Verbal aspect and personal pronouns: The history of aorist markers in north Vanuatu

Alexandre François

Verbal aspect and personal pronouns

The history of aorist markers in north Vanuatu

Alexandre FRANÇOIS

LACITO-CNRS, Paris

Abstract

Among the seventeen languages spoken in the Banks and Torres groups of north Vanuatu, eleven share a TAM category whose functions include sequential, generic, subjunctive, prospective and imperfective. This aspect, labeled here "aorist", also displays cross-linguistic formal similarities: everywhere, the aorist marker shows allomorphic variation depending on the person of the subject. After comparing the eleven languages concerned, I propose to reconstruct their protosystem as a set of four portmanteau proclitics {*gu–*u–*ni–*(k)a} combining aspect and person.

1   Introduction[1]

The subject and object clitics which are reconstructed for Proto Malayo-Polynesian (Blust 1977) and for Proto Oceanic (Lynch, Ross & Crowley 2002:67) have been replaced in Mwotlap, as in many other languages of north Vanuatu, with a unique set of free pronouns: *nɔ* '1sg'; *nɪk* '2sg'; *kɪ* '3sg'… These may be used both in subject and object positions:[2]

---

[1] I am grateful to COOL7 participants for their questions on an oral version of this paper (François 2007), and to Claudia Wegener and Alexis Michaud for their comments on an earlier draft.

[2] All transcriptions use IPA rather than local orthographies, to facilitate comparison. Note that |v| is often bilabial [β]; |j| is the palatal approximant; |ṭ| is a laminal retroflex; all voiced stops are prenasalized (|d| is [ⁿd], |g͡bʷ| is [ᵑg͡bʷ], etc.). Besides abbreviations that follow the Leipzig glossing rules, AO means 'Aorist'; POT 'Potential'; PROSP 'Prospective'; STAT 'Stative'; (P)NCV '(Proto) North-Central Vanuatu'; TAM 'tense-aspect-modality'. The three-letter abbreviations for modern languages are spelled out on *Map 1*.

(1)   *nɔ* m-ɛtsas *kɪ,* *kɪ* m-ɛtsas *nɔ.*
        1sg PRF-see 3sg 3sg PRF-see 1sg
        'I saw her and she saw me.'

Most Mwotlap pronouns are morphologically invariant. However, in subject position, the 1sg pronoun shows allomorphic variation between two forms *nɔ* and *nɔk.* This uncommon alternation depends on the tense-aspect-mood (TAM) marking of the verb. Out of the twenty-five TAM categories in Mwotlap (François 2003), eight allow for free variation between the two forms, whereas in the rest of the system, they come in strict complementary distribution: ten markers require *nɔ* as their subject, while seven require *nɔk.* In fact, as we will see below, *nɔk* itself can be described as a portmanteau form indexing both person and aspect.

This formal variation of the 1sg pronoun depending on the predicate's TAM-marking is typologically original. It also constitutes a morphological puzzle, which I will take as the starting point for this paper. *Section 2* will begin with a synchronic approach, by describing the semantic motivation of the *nɔ–nɔk* contrast in Mwotlap; this will lead to the functional definition of an aspect category labeled "Aorist".[3] Based on this definition, *section 3* will investigate the geographical distribution and the formal characteristics of similar aorist markers across the seventeen languages of the Banks and Torres islands. Finally, *section 4* will take a historical perspective, and attempt to unravel the development of aorist markers in north Vanuatu languages.

2   The Aorist in Mwotlap

The first question I will tackle is the functional distribution of the two allomorphs *nɔ* and *nɔk* in Mwotlap.

2.1   A special pronoun for the Aorist

The word order of constituents in Mwotlap is as follows:

  Subject NP – (TAM clitic/prefix) – Predicate – (TAM postclitic) – (Object NP)

As far as the 1sg pronoun is concerned, its unmarked, default form is clearly the shorter allomorph *nɔ.* It is the only one found in non-subject positions – see (1) – as well as for the subject of non-TAM predicates (e.g., *nɔ na-vatγɔ* 'I'm a teacher', *nɔ itʊk* 'I'm fine'). As for tense-marked predicates, *nɔ* combines with realis (Stative, Perfect, Completive…) as well as irrealis markers (Future, Potential, Counterfactual…):

(2)   *nɔ* mɪ-wɪl nu-suk.
        1sg PRF-buy ART-sugar
        'I've bought some sugar.'

---

[3] About the term "aorist", see the end of 2.2.2. Note that I use capitalization, following Comrie (1976:10), whenever a given term is to be understood as a labeling convention for a morphosyntactic category specific of a given language, rather than a typological concept.

(2')   *nɔ*   tɪ-wɪl   vɪh   nu-suk.
    1sg   POT$_1$-buy   POT$_2$ ART-sugar
    'I can buy some sugar.'

