



**HAL**  
open science

# The historical morphology of personal pronouns in northern Vanuatu

Alexandre François

► **To cite this version:**

Alexandre François. The historical morphology of personal pronouns in northern Vanuatu. *Faits de langues*, 2016, Reconstruction et classification : Tendances actuelles, 47, pp.25-60. halshs-01382593

**HAL Id: halshs-01382593**

**<https://shs.hal.science/halshs-01382593v1>**

Submitted on 22 Oct 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The historical morphology of personal pronouns in northern Vanuatu

Alexandre François<sup>1</sup>

## 1. LANGUAGE GENEALOGY AND THE MORPHOLOGY OF PRONOUNS

### 1.1. *The Comparative Method and the Tree model*

The Comparative Method is commonly hailed as a solid methodology for comparing genetically related languages, and for reconstructing the history of their linguistic systems.<sup>2</sup> Equally common is the assumption that the results of its analyses are best displayed in the form of a tree, or *Stammbaum*: starting from a common protolanguage, its linguistic descendants should form neatly separated branches and subgroups, each of which should be defined by a set of exclusively shared innovations. The expectation – or at least the hope – is that the historical innovations reflected in modern members of a family should be distributed in nested patterns, so as to fit a cladistic representation of that family. This belief is reflected in the vast popularity of the tree model in works of historical linguistics up to this day.

The present paper aims at separating these two lines of thought, by showing that the strength of the Comparative Method does not necessarily entail the validity of the tree model which has been so often associated with it since the Neogrammarians. In fact, I will even propose that the CM provides precisely the analytical tools necessary to demonstrate the limitations of the tree model. Indeed, the method rests on principles of consistency and regularity of sound change, which allow the linguist to conduct rigorous demonstrations in the identification of innovations for each language, and in the reconstruction of words' histories. As each innovation is assigned a set of modern languages, it becomes possible to assess how nested (and thus how tree-like) their distribution is in the family.

The argument of this paper forms part of a broader debate showing the theoretical limitations of the cladistic approach in linguistics (see Heggarty *et al.* 2010; Drinka 2013; François 2014, *f/c*; Kalyan & François *f/c*). These recent contributions, as well as numerous earlier publications by various authors (e.g.

---

<sup>1</sup> Langues et Civilisations à Tradition Orale (CNRS–LACITO); Australian National University. Email <francois@vjf.cnrs.fr>.

<sup>2</sup> This work is part of the program *Investissements d'Avenir* overseen by the French Agence Nationale de la Recherche, ANR-10-LABX-0083 (Labex *EFL*) – and of its axis *Typology and dynamics of linguistic systems*. I wish to thank Malcolm Ross and Siva Kalyan, as well as the anonymous reviewers, for their comments on earlier drafts of this chapter.

Saussure 1917; Bloomfield 1933; Ross 1988), have shown that the tree model is incapable of dealing with the common situations of dialect continuum or sustained contact, in which isoglosses frequently intersect. The nested isoglosses assumed by a tree structure may only arise in the odd case of population splits with loss of contact; but they do not constitute the usual pattern in situations of *linkages* (Ross 1988), i.e. the normal case where related languages share innovations in chaining patterns. *Figure 1*, adapted from François (2014), illustrates the sort of configuration typical of a dialect continuum or a linkage, in which innovations are distributed in crosscutting patterns. (Letters represent dialects or languages; numbered isoglosses refer to their shared innovations.) Such a configuration is incompatible with a tree representation.

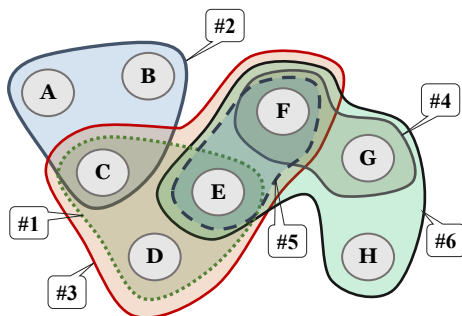


Figure 1 – *Intersecting isoglosses in a dialect continuum or a linkage*

### 1.2. A case study of pronouns in Vanuatu

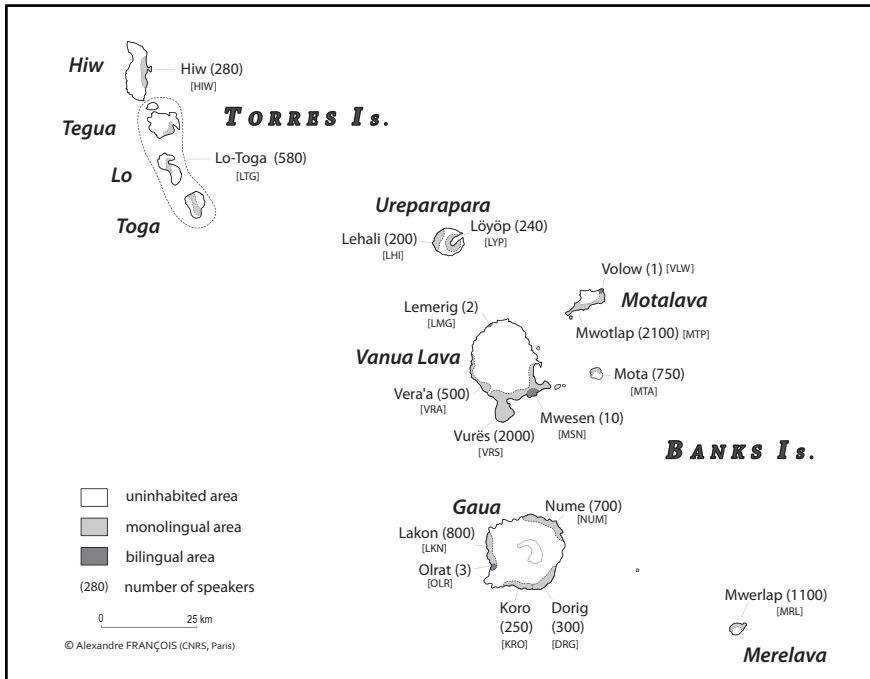
The theoretical point just defined – namely, that the Comparative Method can be used to show the failure of the tree model – will be demonstrated through a case study: the *historical morphology of personal pronouns in northern Vanuatu*.

The notion underlying this choice is that morphological paradigms of personal pronouns tend to evolve language-internally; but they are seldom borrowed through late contact across separate languages (see Nichols & Peterson 1996; Ross 2005:58). As a result, within a set of related languages, any morphological innovation shared by two modern members is likely to reflect the spread of a linguistic feature at an early time when the ancestors of these modern languages were still mutually intelligible, and were dialects in a continuum. In other words, shared innovations in pronominal morphology generally constitute solid diagnostic evidence for understanding the genealogical structure (internal subgrouping) of a given family.

In addition, the rich inventories of personal pronouns found in Oceanic languages – with typically 15 members in each paradigm [§3.2] – provide a critical mass of morphological forms, whose combined histories give a fair approximation of the historical relations between grammatical systems as a whole. In that sense,

they offer a valuable vantage point for assessing theoretical issues of the Comparative Method with some reasonable level of detail.

Map 1 – The seventeen languages of northern Vanuatu



Between 2003 and 2007, I collected linguistic data on the 17 languages still spoken in the Torres and Banks groups of islands, in the northernmost area of the Republic of Vanuatu. (I will henceforth designate these languages, interchangeably, as the “Torres–Banks languages” or the “Northern Vanuatu [NV] languages”.) Map 1 is a linguistic map of the archipelago, showing the name of each language together with a three-letter abbreviation, and the number of its speakers. With 17 distinct languages for a population of 9,300, this area features one of the highest levels of linguistic density in the world (François 2012); this is also true for Vanuatu as a whole, which hosts as many as 138 languages for 0.23 million people (François *et al.* 2015).

In terms of affiliation, NV languages belong to the Oceanic family – a subgroup of the Austronesian phylum that comprises about 500 languages scattered in the Pacific. Their common ancestor is known as Proto Oceanic (abbrev. *POc*), a protolanguage whose reconstruction is relatively solid. Current understanding

suggests that POc was the language spoken by the first inhabitants of Vanuatu, about 3100 BP, and that the country's modern languages result from the slow breakup, over three millennia, of that initial linguistic unity (Pawley 2003; Bedford & Spriggs 2008) – in a way similar to the breakup of Latin into a mosaic of Romance dialects and languages.

This study will unfold as follows. Because one of the mainstays of the Comparative Method is the regularity of sound change, Section 2 will provide an overview of sound change in northern Vanuatu; this will prove useful later, as we examine the detailed evolution of pronominal forms. Section 3 will then list the forms of the pronouns under study, starting with a reminder of Proto Oceanic reconstructions, and followed by the full inventories of independent pronouns in the 17 Torres–Banks languages, presented here for the first time.

The following chapters will reconstruct the detailed history of each personal pronoun in northern Vanuatu, starting with singular forms (Section 4) followed by plural (Section 5), dual (Section 6) and trial (Section 7) forms. For each pronoun, I will identify the main morphological innovations it has undergone in the NV area. Finally, I will examine their distribution across the subfamily, and show in the discussion (Section 8) that it is not perfectly compatible with a tree.

## 2. PRINCIPLES OF SOUND CHANGE IN NORTHERN VANUATU

One of the main forces driving the diversification of Northern Vanuatu languages has been sound change. Before we can start reconstructing the historical morphology of personal pronouns in the area, it is indispensable to present the main principles of sound change in these languages.

When discussing sound change in Vanuatu, the best point of reference is Proto Oceanic [POc]. Comparative work initiated by Dempwolff (1938), followed by authors such as Pawley (1974), Ross (1988, 1998), or Ross, Pawley & Osmond (1998–2011), has made it possible to reconstruct a fair deal of the lexicon, phonology and morphosyntax of this protolanguage. Attempts to reconstruct the ancestor shared by all Vanuatu languages result in a protolanguage which considerably resembles Proto Oceanic: this strongly suggests that any comparison should be carried out with respect to POc.

### 2.1. *Vowels*

#### 2.1.1. *Principles*

Proto Oceanic is reconstructed with five short monophthongs, /i e a o u/. While Mota has kept these five original vowels, the remaining Northern Vanuatu languages have increased their inventories, with between 7 and 14 phonemes in each language – including long vowels and diphthongs. Thus, Lemerig has 11

vowels { i ɪ ε æ a ɒ ɔ ø ɔ u }; the Lo dialect of Lo-Toga has { i e ε a ə ɔ o u ī ē ī iā ō ɔ̄ }.<sup>3</sup>

François (2005) explains the historical reason behind this increase in vowel phonemes. The key process was a form of umlaut, or metaphony, as the final unstressed vowel influenced the stressed one before disappearing. When this change took place, words originally consisted of open syllables CV, and were stressed on the penultimate syllable – e.g. POc \*páRi ‘stingray’, \*RapiRápi ‘evening’. In most cases, what was initially a sequence of syllables \*(C)V<sub>1</sub>(C)V<sub>2</sub> is reflected in modern languages by a single syllable (C)V(C), with a new vowel quality that somehow combines the properties of the two protovowels V<sub>1</sub> and V<sub>2</sub>: e.g. \*páRi ‘stingray’ > HIW *vɔʔL*, LTG *vɛr*, LHI *n-væj*, etc.

Sound change involving vowels was diverse across languages, but quite regular within each language. Thus, a sequence \*-á(C)i is always reflected as /ɔ/ in Hiw, as /ɛ/ in Lo-Toga, as /æ/ in Lehali...<sup>4</sup> This regularity, if taken into account rigorously, proves valuable when doing reconstruction. For example, a vowel /ɒ/ in Lehali always points to an etymon in \*-á(C)u or \*-ó(C)a; but the same vowel /ɒ/ in Lemerig can only reflect a protoform of the form \*-áCa.

A former quadrisyllable, composed of open CV syllables and structured in two right-aligned iambic feet, would typically reduce to two CVC syllables: e.g. POc \*RapiRápi ‘evening’ > LMG *ɛvɛvɛv*, LKN *rævrvæv*, DRG *ra.vrvv*. As for etyma with an odd number of syllables, they would have a “pretonic” vowel, i.e. an unstressed vowel preceding a stressed syllable. Historically, that pretonic position showed various signs of weakness. Its vowel either disappeared altogether, or was changed to schwa, or lost its own quality and assimilated to the next vowel. The data set in (1) illustrates this point with the noun \*panúa ‘inhabited land’: the list of its modern reflexes in modern Torres–Banks languages (ranked geographically from NW to SE; see *Map 1*) shows that the pretonic vowel \*a had weakened reflexes almost everywhere, and was only preserved in Mota and Lakon (François 2005: 470).<sup>5</sup>

- (1) The reflexes of pretonic vowels:  
 POc \*(na) panua ‘island, village, inhabited land’  
 > HIW *vəniə*; LTG *vəniə*; LHI *n-vəniə*; LYP *n-vəniə*; VLW *n-vəniə*;  
 MTP *na-pniə*; LMG *n-vəniə*; VRA *funuə*; VRS *vəniə*; MSN *vəniə*; MTA *vanua*;  
 NUM *funu*; DRG (*vno*); KRO *vəniə*; OLR *vəniə*; LKN *vanə*; MRL (*vəniə*).

### 2.1.2. Relevance

These observations will be relevant to our study of personal pronouns – as we shall see in later sections.

A general rule is that modern forms always reflect a longer protoform. Thus, a modern monosyllable must come from a former disyllabic form: e.g. LKN /ɣɪ:/

<sup>3</sup> François (2011:194) lists the full vowel inventories for the 17 NV languages.

<sup>4</sup> The regular patterns of correspondences are spelled out in François (2005).

