This is a preprint version of the following paper:
François, Alexandre. forthcoming.
Noun articles in Torres and Banks languages: Conservation and innovation.
In Jeff Siegel, John Lynch and Diana Eades (eds.), *Linguistic Description and Linguistic Applications: Studies in Memory of Terry Crowley*.
New York: John Benjamins.

Noun articles in Torres and Banks languages: Conservation and innovation

Alexandre FRANÇOIS

Langues et Civilisations à Tradition Orale – CNRS
⟨Alexandre.Francois@vjf.cnrs.fr⟩

Abstract

Unlike most Vanuatu languages further south, the seventeen languages spoken in the Torres and Banks Is still productively make use of noun phrase articles. This paper describes the morphology, syntax and semantics of these article systems. While these languages have preserved the contrast between "common" and "personal" phrases – often preserving the POc markers *na and *i/*e – they also have undergone several change patterns. Some articles, while remaining syntactically free, have become prefixes; in a few cases, total accretion to the noun root eventually took place. Also noteworthy is the creation of a new common article *wo: in five languages, *na has been restricted to inalienable non-human nouns with a human specific possessor – which makes it typologically quite an unusual article.

Noun articles in Torres and Banks languages: Conservation and innovation

1 Presentation

In his article "Common noun phrase marking in Proto Oceanic" (1985), Terry Crowley described the various ways in which the common noun article *a/*na evolved across the modern languages of Oceania. While some languages are conservative, others have lost all traces of the original NP marker, or have only retained it partially. Quite often, he showed, what was once a fully productive article became accreted to the noun root, progressively losing its syntactic status as an article.

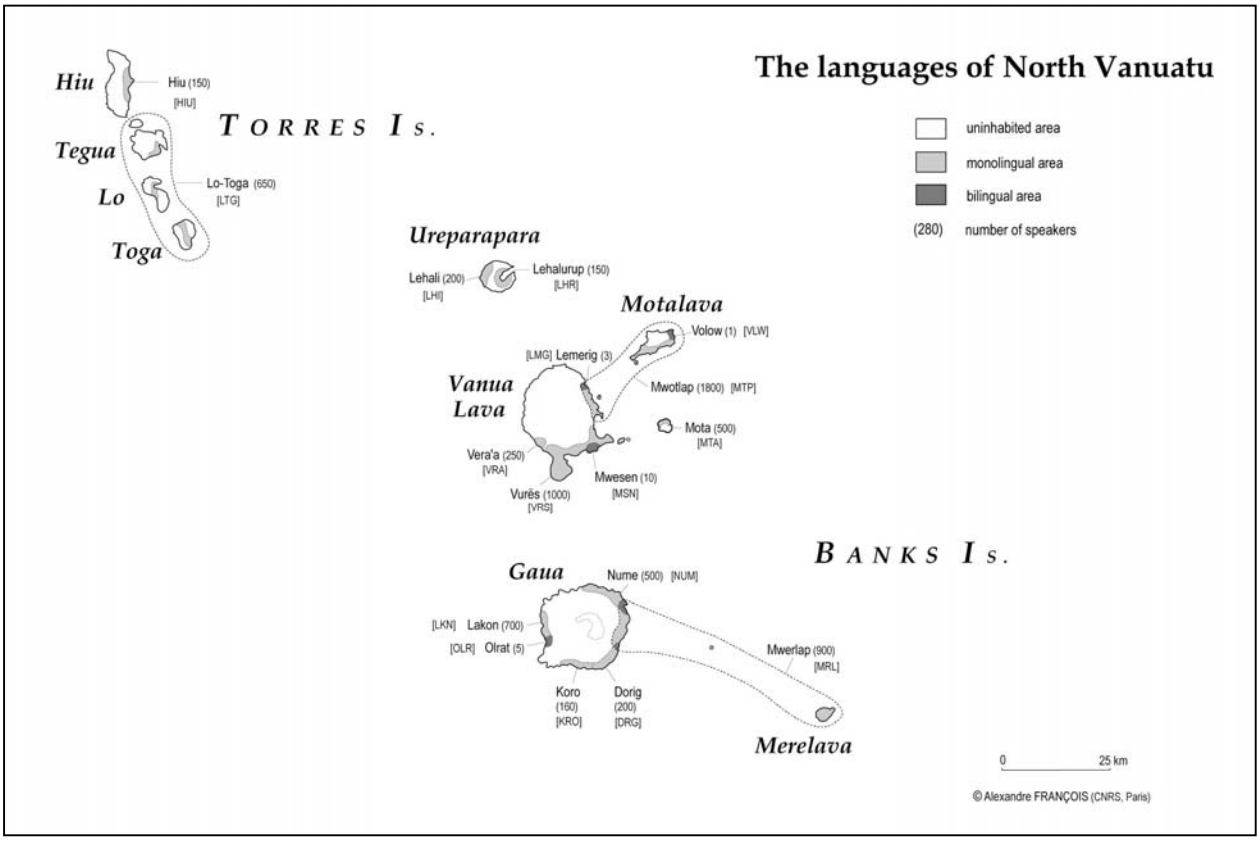
The languages of Vanuatu are sometimes described as having essentially followed this evolution pattern of article loss, whether involving accretion to the noun root or not. Lynch (2001) says: "Very few languages of North and Central Vanuatu have articles per se", and Lynch, Ross & Crowley (2002: 38): "What was historically an article has in many of the languages of Vanuatu (…) been fused with the noun root, being morphologically inseparable in all, or at least most, morphosyntactic contexts".