The seven TAM categories requiring the marked form *nɔk* are the Aorist proper, the Permansive, the Prioritive, the two Presentatives (static and kinetic), the Polite Imperative and the Prospective. Despite their semantic differences, the latter six categories are related, as they are all formally derived from the Aorist, through combination with some secondary morpheme. In other words, the marked form *nɔk* is required whenever the 1sg pronoun is the subject of a tense-marked predicate belonging to the domain of the Aorist, in the wide sense of the term. Therefore I will hereafter gloss it '1sg:AO'.

Crucially, when the predicate is an Aorist strictly speaking – as opposed to one of its derivatives – the pronoun *nɔk* is in fact the only formal TAM marking in the clause. This confirms its interpretation as a portmanteau morpheme, combining person and TAM marking:

(3)   *nɔk*   ŋ͡mʷʊl
    1sg:**AO**   return
    'Let me go back!' (…)

While the first person encodes the Aorist through variation of the pronoun, the 3sg does this with a prefix *ni-* on the verb, in the slot usually devoted to other aspect markers – compare (3') with (2).

(3')   kɪ   *ni-*ŋ͡mʷʊl
    3sg   **AO**:3sg-return
    'Let him go back!' (…)

Finally, all persons other than 1sg and 3sg encode the Aorist with a zero:

(3")   ɣɪn   (Ø-)ŋ͡mʷʊl
    1inc:pl   (**AO**-)return
    'Let's go back!' (…)

The Aorist and its derivatives are the only TAM categories of Mwotlap whose marking depends on the person.

2.2 The semantics of the Aorist

2.2.1 The various uses of the Aorist

Like several other TAM categories of Mwotlap, the Aorist is only compatible with semantically dynamic events. Its combination with a stative predicate – whether a stative verb, an adjective or a noun – forces a dynamic interpretation [see (9) and (13) below].

But the semantic information which the Aorist gives about that dynamic event is widely polysemous (François 2003:165-199). For one thing, the Aorist encodes events that come in sequence. This applies equally in past or future contexts:

(4) *nɔk* hajvɛɣ l-ɪŋ͡mʷ nɔnɔn ɛ kɪ *ni-ɛtsas* nɔ.
1sg:AO enter in-house his TOPIC 3sg AO:3sg-see 1sg
[*past context*] 'I came into his house and (then) he saw me.'
[*future context*] 'I'll come into his house and (then) he'll see me.'

Crucially, the same sentence (4) may be translated in English either as past or as future. This shows that the Aorist is not a tense, but an aspect, which may attach either to a realis or to an irrealis situation. In itself, (4) says nothing more than '(Let there be) my coming into his house, *and then* him seeing me…'. What is relevant here is the relation of sequence or implication between the two successive events, regardless of how they happen to relate to the speech coordinates. The Aorist is commonly found in narratives, for any chain of events:

(5) kɪ *ni-jɛm* haɣ lɛ-vɛt tʊ kɪ *ni-k͡pʷɪsdi* hʊw tʊ *ni-*mat.
3sg AO:3sg-climb up on-stone then 3sg AO:3sg-fall down then AO:3sg-die
'He climbed up the rock, then he fell down and died.'

This use as a sequential marker in a string of events is ubiquitous in Mwotlap. Yet this form is also required in many other contexts which cannot be reduced to this explanation.

The Aorist is used for generic sentences, such as definitions or procedure descriptions – that is, utterances referring to a timeless event that bears no connection with any specific situation:

(6) "nɛ-ŋ͡mʷjajaj" ɛ, nɪk ɛtɛt hɛjlʊ van aɪ.
STAT-transparent TOPIC 2sg AO:look~IPFV through thither in.it
'*Transparent* (means that) you see through it.'

(7) na-mtɛ *ni-*jɔj, na-tak͡pʷŋ͡mʷɪ *ni-*jɛjɛj,
ART-your.eyes AO:3sg-sink ART-your.body AO:3sg-shiver

nɪk mat ŋ͡mʷʊl…
2sg AO:die return
'[with malaria] your eyes sink, your body shivers, you faint…'

Another example where Aorists point to virtual events whose time coordinates are left indefinite, is their use in conditional clauses (note that (8) is identical to (4) above).

(8) nɔk hajvɛɣ l-ɪŋ͡mʷ nɔnɔn ɛ kɪ *ni-ɛtsas* nɔ.
1sg:AO enter in-house his TOPIC 3sg AO:3sg-see 1sg
'(Suppose) I came into his house (then) he would see me!'

In many cases however, the Aorist does relate to a specific situation, which may be the moment of utterance. This happens, for instance, when it represents an event as imminent:

(9) mahɪ *ni-k͡pʷʊŋ* ɪɣɪn.
place AO:3sg-night now
'Night is about to fall.'

The imminence of the event is sometimes factual, as in (9), but quite often it is the speaker's own projection. The Aorist thus takes on modal values, and encodes intent,

optative, instructions or commands:[4]

(10) nɪk ɣɛn mɛj nʊk, nɔk ɣɛn mɛj ɣɪn.
2sg AO:eat the.one there 1sg:AO eat the.one there
'You eat this one, I'll eat that one.'