<sup>5</sup> Bracketed forms correspond to those reflexes whose meaning has changed from that of their etymon.

‘3pl’ < \*ɣɪr < \*kɪra. Likewise, LMG /ɣætru/ ‘1incl:dual’, with its structure CVCCV, necessarily points to a quadrisyllabic etymon \*kita-rua (‘1incl:pl-two’).

Another point is that a former pretonic vowel can be reconstructed only based on Mota or Lakon, because its identity was lost or blurred in all other languages. Example (2), parallel with (1) above, shows the modern reflexes of the 2pl pronoun of POC, \*kamiu:

- (2) POC \*kamiu ‘2pl’  
 > HIW *kəmi*; LTG *kəmi*; LHI *kimi*; LYP *kimi*; VLW *ʔgimi*; MTP *kimi*; LMG *kimi*;  
 VRA *kimi*; VRS *kimi*; MSN *kimi*; MTA *kamiu*; NUM *kimi*; DRG *kmi*; KRO *kimi*;  
 OLR *kimi*; LKN *yamu*; MRL *kemi*.

## 2.2. Consonants

The consonant inventory of Proto Oceanic, as it was reconstructed by Ross (1998), is shown in *Table 1*. *Table 2* lists the regular reflexes of some POC consonants in the 17 NV languages; it is commented on in the next section.

*Table 1 – Consonant inventory for Proto Oceanic (after Ross 1998)*

	labial-velar	labial	alveolar	palatal	velar	uvular
voiceless stop [ORAL GRADE]	p <sup>w</sup>	p	t	c	k	q
prenasalised stop [NASAL GRADE]	<sup>m</sup> b <sup>w</sup>	<sup>m</sup> b	<sup>n</sup> d	<sup>ɲ</sup> ʃ	<sup>ŋ</sup> g	
nasal	<sup>m</sup> w	m	n	ɲ	ŋ	
constrictive			s			
lateral			l			
trill			r			
prenasalised trill			<sup>n</sup> r			
tap			R			
semi-consonant	w			y		

### 2.2.1. Nasal vs. Oral grade

POC is reconstructed with two series of stops: prenasalised voiced stops, and voiceless oral stops (see *Table 1*). The tradition among Oceanic linguists is to refer to these two series of consonants respectively as *nasal grade* and *oral grade* (see Grace 1959:32, Lynch 1975; Ross 1988:32-47). In this paper, I will keep these two terms, in small capitals.

NV languages usually keep the phonemic distinction between the two series, but with some phonetic changes.<sup>6</sup> Some ORAL-GRADE consonants underwent

<sup>6</sup> The palatal stops \*c and \*ɲʃ were merged with \*s. The uvular \*q disappeared with no trace.

lenition to a fricative, with or without voicing (\*p > /v/ or /f/; \*k > /ɣ/). In some languages, this lenition of the ORAL series triggered a “pull chain”, whereby NASAL-GRADE stops lost their prenasalisation and became oral, voiceless stops. For instance, the slot for a voiceless ⟨k⟩, which had been emptied as a result of the lenition \*k > /ɣ/, became filled again by a second change \*ŋg > /k/; the phonemic contrast NASAL–ORAL was thus maintained, but instead of contrasting \*{ ŋg ≠ k }, it now took the form of a contrast { k ≠ ɣ }, respectively. Likewise, the former opposition \*{ mb ≠ p } is now realised as an opposition { mb ≠ v } in some languages, as { p ≠ v } in others.

Table 2 – Reflexes of some POc consonants in modern Torres–Banks languages

POc	* <sup>m</sup> b <sup>w</sup>	* <sup>m</sup> b	* <sup>n</sup> d	*ŋg	*p	*t	*k	*r
HIW	k <sup>w</sup>	p	t	k	v	t	ɣ	<sup>9</sup> L
LTG	k <sup>w</sup>	p	t̥   -t	k	v	t, s	ɣ	r
LHI	k <sup>w</sup> -   -k	p	<sup>n</sup> d-   -n	k	v-   -p	t	ɣ-   -∅	j
LYP	k <sup>p̄</sup> -   -ŋ	p-   -m	<sup>n</sup> d-   -n	k-   -ŋ	v-   -p	t, tʃ	ɣ-   -∅	j
VLW	<sup>ŋ</sup> g <sup>b̄</sup> <sup>w</sup> -   -ŋ <sup>m̄</sup> <sup>w</sup>	<sup>m</sup> b-   -m	<sup>n</sup> d-   -n	<sup>ŋ</sup> g-   -ŋ	v-   -p	t	ɣ	j
MTP	k <sup>p̄</sup> <sup>w</sup>	<sup>m</sup> b-   -m	<sup>n</sup> d-   -n	k	v-   -p	t	ɣ	j
LMG	k <sup>p̄</sup> <sup>w</sup> -   -k	p	t	k	v	ʔ, t	ɣ	r
VRA	k <sup>p̄</sup> <sup>w</sup>	<sup>m</sup> b-   -m	<sup>n</sup> d-   -n	k	f	ʔ, t	ɣ	r
VRS	k <sup>p̄</sup> <sup>w</sup>	<sup>m</sup> b-   -m	<sup>n</sup> d-   -n	k	v	t	ɣ	r
MSN	k <sup>p̄</sup> <sup>w</sup>	p	n	k	v	t	ɣ	r
MTA	k <sup>p̄</sup> <sup>w</sup>	p	n	k	v	t	ɣ	r
NUM	k <sup>p̄</sup> <sup>w</sup>	<sup>m</sup> b-   -m	<sup>n</sup> d-   -n	k	f	t	ɣ	r
DRG	k <sup>p̄</sup> <sup>w</sup>	<sup>m</sup> b	<sup>n</sup> d	k	v	t	ɣ	r
KRO	k <sup>p̄</sup> <sup>w</sup>	<sup>m</sup> b-   -m	<sup>n</sup> d-   -n	k	v	t	ɣ	r
OLR	k <sup>p̄</sup> <sup>w</sup>	p	tʃ	k	v	t	ɣ-   -V:	r-   -j
LKN	k <sup>p̄</sup> <sup>w</sup>	p	tʃ	k	v	t, s	ɣ	r-   -V:
MRL	k <sup>w</sup> -   -k	<sup>m</sup> b-   -m	<sup>n</sup> d-   -n	k	v-   -p	t	ɣ	r

This point will be relevant in our study of personal pronouns. For example, ex. (2) above listed the modern forms for 2pl. Despite appearances, the initial consonant /k-/ found in most languages is not the expected reflex of POc \*k; that reflex should have been /ɣ/, which is only found in Lakon. All other languages have undergone an irregular change<sup>7</sup> from ORAL to NASAL GRADE here (a fact already noted by Lynch & Ozanne-Rivierre (2001) and Clark (2009) for the rest of

<sup>7</sup> Throughout this chapter, by convention I will use the sign ‘>’ when dealing with REGULAR sound change, and ‘→’ to represent an IRREGULAR change, which does not fit expected sound correspondences in the language considered: e.g. { \*k<sup>m</sup>amiu → \*<sup>ŋ</sup>g<sup>m</sup>amiu > /k<sup>m</sup>ami/ }. That is, ‘>’ captures purely phonological change, whereas ‘→’ refers to a morphological innovation.



Vanuatu). Lakon here constitutes a pocket of retention amidst a set of innovative languages [see §5.2.1].

### 2.2.2. Onset vs. Coda reflexes

As I explained in §2.1.1, NV languages went historically through a radical change in phonotactics, as a sequence of two open syllables  $*C_1V_1C_2V_2$  lost its unstressed vowel, and changed to a closed syllable  $C_1VC_2$ .

Some modern languages reflect a POC consonant differently, depending on its position in the syllable of the modern word. For example, prenasalised stops often lose their oral component in coda position, and are merged with the corresponding nasals. Thus, the POC word *\*mbembe* ‘carry s.o. piggyback’ became /**mbe<sup>m</sup>b**/ in Dorig, but /**mbem**/ in Mwotlap, and /**pem**/ in Löyöp. These differential reflexes are shown in *Table 2* above, and indicated with hyphens. For example, *\*mb* > {p- | -m} in Löyöp means that the regular reflex of *\*mb* in this language is /p/ in syllable onset position, but /m/ in coda.

These observations will prove useful when reconstructing some pronominal forms. For example, MTP /**nik**/ and VLW /**niŋ**/ ‘2sg’ are both regular reflexes of a protoform *\*niŋgo* with a prenasalised (NASAL GRADE) stop. Likewise, forms for ‘incl:pl’ like MTP /**ɣim**/, LMG /**ɣæt**/, DRG /**ɣi<sup>m</sup>d**/, LKN /**ɣi<sup>f</sup>**/, all point unambiguously to the same protoform *\*ki<sup>n</sup>da*.

### 2.2.3. Other consonants

For reasons of space, I will only discuss here one other consonant, which will be relevant to our discussion of pronouns. This is the rhotic *\*r*, which in POC was probably an alveolar trill.

*Table 2* showed the reflexes of *\*r* in modern NV languages. The rhotic became a preploded velar lateral /**ʁ<sup>l</sup>**/ in Hiw (François 2010), and a palatal glide /**j**/ in four other languages. Two languages of Gaua show differential reflexes depending on their position in the modern syllable [§2.2.2]. In Olat, *\*r* became a glide /**j**/ in coda position. In Lakon, *\*r* disappeared from codas, yet triggered compensatory lengthening of the preceding vowel (François 2005:461). This is represented in *Table 2* as { r- | -j } for Olat, and { r- | -V: } for Lakon.

These observations will be useful when we track the reflexes of two number morphemes: *\*-ra* ‘plural’ and *\*=rua* ‘dual’. If we hadn’t sorted out regular sound correspondences, it would have been difficult to unveil the presence of *\*-ra* in Lakon’s 3pl pronoun /**ɣi:**/ < *\*ɣir* < *\*kira*; or to detect *\*=rua* in its 2du pronoun /**ɣamu:**/ < *\*ɣamur* < *\*ɣamúru* < *\*kamu-rua* [§6.3].

With this knowledge in mind, we can now move on to the examination of pronominal paradigms, and to the reconstruction of their morphological history.

## 3. PRONOMINAL PARADIGMS: THE FORMS

## 3.1. Proto Oceanic

The pronoun paradigm of Proto Oceanic can be reconstructed with reasonable confidence, by applying the Comparative Method to modern Oceanic languages. Ross (1988:367; 2002:67) reconstructs four distinct paradigms, based on their syntactic properties: independent pronouns; subject proclitics; object enclitics; possessor suffixes. *Table 3* shows the reconstructions he proposes.

*Table 3 – Reconstructed forms for POC pronouns (after Ross 2002:67)*

	Independent	Subject clitics I	Subject clitics II	Object enclitic	Possessor suffix
1sg	[i]au	ku=	au=	=au	- <sup>o</sup> gu
2sg	[i]ko[e]	mu=	ko=	=ko	-mu
3sg	ia	ña=	(y)a=, i=	=a	-ña
1inc:pl	kita	ta=	ta=	—	- <sup>o</sup> da
1exc:pl	ka[m]i, kamami	—	—	—	-ma[m]i
2pl	ka[m]u, kamiu	—	—	—	-m[i]u
3pl	[k]ira	ra=	ra=	=ra	- <sup>o</sup> ra

In this table, square brackets mean that modern languages point to two equally plausible reconstructions at the level of Proto Oceanic. For example, {*ka[m]u*, *kamiu*} means that three protoforms can be reconstructed for the 2pl pronoun: *\*kau*, *\*kamu*, *\*kamiu*; the available evidence does not allow us to privilege any of these reconstructions over the others.

## 3.2. Free pronouns in Northern Vanuatu

The modern languages of Northern Vanuatu reflect these four pronominal paradigms diversely. In this article, for reasons of space, I will focus on the set of independent (or ‘free’) pronouns. *Tables 4 (a–d)* list the forms in the 17 languages under study; for technical reasons, each of the four numbers is assigned a separate table.

The first two rows of each table include protoforms. The row for POC provides the Proto Oceanic reconstructions proposed by Ross (2002), and listed in *Table 3* above. The row for PTB, or “Proto Torres–Banks”, indicates the protoforms I reconstruct for the common ancestor of the 17 Torres–Banks languages, based on the modern forms listed underneath. The arguments for these PTB reconstructions will be given in sections 4–7 of this study.