While such statements are certainly true for other languages of Vanuatu, I will show in this paper that they hardly represent the two northernmost island groups of the archipelago, the Torres and Banks Is (a province usually abbreviated as "Torba"). The seventeen languages spoken in this area (*Figure 1*), which were still little known until recently,[1] present diverse but essentially similar systems of noun articles. Overall, they show few examples of complete article loss, and only four genuine cases of article accretion. On the contrary, what I observed is that Torba languages still make regular use of noun articles in a way reminiscent of their POc ancestor, albeit in a different manner.

Remarkably, this was Terry Crowley's own insight when he drew his map of article

[1] The data cited in the present paper were collected by the author during three field surveys: May-July 1998 for Mwotlap, Vurës and Mwesen; July-September 2003 for Volow, Vera'a, Lemerig, Nume, Dorig, Koro, Olrat, Lakon and Mwerlap; July-August 2004 for Mota, Lehali, Lo-Toga and Hiu. Information on Lehalurup comes from Codrington (1885). Note that I use here the term 'languages' in the broad sense of 'speech varieties', regardless of whether some can be grouped together as dialects of a single language.

Figure 1 – The languages of North Vanuatu



retention in Oceania (1985: 162), despite terrible gaps in language documentation in those times. In a way, the present paper can be seen as an occasion to confirm and refine his correct intuition, by making available the first-hand data he didn't have access to twenty years ago. In addition, I will include here information on other articles, and describe certain patterns of innovation, whether morphological or syntactic, that these Torba languages have gone through.

After describing the form of NP articles (*Section 2*), I will analyse their distribution across noun categories (*Section 3*), and finally discuss their syntactic functions (*Section 4*).

2 The morphology of articles

The basic pattern attested in north Vanuatu, as in various other parts of Oceania, is for nouns – whether definite or not – to require a preposed article in order to form a valid NP (see *Section 4*). Thus in Lo-Toga, common nouns must be preceded by the invariable article /nə/: e.g. ***nə*** *rəɣa* 'a/the tree, Ø/the trees, Ø/some/the wood'.[1]

[1] Throughout this article, forms are transcribed phonemically rather than using standard orthographies, to enable comparison. Note that /v/ = [β], and that all voiced stops are prenasalised: /b/ = [^{m}b], etc.

For reasons of length, I will choose to focus my study on major NP articles, that is, those shared by entire categories of nouns in each language, and made obligatory in most syntactic contexts.[1] After an inventory of these articles (§2.1), I will describe the diversity of their morphological statuses (§2.2).

2.1 Inventory of articles across Torba languages

The choice of articles is governed by the semantics of the following noun, according to such criteria as *singular* vs *plural*; *common* vs *personal*; *directly possessed* vs *non-directly possessed* nouns (see *Section 3*). A comparative list of these articles is proposed in *Table 1*.[2]

Table 1 – Inventory of noun articles in the Torba languages

abbr.	*language name*	*directly possessed common N*	*other common N*	*human personal singular N*	*human plural N*
HIU	Hiu	nə		Ø	təkŋ^{w}a
LTG	Lo-Toga	nə		Ø	hək^{w}ərə
LHI	Lehali	n-		Ø	k^{w}ɔy
LHR	Lehalurup	n(V)-		i (?)	dɛ (?)
VLW	Volow	n(V)-		Ø	iɣɛ
MTP	Mwotlap	nV-		Ø [i/]	iɣɛ
LMG	Lemerig	n-		Ø	irɣɛ
VRA	Vera'a	(ɪ)n [n/]		ɛ	ɛ raɣa
VRS	Vurës	na	ɔ	i	i rɛɣɛ
MSN	Mwesen	ɔ		ɛ	ɛ ra
MTA	Mota	na	o	i	i ra(ɣai)
NUM	Nume	na-	u [w/]	Ø	ra mɛl
DRG	Dorig	na	ɔ	i	i ra
KRO	Koro	na	ɔ	i	i rat
OLR	Olrat	Ø [n/]		Ø	nɪy wʊ
LKN	Lakon	(ɪ)n		i	ɣɪː wɪː
MRL	Mwerlap	nV-		i	rɛ

[1] I will therefore leave unmentioned here certain (quasi) articles with a more limited scope. For example, (1) certain proper names and kinterms reflect a vestigial feminine "article" *ra/ro; (2) Mwotlap has a partitive *tɛ* < *tewa 'one'; (3) four Gaua languages possess an indefinite article, grammaticalised from *tuara 'other'; (4) Torres languages have a set of human markers, used both as pronouns and as articles; etc.

[2] Hyphenated forms indicate prefixes, otherwise the article is a clitic. Forms in square brackets and followed by '/' correspond to reflexes that are only vestigial, taking the form of a phoneme that is now incorporated into certain noun roots (see §2.2.2).