(11) kɪ *ni-ŋ͡mʷʊl* lɛ-pnʊ nɔnɔn.
3sg AO:3sg-return in-village his
'Let him return to his village!' [*or*: 'he returned…', see (5)]

Strictly speaking, the Aorist cannot be said to inherently entail such illocutionary forces as desiderative or imperative, because it is also used in plain declarative sentences. In other words, just as it does not by itself convey any indication of time, it is also underspecified with regard to modality: it is found in statements as much as in hypotheses, commands or optatives. Both the time coordinates and the modal value of the Aorist thus need to be inferred from prosodic clues, and from the discourse context.

This semantic underspecification with regard to time and mood explains why the Aorist (or its derivative the Prospective) is required in modality-bound subordinate clauses: e.g., clause complements of verbs of will or manipulation, as well as purposive and consecutive clauses.

(12) nɔ nɛ-mjʊs sɔ *nɔk* (sɔ) in ni-ti.
1sg STAT-want COMP 1sg:AO (PROSP) drink ART-tea
'I want to drink some tea.' [lit. 'I want that I drink_AO…']

(13) nɔ mʊ-mʊk nɛ-vɛt l-ɛp tʊ kɪ *ni-vɛj*.
1sg PRF-put ART-stone in-fire then 3sg AO:3sg-red.hot
[*purposive*] 'I laid the stones on the fire so that they become_AO red-hot.'
[*consecutive*] 'I laid the stones on the fire so they became_AO red-hot.'

The semantic incompleteness of the Aorist thus makes it particularly compatible with certain forms of syntactic dependency, in a way reminiscent of the subjunctive of Indo-European languages.

If a dynamic verb is reduplicated, it acquires imperfective aspectual properties, including when combined with the Aorist. This means (Comrie 1976) it may take either a *habitual* reading or a *progressive* one:

(14) *nɔk* jap hij tita mino.
1sg:AO write to mother my
                simple verb: perfective interpretation

[*sequential*]   '(then) I wrote to my mother.'
[*intent/optative*]   'Let me write to my mother!' …

(14') *nɔk* japjap hij tita mino.
1sg:AO write~IPFV to mother my
                reduplicated verb: imperfective interpretation

---

[4] This modal function accounts for the formal links between the Aorist, and the three modal markers derived from it (Prospective; Prioritive; Imperative).

[*habitual*]   'I write to my mother (every day…)'
[*progressive*]   'I'm writing to my mother.'

The absence of reduplication in (14) gives the verb a perfective reading, which makes it compatible with the various aspectual values reviewed so far for the Aorist: sequential, intent, etc. Conversely, reduplication in Mwotlap (François 2004) has the power to convert a perfective into an imperfective, which disrupts the impact of the Aorist marking altogether. Arguably, the latter then functions as a neutral aspect marker, whose role is simply to state the imperfective process (habitual or progressive) in relation to the context.

2.2.2   *Defining the underlying mechanism*

Despite the impressive polyfunctionality of this TAM category, it is possible to identify a constant aspectual pattern behind the variety of its contextual meanings. In all cases, the Aorist consists in representing a new event considered in itself, that is, regardless of its deictic coordinates in terms of tense or modality.

Precisely because it lacks any inherent deictic reference, this indeterminate event needs to be connected to an external point of reference – its "anchor" – in order to receive proper pragmatic interpretation. Quite often, the anchoring situation is easy to retrieve from the context. For example, in a string of successive events, it corresponds to the end of the previous event (4, 5). In a subordinate pattern, the dependent event will hook onto the coordinates of the main clause (12, 13). In many cases, the default reference point will be the utterance situation, whether the new event that is supposed to cling to it is presented as a statement of fact (9, 14') or as the speaker's projection (10, 11, 14). Finally, it sometimes happens that this "orphan" event in search of situational anchoring finds none, and remains suspended in time: this is what happens when the Aorist points to a timeless event with no connection to any specific situation, as in generic statements (6, 7) or hypotheses (8).

Typologically speaking, the term "aorist" has been used with various senses, and sometimes inconsistently, across language descriptions. However, the aspectological tradition that has developed, especially in France, after Benveniste (1966) and Culioli (1978), has now solidly established the notions of "*aoriste*" or "*aoristique*", as a verbal aspect whereby the depicted event is disconnected from the situation of utterance. Similar examples of "aorist" have been described for several languages, such as Coptic (Depuydt 1993), Wolof (Robert 1996) or Berber (Galand 2003). A full typological survey of the aorist aspect still needs to be carried out.

3   The morphology of the Aorist in northern Vanuatu languages

Now that the semantics of the Aorist have been observed for Mwotlap on a synchronic, language-internal basis, it becomes possible to observe whether its neighbors of north Vanuatu possess a similar aspect category, and if so, how they encode it morphologically. This observation might help trace the formal history of Mwotlap's

Aorist, and especially of the unusual alternation between the two 1sg pronouns *nɔ* and *nɔk*.