Table 4a – Independent pronouns in north Vanuatu languages: Singular

		1SG	2SG	3SG
POc	<i>p Oceanic</i>	*[i]au	*[i]ko[e]	*ia
PTB	<i>p Torres-B</i>	*nau	*(n)i <sup>o</sup> go	*nia
HIW	<i>Hiw</i>	nəkə	ikə	ninə
LTG	<i>Lo-Toga</i>	nekə, nəkə	nikə	niə
LHI	<i>Lehali</i>	nɔ   nɔk	nek	kɛ
LYP	<i>Löyöp</i>	nø   nɔk	niŋ	kje
VLW	<i>Volow</i>	nɛ   nɛŋ	nŋ	ʔgr
MTP	<i>Mwotlap</i>	nɔ   nɔk	nik	kɪ
LMG	<i>Lemerig</i>	nœ	næk	ti
VRA	<i>Vera'a</i>	nɔ	nik(t)	ʔdi(t)
VRS	<i>Vurës</i>	nɔ   na	nik	nɪ
MSN	<i>Mwesen</i>	nɔ   na	nik	nɪ
MTA	<i>Mota</i>	na   nau	ka   ko   niko	ni   nia
NUM	<i>Nume</i>	na	nik	ni
DRG	<i>Dorig</i>	na	nik	ni
KRO	<i>Koro</i>	na	nik	ni
OLR	<i>Olrat</i>	na	nik	nɪ
LKN	<i>Lakon</i>	na	nik	ni ~ nɪ
MRL	<i>Mwerlap</i>	nɔ ~ nœ	nɛāk	(ki)sɛān

Table 4b – Independent pronouns in north Vanuatu languages: Dual

	1INC.DU	1EXC.DU	2DU	3DU
POc	*kita=rua	*ka[m]i=rua	*kamiu=rua	*[k]ira=rua
PTB	*ki <sup>o</sup> da-rua	*Kama-rua	*Kamu-rua	*[i]ra-rua
HIW	tə <sup>o</sup> Lə	kama <sup>o</sup> Lə	kimi <sup>o</sup> Lə	sə <sup>o</sup> Lə
LTG	ʔor	kəmɔr	kəmɔr	hor
LHI	ʔinjɔ	mæjɔ	mɔjɔ	kɔjɔ
LYP	je <sup>o</sup> dɔ	mɔmjɔ	mɔjɔ	kjejɔ
VLW	ʔdɔ   ʔdojɔ	ʔgemjɔ	ʔgɔmjɔ	ʔgɔjɔ
MTP	ʔdɔ   ʔdojɔ	kamjɔ	kɔmjɔ	kɔjɔ
LMG	ʔætru	kamaru	kumru	tæru
VRA	ʔdu   ʔi <sup>o</sup> du(ɔ)	kama <sup>o</sup> du(ɔ)	kumru(ɔ)	ʔduru(ɔ)
VRS	ʔdɔrɔk	kɔmɔrɔk	kɔmɔrɔŋ	rɔrɔ
MSN	ninrɔ	kɛmemrɔ	kɔmɔrɔ	nɔrɔ(rɔ)
MTA	nara	kara	kamra	rara
NUM	ʔduru	kamar	kumur	ruru
DRG	ʔda:r	kma:r	kmur	ra:r
KRO	ʔduru	kɛmɛār	kumur	(i)rru
OLR	ʔʔrɔ	kɪmj	kumuj	nɔrɔ
LKN	wɔʔʔ	ʔama:	ʔamu:	wɔrɔ
MRL	ʔdoro	kamar	kamrɔ	karar

Table 4c – Independent pronouns in north Vanuatu languages: Trial

	1INC.TRI	1EXC.TRI	2TRI	3TRI
POc	*kita=tolu	*ka[m]i=tolu	*kamiu=tolu	*[k]ira=tolu
PTB	*k <sup>i</sup> da-tolu	*Kama-tolu	*Kamu-tolu	*[i]ra-tolu
HIW, LTG	—	—	—	—
LHI	yɪntɪl	mæɪtɪl	mɪtɪl	keʒɪtɪl
LYP	ʒɛɲʃøɭ	møɲʃøɭ	møʃøɭ	kjeʒʃøɭ
VLW	ʰdɪtɪl	ʰgɛmtɪl	ʰgɔmtɪl	ʰgɪʒɪtɪl
MTP	ʰdɪtɪl ~ ɪntɪl	kamtɪl	kɪmtɪl	kɪʒɪtɪl
LMG	ɣætʔøɭ	kœmœmʔøɭ	kimiʔøɭ	tærʔøɭ
VRA	(yɪ)ʰdɔʔøɭ	kamamʔøɭ	kimiʔøɭ	ʰdirʔøɭ
VRS	nɪntøɭ	kɛmɛktøɭ	kimitøɭ	nøɾtøɭ
MSN	nɪntøɭ	kɛmɛmtøɭ	kimitøɭ	nøɾtøɭ
MTA	nətøɭ	katøɭ	kamtøɭ	ratøɭ
NUM	ʰdøtøɭ	kamatøɭ	kømtøɭ	røtøɭ
DRG	tøɭyɪn	tøɭkma	tøɭkmi	tøɭnɪr
KRO	tɪɪn	kɪɪkama	kɪɪkimi	tɪɪr
OLR	ʃɪɪɪt	kɪɪkama	kɪɪkimi	tɪɪj
LKN	t(ʃ)ɪɪt(ʃ)	tɪɪmæ	tɪɪɪmu	tɪɪ:
MRL	—	—	—	—

Table 4d – Independent pronouns in north Vanuatu languages: Plural

	1INC.PL	1EXC.PL	2PL	3PL
POc	*kita	*ka[m]i	*kamiu	*[k]ira
PTB	*k <sup>i</sup> da	*Kama[m]i	*Kamiu	*[k]ira
HIW	tɪtə	kama	kimi	sisə
LTG	yɪʔə, ʔəyə	kœmɛm	kœmi	nɪhə
LHI	ɣɛn	kɪmæm	kimi	kej
LYP	ʒɛn	køɲøɲ	kimi	kjeʒ
VLW	ɣɪn	ʰgɛm(ɛm)	ʰgɪmi	ʰgɪj
MTP	ɣɪn	kɛm(ɛm)	kimi	kɪj
LMG	ɣæt	kœmœm	kimi	tær
VRA	ʰdɪ   ɣɪʰdɪ	kamam	kimi	ʰdir(t)
VRS	nɪn	kɛmɛk	kimi	nɪr
MSN	nɪn	kɛmɛm	kimi	nɪr
MTA	nɪna	kamam	kamiu	n(e)ira
NUM	ɣɪn	kama	kimi	nɪr
DRG	ɣɪn	kma	kmi	nɪr
KRO	ɣɪn	kama	kimi	nɪr
OLR	ɣɪʃ	kama	kimi	nɪj
LKN	ɣɪʃ	ɣamæ	ɣamu	ɣɪ:
MRL	ɣɛān	kɛmɛm	kɛmi	kɛār

The following are typographical conventions used in *Tables 4a-d*:

- a tilde sign (‘~’) is used when two forms are in free variation in all contexts: e.g. LKN /ni/ ~ /ni/ ‘3sg’.
- simple brackets indicate a segment that can optionally be deleted: e.g. in MSN ‘3du’ is /noro/ or /noroꝛo/; LKN ‘1inc:tri’ can surface as /ʔilitʃ/ or /tilitʃ/ or /ʔilt/.
- pointy brackets in Vera’a forms indicate a vowel that tends to elide in subject position (François 2005:464): e.g. ‘1inc:pl’ /ʔi<sup>ᶇ</sup>duo/ becomes /ʔi<sup>ᶇ</sup>du/ when subject.
- a comma indicates dialectal variation, as in Lo-Toga: e.g. ‘1inc:pl’ is /ʔitə/ in the Toga dialect, /təʔə/ in the Lo dialect.
- a pipe sign (‘|’) means that two forms are used in complementary distribution, depending on the syntactic context: e.g. in Löyöp, /nok/ is the 1sg subject marker with an aorist verb, /nø/ is the 1sg pronoun everywhere else.

The purpose of this paper is to analyse the synchronical paradigms of free pronouns, and reconstruct the accumulation of innovations each language has been through.

### 3.3. *The distribution of pronominal forms*

Northern Vanuatu languages have several pronominal paradigms. While this article will focus on free pronouns, it is useful to briefly outline the syntactic distribution of its other pronoun paradigms.

#### 3.3.1. *Possessive suffixes*

The set of possessive suffixes in Proto Oceanic (*Table 3* p.9) has survived in all NV languages. Its rich historical morphology would warrant a study of its own; for reasons of length, I will refrain from mentioning it, as it shows limited interference with free pronouns.

#### 3.3.2. *Object suffixes*

As *Table 3* showed, POC had a defective set of enclitics for encoding the object of transitive verbs, and of certain prepositions. These clitics have been preserved in the form of suffixes, in four NV languages: Hiw, Lo-Toga, Mota, Mwerlap. *Table 5* illustrates them with the preposition ‘with’, which takes these object suffixes. Forms in bold include the modern reflexes of POC object clitics; plain characters show the use of free pronouns instead.

The common strategy in NV languages – just like in the rest of Oceanic – is to replace missing object suffixes with free pronouns (Evans 2003). This substitution is often an option even in cases where the suffix has been preserved: see the examples in Hiw. The majority of NV languages have generalised this pattern to all persons, thereby losing all traces of the former object pronouns of POC. *Table 5* illustrates this case with Mwotlap.

Table 5 – Reflexes of POc object clitics in some NV languages

	‘with me’	‘with you’	‘with him/her’	‘with them’
<i>POc</i>	= <i>au</i>	= <i>ko</i>	= <i>a</i>	= <i>ra</i>
<i>Hiw</i>	mi nəkə	<b>mi-kə</b>	<b>mi-ə</b> / mi ninə	<b>mi-sə</b> / mi sisə
<i>Toga</i>	mi nəkə	<b>mə-kə</b>	<b>me</b>	<b>me-hə</b>
<i>Mwotlap</i>	mi nə	mi nɪk	mi kɪ	mi kɪj
<i>Mota</i>	amen nau	<b>amai-ko</b>	<b>amai-a</b>	<b>amai-ra</b>
<i>Mwerlap</i>	<b>mi-ɔ</b>	<b>mi-ak</b>	<b>mi-a</b>	<b>mi-ar</b> / mi kɛr

### 3.3.3. Subject clitics

Finally, POc had two sets of subject proclitics (*Table 3* p.9). Just like we saw for object clitics, the tendency has been for NV languages to lose those special subject markers, and replace them with free pronouns everywhere.

That said, the northernmost languages have kept traces of these subject clitics (more precisely those of set I) in their morphology. Thus for the singular, POc \**ku*=, \**mu*=, \**na*= are reflected, at the level of Proto Torres–Banks, by three clitics that can be reconstructed, respectively, as \**ḡu*= ‘1sg’, \**u*= ‘2sg’, \**ni*= ‘3sg’. This has been the topic of a specific study (François 2009), to which the reader is referred. These subject clitics will be mentioned later in this study, as they have left some traces in the forms of the free pronouns in several languages of the northern Banks and of the Torres islands [see §4.1.1, 4.3].

## 4. SINGULAR PRONOUNS

### 4.1. First singular

All modern NV forms for 1sg reflect a form which can be reconstructed as \***nau**. This corresponds to POc \**au*, augmented with a non-etymological consonant /n-/, whose origin is unclear.

#### 4.1.1. The trace of former clitics

In the six northernmost languages, that pronoun sometimes shows an extra velar, evidently reflecting a prenasalised stop \**ṅ* [§2.2.1]. In the two Torres languages, that velar is part of the modern form of the 1sg pronoun in all contexts: *Hiw*, /nəkə/, *TGA* /nəkə/. In four languages (*LHI* /nək/, *LYP* /nək/, *VLW* /nɛŋ/, *MTP* /nək/), this longer form is only found in subject position when the verb takes the so-called ‘aorist’ aspect: this is a TAM category mostly found in dependent clauses, yielding such values as *subjunctive*, *sequential* or *narrative* (François 2003:165-199; 2009).

François (2009) showed that these forms result from the accretion of \**nau* with a former clitic \**ḡu*=, formerly used to encode the 1sg subject in a dependent clause: *VLW* /nɛŋ/ < \**nɛ+ṅ*- < \**nau* + *ḡu*=. This clitic \**ḡu*= was itself part of a paradigm of subject clitics [§3.3.3] which can be reconstructed for Proto Torres–

Banks (pTB): \**ʔgu*= ‘1sg’, \**u*= ‘2sg’, \**ni*= ‘3sg’, \*(*k*)*a*= ‘non.sg’. These forms have nowhere survived as full pronouns, but have left vestigial traces in the modern form of pronouns: this explains, for example, the alternation between /*nɛ*/ and /*nɛŋ*/ in Volow, depending on the aspect of the verb.

It is possible to reconstruct simple sentences at the level of Proto Torres–Banks (François 2009:191). (3) shows how a string of serialised verbs would typically involve a free pronoun (here \**nau*), followed by a person-indexing TAM clitic (here \**ʔgu*=):

- (3) pTB      \***nau**    *ʔgu*=    mule    *ʔgu*=    maturu  
                  1sg:FREE   AO:1sg=   return   AO:1sg=   sleep  
                  ‘So I went back and slept.’

A modern translation of (3) in the Hiw language would be sentence (4):

- (4) Hiw      **nəkə**    ɲʷujə    **kə**=    miti<sup>9L</sup>.  
                  1sg:FREE   return   AO:1sg=   sleep  
                  ‘So I went back and slept.’

The /*kə*=/ clitic of Hiw is a regular reflex of PTB \**ʔgu*= : it encodes both the aorist aspect and the 1sg person of the subject. But the crucial point, here, concerns the first word of the sentence: the form /*nəkə*/ of the 1sg pronoun reflects the accretion of { \**nau* + *ʔgu*= }.

In northern Banks, the accreted form is still restricted to clauses marked as aorist. In the Torres languages, the reflex of { \**nau* + *ʔgu*= } has been generalised to all contexts (see *Table 5* above).<sup>8</sup>

Finally, three central Banks languages – Mota, Mwesen, Vurës – have a 1sg pronoun /*na*/, along with the more regular reflex of \**nau*. It can be shown that the /*na*/ form reflects a similar form of accretion as the one mentioned above, except it involves an aorist clitic \**a*= < \*(*k*)*a*= rather than \**ʔgu* (François 2009:194). Thus, VRS-MSN /*nə*/ reflects \**nau*, but /*na*/ reflects { \**nau* + *a*= }.

#### 4.1.2. Summary

This section can be summarised by listing the innovations that have affected the 1sg pronoun (POC \**au*), and checking which languages reflect them.

