. *Atlas linguistique du sud Malakula / Linguistic atlas of south Malakula*. Paris: SELAF.

Crowley, Terry. 1985. Common noun phrase marking in Proto-Oceanic. *Oceanic Linguistics* 34:135–93.

———. 2006. *The Avava language of Central Malakula (Vanuatu)*. [Ed. by John Lynch.] Canberra: Pacific Linguistics.

François, Alexandre. 2000. Vowel shifting and cloning in Motlav: Historical explanation vs. formal description. In *Proceedings of the Seventh Annual Congress of AFLA (Austronesian Formal Linguistics Association)*, ed. by Marian Klamer, 49-68. Amsterdam: Vrije Universiteit.

Lynch, John. 2001. Article accretion and article creation in Southern Oceanic. *Oceanic Linguistics* 40:224–46.

———. 2003. Low vowel dissimilation in Vanuatu languages. *Oceanic Linguistics* 42:359–406.

———. 2016a. Malakula internal subgrouping: Phonological evidence. *Oceanic Linguistics* 55:399–431.

———. 2016b. Numeral systems, internal subgrouping, and language contact in Malakula. *Language and Linguistics in Melanesia* 33(1):95–106. http://www. langlxmelanesia.com/LLM%20Vol.%2034%202016_Lynch%20JUSTIFIED. pdf. [Accessed July, 2017.]

Lynch, John, and Terry Crowley. 2001. *Languages of Vanuatu: A new survey and bibliography*. Canberra: Pacific Linguistics.

Lynch, John, Malcolm Ross, and Terry Crowley. 2002. *The Oceanic languages*. Richmond, UK: Curzon Press.

Pearce, Elizabeth. 2007. The reflexes of Proto-Oceanic *na in Unua. In *Language description, history and development: Linguistic indulgence in memory of Terry Crowley*, ed. by Jeff Siegel, John Lynch, and Diana Eades, 327–39. Amsterdam: Benjamins.

Shimelman, Aviva. n.d. Malakula word lists. http://www.soundcomparisons.com/#/en/Malakula/map/Lgs_All.

Tryon, D. T. 1976. *New Hebrides languages: An internal classification*. Canberra: Pacific Linguistics.