Since 2003, my field investigations have precisely involved the firsthand study of all the languages of the Banks and Torres groups, of which basically nothing was known to date. *Map 1* locates these seventeen languages; it indicates their current number of speakers, together with the three-letter abbreviations I propose to use for them. The remainder of this section will summarize the results of this survey with regard to the Aorist aspect.

Map 1 – *The languages of north Vanuatu*
<add map here>

3.1   South Banks

The TAM systems observed in the five languages of Gaua, together with Mwerlap, differ significantly from that of Mwotlap. In particular, the semantic spectrum of Mwotlap's Aorist, instead of being encompassed by a single marker, is divided in these languages into two, three or even four distinct categories, each language showing its own particular distribution (*Table 1*).

Table 1 – *Equivalents to Mwotlap's Aorist in the six south Banks languages*

	MTP ex	MRL	NUM	DRG	KRO	OLR	LKN
Sequential	(4, 5)	ti	tɔv	sɔ		Ø	tɪ
Generic (definitions, procedures)	(6, 7)		vɛ		v-		
Subjunctive (dependent, conditional)	(8, 12, 13)	sV-	tɛ	s-		(ɣ)a	(ɣ)a
Prospective (optative, commands…)	(9, 10, 11, 14)				s-		
Imperfective (habitual, progressive)	(14')		tɛ… ti	t-… ti	t-… ti	tɪ-… ti	tɪ-… tɔ

It would be a matter for complex discussion to decide which of these morphemes should be properly labeled 'Aorist', and which ones should receive a name of their own. For example, in Dorig, it is safe to call *sɔ* a Sequential, and *t-… ti* an Imperfective. As for *s-*, the union of 'generic', 'subjunctive' and 'prospective' could be tagged Aorist, in the sense of "deictically indeterminate new event", as defined above for Mwotlap. Yet it could as well, and perhaps more accurately, be called *Irrealis* or *Virtual* – a choice impossible in Mwotlap due to both the sequential and the imperfective uses.

In sum, none of these languages possess a proper aorist, in the sense defined for Mwotlap. Furthermore, all markers in *Table 1* are invariable prefixes or proclitics, used for all persons. Their forms resemble neither MTP *nɔk* nor *ni-*, the origins of which will have to be sought elsewhere.

3.2 Central Banks

The ten remaining languages of the Banks and Torres are more promising. Indeed, each of these languages possesses a TAM category which essentially matches the Aorist of Mwotlap, encompassing all the functions of *Table 1*, from 'sequential' to 'imperfective';[5] I shall therefore use the label 'Aorist' everywhere. And, crucially, in each language, its formal marking depends on the person of the subject, in a way reminiscent of Mwotlap.

Let us first observe the three languages located in the central part of the Banks Islands: Mota, Mwesen and Vurës. Taking the verb 'see' (MTA *ilo*, MSN-VRS *il*) as an example, *Table 2* illustrates the behavior of subject markers for the Aorist, in comparison with an ordinary TAM marker – in this case, the Perfect.[6] The pattern for 1inc:pl, which is given here, exemplifies the twelve non-singular forms.

Table 2 – *Aorist inflections in three central Banks languages*

language	*TAM*	*1sg*	*2sg*	*3sg*	*1inc:pl*
Mota	*Perfect*	nau *me* ilo	ko *me* ilo	ni *me* ilo	nina *me* ilo
	Aorist	**na** ilo	**ka** ilo	**ni** ilo	nina **a** ilo
Mwesen	*Perfect*	na *mɛ* ɪl	nɪk *mɛ* ɪl	nɪ *mɛ* ɪl	nin *mɛ* ɪl
	Aorist	na **na** ɪl	nɪk **a** ɪl	nɪ **ni** ɪl	nin **a** ɪl
Vurës	*Perfect*	nɔ *mɪ*-ɪl	nɪk *mɪ*-ɪl	nɪ *mɪ*-ɪl	nɪn *mɪ*-ɪl
	Aorist	**na** ɪl	nɪk **i** ɪl	nɪ **ni** ɪl	nɪn **a** ɪl

Taking only the non-singular forms, we would simply have an invariant clitic *a* 'Aorist' behaving like other TAM markers. But the singular makes the description more complex, because the marking of the Aorist differs according to the person of the subject. To be precise, two distinct cases are attested:

a) The pronoun itself remains unchanged, but the Aorist clitic presents allomorphic variation according to the person of the subject. Thus for Mwesen, the Aorist is *a* for all persons, but *na* for 'AO:1sg' and *ni* for 'AO:3sg'.

b) The sequence {pronoun + TAM marker} found with other tenses is replaced by a single portmanteau clitic that incorporates person- and TAM-marking. Thus in Mota, *na* should be properly glossed '1sg:AO', and *ka* '2sg:AO'.[7]

---

[5] In order to demonstrate this, a full set of examples should ideally be provided for each language. Unfortunately, this is impossible here due to considerations of space.