- (i1)    \**au*    → \**nau*  
 (i2)    accretion of { \**nau* + *ʔgu*= } before aorist verb  
 (i3)    accretion of { \**nau* + *ʔgu*= } in all contexts  
 (i4)    accretion of { \**nau* + (*k*)*a*= } before aorist verb

<sup>8</sup> We will see that Hiw has gone through a similar process of accretion for 3sg [§4.3], 1inc:pl [§5.1], 3pl [§5.4.2].

<i>i</i>	Hw	LO	TGA	LHI	LTP	VLW	MTP	LMG	VRA	VRS	MSN	MTA	NUM	DRG	KRO	OLR	LKN	MRL
(i1)	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
(i2)	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-
(i3)	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
(i4)	-	-	-	-	-	-	-	-	-	+	+	+	-	-	-	-	-	-

#### 4.2. Second singular

The free pronoun for second singular exhibits few changes from the POc form \*[i]ko[e].

One innovation is the irregular change ORAL → NASAL GRADE in the velar stop: all languages point to a NASAL GRADE protophoneme \*<sup>ng</sup>, e.g. VLW /niŋ/ < \*ni<sup>ng</sup>o. Interestingly, while this process of fortition { \*k → \*<sup>ng</sup> } is irregular when compared to the lexicon (see Table 2 p.7), it is observed for all velar stops in the pronoun paradigm (\*kamami → \*<sup>ng</sup>gamami [§5.2.1]; \*ku= → \*<sup>ng</sup>u= [§3.3.3]...) except for 1inclusive (\*kita). The motivation of this fortition is unclear.

While POc may be reconstructed with a 2sg pronoun \*[i]ko or \*[i]koe, all NV languages unambiguously point to the shorter variant \*[i]ko. Only Mota has kept traces of a monosyllabic \*ko; all other languages reflect a dissyllabic form \*iko, with penultimate stress on /i/: \*iko → \*i<sup>ng</sup>o > Hw /íkə/. In terms of vowel change [§2.1], all languages show the expected result of a sequence \*i(C)o.

Hiw is the only language to have kept a vowel-initial etymon \*iko/\*i<sup>ng</sup>o. All other languages show the addition of a non-etymological consonant \*n-: \*iko → \*i<sup>ng</sup>o → \*ni<sup>ng</sup>o > LTG /nikə/, LHI /nək/, VLW /niŋ/, etc. The change is parallel with 1sg (\*au → \*nau), as well as 3sg (\*ia → \*nia), though it is difficult to know which pronoun influenced the others.

In modern Mota, /ko/ is the standard, unstressed free pronoun;<sup>9</sup> /niko/ is a stressed pronoun used in focal position:

- (5) MTA **Niko** iloke **ko** me pal ŋm<sup>w</sup>o-k o yasal ti?  
 2sg:FOC here 2sg PFT<sub>1</sub> steal POSS-1sg ART knife PFT<sub>2</sub>  
 ‘Is it you who (you) stole my knife?’

It is likely that Mota is here conservative of an earlier configuration, which has been lost everywhere else: namely, a contrast between an unstressed (proclitic) pronoun \*[i]ko, and an augmented form \*niko used in stressed or focal contexts. In all NV languages other than Mota, the stressed variant became the unmarked form for the 2sg pronoun, replacing the shorter \*[i]ko.<sup>10</sup>

<sup>9</sup> As for the form /ka/, it is a portmanteau morpheme encoding ‘2sg:subject’ and ‘Aorist’; its form accretes \*ko and the former Aorist clitic \*(k)a= (François 2009:194, cf. §4.1.1).

<sup>10</sup> Such a process of “markedness shift” (Dik 1989: 44) is frequent in pronominal paradigms. Thus Latin *ego*, initially a focal pronoun for 1sg, became the unstressed clitic *je* in modern French, and was replaced by *moi* in stressed contexts.



#### 4.2.1. Summary

The table below lists all the innovations affecting the 2sg pronoun (POc \*[i]ko).

- (i5) \*[i]ko → \*[i]<sup>o</sup>go  
 (i6) \*i<sup>o</sup>go → \*ni<sup>o</sup>go (at least in pragmatically focal contexts)  
 (i7) \*ni<sup>o</sup>go used even in non-focal contexts

<i>i</i>	Hiw	Lo	TGa	LHI	LYP	VLW	MTP	LMG	VRA	VRS	MSN	MTA	NUM	DRG	KRO	OLR	LKN	MRL
(i5)	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
(i6)	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
(i7)	-	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+

#### 4.3. Third singular

The 3sg pronoun of POc, reconstructed as \*ia, is not reflected as such in modern NV languages. A number of them reflect an augmented form \*nia, featuring the same additional consonant as 1sg \*nau and 2sg \*ni<sup>o</sup>go. It is quite likely that this \*nia reflects the influence of \*-ñā, the 3sg possessive suffix of POc (Table 3 p.9).

The form \*nia was not preserved in all languages; however, its geographical distribution (in the Torres Is on the one hand, in the southern Banks Is on the other, not to mention other Vanuatu languages further south) strongly suggests that this innovative protoform was once shared by the whole Northern Vanuatu area. At some later point, some languages went through further innovations.

Several languages show evidence of a lighter pronoun \*ni (presumably an abbreviated form of \*nia). This is true, first, of Mota, in which /ni/ is to /nia/ as /ko/ is to /niko/, i.e. light vs. heavy pronouns, respectively. In other languages, \*ni= (> e.g. Hiw /nə=/, MTP /ni-/) is the subject clitic coding for aorist aspect.

While Lo-Toga /niə/ reflects \*nia regularly, the Hiw form /niə/ evidently shows the accretion of two former pronouns: the heavy \*nia and the light \*ni. The combination { \*nia + ni= } became \*niə > /niə/ [see fn.8 p.14].

Other languages show innovations regarding mostly the initial consonant of the 3sg pronoun. Vera'a /<sup>h</sup>dii/ and its neighbour Lemerig /ti/ share a morphological innovation whereby \*nia became \*<sup>h</sup>dia. The origin of this /<sup>h</sup>d/ is unknown: it might result from the fortition of /n/.

Four languages have a pronoun that starts with a velar stop: LHI /kɛ/, LYP /kʲɛ/, VLW /<sup>h</sup>gɪ/, MTP /ki/. These are all regular reflexes of a protoform \*<sup>h</sup>gia. The origin of this form is unclear; but it is possible that the velar stop was imported from the POc 3pl pronoun, \*kīra [see §5.4]. Finally, Mwerlap has a 3sg pronoun /sɛān/ or /kīsɛān/. The short form /sɛān/ is a demonstrative; as for the /ki/ prefix, it is possibly related to \*<sup>h</sup>gia.

##### 4.3.1. Summary

The table below lists all the innovations affecting the 3sg pronoun (POc \*ia).

- (i8) \*ia → \*nia  
 (i9) \*nia → \*ndia  
 (i10) \*nia → \*ngia  
 (i11) accretion of { \*nia + mi= } in all contexts

<i>i</i>	HW	LO	TGA	HFI	LTP	VW	MTP	LMG	VRA	VRS	MSN	MTA	NUM	DRG	KRO	OLR	LKN	MRL
(i8)	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
(i9)	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-
(i10)	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	?
(i11)	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

## 5. PLURAL FORMS

Like virtually all Oceanic languages, those of the Torres–Banks area have at least three numbers: singular, dual and plural. Most of them even have a trial number (for groups of three referents), even though the latter has been lost in the Torres languages and in Mwerlap.

It is unclear whether POC had dual and trial numbers; but if it did, its dual and trial pronouns were evidently derived morphologically from the plural forms, through the transparent addition of a numeral (Ross 2002:69): e.g. \*kita=rua ‘1inc:du’ < \*kita ‘1inc:pl’ + \*rua ‘two’. For this reason, it is better to examine plural pronouns first; I shall come back to other numbers in §6–7.

### 5.1. First inclusive plural

#### 5.1.1. A widespread innovation

The first inclusive plural pronoun is reconstructed as \*kita for POC.

All NV languages show evidence of an irregular change from \*kita to \*ki<sup>n</sup>da.<sup>11</sup> This is made evident by the second consonant of each modern form, which systematically reflects a \*<sup>n</sup>d (see Table 2 p.7): \*ki<sup>n</sup>da > TGA /yit̪ə/, MTP /yim/, LMG /yæt̪/, VRA /yid̪i/, DRG /yid̪/, LKN /yit̪/, MRL /yēan̪/... The change from ORAL-GRADE \*t to NASAL-GRADE \*<sup>n</sup>d was arguably due to the analogy with the possessive suffix for 1inc:pl, POC \*<sup>n</sup>da (Table 3 p.9).

Everywhere, the modern vowels are the regular reflexes of a sequence \*i(C)a, so no sound change affected the original vowels. In Löyöp, the regular outcome of \*i(C)a is a diphthong /iē(C)/; \*ki<sup>n</sup>da was first reflected regularly as \*yīēn̪, but the yod resulted in the deletion of the velar fricative: \*ki<sup>n</sup>da > \*yīēn̪ > /jēn̪/.

<sup>11</sup> The change was already noticed by Clark (1985) and by Lynch & Ozanne-Rivierre (2001) for North-Central Vanuatu languages more generally.

### 5.1.2. The Torres languages

In conformity with principles of regular sound change in the region (François 2005:462), *\*ki<sup>r</sup>da* is reflected by monosyllables in most NV languages, but remains a disyllable in Vera'a /*γi<sup>n</sup>di*/ as well as in the Torres islands.

Among Torres languages, the Toga dialect of Lo-Toga is straightforward: *\*ki<sup>r</sup>da* > *\*γi<sup>n</sup>də* > /*γi<sup>n</sup>tə*/. However, the Lo variety of Lo-Toga, as well as Hiw, reflect interferences between the free pronoun and the set of subject clitics inherited from Proto Torres–Banks [§3.3.3].

The free pronoun /*γi<sup>n</sup>tə*/ of Toga is commonly followed by a proclitic /*γə=*/, the aorist marker for plural subjects:

- (6) TGA    *wə*        *ven mə*,    *γi<sup>n</sup>tə*    *γə=*    *ŋ<sup>w</sup>ulə!*  
 AO:2sg=    go    hither    linc:pl    AO:pl=    return  
 ‘Come, let us all go back!’

The Lo dialect has coalesced these two morphemes /*γi<sup>n</sup>tə* + *γə*/ into a truncated form /*təγə*/:

- (7) LO        *wə*        *ven mə*,    *təγə*    *ŋ<sup>w</sup>ulə!*  
 AO:2sg=    go    hither    linc:pl    return  
 ‘Come, let us all go back!’

Finally, Lo has generalised this shorter form /*təγə*/ as the linc:pl pronoun not just with aorist clauses, but in all contexts (François 2009:190):

- (8) TGA    *γi<sup>n</sup>tə*    *mi nə=*    *tə<sup>n</sup>aləwo*    *mərən*    *li*    *vənə*.  
 LO        *təγə*    *mi nə=*    *tə<sup>n</sup>aləwo*    *mərən*    *li*    *vənə*.  
           linc:pl with ART=    celebration    tomorrow    LOC    village  
 ‘We’re having [*lit.* We are with] a celebration tomorrow in the village.’

The process of fusion {*\*ki<sup>r</sup>da* + *ka=*} > /*təγə*/ is analogous to the one at the source of the 1sg pronoun in the same language, /*nəkə*/ < { *\*nau* + *\*gu=* } [§4.1.1].

The form /*titə*/ in neighbouring Hiw has a similar origin. In Hiw, the aorist clitic for linc:pl is /*tə=*/ (reflecting POc *\*ta=*, see Table 3 p.9). The expected reflex of *\*ki<sup>r</sup>da*, namely *\*γitə*, was evidently fused with the clitic /*tə=*/: *\*(γi)tə*+*tə=* → /*titə*/ [see fn.8 p.14].

### 5.1.3. Three Banks languages

The linc:pl pronoun shows a further innovation in three languages of central Banks islands: MTA *nina*, VRS/MSN *nin*.

According to regular sound correspondences (Table 2 p.7), the expected reflex of *\*ki<sup>r</sup>da* in these languages would have been MTA *\*γina*, VRS/MSN *\*γim*. The forms *nina* and *nin* thus reflect an irregular change in the pronoun’s first consonant. The reason is unknown; it may be due to analogy with 3<sup>rd</sup> person pronouns, which also start with /n-/.

### 5.1.4. Summary

The table below lists the innovations affecting the 1inc:pl pronoun (POc \**kita*).

- (i12) \**kita* → \**ki<sup>n</sup>da*
- (i13) \**ki<sup>n</sup>da* → \**ni<sup>n</sup>da*
- (i14) accretion of { \**ki<sup>n</sup>da* + *ka*= } in all contexts
- (i15) accretion of { \**ki<sup>n</sup>da* + *ta*= } in all contexts

<i>i</i>	Hw	LO	TGA	LHI	LVP	VWV	MTT	LMG	VRA	VRS	MSV	MTA	NDM	DRG	KRO	OLR	LKV	MRL
(i12)	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
(i13)	-	-	-	-	-	-	-	-	-	+	+	+	-	-	-	-	-	-
(i14)	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
(i15)	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

## 5.2. First exclusive plural

### 5.2.1. Initial consonant

The forms of the first exclusive and second person non-singular were similar in POc: 1exc:pl \**ka[m]i* ~ \**kamami*; 2pl \**ka[m]u* ~ \**kamiu*. In modern languages, these pronouns tend to show parallel evolution.

Most NV languages reflect an irregular change of the first consonant, from an ORAL-GRADE stop \**k* to a NASAL-GRADE one (\**g*): 2pl \**kamiu* → \**gamiu* > Hw /*kəmi*/, VLW /*gimi*/, MTA /*kamiu*/, DRG /*kmi*/, MRL /*kəmi*/... In all languages, this fortition affected the two pronouns equally, for all non-singular numbers. Incidentally, this morphological change is reflected in a vast area – covering virtually all of Vanuatu (Clark 2009:58) – and must have diffused at an early historical time, when intelligibility was still high across the archipelago.