As *Table 1* shows, the number of articles in each language is variable. Olrat has basically no article at all, except for the plural marking on human nouns. Vurës, Mota, Dorig and Koro possess as many as four distinct articles. Other languages have either two or three.

The various forms listed here may be reduced to a small number of etyma (see Lynch *et al.* 2002: 71).

The POc article *a/*na is reflected – always under its form *na – in essentially all the languages of the area. Most often it is still a productive full-fledged article, whether a clitic or a prefix. In Olrat, it is only reflected as an accretion on certain nouns. Remarkably, Vera'a reflects it both as an accretion and as a free article (§2.2.2). Only Mwesen has lost all traces of *na (§3.1).

Six languages make use of an article of the form *ɔ*, *o* or *u*. Available evidence suggests it should be reconstructed as *wo.[1] This article *wo is obviously an innovation, both in formal and functional terms: see §3.1.

The POc personal article *i/*e is represented in as many as ten languages, reflecting either *i or *e. Out of these ten reflexes, nine are still productive, while one (Mwotlap) only exists as an accretion on certain nouns (§2.2.2). Most languages have lost *i/*e altogether: their personal NPs take zero article. Note that no language retains the other POc personal article *qa.

Finally, the maximum formal diversity is found with human plurals. However, knowledge of the regular phonetic correspondences throughout the Torba area (François 2005b) suggests all these attested forms – except HIU *təkŋʷa* and LHR *dɛ* – include a reflex of a syllable *ra, albeit in a hidden way (e.g. LKN *ɣɪː* < **ɣ*ɪr < *ki**ra**). This obviously corresponds to POc *ra marking 3rd person (normally human) plural. In six languages (Volow, Mwotlap, Lemerig, Vera'a, Vurës, Mota), the form regularly reflects a phrase **i/e ra kai*, which can be analysed as Personal *i + plural *ra + POc *kai 'native, person' (Pawley 1976): e.g. VLW/MTP *iɣɛ* < *iyɣɛ* < *irɣɛ* < *i-raɣai < *i ra kai.

2.2 Morphological status

2.2.1 From clitic to prefix

All the articles cited in *Table 1* are immediately followed by the noun which is the head of the NP. Not surprisingly, the two morphemes together form a prosodic unit. The whole noun phrase follows a single intonational contour, whereby the article lacks its own primary stress,

[1] The reasons for reconstructing an initial consonant *w include: (a) the accretion of /w-/ to certain roots in Nume (§2.2.2); (b) the existence in Mwotlap of an article-like prefix forming honorific nouns or nicknames, with the form *wo-* (François 2001: 242); (c) the form of the noun article *we* ~ *wu* in the Banks "song dialect" (Codrington 1885: 309), an archaic poetic language common to the whole area.

being prosodically integrated to the following noun: e.g. LTG *nə=ɣəˈhuwə* ‘rat’. This status of noun articles as proclitics, which must probably be reconstructed also for POc (e.g. *na=kaˈsupe), is still well attested in Torba languages. For example, all the human articles – except obviously in their accreted versions – whether singular or plural, still behave as clitics; and so do the reflexes of *wo. But the situation is more complex regarding *na.

In six languages (Hiu, Lo-Toga, Vurës, Mota, Dorig, Koro), *na still shows all the properties of a clitic. In Vera’a and Lakon, this clitic has the basic form /ɪn/, but undergoes inverse elision (aphaeresis) after a vowel-ending word: compare VRA *ɣɛn* ***ɪn*** *mɛs* ~ LKN *ɣœn* ***ɪn*** *mœh* ‘eat a fish’ with VRA *lɛ͡* ***n*** *mɛs* ~ LKN *lœ͡* ***n*** *mœh* ‘take a fish’. Despite this rule of sandhi depending on the preceding word, the syntactic scope of this article /ɪn/ is clearly the following noun: it can still be described as a proclitic.

In seven other languages (Lehali, Lehalurup, Volow, Mwotlap, Lemerig, Nume, Mwerlap), the prosodic incorporation of the article eventually triggered also its phonological integration to the following noun, so that it must now be considered a prefix. There are three ways for the article to form a single phonological word with the following noun (see François 2000; 2005b):

1. VOWEL HARMONY [Lehalurup, Volow, Mwotlap, Mwerlap]: the vowel of *na regularly undergoes partial or complete assimilation to the first vowel of the noun: e.g. *na kutu ‘louse’ > VLW/MTP *n****i****-ɣit*.
2. VOWEL ELISION [Lehali, Volow, Lemerig]: the vowel of *na is deleted, so that the article is reduced to the consonant *n*-. These three languages don't only allow this elision before another vowel, which also happens elsewhere (e.g. *na ikan ‘fish’ > LTG *n’ iɣə* ~ MRL *n-ɛ͡aɣ*), but also before a consonant (e.g. *na patu ‘stone’ > LMG *n-vɛʔ*).
3. PHONOTACTIC INTEGRATION [Mwotlap, Nume, Mwerlap]: *na is regularly followed by consonant clusters, which otherwise never occur word-initially in the language. Thus in Mwotlap, compare the form with article *na-ɣhʊw* ‘rat’ < *na kasupe with the bare noun *ɣ****ʊ****hʊw* showing epenthesis.