[6] For each language, the first row translates as 'I have seen'…; the second row as 'Let me see', etc.

[7] For 3sg in Mota, the Aorist can be identified either as a Ø clitic (commuting with *a*) or as incorporated in *ni* (commuting with *ka*).

Vurës combines the two patterns: (*a*) for 2sg and 3sg, but (*b*) for 1sg. In fact the same complexity was found in Mwotlap, where *nɔk* was to be analyzed as an aspect-indexed pronoun ('1sg:AO'), but *ni-* as a person-indexed aspect prefix ('AO:3sg').

Now, MTP *ni-* is clearly the same morpheme as *ni* in these three languages. Furthermore, a connection can be drawn between that *ni* 'AO:3sg' and the form of the free pronoun for 3sg *ni* in Mota, Nume, Dorig and Koro. In several languages of north Vanuatu, the 3sg pronoun (*nɪ, niə*…) reflects an earlier form *nia '3sg', itself connected with *ni*. This formal connection has been blurred in Mwotlap, where the 3sg pronoun is now an innovative *kɪ*.

These first findings thus shed light on our initial puzzle. Yet still nothing can be said about the strange form *nɔk* in Mwotlap: where does this /k/ come from? The answer will appear as we continue our survey further north.

3.3 North Banks

Not surprisingly, a system much closer to Mwotlap can be found in Volow, an extinct dialect formerly spoken on the same island, and passively remembered by a handful of people. The structures of the two dialects are so parallel that the only differences lie in the phonological forms of the markers: to the alternation between MTP *nɔ* and *nɔk* corresponds a pair of forms *nɛ* '1sg' vs *nɛŋ* '1sg:AO' (see *Table 3* below).[8] But precisely because Volow is so close to Mwotlap, it is of little help in our investigation.

More instructive findings come from the four languages of the northwest Banks area: Lehali, Löyöp, Lemerig and Vera'a. Unlike the three languages of *Table 2*, they do possess a trace of the velar /k/ which is found in the 1sg pronoun of Mwotlap. But, interestingly, instead of being part of the pronoun itself, the consonant /k/ is separable from it, and prefixed to the verb. This becomes obvious when the pronoun and the verb are separated by another morpheme. Compare the Prospective of Mwotlap with its form in Lehali and Vera'a:

(15) MTP *nɔ**k**   sɔ   mitij.
            1sg:AO   PROSP   sleep

      LHI   nɒ   dɛ   **k-**   mutuj.
            1sg   PROSP   AO:1sg-   sleep

      VRA   nɔ   sɔ   **k-**   miʔir.
            1sg   PROSP   AO:1sg-   sleep
            'I'd like to sleep.'

This syntactic test makes it easy to define the boundary between the personal pronoun proper and the (person-conditioned) TAM-marker. Unlike Mwotlap, these northwest Banks languages thus treat the 1sg Aorist marker in the same way as their 3sg, as a prefix to the verb:

---

[8] Note that the correspondence between MTP /k/ and VLW /ŋ/ syllable-finally is regular, and reflects a former prenasalised voiced stop [ᵑg] (noted *g*).

(16) MTP   *kɪ*   sɔ   ***ni*-**   mtij.
           3sg   PROSP   AO:3sg-   sleep

     LHI   *kɛ*   dɛ   ***n*-**   mutuj.
           3sg   PROSP   AO:3sg-   sleep

    VRA   *di*   sɔ   ***nɛ*-**   miʔir.
           3sg   PROSP   AO:3sg   sleep
           'He'd like to sleep.'

Lemerig does not allow any element between the pronoun and the (inflected) verb, which makes it impossible to conduct the test illustrated by (15). For example, 'Let me sleep' will take the ambiguous surface form /nœkmiʔir/, which could be parsed *nœk miʔir* <1sg:AO|sleep> as in Mwotlap, or *nœ k-miʔir* <1sg|AO:1sg-sleep> as in Vera'a. My Lemerig corpus shows 38 instances of a 1sg Aorist, out of which 37 show this ambiguity. Luckily, one sentence has two Aorists chained together, a context where the personal pronoun may be dropped. This single example gives the solution to the puzzle, and highlights the structural difference between Lemerig and Mwotlap:

(17) LMG   *nœ*   ***k*-œn**   sur   ɛ   (*nœ*)   ***k*-miʔir.**
           1sg   AO:1sg-lie   down   LNK   (1sg)   AO:1sg-sleep

    MTP   *nɔk*   ɛn   hij   ɛ   (*nɔk*)   mitij.
           1sg:AO   lie   down   LNK   (1sg:AO)   (AO:)sleep
           'Let me lie down and sleep.'