Lakon is the only language in the area which does not show this change in the first consonant: its forms /*yamə*/ ‘1exc:pl’ and /*yamu*/ ‘2pl’ point to an ORAL-GRADE consonant \**k*. There are two ways to interpret this.

Lakon could simply be the only language to have preserved the original consonant of POc: amidst an area where all other languages have changed \**k* to \**g*, Lakon would constitute a pocket of retention. This scenario is not implausible, especially considering how Lakon lies at the end of the Gaua dialect chain, relatively isolated from the main roads of linguistic diffusion in North Vanuatu (François 2014:182). Such a hypothesis requires that pre-Lakon already existed as a separate dialect capable of forming a pocket of retention independent of its neighbours when the fortition (\**k* → \**g*) took place in Vanuatu. Surprising though this hypothesis may be,<sup>12</sup> it is made plausible by the existence of other conservative features of Lakon.

<sup>12</sup> For example, Lakon is one of the few languages to have preserved certain etymological word-final consonants, when all northern Vanuatu languages have lost them (François 2005:479, 2011:200).

Alternatively, it may be the case that Lakon underwent the same innovation as all the other languages in the region (\*k → \*<sup>ng</sup>k > /k/, yielding forms like \**kamae* and \**kamu*); but that it later went through a second innovation of its own, whereby its first consonant changed back from NASAL to ORAL GRADE. That change, whereby putative \**kamae* and \**kamu* would have changed their initial consonant to /ɣamæ/, /ɣamu/, could be due to analogy with other plural pronouns, thus yielding the plural pronoun paradigm /ɣɪŋ/, /ɣamæ/, /ɣamu/, /ɣɪ:/.

Both hypotheses seem equally costly, and equally plausible (see Clark 2009: 58).

### 5.2.2. Ending of 1 exclusive plural

Let us now look at the ending of the 1<sup>st</sup> exclusive plural form. In order to avoid being distracted by the issue of the initial consonant being ORAL or NASAL GRADE [§5.2.1], I will occasionally transcribe it using uppercase K (e.g. \**Kamami*).

POc is reconstructed with three possible protoforms for 1exc:pl: \**kai*, \**kami*, \**kamami*. As for Northern Vanuatu languages, they all point to two possible protoforms: \**Kamami* and \**Kamai*. \**Kamai* is reflected in Hiw /kama/, in Vurës /kɛmɛk/ (see below), and in the five Gaua languages; otherwise \**Kamami* is found. Examination of other Vanuatu languages (Clark 2009:161) shows that the two cognate sets are scattered throughout the archipelago, and are likely to be both ancient. Yet if Ross' POc reconstructions are correct, \**Kamai* is the innovative form among the two.

The final /-k/ in Vurës demonstrably results, once again, from the accretion of the free pronoun \**Kama(m)i* with an aorist clitic – etymologically \**ngu* '1sg', which Vurës has generalised to other numbers (François 2009:194). It is difficult to assess whether the original form before that fusion was \**Kamai* or \**Kamami*.

### 5.2.3. Summary

The table below lists the innovations affecting the 1exc:pl pronoun.

(i16) \*kam- → \*<sup>ng</sup>gam-

(i17) distribution of (innovative?) \**Kamai* instead of \**Kamami*

<i>i</i>	HW	LO	TOA	LH	LTP	VLW	MTP	LMG	VRA	VRS	MSN	MTA	NUM	DRG	KRO	OLR	LKN	MRL
(i16)	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	?	+
(i17)	+	-	-	-	-	-	-	-	-	?	-	-	+	+	+	+	+	-

### 5.3. Second plural

The 2pl pronoun is reconstructed as \**ka[m]u* or \**kamiu*. All NV languages, except Lakon, reflect a change to a NASAL-GRADE initial consonant, just like for 1exc:pl, and point to a protoform \**ngamiu*.

The Lakon form /ɣamu/ is noteworthy on several regards: first, because it is the only form which has preserved the ORAL GRADE of the initial consonant /ɣ/ [see

§5.2.1]; second, because it is one of the few languages that has preserved the pretonic vowel /a/ [see §2.1.2]; third, because it shows an unexpected final vowel /u/. This form /*yamu*/ evidently reflects, not \**Kamiú*, but a form like \**Kamúi* (with vowel metathesis), \**Kamúu* or \**Kamúyu*. Interestingly, \**gamuyu* is the form reconstructed by Clark (2009:161) at the level of “Proto North Central Vanuatu”; indeed, other cases like Lakon are found further south in Vanuatu (e.g. Ninde *kamu* on Malakula island, Eton *kamus* on Efate). One possible scenario is that all NV languages once had such a diphthong \**uy*, which was most often resolved into a front vowel /i/ (\**-uyu* > \**-iu* > /-i/); if so, Lakon would be the only language in the northern area to preserve a trace of the back vowel (\**-uyu* > /-u/). A further argument for reconstructing a vowel /u/ at least for Gaua languages, is the fact that it is also present in dual pronouns [§6.3]: forms like OLR *kumu*<sub>j</sub>, LKN *yamu*: ‘2dual’ point to an etymon \**Kamu-ru* (not \**Kami-ru*).

That said, the historical scenario remains speculative here. This is because, among the NV languages whose 2pl pronoun ends in /-i/, we have no way to determine which languages inherited their form directly from POc \**Kamiu*, and which ones went through the reversal process \**Kamiu* → \**Kamuyu* > \**Kamiu*.

#### 5.4. Third plural

##### 5.4.1. Banks languages

Proto-Oceanic is reconstructed as having a 3pl pronoun of the form \**ira* or \**kira*. A first observation is that the 15 Banks languages all have 3pl pronouns that point to a former disyllabic etymon of the form \**Cira*, even if this has been often obscured by regular sound change. For example, MRL /*kɛ̄aɾ*/ and VLW /<sup>h</sup>*gɪj*/ presuppose \**gira*; LMG /*tæɾ*/ points to \**nɔ̄dira*; VRS /*nɪɾ*/ and OLR /*nɪj*/ reflect \**nira*...

Among Banks languages, Lakon /*ɣɪ*/ is the only form which unambiguously points to \**kira* [see §2.2.3], and is possibly conservative of the POc form; all other languages have changed the pronoun’s first consonant. Five languages (LHI, LYP, VLW, MTP, MRL) reflect a NASAL-GRADE velar \**g*, which could well result from the fortition of the etymological \**kira* (→ \**gira*). If so, this could also be the source of the spread of initial \**g* to 3sg in these same languages, which have all replaced \**nia* with an irregular \**gia* [§4.3].

VRA /<sup>n</sup>*dirɪ*/ and LMG /*tæɾ*/ reflect \**nɔ̄dira*, with an initial consonant also found for 3sg (\**nɔ̄dia* > VRA /<sup>n</sup>*dii*/, LMG /*ti*/).

Most NV languages show an initial /n/ for 3pl (\**nira*). It is likely that this results from analogy with 3sg (\**nia*), whose initial /n/ was most probably inherited from the 3sg possessive suffix \**-ña* of POc [§4.3]. Notice here that I am suggesting two reverse directions for analogical levelling (see Table 6): languages with an initial velar for their 3<sup>rd</sup> person pronouns would show an influence of the plural form upon the singular; whereas languages with a nasal would illustrate the reverse alignment.

Table 6 – Initial consonant in 3<sup>rd</sup> person pronouns: what direction for the analogy?

Language	3sg	direction	3pl
Proto Oceanic	*ia (PRON) *-ña (POSS)		*[k]ira
Proto Torres–Banks	*nia	≠	*kira
Lakon	/ni/	≠	/yɪ:/
Mwotlap	/ki/	←	/kiɟ/
Dorig	/ni/	→	/niɪ/

#### 5.4.2. Torres languages

The two Torres languages are exceptional, in that their 3pl pronoun reflects a protoform \*Cisa rather than \*Cira: H1W /sisə/, LTG /nihə/. At first sight, the reason for this consonant change is unclear.

However, one crucial piece of information is that the prenasalised trill of POC \**nɾ* has apparently – and surprisingly – merged with \*s in the two Torres languages.<sup>13</sup> As Table 3 (p.9) showed, POC associated 3pl usually with a form \*ra, except for the possessor suffix where it had a prenasalised suffix \**<sup>n</sup>ra*. The latter is regularly reflected in the Torres languages as H1W /-sə/, LTG /-hə/: e.g. POC \**<sup>m</sup>b<sup>w</sup>atu-<sup>n</sup>ra* ‘their heads’ > \**k<sup>w</sup>atu-sa* > H1W /k<sup>w</sup>iti-sə/, LTG /k<sup>w</sup>ətu-hə/.

The Torres languages have generalised that { \**<sup>n</sup>ra* > \**-sa* } form to all instances of 3pl pronouns. For example, they have an object suffix of the same form (see Table 5 p.13). It is therefore very likely that what seems to be a \*Cisa reconstruction for 3pl is really the reflex of an earlier \*C<sup>i</sup>ra, itself ultimately modelled, by analogy, on the 3pl \**-ra* possessive suffix of POC.

The initial consonant of that 3pl pronoun differs between the two languages. Lo-Toga has /nihə/ ‘3pl’, which is obviously parallel with /niə/ ‘3sg’. This is reminiscent of the \*nira form reconstructible for various languages of the Banks, except that Lo-Toga /nihə/ really points to a local variant \*ni<sup>n</sup>ra. Using the symbols ‘→’ for *morphological innovations* and ‘>’ for *regular sound change* [see fn.7 p.7], I reconstruct: POC \*[k]ira → \*nira → \*ni<sup>n</sup>ra > \*nisa > LTG /nihə/.

As for the initial consonant of H1W /sisə/, it is due to a process of analogy with the pronouns of 3sg /ninə/ and of 1inc:pl /titə/. The sequence of change for H1w could thus be POC \*[k]ira → \*nira → \*ni<sup>n</sup>ra > \*nisa > \*nisə → /sisə/.

#### 5.4.3. Summary

The table below lists the innovations affecting the 3pl pronoun (POC \*[k]ira).

- (i18) \*[k]ira → \*<sup>n</sup>gira
- (i19) \*[k]ira → \*nira
- (i20) \*nira → \*<sup>n</sup>dira
- (i21) \*nira → \*ni<sup>n</sup>ra > \*nisa
- (i22) \*ni<sup>n</sup>ra > \*nisa → /sisə/

<sup>13</sup> Thus, the tree *Pipturus argenteus*, whose name is locally reconstructible as \**a<sup>n</sup>rom<sup>w</sup>ea*, becomes LHI /<sup>n</sup>dɪŋ<sup>w</sup>je/, but H1W /asəŋ<sup>w</sup>je/ and LTG /həŋ<sup>w</sup>i/.

<i>i</i>	Hiw	Lo	TGA	LHI	LYP	VLW	MTP	LMG	VRA	VRS	MSN	MTA	NUM	DRG	KRO	OLR	LKN	MRL	
(i18)	–	–	–	+	+	+	+	–	–	–	–	–	–	–	–	–	–	–	+
(i19)	?	+	+	–	–	–	–	?	?	+	+	+	+	+	+	+	+	+	–
(i20)	–	–	–	–	–	–	–	+	+	–	–	–	–	–	–	–	–	–	–
(i21)	+	+	+	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
(i22)	+	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–

## 6. DUAL FORMS

There is clear evidence that the dual pronouns were historically – at the level of Proto Oceanic – formed by combining the plural radical with the numeral *\*rua* ‘two’: e.g. *\*kita=rua* ⟨1inc:pl=two⟩ ‘1inc:du, you and I’. However, dual pronouns in modern NV languages are not compositional any more, and rest on etymons which have evidently long become opaque. This is clear from the fact that the morphological connection between plural and dual forms has usually been lost: for example, the etymon *\*kitarua* has evolved in modern languages as a single word, more or less independently of the evolution of *\*kita* alone, and/or of *\*rua* alone.

The data set in (9) shows the form of the numeral ‘two’ in NV languages. While these numerals take different prefixes (not shown here), their radicals are all regular reflexes of POC *\*rua* ‘two’:

- (9) Reflexes of the numeral ‘two’:  
 POC *\*rua* > Hiw *-ʉlɔ*; LTG *-ruɔ*; LHI *-jɔ*; LYP *-jɔ*; VLW *-jɔ*; MTP *-jɔ*;  
 LMG *-ru*; VRA *-ruw*; VRS *-ru*; MSN *-ru*; MTA *-rua*; NUM *-ru*; DRG *-ru*;  
 KRO *-ru*; OLR *-ru*; LKN *-ru*; MRL *-ru*.

This set will be useful below, as we examine the varying degree of compositionality shown by modern dual pronouns in the area. For example, LMG /ɣætru/ ‘1inc.dual’ could potentially parse as /ɣæɬ/ ‘1inc.pl’ + /-ru/ ‘two’; whereas MTA /nara/ cannot be analysed synchronically as the combination of /nina/ and /-rua/.

### 6.1. First inclusive dual

All modern forms of the first inclusive dual ultimately reflect *\*kita=rua*.<sup>14</sup> To be precise, they all show a trace of the irregular change *\*kita* > *\*ki<sup>n</sup>da* which is also characteristic of 1inc:pl [§5.1.1]. In other words, the common protoform underlying all 1inc:du pronouns in NV would be *\*ki<sup>n</sup>darua*. At an early stage, that form would have been recognisable as the combination of *\*kita* ~ *ki<sup>n</sup>da* ‘1inc:pl’ plus

<sup>14</sup> The dialectology of 1inc:du pronouns in northern Vanuatu is also discussed in François (2011:201 sqq.).



the numeral ‘two’, but processes of sound change have blurred this morphological connection in most languages.