If an article meets any of these criteria, then it is a prefix. Conversely, when the phonological form of the article and that of the noun are independent from each other (e.g. LTG *nə=ɣəhuwə* ‘rat’), one may still speak of a clitic.

Crucially, the change in morphological status – from clitic to prefix – doesn't necessarily involve any change in syntactic behaviour. Thus, the prefix *nV*- in Mwotlap can still be analysed as a genuine article, just as much as its clitic ancestor *na. The lack of a clear distinction between the morphological and syntactic levels has led certain scholars to confusion. For example, Crowley (2002: 591) had this erroneous analysis about Mwotlap: “There are no articles in Mwotlap. The original prenominal article /*na/ has been reanalysed as part of the citation form of the noun”.

Only syntactic properties should define the status of a morpheme as a productive article, such as:

- its capacity to affect the whole noun lexicon, or at least entire, definable, categories of this lexicon (e.g. inanimate nouns), including new lexical items when they fall into these categories
- the existence of productive rules governing the presence vs absence of this article, depending on the syntactic context.

The remainder of this paper (*Sections 3* and *4* respectively) will demonstrate that these two requirements can be fulfilled by a prefix just as much as a clitic.

2.2.2 Article accretion

The historical phenomenon of "article accretion" brings about a different situation altogether. In this case, what was once a free article has been attracted not only to the following phonological *word*, but to the following *root* itself, to such an extent that it has lost its status as an article. While this process seems to have occurred quite often throughout Vanuatu (Crowley 1985, Lynch 2001) and in New Caledonia (Ozanne-Rivierre 1992), in the Torba area it is only attested in a few cases.

In Vera'a and Olrat, the article *na has been accreted to vowel-initial noun radicals. To take reflexes of POc *(na) quraŋ 'lobster', one must carefully distinguish between, on the one hand, such forms as MTP *n-ɪy* ~ MRL *n-o͡əɾ* where the article is still syntactically a free prefix; and, on the other hand, VRA *n/irɪ* ~ OLR *n/urɪŋ*, where it has become an inseparable part of the noun radical.[1] Examining such forms in the light of the two criteria stated above (end of §2.2.1) makes it clear we are no longer dealing with articles. First, this consonant /n/ does not affect productively any noun category that would be definable in synchrony, but is arbitrarily present in certain items of the lexicon. Second, there is essentially no syntactic context that allows for regular deletion of this /n/. These words now behave exactly the same as any **n*-initial noun. For example, and quite remarkably, Vera'a allows them to take the common article *ɪn*, also a reflex of *na (§2.2.2): e.g. ***ɪn n***/*irɪ* 'a/the lobster'.

Similar cases of accretion occurred with other articles. In Nume, vowel-initial noun radicals also underwent article accretion, but this time with *wo, under the form /w/: e.g. NUM ***wɛv*** 'fire' < POc *api; ***wʊw*** 'turtle' < PNCV *ʔavua; ***wɪɣ*** 'fish' < POc *ikan; ***wɛm*** 'mat bed' < POc *qebal; ***wɪm*** 'earth oven' < POc *qumun; ***wʊr*** 'lobster' < POc *quraŋ; ***wak*** 'canoe' < POc *waga. In this case too, the accreted /w/ may co-occur with the free article *u*, itself a reflex of *wo: e.g. ***u w***/*ʊr* 'a/the lobster'.

[1] Other examples include: VRA ***nɛv*** 'fire' < POc *api; VRA ***nɛr*** 'Casuarina' < POc *aRu; VRA ***nuwʊ*** ~ OLR ***nʊw*** 'turtle' < PNCV *ʔavua; VRA ***niŋ͡m͡ʷɪ*** 'house' < POc *Rumaq; VRA ***nʊn*** 'sand' < POc *qone; VRA ***nur*** 'Spondias cytherea' < POc *quRis; OLR ***num*** 'earth oven' < POc *qumun; VRA ***naka*** ~ OLR ***nak*** 'canoe' < POc *waga.

Finally, Mwotlap only reflects the personal article *i as a fossil vowel at the beginning of about twenty human nouns, essentially kinterms and a few proper names (François 2001: 208-213): MTP ***i****thi-k* ‘my brother’ < POc *i taci-gu; ***i****mam* ‘Dad’ < *i mama; ***I****k͡pʷɛt* ‘cultural hero’ (MTA *i K͡pʷat*) – see also ***i****ŋɛ* ‘plural article’ (§2.1).

3 Noun categories and their articles

Section 2 has shown both the unity and diversity of articles across the Torba area, at least regarding their form. As for the distribution of articles across noun categories in each language, it also shows some variety, yet allows for a general description. I will first describe the contrast between *na and *wo in the languages that have both (§3.1), and later will delineate the categories of common vs personal nouns (§3.2).