Finally, the case of Löyöp is slightly tricky, because it is a hybrid of both patterns. On the one hand, just like its neighbours but unlike Mwotlap, Löyöp has kept a 1sg Aorist prefix of the form *k-*. This form appears when the clause lacks the free pronoun *nø*:

(18) LYP   nø   nɣɛ   møjøs   sɛ   ***k*-sʊwjɪ**   n-k͡pʷɛŋ.
           1sg   STAT   want   SUB   AO:1sg-cast   ART-net
           'I want to go net-casting.'   [lit. I want that I$_{AO}$ cast the net]

However, Löyöp has also taken the same path as Mwotlap, in that the combination of the free pronoun *nø* and of the *k-* prefix has been resegmented, giving rise to a new, unanalysable pronoun *nʊk*, with an unpredictable vowel:

(19) LYP   *nø*   tɛ   pɪl   tʃɛ,   ***nʊk***   dɔn   sɛ   n-jø-k.
           1sg   NEG$_1$ steal   NEG$_2$ 1sg:AO   think   SUB   ART-CLPOS-1sg
           'I didn't steal it, I thought it was mine.'

Taking the verb meaning 'see' again as an example,[9] *Table 3* shows the Aorist morphology for the six languages under comparison here – including Mwotlap and Volow – thereby covering the whole 'north Banks' area. The hybrid case of Löyöp appears in the middle.

---

[9] Once again, '1inc:pl' stands for all non-singular forms.

Table 3 – *Aorist inflections in six north Banks languages*

language	1sg	2sg	3sg	1inc:pl
Lemerig	nœ **k**-ʔɛt	næk (ø-) ʔɛt	ti **n**-ʔɛt	ɣæt (ø-) ʔɛt
Vera'a	nɔ **k**-ʔɪn	nikɪ (ø-) ʔɪn	di **nɛ**-ʔɪn	ɣidɪ **k**-ʔɪn
Lehali	nɒ **k**-ɛt	nɛk (ø-) ɛt	kɛ **n**-ɛt	ɣɛn (ø-) ɛt
Löyöp	**k**-ɛt / nʊ**k** ɛt	niŋ (ø-) ɛt	kjɛ **n**-ɛt	jɛn (ø-) ɛt
Mwotlap	nɔ**k** (ø-) ɛt	nɪk (ø-) ɛt	kɪ **ni**-ɛt	ɣɪn (ø-) ɛt
Volow	nɛ**ŋ** (ø-) ɛt	nɪŋ (ø-) ɛt	gɪ **n**-ɛt	ɣɪn (ø-) ɛt

The problem raised by *k-* in Vera'a non-singular forms will be addressed in 4.2.

3.4 The Torres Islands

3.4.1 Two sets of personal markers

I will end this survey of Aorist markers in north Vanuatu with the two languages of the Torres group. In comparison with the Banks languages, the Aorist inflection in Hiw and Lo-Toga is morphologically richer. Not only are there specific (non-zero) morphemes for each person in the singular, but also for non-singular subjects, including different forms for the dual and for the plural (the Torres languages have lost the trial).

Table 4 lists the complete sets of personal subject markers for the two languages. On the left are given the free pronouns; on the right, the set of person-indexed Aorist clitics.

Table 4 – *Full pronouns vs Aorist clitics in the two Torres languages*

	Hiw		Lo-Toga	
	full pronoun	*Aor. clitic*	*full pronoun*	*Aor. clitic*
1sg	nɔkə	kə	nɔkə(LO)/ nɛkə(TGA)	kə
2sg	ikə	wɵt ~ wɵk	nikə	wə
3sg	ninə	nə	niə	ni
1in:du	tɵᵍʟɵ		ṭor	
1ex:du	kamaᵍʟə	ᵍʟə	kəmɔr	or
2du	kimiᵍʟə		kəmor	
3du	sɵᵍʟɵ		hor	
1in:pl	titə	tə	ṭəɣə(LO)/ ɣiṭə(TGA)	
1ex:pl	kama	–	kəmɛ(m)	ɣə
2pl	kimi	–	kəmi	
3pl	sisə	sə	nihə	

The first obvious observation is that these two languages possess interesting clues for our study: the form of the 3sg clitic (*nə*/*ni*) recalls the prefixes *n-* or *ni-* we saw in

the Banks languages; and the 1sg clitic *kə* is reminiscent of the prefix *k-* shown in *Table 3* above. Finally, the 1sg pronoun *nɔkə* strikingly resembles Mwotlap *nɔk*, a point which warrants a discussion of its own (see 4.2.2).

3.4.2   TAM markers or light pronouns?

Just as in Banks languages, the function of the clitics of *Table 4* is essentially to encode a TAM category, the Aorist. This status is proven by the comparison of (20) and (20'). Both sentences show serial verb constructions, one with the Potential, the second with the Aorist.

(20) LTG   kəmi ***si***   ŋʷulə   ***si***   mətur.
          2pl   POT   return   POT   sleep
          [*Potential*]   'You (*pl.*) may go back and sleep.'