### 6.1.1. One etymon, many changes

Only two modern NV forms are still compositional: LMG /ɣæt-ru/, MSN /nin-ro/. It is quite possible that these are not conservative, but reflect a recent recreation of dual pronouns based on plural forms. Other languages have more opaque forms. For example, LHI /ɣinjɔ/ reflects *\*ki<sup>n</sup>darua*, but shows apophony in comparison with the plural /ɣen/.

Some languages went through syncope, as *\*ki<sup>n</sup>darua* was shortened to *\*ki<sup>n</sup>dua* > LYP /jɛ<sup>n</sup>du/; VRA /ɣi<sup>n</sup>duo/. This is also the etymon of LKN /wɔfɔ/, which reflects a series of regular changes: *\*ki<sup>n</sup>dua* > *\*ɣə<sup>n</sup>du* > *\*ɣəfɔ* > *\*ɣofɔ* > /wɔfɔ/.

All other languages reflect a protoform showing loss of the initial (unstressed) syllable: *\*ki<sup>n</sup>darua* > *\*n<sup>n</sup>darua* > MRL /<sup>n</sup>dorɔ/, KRO /<sup>n</sup>duru/, MTP /<sup>n</sup>dɔjɔ/, OLR /tʃɔrɔ/, HIW /tə<sup>9</sup>Lə/... Notice how the vowel in *\*n<sup>n</sup>da-* systematically assimilated to the (stressed) final vowel, a regular phenomenon in the area [§2.1.1].

VRS /<sup>n</sup>dorok/ includes a non-etymological consonant /k/. This corresponds to an earlier TAM marker (see §5.2.2, and François 2009:194).

MTA /nara/ also reflects *\*n<sup>n</sup>darua*, but with an irregular simplification of *\*-rua* to *\*-ra* – a change found for all its dual pronouns: /nara/, /kara/, /kamra/, /rara/.

DRG /<sup>n</sup>da:r/ cannot reflect *\*n<sup>n</sup>darua*, and points to an irregular protoform *\*n<sup>n</sup>dāru*. This likely results from analogy with lex:du /kma:r/ < *\*<sup>9</sup>gamāru* – a truncated form which is attested in other languages [§6.2].

Lo-Toga has an irregular linc:du form /tɔr/. It also points to an etymon of the form *\*n<sup>n</sup>dāru*, but with an unexpected vowel (*\*n<sup>n</sup>dāru* would normally have yielded a lower vowel *\*\*/tɔr/*).

Finally, MTP-VLW /<sup>n</sup>du/ and VRA /<sup>n</sup>du/ are cliticised variants of the pronoun, restricted to subject position. They are both shortened forms of the free pronoun – respectively /<sup>n</sup>dɔjɔ/ and /ɣi<sup>n</sup>duo/. Surprising though it seems, /<sup>n</sup>du/ and /<sup>n</sup>du/ thus ultimately descend from *\*kita=rua*, through layers of phonological and morphological change.

### 6.1.2. Summary

<i>i</i>	HIW	LO	TGA	LHI	LYP	VLW	MTP	LMG	VRA	VRS	MSN	MTA	NUM	DRG	KRO	OLR	LKN	MRL
(i23)	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
(i24)	–	–	–	–	+	–	–	–	+	–	–	–	–	–	–	–	–	+
(i25)	+	+	+	–	–	+	+	+	–	+	?	+	+	+	+	+	–	+
(i26)	–	+	+	–	–	–	–	–	–	–	–	–	–	+	–	–	–	–
(i27)	–	–	–	–	–	–	–	–	–	–	–	+	–	–	–	–	–	–

Let us list all the innovations affecting the linc:du pronoun (POc *\*kita=rua*).

- (i23) *\*kitarua* → *\*ki<sup>n</sup>darua*  
 (i24) *\*ki<sup>n</sup>darua* → *\*ki<sup>n</sup>dua*

- (i25) \*ki<sup>n</sup>darua → \*<sup>n</sup>darua  
 (i26) \*<sup>n</sup>darua → \*<sup>n</sup>daru  
 (i27) \*<sup>n</sup>darua → \*<sup>n</sup>dara

## 6.2. First exclusive dual

### 6.2.1. Micro-paradigms in each language

Almost all languages in the area show very similar forms for 1exc:du and 2du pronouns, often yielding minimal pairs; yet they do so in unpredictable ways, which cannot be explained easily using a single historical scenario. For example :

- LTG /kəmər/ ‘1exc:du’ vs. /kəmər/ ‘2du’ only differ in the height of the last vowel
- LHI /mæjə/ vs. /mɔjə/ (respectively) differ in the quality of the first vowel
- LYP /mɔmjə/ vs. /mɔjə/ differ in the presence of a consonant /m/ for 1exc:du
- MTA /kara/ vs. /kamra/ also contrast by the presence of /m/, but with a reverse pattern from Löyöp
- VRS /kómorək/ vs. /kómorəŋ/ show a different final (non-etymological) consonant...

Evidently, each language has created its own micro-paradigm for dual pronouns, in which 1exc:du and 2du have tended to coevolve and influence each other so as to enter in a minimal-pair relationship. Morphological levelling and analogy must have taken place, not only between these two pronouns in each language, but also between the dual, trial and plural forms in the same system. Therefore, it makes limited sense to compare languages by examining each pronoun set individually as we’ve been doing in previous sections; this approach may allow us to propose protoforms for each pronoun, but the intricate history of forms in each language would take us beyond the limits of this study.

### 6.2.2. Paths of evolution

Comparison of the various 1exc:du pronouns points to an ultimate protoform \*kama=rua. The initial consonant is preserved as \*k >/ɣ/ in Lakon (see our discussion of the plural forms in §5.2.1) and changed to \*g everywhere else. Again, this initial consonant will be henceforth represented as \*K (whether ORAL or NASAL GRADE).

The etymon \*Kamarúa is still transparently preserved, for example, in MTP /kamjə/, LMG /kamaru/. VRS /kómorək/ adds to this an extra consonant /k/ found elsewhere in its non-singular pronouns [§5.2.2, 6.1.1]. VRA /kama<sup>n</sup>duə/ shows an irregular change of \*r to \*<sup>n</sup>d, probably by analogy with /ɣi<sup>n</sup>duə/ ‘1inc:du’.

Several languages in the Torres and in southern Banks show evidence of a truncation \*Kamarúa → \*Kamáru, with subsequent shift in stress placement:

- (10) Irregular change for 1exc:dual: \*Kama-rúa → \*Kamáru  
 > LTG kəmər; NUM kamar; DRG kma:r; KRO kemeār; OLR kmiij; LKN yama:;  
 MRL kamar.

Remember that the long vowel of Lakon is the regular reflex of a final /r/ [§2.2.3]: \*Kamarúa → \*Kamáru > \*yamar > /yama:./

We saw that 1exc:pl pronouns had two protoforms: *\*Kamai* and *\*Kamami* [§5.2.2]. Likewise, some dual pronouns show an extra /m/, surely by analogy with their plural forms: LYP /mómjǝ/ (cf. plural /kómóm/), MSN /kémémrǝ/ (cf. plural /kémém/).

Two languages show apocope of the dual pronoun's initial syllable: LHI /mæjǝ/ (cf. plural /kímæm/), LYP /mómjǝ/ (cf. plural /kómóm/).<sup>15</sup>

Finally, Mota shows a drastic reduction of the pronoun, from *\*Kamarua* to /kara/. This is specific to this language, and involves no shared innovation with its neighbours.

### 6.2.3. Summary

Below are the most significant innovations affecting the 1exc:du pronoun (POc *\*Kama-rua*).

- (i28) *\*Kamarua* → *\*Kamaru*  
 (i29) *\*Kamarua* → *\*Kamami-rua*  
 (i30) *\*Kama(mi)rua* → *\*ma(mi)rua* (*apocope*)

<i>i</i>	HW	LO	TGA	LHI	LYP	VLV	MTP	LMG	VRA	VRS	MSN	MTA	NUM	DRG	KRO	OLR	LKN	MRL
(i28)	-	+	+	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+
(i29)	-	-	-	-	+	-	-	-	-	-	+	-	-	-	-	-	-	-
(i30)	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-

### 6.3. Second dual

The protoform which can be locally reconstructed for the 2du pronoun is *\*Kamurua*.

The quality of the first vowel is preserved, as usual [§2.1.2], in Mota and Lakon, but also in Mwerlap. All other languages have weakened that first vowel either to a schwa, or to a copy of the following high vowel: LTG /kəmor/, NUM /kumur/, NUM /kómorǝ/. In several languages, the latter high vowel, which was unstressed, subsequently disappeared: *\*Kamurúa* > *\*Kəmurúa* > *\*Kumurúa* > VRA /kumruǝ/, MTP /komjǝ/.

Just like 1exc:du *\*Kamarúa* was sometimes truncated to *\*Kamáru*, likewise the 2du *\*Kamurúa* was evidently shortened to *\*Kamúru*. This process affected essentially the same languages (except Mwerlap) :

- (11) Irregular change for 1exc:dual: *\*Kamu-rúa* → *\*Kamúru*  
 > LTG *kəmor*; NUM *kumur*; DRG *kmur*; KRO *kumur*; OLR *kumuj*; LKN *yamu*.

<sup>15</sup> See a similar apocope for 1inc:dual, *\*kí<sup>n</sup>darua* > *\*<sup>n</sup>darua* [§6.1.1].

Another form of truncation is initial apocope. Again, it took place in the same languages as we saw for 1exc:du, namely Lehali and Löyöp.

The main innovations affecting the 2du pronoun (PTB *\*Kamu-rua*) are summarised below.

- (i31) *\*Kamurua* → *\*Kamuru*
- (i32) *\*Kamurua* → *\*murua*

<i>i</i>	HW	LO	TGA	LHI	LYP	VLW	MTT	LMG	VRA	VRS	MSN	MTA	NUM	DRG	KRO	OLR	LKN	MRL
(i31)	-	+	+	-	-	-	-	-	-	-	-	-	+	+	+	+	+	-
(i32)	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-

### 6.4. Third dual

#### 6.4.1. Reconstructing a likely protoform

The morphology of 3dual pronouns is closely linked to developments in the 3<sup>rd</sup> plural pronouns. Thus, an innovation affecting the initial consonant of a plural pronoun was almost always reflected in the dual and the trial pronouns.

The diversity of modern forms makes it difficult to reconstruct a common protoform. The answer may come from the southern Banks languages, whose dual pronouns have been kept distinct from their plural counterparts. Thus Vurës has /nir/ for 3pl (from *\*nira*, itself altered from POC *\*[k]ira*), but /roro/ for 3du. The latter would be the regular reflex of an etymon *\*rarúa*. If indeed 3pl in POC had the form *\*kira* or *\*ira*, then it is reasonable to reconstruct *\*ira-rua* as the oldest possible form of the 3du pronoun in northern Vanuatu.

- (12) Reflexes of *\*(i)ra-rua* ‘3 dual’:  
 > VRS *roro*; MTA *rara*; NUM *ruru*; DRG *ra:r*; KRO (*irru*).

Among these, Dorig shows evidence of a truncation *\*rarúa* → *\*ráru*, by analogy with other persons; this is the only language of Gaua where this truncation affected all four dual pronouns.

Other languages usually aligned their 3<sup>rd</sup> dual with their 3<sup>rd</sup> plural pronoun in one way or another; this is usually reflected in the pronouns’ initial consonant. Based on 3pl *\*nira*, Mwesen and Ólrat have evidently created an analogous form *\*nirarua* or *\*narua* for 3du: *\*nirarúa* > *\*nærəro* > MSN /nŋrŋrŋ/; *\*narúa* > *\*nəro* > OLR /norŋ/.

Lemerig and Vera’ a point to a 3du form *\*ndarua* > LMG /tæru/, VRA /<sup>n</sup>duruŋ/.

Lakon did the same with its ORAL-GRADE consonant /ɣ-/ (<*\*k*-). Compare its 3pl /ɣi:/ < *\*kira* with its 3du /woro/ < *\*yoro* < *\*yərŋ* < *\*kərŋ* < *\*karua*.

Those languages in northern Banks that have NASAL-GRADE *\*ŋg-* for their plural show the same initial for their dual: *\*ŋgarua* > VLW /ŋgŋjŋ/, LYP /kjejŋ/... As for MRL /karar/, it reflects *\*ŋgarúru*, evidently the result of a local reanalysis.

Finally, the two Torres languages point to an initial \*s also found in their 3pl. As we saw in §5.4.2, this sibilant may ultimately reflect a prenasalised trill \*nr: hence \*(i)ra-rua → \*nrarúa > \*sarúa > \*səɾə > Hiw /sə<sup>9</sup>Lə/. As for LTG /hor/, it reflects later truncation.

#### 6.4.2. Summary

The table below lists the innovations affecting the 3du pronoun \*irarua.

(i33)	* (i)rarua	→ *(i)raru
(i34)	* (i)rarua	→ *n(ir)arua
(i35)	* (i)rarua	→ *ndarua
(i36)	* (i)rarua	→ *karua
(i37)	* (i)rarua	→ *ngarua
(i38)	* (i)rarua	→ *nrarua > *sarua

<i>i</i>	HW	LO	TGA	LH	LTP	VW	MTP	LMG	VRA	VRS	MSV	MTA	NDM	DRG	KRO	OLR	LKN	MRL
(i33)	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-
(i34)	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	+	-	-
(i35)	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-
(i36)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-
(i37)	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	+
(i38)	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

## 7. TRIAL FORMS

Our detailed examination of northern Vanuatu pronouns will end with the trial forms. This section will be shorter, because most of the processes involved in the morphological history of trial forms are similar to the ones we have seen for other numbers.