3.1 A specific article for inalienable possession

Table 1 showed the existence of an article *wo in six geographically adjacent languages. Not only is this form unknown outside this small “central Banks” area; but its precise distribution is also, to my knowledge, unusual both from an Oceanic and from a typological perspective. It clearly results from a local innovation, which either never took place in the neighbouring languages, or did and was later reversed.

In five of these languages (Vurës, Mota, Nume, Dorig, Koro), the article *na has been restricted to only one category of nouns, namely, those [-human] nouns that are inalienably possessed, i.e. directly followed by a possessor (whether a suffix or an NP). The innovative article *wo is used otherwise, that is, with [-human] nouns that are *not* directly possessed: contrast VRS ***na*** *ŋærvü-k* ‘my house’ vs ***ɔ*** *ŋøvür* ‘a/the house’.

To be precise, a noun normally combines with *na* only if it has a specific human possessor. Thus compare for Dorig:

(1) DRG
a. ***na*** *ssa* *i* *vvɪ-k*
ART:DIR.POSS name ART:PERS mother-1SG
‘my mother's name’ (possessor [+spec] [+hum])

b. ***ɔ*** *ssa* *rk͡pʷa*
ART:COMM name woman
‘a woman's name’ (possessor [–spec] [+hum])

c. ***ɔ*** *ssa* *wasŋinnɛŋ*
ART:COMM name place that
‘the name of that place’ (possessor [+spec] [–hum])

Likewise, when an inalienable noun is marked for a generic possessor, it must take the *wo article: e.g. MTA ***na*** *pane-ŋm͡ʷa* ‘your hand’ vs ***o*** *pane-i* ‘one's/a hand’.

Furthermore, each of these languages has between four and six possessive classifiers, which behave as a subclass of inalienable nouns. Since their function consists in indexing a possessor, they almost always occur with the article *na*: MTA ***na*** *ma-k* ‘my X (Drink possession)’; ***na*** *ŋm͡wo-ŋm͡wa* ‘your X (General possession)’… Remarkably, the noun X which is modified by this classifier, although it is semantically possessed, must bear the *wo article, because it does not receive itself the possessive morphology: it is *indirectly* possessed. In these five languages, this regularly results in quite paradoxical NP structures, where the noun combines with *wo while its classifier takes *na:

(2) MTA ***o*** *tkp͡wei* ***na*** *ŋm͡wo-ra*
ART:COMM garden ART:DIR.POSS GeneralPoss-3PL
‘their garden’

(3) VRS ***na*** *ka-ŋ* ***ɔ*** *ak*
ART:DIR.POSS VehicPoss-2SG ART:COMM canoe
‘your canoe’

Obviously, the languages that only retain *na have no such asymmetry. Thus the equivalent of (2) in Volow would be ***n****-tɪgb͡wɪ* ***nɔ****-yɔ-y*, with two instances of *na.

In summary, these five languages allow *na only with a suffixable word – whether a noun or a classifier – that is directly possessed, either followed by a personal suffix or a semantically [+human] [+specific] NP possessor. In all other cases, the article is *wo.

Finally, out of the six languages that reflect *wo, Mwesen is original in having lost all traces of *na, and generalised *wo to all contexts. Thus Mwesen has ***ɔ*** *ɪŋm͡w* ‘a house’; ***ɔ*** *ɪŋm͡wɔ-k* ‘my house’; ***ɔ*** *sa-n ɛ rɪtnɔ-k* ‘my mother's name’; ***ɔ*** *pini-m* ‘your hand’; ***ɔ*** *mɔɣɔ-nɪr* ***ɔ*** *tukp͡wɪ* ‘their garden’.

3.2 Personal vs Common articles

Another functional notion that proves relevant for the description of NP articles in Torba languages, as indeed elsewhere in Oceania (Pawley 1972: 32), is the contrast between *personal* and *common* NPs. Formally speaking (*Table 1*), personal NPs are either marked by *i/*e or by *zero*; they contrast with common NPs, which take *na or *wo.

On the semantic level, a “personal” NP normally has a *human specific* referent. This is typically the case with proper names, kinterms, pronouns or deictics with human reference. Thus one finds MSN ***ɛ*** *Tɛvɪt* ‘David’; KRO ***i*** *mam* ‘Dad’; VRS ***i*** *kp͡wœlɣœ-k* ‘my father-in-law’; MTA ***i*** *nau* ‘I (1sg pronoun)’; VRA ***ɛ*** *sɪ* ‘who?’; DRG ***i*** *at nɛŋ* ‘the one there’… The plural articles given in *Table 1* often include the personal article *i/*e, because they are only used with human referents: MSN ***ɛ*** *tɛsɛ-n* ‘his brother’ ~ ***ɛ ra*** *tɛstɛsɛ-n* ‘his brothers’. Non-human nouns never take a personal article, except in stories where animals or objects are personified: MSN ***ɛ*** *ɣʊsʊw min* ***ɛ*** *ɣʊtʊ* ‘Rat and Hermit-Crab’.