(20') LTG   kəmi ***ɣə***   ŋʷulə   ***ɣə***   mətur.
          2pl   AO:PL   return   AO:PL   sleep
          [*Aorist*]   'You (*pl.*) go back and sleep!'

But there is further complexity. Amongst the clitics of *Table 4*, only two (*ni* and *ɣə* in Lo-Toga, none in Hiw) may be immediately preceded by a free pronoun, as in (20') *kəmi ɣə*. All other clitics must be deleted in presence of the free pronoun, in which case the latter is directly followed by the verb. As a result, most Aorist sentences, when they include the free pronoun, appear to be unmarked (or zero-marked) for TAM. Conversely, the clitics are restricted to those clauses that lack a free pronoun. This happens typically in a string of clauses, when the pronoun is mentioned only with the first verb [see (17)]:

(21) LTG   *nikə*   (Ø)   ŋʷulə   ***wə***   mətur.
          2sg   (AO)   return   AO:2sg   sleep

     HIW   *ikə*   (Ø)   ŋʷujə   ***wɵt***   mitiᵍL.
          2sg   (AO)   return   AO:2sg   sleep
          'You (*sg*) go back and sleep!'

A superficial analysis of (21) would probably have posited only one marker for the Aorist (*zero*), and then two sets of pronouns: 'heavy' pronouns for the main or first clause in a string, followed by 'light' pronouns in secondary and other dependent clauses. In that framework, it would have made sense to label these clitics "secondary subject pronouns". This interpretation is appealing, and could perhaps be proposed for Hiw; but in Lo-Toga, it seems to be contradicted by (20'). For the sake of consistency, it is thus safer to analyze *wə* in (21) not as a personal pronoun, but as a (person-indexed) aspect clitic. For most subjects,[10] deletion rules must be formulated, whereby the sequence {pronoun+clitic} simplifies to {pronoun}, e.g. *nikə wə V ⇒ *nikə* V.

---

[10] In Lo-Toga, this deletion rule applies to 1sg, 2sg, and dual forms. To this list, one must add *ṭəɣə* '1inc:pl' in the Lo dialect of Lo-Toga: compare Toga *ɣiṭə ɣə ŋʷulə* with Lo *ṭəɣə* (Ø) *ŋʷulə* 'Let's go back'.

Because Hiw operates this deletion rule for all its pronouns, its Aorist clitics seldom show up in fluent speech, as they are restricted to subordinate or secondary clauses; and even in that case they are optional, being often replaced by the full pronouns. The situation is very different in Lo-Toga, where the clitics are extremely productive, and massively represented in my corpus. This productivity of Lo-Toga clitics has two reasons: first, the two clitics *ni* and *ɣə* cannot be deleted, and are pervasive in speech; second, each clitic also appears as a constituent element in three compound TAM markers historically derived from the Aorist: Prospective ⟨*tɛ* + Cl.⟩, Time Focus ⟨Cl. + *akə*⟩, and Future ⟨*tɛ* + Cl. + *akə*⟩ – e.g., *nikə **tɛ w' akə** mətur* 'you will sleep'.

In sum, in Hiw and Lo-Toga, one identifies a clause as Aorist either because it displays an Aorist clitic, or because it consists of {free pronoun + zero-marked verb}. For example, the Lo-Toga sentence (21) shows two Aorist predicates: **nikə ŋʷulə** 'you return[AO]' and **wə mətur** 'you sleep[AO]'.

4 The historical perspective

The eleven languages endowed with a genuine aorist (3.2 to 3.4) show such solid formal similarities that they obviously share a common history. In this section I will endeavor to reconstruct a set of aorist markers for their common (post-POc) ancestor language. Logically, this protolanguage should be Proto North-Central Vanuatu, the proposed ancestor for the majority of Vanuatu languages (Clark 1985). But since my reconstruction is reflected only in the northernmost languages of this family – those spoken in the Banks and Torres groups – it could well represent a subgroup within NCV, the precise limits of which would need to be confirmed by further diagnostic evidence.

4.1 Reconstructing the set of Aorist proclitics

First, this pre-modern system possessed a set of free pronouns. They can be reconstructed for these northern Vanuatu languages (Clark 1985; Lynch & Ozanne-Rivierre 2001:38): **nau* '1sg'; **nigo* '2sg'; **n(a)ia* '3sg'; **kida* '1inc:pl'; **ga(ma)mi* '1exc:pl'; **gamuyu* '2pl'; **n(a)ira* '3pl'.

The system can also be reconstructed with a set of person-indexed aorist markers, consisting of proclitics preceding the verb. Taking into account the phonological history of this area – especially the phenomenon of vowel reduction (François 2005) – the most plausible reconstructions are as follows:

- *1sg*: all languages point to the 'nasal grade' *g [ᵑg] (see fn.8 p.2). As for the vowel that followed this consonant, comparative evidence (see below) suggests it must have been /u/, hence a protoform ***gu**.