### 7.1. General comments on trials

#### 7.1.1. Introduction

Trials – a number referring strictly to three referents – are quite rare typologically (Corbett 2000), but relatively common within the Oceanic family. One first observation is that not all languages in our sample have a trial: the two Torres languages, as well as Mwerlap in the south, have only three numbers (singular–dual–plural). Because trials are common in Oceanic languages, and can probably be reconstructed for POC, a likely interpretation is that the NV languages with a trial are here conservative, while the ones that lack it are innovative.

Just like we saw for dual, trial pronouns originate in what was initially a transparent compound {plural pronoun + ‘three’}: e.g. \*kita + tolu > \*kita=tolu

‘1inc:trial’. Yet these compounds have become progressively opaque, and have taken up a life of their own. The data set in (13) shows the form of the numeral ‘three’ in NV languages. While these numerals take different prefixes (not shown here), their radicals are all regular reflexes of POc \*tolu ‘three’:

- (13) Reflexes of the numeral ‘three’:  
 POc \*tolu > Hiw -tøj; LTG -təl; LHI -til; LYP -fjəl; VLW -til; MTP -til;  
 LMG -ʔəl; VRA -ʔol; VRS -təl; MSN -tol; MTA -tol; NUM -tol; DRG -tol;  
 KRO -tol; OLR -til; LKN -til; MRL -təl.

### 7.1.2. The trial metathesis

The etymological order {plural pronoun + ‘three’} is preserved in most modern languages of North Vanuatu: thus VLW /<sup>h</sup>gɛmtil/ ‘1exc:tri’ clearly reflects \*<sup>h</sup>gami + tolu, in this order. But a conspicuous innovation is found in the languages of Gaua, in the form of a morphological metathesis.

This metathesis inverted the order of the two elements, so that {pronoun + \*tolu} was reshaped to {\*tolu + pronoun}. This metathesis affected trial pronouns in all Gaua languages except Nume, that is: Dorig, Koro, Oirat and Lakon. Thus ‘1inc:trial’ in Nume is /<sup>n</sup>dətəl/ < \*<sup>n</sup>datolu < \*<sup>ki</sup>da=tolu, but Dorig has /tolɣɪn/ < \*tolu=<sup>ki</sup>da. In the four languages concerned, metathesis affected all four persons; it will therefore be mentioned in each of the sections below, but counted only once when we list innovations.

### 7.2. First inclusive trial

All languages ultimately point to an etymon \*<sup>ki</sup>da=tolu – itself from POc \*kita ‘1inc:p1’ and \*tolu ‘three’.

This form has followed different paths of evolution, usually in a way parallel to its dual counterpart \*<sup>ki</sup>da=rua [§6.1]. The analogy of evolution is obvious, for example in Lemerig, if we compare the inclusive dual /<sup>y</sup>ætru/ and the trial /<sup>y</sup>æʔəl/; or in Mwesen with, respectively, /ninrɔ/ and /nintəl/; in Mota with /nara/ and /natəl/; or in Nume between /<sup>n</sup>duru/ and /<sup>n</sup>dətəl/...

Several languages reflect the apocope of the first syllable: \*<sup>ki</sup>datolu > \*<sup>n</sup>datolu > VLW-MTP /<sup>n</sup>dɪtɪl/, VRA /(<sup>y</sup>i)<sup>n</sup>dʊʔəl/, MTA /natəl/, NUM /<sup>n</sup>dətəl/.

The Gaua languages that went through trial metathesis reflect a form \*tolu=<sup>ki</sup>da. This is transparent in Dorig /tolɣɪn/, but slightly more hidden in the three other languages, because they evidently went through an intermediate form \*<sup>tol</sup>da: e.g. \*tolu=<sup>ki</sup>da → \*<sup>tol</sup>da > \*<sup>til</sup>da > KRO /tɪln/.

Oirat and Lakon have gone even further in their evolution. The affricate /tʃ/, which is the regular reflex of \*<sup>n</sup>d, has shown interference with the word-initial \*<sup>t</sup>, sometimes resulting in another metathesis, this time phonological. Thus, \*<sup>tol</sup>da is regularly reflected in Lakon as /tɪltʃ/, but other variants are also common in the same language, including /tʃiltʃ/, /tʃilit/ and /tʃilis/. For these two languages, we thus have a sequence \*tolu=<sup>ki</sup>da → \*<sup>tol</sup>da > \*<sup>til</sup>da > /tɪltʃ/ → /tʃiltʃ/.

In sum, a form like LKN /ʃilit/ shows two layers of metathesis: one morphological (\**ki<sup>r</sup>da=tolu* > \**tolu=ki<sup>r</sup>da...*), one phonological (\**til<sup>r</sup>d* > /tɪltʃ/ → /ʃilit/). Forms like LKN /ʃilit/ or VRA /<sup>n</sup>dʊʔol/ ultimately reflect POc \**kita* – a connection which is far from obvious at first glance.

### 7.3. First exclusive trial

Modern languages point to a set of protoforms \**ḡami-tolu* (> VLW /<sup>n</sup>ḡemɪtɪl/), \**ḡamai-tolu* (> NUM /kamatɔl/), \**ḡamami-tolu* (> LMG /kæmɕemʔəl/, LYP /mɕm-ʃəl/). Their distribution is generally consistent with the dual and plural forms of the same languages.

Among metathesising languages, DRG /tɔl-kma/ is transparent. Koro and Olrat show the assimilation of the initial \**t* to /k/: \**tolu-ḡamái* > \**tɪl-kama* > /kɪlkama/. As for Lakon /tɪlmæ/, it reflects the loss of the velar consonant: \**tolu-yamai* > \**tɪlymæ* > /tɪlmæ/.

### 7.4. Second trial

Trial pronouns of the 2<sup>nd</sup> person point sometimes to a protoform \**ḡámu-tólu* (> VLW /<sup>n</sup>ḡɔmtɪl/), sometimes to a longer variant \**ḡamú-tólu* (> VRS /kɪmɪtɔl/).

All the developments of the second trial pronoun are parallel with the first exclusive trial we just saw. This is also true of the languages showing metathesis: the only difference between DRG /tɔlkma/ and /tɔlkmi/, or LKN /tɪlmæ/ and /tɪlmu/, is in the nature of the vowels.

### 7.5. Third trial

For the third dual pronoun, we had reconstructed a likely etymon \**ira-rua*; it is likely that the trial equivalent was simply \**ira-tolu*, as witnessed by MTA /ratol/ and NUM /rotɔl/.

This pronoun later went through the same morphological changes as the third person pronouns of other numbers [§5.4, 6.4.1]. Thus most languages reflect a pronoun \**Cira-tolu*, with a different consonant: \**nɪra-tolu* (> LMG /tærʔəl/), \**ḡira-tolu* (> LYP /kɛjɛʃəl/), \**nira-tolu* (> VRS /nɔrtɔl/)...

Among the metathesising languages of Gaua, three reflect a form \**tolíra* > KRO /tɪlɪr/, OLR /tɪlɪj/, LKN /tɪlɪ:/. It is difficult to know if this \**tolíra* form results directly from the metathesis of \**ira-tolu*; or if these languages' ancestor once went through a longer form \**tolu-nira*. DRG /tɔlnɪr/ apparently supports that latter hypothesis, but it might as well reflect a more recent recreation of a regular trial paradigm in that language, based on plural forms. Conversely, Lakon shows no evidence of any 3<sup>rd</sup> person pronoun beginning in /n-/ anywhere in its paradigm, so it would be costly to reconstruct an intermediate phase \**tolu-nira* (introducing an /n/) before \**tolira* (showing subsequent deletion of that /n/). Following Occam's razor, it is more economical to propose that \**ira-tolu* metathesised directly to \**tolíra*, a form which is reflected in all modern languages except Dorig (the latter having gone

through analogy with its plural forms). That trisyllabic pronoun *\*tolíra* would then have become a model which attracted the irregular change of 1inc:trial *\*tolu-ki<sup>n</sup>da* → *\*tolí<sup>n</sup>da* [§7.2].

7.5. Summary

The table below lists the principal innovations affecting the trial pronouns.

- (i39) loss of trial number
- (i40) *Plur.* + *\*tolu* → *\*tolu* + *Plur.* (metathesis)
- (i41) *\*toluki<sup>n</sup>da* → *\*tolí<sup>n</sup>da*
- (i42) *\*ki<sup>n</sup>datolu* → *\*<sup>n</sup>datolu*
- (i43) *\*(i)ratolu* → *\*<sup>n</sup>giratolu*
- (i44) *\*(i)ratolu* → *\*<sup>n</sup>díratolu*
- (i45) *\*(i)ratolu* → *\*<sup>n</sup>niratolu*
- (i46) *\*(i)ra-tolu* → *\*tolíra*
- (i47) *\*tolu-Kamai* → *\*Kolu-Kamai*

<i>i</i>	HW	LO	TGA	LHI	LYP	VLW	MTP	LMG	VRA	VRS	MSN	MTA	NUM	DRG	KRO	OLR	LKN	MRL	
(i39)	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+
(i40)	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	-	-
(i41)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	-
(i42)	-	-	-	-	-	+	+	-	+	-	-	+	+	-	-	-	-	-	-
(i43)	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-
(i44)	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-
(i45)	-	-	-	-	-	-	-	-	-	+	+	-	-	+	-	-	-	-	-
(i46)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	-
(i47)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-

8. DISCUSSION: TREE OR LINKAGE?

8.1. Comparative Method and family structure

The previous sections reconstructed the history of 15 independent personal pronouns in the 17 languages of north Vanuatu. If we count variant forms, we discussed 260 pronominal forms in total. For each pronoun form or subparadigm, I reconstructed one or more individual innovations, and identified which languages took part in them.

Every step of the demonstration has been faithful to the tenets of the Comparative Method. One important principle – first formulated by Leskien (1876) –



was to distinguish shared retentions from *shared innovations*, as only the latter can tell us anything about the genealogical history of a linguistic family. In doing so, I took as my reference the reconstructions proposed for Proto Oceanic, which scholars have established based on a large number of Oceanic languages, also following the Comparative Method. As I reconstructed these innovations, another important rule of thumb was to adhere to the principle of regular sound change within each language – what Neo-Grammarians called *sound laws*. Throughout this study, I thus endeavoured to apply the Comparative Method with utmost rigour.

The question that was raised in the introduction concerned the next step, namely the best way to interpret the distribution of innovations in terms of a family structure. The Comparative Method tends to be associated with the tree model, but is this a legitimate view? Specifically, does the genealogical history of northern Vanuatu languages fit a tree structure?

In other publications (François 2014, *f/c*; Kalyan & François *f/c*), I discussed the advantages and disadvantages of the tree model, and highlighted its inability to deal with intersecting isoglosses, though these are common in dialect networks and linkages [see §1.1]. I illustrated this principle with the 17 languages of northern Vanuatu, for which I reconstructed 474 innovations of various natures: regular or irregular sound change, morphology, syntax, lexicon.

The resulting matrix (with a structure analogous to the tables in §4.1.2 sqq. in this study) was then analysed using an approach called *Historical Glottometry*, the purpose of which is to identify genealogical subgroups and quantify their level of support.

One possible way to represent the results of that study took the form of a “glottometric diagram” (Figure 2). This diagram shows the best supported subgroups, using line thickness to visually reflect their respective strength. The family shows various cases of intersecting subgroups: for example, the Lemerig language (LMG) belongs to a subgroup LMG–VRA, but also to a larger Northern-Banks subgroup LHI–LYP–VLW–MTP–LMG which crosscuts it. Likewise, the bottom of the diagram shows how the languages of Southern Banks form a perfect dialect chain (François 2014:182). That 2014 glottometric study thus demonstrated that the Northern Vanuatu family cannot be appropriately rendered using the tree model, which would have assumed that all subgroups must be discrete and embedded.

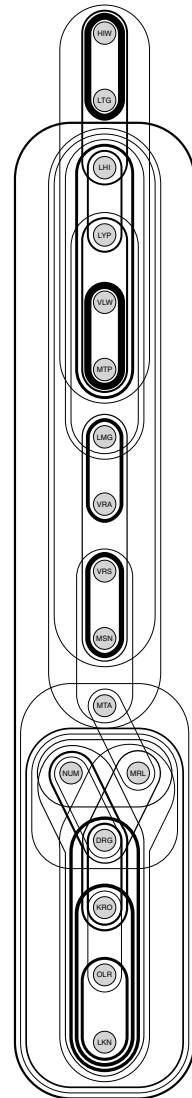


Figure 2 – A glottometric diagram of Northern Vanuatu languages (from François 2014)

## 8.2. The intersecting isoglosses of pronoun morphology

As for the present study, its purpose was to focus on one particular domain of historical evidence, the morphology of personal pronouns. I identified a total of 47 innovations – several of which also featured among the 474 that were used in the 2014 glottometric study underlying *Figure 2*. While there are sometimes debates on the diagnostic value, for subgrouping purposes, of processes such as lexical replacement, the idea in studying pronoun morphology was that the sort of innovations it defines are not easily borrowable across separate languages, and should thus be viewed as uncontroversially solid evidence for grouping languages together in a genealogical sense [§1.2].

At this point of our discussion, the question that arises is the following. Are the morphological innovations concerning pronouns distributed along a tree structure, or do they form entangled isoglosses?