To be precise, the contrast personal vs common somehow constitutes a semantic continuum, of which only the two ends are clearly defined. On the one hand, proper names, or kinterms with individual reference, must be treated as personal; on the other hand, non-human NPs must be treated as common. The situation is less clearcut in the intermediate zone – that is, non-kin human noun phrases. As a tendency, a given noun will be treated as personal if it points to a definite individual in the given context – e.g. MRL ***i*** *vatɔɣ* 'the teacher', DRG ***i*** *maɣtɛ* 'the (aforementioned) old lady', MSN ***ɛ*** *maranaɣ* 'the chief', VRS ***i*** *birŋi-k* 'my partner' – but as common if it points to a generic or indefinite referent, or to the notional quality of the noun (e.g. predicate 'be a N'):

(4) DRG *na* *m-tɛk* ***ɔ*** *maɣtɛ* *s-rʊ.*
1SG PRF-see ART:COMM old.woman NUM-two
'I saw two old ladies.'

(5) MSN *ɛ* *sɪ* ***ɔ*** *maranaɣ* *ɛlilɛ?*
ART:PERS who ART:COMM chief here

– ***ɛ*** *maranaɣ* *ɛ* *nɔ.*
ART:PERS chief ART:PERS 1SG
'Who's (the) chief here? – The chief, that's me.'

No clearcut principle can really be asserted here. First, discrepancies are common, whereby the same noun can equally be treated as personal or common – including in the same sentence (e.g. MRL ***i*** *bulsala-n* ~ ***nu****-bulsala-n* 'her boyfriend').

Second, certain human nouns appear to be just incompatible with the personal article, whatever their actual reference. This is especially the case of the four common nouns 'person', 'man', 'woman', '[non-relational] child', perhaps because these lexical items are statistically most often used with non-referential or qualitative value – whether as a generic NP, a predicate or an attribute. These nouns are systematically treated as common in all Torba languages, including when they clearly designate a specific individual:

(6) DRG ***i*** *ntʊ-ŋ* *nɛŋ* *sa,* ***ɔ*** *ŋm͡wɛrat* *sa…*
ART:PERS son-2SG that there ART:COMM man there
'that son of you there, the boy there…'

In other languages, this structural asymmetry takes the form of a contrast Ø (personal) vs *na (common):

(7) VLW (**Ø**) *niti* *na,* *gb͡wɛ* ***n****-taŋm͡wan* *na…*
(ART:PERS) son:2SG there that.is ART:COMM-male there
'that son of you there, the boy there…'

Although these four exceptional nouns always take a common article in singular, they become compatible again with personal marking in plural: e.g. VRS ***ɔ*** *ŋm͡wirŋm͡wiar* 'a child' → ***i rɛɣɛ*** *ŋm͡wirŋm͡wiar* 'children'. Finally, a similar paradox can be illustrated with the phrases meaning 'my wife'. When a language possesses a dedicated kinterm, then it is encoded as Personal: VRS ***i*** *ɣünø-k*; MTA ***i*** *rasoa-k*; MRL ***i*** *rənətu-k*. Otherwise, a periphrasis

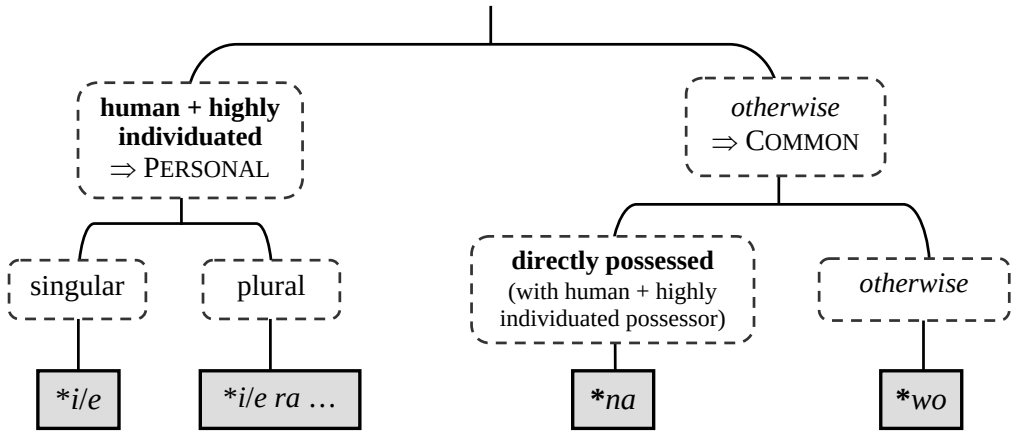
will be used with the noun 'woman' – in which case common articles are required: KRO ***na*** *mu-k ɔ rakp͡ʷa*; HIU ***nə*** *yəkʷen əkiə*.

In sum, personal articles (*i/*e or zero) are restricted to highly individuated human referents. Common articles (*na or *wo) are required in all other cases: that is, for non-human, non-specific, poorly individuated referents; and by extension, with certain nouns that are statistically seldom referential – even when they actually are.

3.3 Synthesis

This section has delimited the various noun categories that are relevant to explain the distribution of major articles (*na, *wo, *i/*e, zero) in Torba languages. One particular semantic class deserves attention here, namely that of *human specific, highly individuated referents*, as opposed to all the rest (François 2005a). Not only does this feature account for the contrast between personal and common articles; but it also helps define the type of possessor that requires *na vs *wo in those languages that possess two common-noun articles. The architecture of the article system in Torba languages – or better, of those systems with the maximum number of distinctions – is summarised in *Figure 2*.

Figure 2 – The maximal architecture of Torba article systems



4 The syntax of noun articles

The preceding sections have examined the form and nature of noun articles attested in the Torba area, as well as the semantic properties of nouns that are canonically associated with them. Each NP was cited with its article, as though it just had to be there. In reality, the status of these morphemes as true articles entails the possibility of observing which contexts require their presence vs their absence (see §2.2.1). This issue should help work out the syntactic *raison d'être* of these articles.

4.1 Syntax of the personal article

Even if a noun fulfills all the semantic requirements to be compatible with a personal (rather than a common) article, this doesn't mean this article will always be there.

First, most languages seem to treat the *i/*e article as more or less optional, even in those contexts where its presence would be expected. For example, in Dorig, 'their grandfather' in subject position is normally ***i*** *tbiraːr*, but my oral corpus shows several instances of just *tbiraːr* without the article.

Furthermore, while most syntactic functions require – or at least are compatible with – the personal article (subject, direct object, object of a preposition, possessor, predicate), at least one excludes it: the vocative. Following the framework developed by Lemaréchal (1989: 37) based on Tagalog data, one could analyse these patterns in terms of change in syntactic status (François 2001: 209):

- Considered in its bare form, the status "personal phrase" is compatible with just one function – **address**: KRO *tsi-k!* 'brother!', MSN *mam!* 'Dad!', MTA *Teri!* 'Terry!'.[1]
- The personal article *i/*e then has the power to "transfer" (Fr. "translater") this phrase towards a new status, that of a **referential** phrase: KRO ***i*** *tsi-k* 'my brother', MSN ***ɛ*** *mam* 'Dad', MTA ***i*** *Teri* 'Terry'. This status allows it to play all the syntactic functions that are open to referential phrases (subject, object, possessor, etc.).

4.2 Syntax of the common articles

Establishing a similar rule for the common articles (*na/*wo) is slightly more difficult, due to the diversity observed from one language to another. Especially, languages appear to vary again as to how optional the presence of the article is in those contexts where it is allowed. For example, LHI *n-*, NUM *u*, VRA/LKN (*ɪ*)*n*, MRL *nV-* are dropped at will in spontaneous speech, even when they could be present. In contrast, the rules regarding HIU/LTG *nə*, MTP *nV-*, MSN/VRS/DRG *ɔ*… are much more constraining: basically, if the article is allowed by the context, then it must be there. Obviously, the latter languages are a more reliable source of observation regarding rules for the presence/absence of the common article.

4.2.1 When is the article excluded?

For the sake of consistency, I will cite here data from one language, Mwotlap (François 2001: 187-214), taking it as essentially representative of the whole area. Incidentally, this

[1] These bare noun radicals precisely confirm that KRO *i* and MSN *ɛ* are still synchronically productive articles. In contrast, the fossilised *i of Mwotlap (§2.2.2) cannot disappear: e.g. ***i****thi-k!* 'brother!'; ***i****mam!* 'Dad!'.

choice constitutes a belated response to Crowley's regret (1985: 161), that "there is no evidence concerning the separability of this prefix [*nV-*] in Motlav [Mwotlap]".

I will choose the word *nɪ-bɪ* 'fresh water, river' as illustrative. Although corpus-based statistics (François 2001: 204) show that 73% of this word's occurrences – as well as its citation form – include *nV-*, many syntactic contexts actually require the noun in its bare form.

- X modifies another noun
  - X indicates the contents or the substance of N_1:
    na-plastik ***bɪ*** 'bottle of water'; *na-k͡pʷlɪs* ***bɪ*** 'water puddle'
  - X is a distinctive property of N_1:
    na-mya ***bɪ*** 'river eel'; *na-pnʊ* ***bɪ*** 'river island'
  - X is the "possessor" of an alienable noun N_1:
    na-mnɛ ***bɪ*** 'the taste of the water'; *nʊ-tʊti* ***bɪ*** 'the river's source'
  - X is a "possessed" noun following its classifier:
    nɛ-mɛ-k ***bɪ*** 'my water (to drink)'; *minɔ* ***bɪ*** 'my water (for non-drink uses)'
  - X complements a dependent noun N_1:
    na-mathɛɣ ***bɪ*** /craving.for/water/ 'thirst'
- X modifies a dependent morpheme
  - X complements a prenominal element:
    babahnɛ ***bɪ*** 'the last river'; *na-han* ***bɪ*** 'which river?'; *ni-tiy* ***bɪ*** 'genuine water'
  - X complements a preposition (free or prefixed):
    lɛlɔ ***bɪ*** 'inside the water'; *taval* ***bɪ*** 'across the river';
    *(sisɣɔy) lɪ-****bɪ*** '(fall) in the water'; *(vɪyɣɪl) bɪ-****bɪ*** '(argue) about water'
  - X complements the linker /nɛ/ 'of':
    na-mtɛhal nɛ ***bɪ*** 'the course of the river'
  - X complements the partitive /tɛ/ 'some':
    nɔk sɔ in tɛ ***bɪ*** 'I want to drink some water'
- X modifies a verb or a predicate
  - X is a non-referential incorporated object within a verb:
    inin ***bɪ*** 'drink water'; *haha* ***bɪ*** 'draw water (from well)'
  - X is a non-referential incorporated object in a noun compound:
    *tɪytɪy-****bɪ*** 'healer' [lit. 'water-holder'];
    *nɛ-mɛn inin-****bɪ*** 'dragonfly' [lit. 'water-drinking insect']