A first examination of the pronominal data suggests that a tree-like representation of the Northern Vanuatu family should not, in fact, be ruled out completely. The tree given in *Figure 3*, based on the pronoun data presented here, is the most parsimonious tree possible, i.e. the one capable of accommodating most of the isoglosses we have identified, and assign shared innovations to nodes in the most economical way. Those innovations which fit the tree have their number code indicated on the relevant node: for instance, the subgroup/node *VRS-MSN-MTA* is captured by two innovations, (*i4*) and (*i13*).<sup>16</sup>

If the genealogy of Northern Vanuatu languages had to be forced into a tree structure, then it would likely be very similar to *Figure 3*.<sup>17</sup> However, the trouble with such a tree is that it only accommodates a portion of the empirical evidence we have. Among the 47 innovations related to pronouns, six took place in just a single language (e.g. (*i11*, 22) for Hiw, (*i33*) for Dorig...): they can be disregarded here, because they are obviously compatible with any representation, and can't help us determine whether a family is tree-like or not. Among the 41 remaining isoglosses, only 24 are compatible with the tree in *Figure 3*: the other 17 are not, as they straddle across branches.<sup>18</sup> Thus, (*i43*) includes *LHI-LYP-VLW-MTP*, but (*i42*) has *VLW-MTP-VRA-MTA-NUM*. Likewise, Dorig takes part in an isogloss (*i40*) *DRG-KRO-OLR-LKN*, but also in another isogloss (*i45*) that includes *VRS-MSN-DRG*.

<sup>16</sup> In *Figure 3*, the numbers in square brackets correspond to cases when a given node is targeted by an isogloss that also includes an external branch. For example, (*i17*) targets the whole subgroup *NUM-DRG-KRO-OLR-LKN*, but also includes *HIW* (possibly reflecting a case of parallel innovation).

<sup>17</sup> Note, in passing, that every node in this tree corresponds to a subgroup in my glotto-metric diagram. This confirms my intuition, that pronominal morphology does indeed constitute a microcosm of the area's linguistic history more generally [§1.2].

<sup>18</sup> The 17 innovations incompatible with this tree structure are: (*i6*)–(*i7*), (*i16*)–(*i19*), (*i24*)–(*i26*), (*i28*)–(*i29*), (*i31*), (*i34*), (*i37*), (*i39*), (*i42*), (*i45*). See also fn.16.

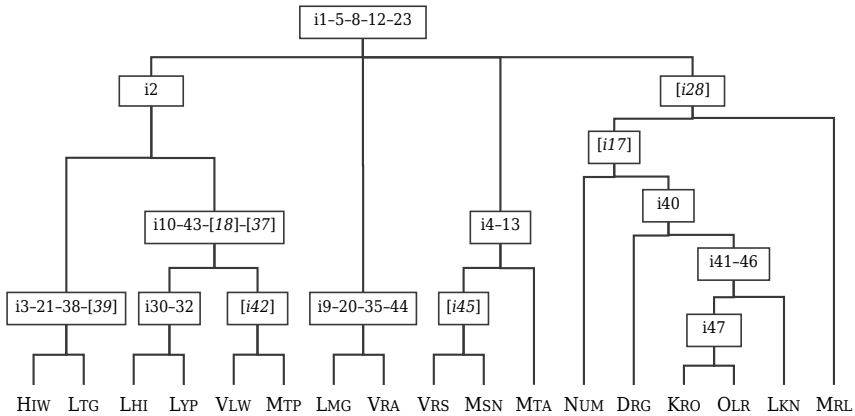


Figure 3 – A plausible (though imperfect) tree of North Vanuatu languages, based on the historical morphology of pronouns

The same conclusion would be reached if we brought in other evidence outside pronoun morphology – such as the patterns of regular sound change discussed in §2 above, or any other linguistic features. While some innovations – perhaps the majority – might confirm the tree structure given in *Figure 3*, several others would provide conflicting evidence. In order to save the tree, the only solution would be to eliminate these problematic isoglosses, a decision which is often ad hoc and illegitimate.

In sum, if we take into account all the evidence available – and not just that which is compatible with a tree – then the pattern we observe is one where sub-groups intersect. The chain-like distribution defined by the shared innovations relating to pronouns is consistent with the overall genealogical structure of the northern Vanuatu linkage, which is best represented in a glottometric diagram (*Figure 2*).

## 9. CONCLUSION

The present study aimed at more than one goal.

One prime objective was to document for the first time the pronominal systems of the 17 languages of the Banks and Torres Is, in northern Vanuatu. Given that many of these languages are endangered or even moribund (François 2012), publishing this set of firsthand data is valuable in itself.

A second purpose was to show that, in spite of the apparent diversity of these pronouns' modern forms, knowledge of the area's linguistic history makes it possible to uncover their former unity, and reconstruct a common protoform for each pronoun. For example, 1<sup>st</sup> inclusive dual pronouns in the region include such forms as /tor/, /ɣætru/, /nara/, /ɣinjɔ/, /<sup>h</sup>da:t/ and /wɔfʊ/; as heterogeneous as these

pronouns may sound, they can all safely be shown to derive from a single etymon \**ki<sup>r</sup>darua* [§6.1]. The path leading from that etymon to modern forms is a combination of (regular) sound change and (irregular) morphological innovations, whose history can be unravelled thanks to the powerful principles of the Comparative Method.

These principles enabled us to reconstruct systematically, step by step, the whole set of innovations that have affected the rich morphology of pronouns in the area, taking Proto Oceanic as a point of departure. The 47 innovations identified for Northern Vanuatu were sufficient in number to be taken as an illustration of the area's linguistic history more generally.

Finally, the morphology of independent personal pronouns in Northern Vanuatu was used as a test case for tackling a theoretical issue of broader significance: namely, the question whether the Comparative Method should be viewed as inseparable from the tree model – as is usually assumed – or if its results can also be compatible with non-cladistic approaches. I showed that the Comparative Method's high-precision tools, if applied rigorously, can be used to assess whether the genealogical structure of a family is tree-like or not. Taking the morphological history of pronouns, it was first possible to propose a tree for northern Vanuatu languages, as a rough approximation of their history [§8.2]; but such a cladistic representation forced us to leave out a large amount of data which did not fit its nested structure.

As an alternative, I proposed that this language family – like most in the world, in fact – constitutes a *linkage*, that is, a set of related languages in which internal subgroups intersect. In order to represent the history of a linkage without distorting or arbitrarily selecting our data, it is safer to forsake the assumptions of the Tree model, and adopt the more encompassing Wave model – or a model derived from it – following common practice in dialectology and sociolinguistic studies (Heggarty *et al.* 2010; François 2014).

Unusual though it may seem at first glance, a genealogical structure made of entangled subgroups simply reflects the fact that a linkage originates in a former dialect network. In the earlier times of mutual intelligibility, each dialect would typically share innovations now with one neighbour, now with another [§1.1]. Such a situation is certainly more common in the world than is usually assumed among historical linguists. Hopefully, this study will inspire readers to develop non-cladistic approaches to language genealogy, while still remaining faithful to the powerful insights of the Comparative Method.

#### LANGUAGE ABBREVIATIONS

The abbreviations for language names appear on *Map 1*, and are repeated below.

DRG	Dorig	LTG	Lo-Toga	OLR	Oirat
HIW	Hiw	LYP	Löyöp	POc	Proto Oceanic
KRO	Koro	MRL	Mwerlap	PTB	Proto Torres–Banks
LHI	Lehali	MTA	Mota	TGA	Toga (dialect of LTG)
LKN	Lakon	MTP	Mwotlap	VLW	Volow

LMG	Lemerig	MSN	Mwesen	VRA	Vera'a
LO	Lo (dialect of LTG)	NUM	Nume	VRS	Vurès

## REFERENCES

- Bedford, Stuart & Matthew Spriggs. 2008. Northern Vanuatu as a Pacific cross-roads: The archaeology of discovery, interaction, and the emergence of the “ethnographic present”. *Asian Perspectives* 47, 95–120.
- Bloomfield, Leonard. 1933. *Language*. New York: Holt.
- Clark, Ross. 2009. \**Leo Tuai: A comparative lexical study of North and Central Vanuatu languages*. Pacific Linguistics, 603. Canberra: Australian National University.
- Corbett, Greville. 2000. *Number*. Cambridge: Cambridge University Press.
- Dempwolff, Otto. 1938. *Vergleichende Lautlehre des Austronesischen Wortschatzes: Austronesisches Wörterverzeichnis*. Liechtenstein: Nendeln.
- Dik, Simon. 1989. *The Theory of Functional Grammar. Part I: The Structure of the Clause*. Functional Grammar Series. Dordrecht: Foris.
- Drinka, Bridget. 2013. Phylogenetic and areal models of Indo-European relatedness: The role of contact in reconstruction. *Journal of Language Contact* 6, 379–410.
- Evans, Bethwyn. 2003. *A study of valency-changing devices in Proto Oceanic*. 539. Canberra: Pacific Linguistics.
- François, Alexandre. 2003. *La sémantique du prédicat en Mwoitlap (Vanuatu)*. Collection Linguistique de la Société de Linguistique de Paris. Leuven-Paris: Peeters. 408 pp.
- . 2005. Unraveling the history of the vowels of seventeen northern Vanuatu languages. *Oceanic Linguistics* 44, 443–504.
- . 2009. Verbal aspect and personal pronouns: The history of aorist markers in north Vanuatu. In Andrew Pawley & Alexander Adelaar (eds.), *Austronesian historical linguistics and culture history: A festschrift for Bob Blust*, 179–195. Canberra: Pacific Linguistics.
- . 2010. Phonotactics and the prestopped velar lateral of Hiw: Resolving the ambiguity of a complex segment. *Phonology* 27, 393–434.
- . 2011. Social ecology and language history in the northern Vanuatu linkage: A tale of divergence and convergence. *Journal of Historical Linguistics* 1, 175–246.
- . 2012. The dynamics of linguistic diversity: Egalitarian multilingualism and power imbalance among northern Vanuatu languages. *International Journal of the Sociology of Language* 214, 85–110.
- . 2014. Trees, Waves and Linkages: Models of Language Diversification. In Claire Bowers & Bethwyn Evans (eds), *The Routledge Handbook of Historical Linguistics*. New York: Routledge, 161–189.
- . f/c. Méthode comparative et chaînages linguistiques : Pour un modèle diffusionniste en généalogie des langues. In J.-L. Léonard (ed.), *Diffusion : implantation, affinités, convergence*. Mémoires de la Société de Linguistique de Paris, XXIV. Louvain : Peeters.
- François, Alexandre; Michael Franjeh; Sébastien Lacrampe; Stefan Schnell. 2015. The exceptional linguistic density of Vanuatu. In A. François; S. Lacrampe; S. Schnell & M. Franjeh (eds), *The Languages of Vanuatu: Unity and*

- Diversity*. Studies in the Languages of Island Melanesia. Canberra: Asia-Pacific Linguistics Open Access, 1–21.
- Grace, George W. 1959. *The position of the Polynesian languages within the Austronesian (Malayo-Polynesian) language family*. Memoir 16 of the International Journal of American Linguistics. Baltimore: Waverly Press.
- Heggarty, Paul; Warren Maguire & April McMahon. 2010. Splits or waves? Trees or webs? How divergence measures and network analysis can unravel language histories. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365, 3829–3843.
- Kalyan, Siva & Alexandre François. forthcoming. Freeing the Comparative Method from the Tree Model: A framework for Historical Glottometry. In R. Kikusawa & L. Reid (eds), *Let's talk about trees: Tackling Problems in Representing Phylogenetic Relationships among Languages* (Senri Ethnological Studies). Osaka: National Museum of Ethnology.
- Leskien, August. 1876. *Die Declination im Slawisch-Litauischen und Germanischen*. Leipzig: Hirzel.
- Lynch, John. 1975. Oral/nasal alternation and the realis/irrealis distinction in Oceanic languages. *Oceanic Linguistics* 14 (2), 87–99.
- Lynch, John & Ozanne-Rivierre, Françoise. 2001. Some shared developments in pronouns in languages of Southern Oceania. *Oceanic Linguistics* 40-1, 33-66.
- Nichols, Johanna & David Peterson. 1996. The Amerind personal pronouns. *Language* 72: 336–371.
- Pawley, Andrew. 1973. Some problems in Proto-Oceanic Grammar. *Oceanic Linguistics* 12, 103-188.
- . 2003. Locating Proto Oceanic. In Malcolm Ross, Andrew Pawley & Meredith Osmond (eds.), *The lexicon of Proto-Oceanic: The physical environment*. Canberra: Australian National University, 17–34.
- Ross, Malcolm. 1988. *Proto-Oceanic and the Austronesian languages of Western Melanesia*. (Pacific Linguistics). Canberra: Australian National University.
- . 1998. Proto Oceanic phonology and morphology. In Malcolm Ross, Andrew Pawley & Meredith Osmond (eds.), *The lexicon of Proto Oceanic. Vol. 1: Material culture*. Canberra: Australian National University, 15–35.
- . 2002. Proto Oceanic. In John Lynch, Malcolm Ross & Terry Crowley (eds.), *The Oceanic languages*. London: Curzon Press, 54–91
- . 2005. Pronouns as a preliminary diagnostic for grouping Papuan languages. In Andrew Pawley, Robert Attenborough, Robin Hide & Jack Golson (eds.), *Papuan pasts: cultural, linguistic and biological histories of Papuan-speaking peoples*. Canberra: Pacific Linguistics, 15–66
- Ross, Malcolm; Andrew Pawley & Meredith Osmond (eds.) (1998–2011). *The lexicon of Proto Oceanic*. 4 volumes. Canberra: Australian National University.
- Saussure, Ferdinand de. 1916 [1995]. *Cours de linguistique générale*. Edited by Tullio De Mauro. Paris: Payot.