- X is the (non-patient) internal complement of a verb:
  vɾhɣi ***bɪ*** ‘be changed into water’
- X is the internal complement of an existential predicate:
  tatɛh ***bɪ*** ‘there's no water’; *takp͡ʷsɛ* ***bɪ*** ‘there are many rivers’

♦ Others

- X is a TAM-marked predicate noun:
  n-aɾs mal ***bɪ*** *lɔk* ‘the ice has [become] water again’

4.2.2 Function of the common noun article

The syntactic contexts listed above share certain essential properties. Typically, the noun appears unprefixed when it constitutes a phrase-internal modifier, pointing semantically towards a generic notion (‘water’) or a quality (‘watery’…) rather than designating a referential entity.

In contrast, the article is required whenever the noun is syntactically the head of an autonomous constituent, whether a verb's argument, a noun predicate, etc. Semantically, the function of this article consists in embodying the noun's quality into a discrete, specific referent: *nɪ-bɪ* ‘some/the water: a specific quantity of water, a river…’.

Following the analysis I proposed for the personal article (§4.1), common-noun articles may be described as a device used to transform **qualifying** noun phrases (MTP *bɪ*) into **referential** noun phrases (MTP *nɪ-bɪ*). Interestingly, this means that personal and common articles operate upon different input entities (vocative NP vs qualifying NP), but make them converge into the same output (referential NP).

5 Conclusion

Overall, Torba languages essentially agree in having kept noun articles alive and productive, as a syntactic device to form referential noun phrases. In each language, selecting the proper article ultimately depends on the syntactic, semantic and pragmatic properties of the NP in its specific context.

The various structures described in this paper may be of some interest to two communities of linguists. Oceanists in quest of historical evidence will notice the retention of several inherited features that have been lost elsewhere in Vanuatu. On the other hand, these formal structures delineate certain cognitive categories that may be of interest to typological linguists, whether they sound universally familiar (e.g. human specific referents) or more unusual (inalienable nouns possessed by a human specific referent). Building bridges between Oceanists and typologists was precisely one of Terry Crowley's major achievements, and a perspective we shall hopefully continue to bear in mind.

Abbreviations

ART:COMM – common article; ART:DIR.POSS – article for directly possessed nouns; ART:PERS – personal article; NUM – numeraliser; GeneralPoss – possessive classifier for general possession; VehicPoss – possessive classifier for vehicles; TAM – tense-aspect-mood.

References

Codrington, R. 1885. *The Melanesian languages*. Oxford: Clarendon Press.

Crowley, T. 1985. "Common noun phrase marking in Proto-Oceanic". *Oceanic Linguistics* 24: 135-193.

——. 2002. "Mwotlap". In Lynch *et al*. (eds), 587-598.

François, A. 2000. "Vowel shifting and cloning in Motlav: Historical explanation vs formal description". In *Proceedings of AFLA7 (The Seventh Meeting of Austronesian Formal Linguistics Association)*, M. Klamer (ed), 49-68. Amsterdam: Vrije Universiteit.

——. 2001. Contraintes de structures et liberté dans l'organisation du discours. Une description du mwotlap, langue océanienne du Vanuatu. Doctoral dissertation. Université Paris-IV Sorbonne: Paris. 3 vol., 1078 pages. [downloadable at http://alex.francois.free.fr/, Jan 2006]

——. 2005a. "A typological overview of Mwotlap, an Oceanic language of Vanuatu". *Linguistic Typology* 9 (1): 115-146.

——. 2005b. "Unraveling the history of vowels in seventeen north Vanuatu languages". *Oceanic Linguistics* 44 (2): 443-504.

Lemaréchal, A. 1989. *Les parties du discours, Syntaxe et sémantique*. Linguistique Nouvelle. Paris: Presses Universitaires de France.

Lynch, J. 2001. "Article accretion and article creation in Southern Oceania". *Oceanic Linguistics* 40 (2): 224-246.

Lynch, J., Ross, M. & Crowley, T. (eds). 2002. *The Oceanic languages*. Richmond: Curzon Press.

Ozanne-Rivierre, F. 1992. "The Proto-Oceanic consonantal system and the languages of New Caledonia". *Oceanic linguistics* 31 (2): 191-207.

Pawley, A. 1972. "On the internal relationships of Eastern Oceanic languages". In *Studies in Oceanic culture history*, R.C. Green and M. Kelly (eds), vol.3, 1-141. [Pacific Anthropological Records 13]. Honolulu: Bernice Bishop Museum.

Pawley, A. 1976. Some new Oceanic comparisons. Unpublished manuscript.