- *2sg*: a single vowel, probably ***u** (reflected as /w-/ in the Torres; /i/ in Vurës; *zero* in most languages).

- *3sg*: *nV, probably ***ni** (reflected as /ni/, /nə/, /nɛ/ or /n-/).

▪ *non-singular*: the modern forms /ɣə/, /a/ and Ø suggest a reconstruction *(k)a.[11] The dual forms of the Torres may result from a local innovation, perhaps **(k)a-ru (?).

The semantic array reconstructible for this set of aorist clitics most probably coincided with the observations made for modern Mwotlap (2.2) and its neighboring languages. Their function was to construe a "deictically indeterminate new event" – a definition which encompasses the functions of *sequential*, *generic*, *subjunctive*, *prospective* and (with reduplication) *imperfective*.

Syntactically, these proclitics {*gu, *u, *ni, *(k)a} occupied the same slot as other TAM markers. They were preceded by the free pronoun in main clauses, or in the first clause of a chain (serialized verbs, narratives); yet they appeared on their own in dependent or secondary clauses (e.g., same-subject sequential clauses). It is thus possible to reconstruct sentences such as (22):

(22) *nau   *gu*=   mule   *gu*=   maturu
      1sg   AO:1sg=   return   AO:1sg=   sleep
      'So I went back and slept.' ~ 'Let me go back and sleep!'…

      *nigo   *u*=   mule   *u*=   maturu
      2sg   AO:2sg=   return   AO:2sg=   sleep
      'So you went back and slept.' ~ 'You go back and sleep!'…

      *nia   *ni*=   mule   *ni*=   maturu
      3sg   AO:3sg=   return   AO:3sg=   sleep
      'So he went back and slept.' ~ 'Let him go back and sleep!'…

      *kida   *(k)a*=   mule   *(k)a*=   maturu
      1inc:pl   AO:non.sg= return   AO:non.sg= sleep
      'So we went back and slept.' ~ 'Let's go back and sleep!'…

As far as their origin is concerned, the singular forms {*gu, *u, *ni} are reminiscent of two sets of personal markers:

▪ the 1st, 2nd and 3rd person singular forms of the possessive suffixes, whose protoforms in north Vanuatu are {*-gu, *-u[12]/-mu, *-na}, from POc {*-gu, *-mu, *-ña};

▪ one of the sets reconstructed for the POc subject proclitics, namely {*ku, *mu, *(y)a/ña}, which ultimately reflect PMP genitive pronouns (Blust 1977; 2003). As for the non-singular prefix *(k)a, it is reminiscent of a subject clitic *ka[i] '1exc:pl' reconstructed for some Oceanic interstage languages (Lynch *et al.* 2002:68).

Given the nature of the Aorist clitics, the subject clitics are a more likely source than the possessive suffixes. There is still some debate about the precise function of these POc proclitics. According to Kikusawa (2005), they retained in POc their earlier function as ergative subjects (i.e., 'A' in divalent clauses) as opposed to intransitive

---

[11] The Sungwadia language of north Maewo also has a series of aorist prefixes (Agnès Henri, pers. comm.). The plural form is *ɣe*, which tends to confirm the consonant in our reconstruction *(k)a.

[12] For 2sg *-u, see François (2005:486). This local variant of the suffix may have played some influence upon the 2sg proclitic *u.

subjects ('S' in monovalent clauses). Lynch *et al.* (2002:68) suggest that this function was probably "being lost when POc broke up".

But one still has to explain how a set of ordinary subject pronouns should have evolved into TAM-marking clitics. A tentative hypothesis would suggest these subject markers once became specialized in subordinate or other dependent clauses while main declarative clauses eventually generalized the use of free pronouns. As a result, what were once genuine subject pronouns in clauses otherwise unmarked for TAM, eventually grammaticalized into subjunctive-like TAM markers. This hypothesis would account for the affinities of Aorist predicates with syntactic dependency, discourse backgroundedness, and TAM indeterminacy. But this is mainly speculation at this stage; the functional connection at stake here definitely warrants further investigation.

4.2   From the protosystem to modern languages

After tentatively reconstructing the protosystem of Aorist clitics in the protolanguage ancestral to the Torres and Banks groups (either PNCV or one of its branches), I will end this study with an overview of the various paths of evolution that historically led to the modern systems.

4.2.1   *Phonological attrition and affixation*

The phonological process of *unstressed vowel deletion*, which massively affected the languages of the whole Banks and Torres area (François 2005), explains why former *CV clitics are generally reflected as a single consonant in modern languages (*gu > /k/ or /ŋ/; *ni > /n/…), and also why the two vowel-only clitics are so often reflected as *zero*. Lemerig shows the expected reflexes in this regard:

	protosystem	>	*Lemerig*
'let me go back'	*ˈnau *gu* = ˈmule	>	nœ **k**-mʊl
'you go back'	*ˈnigo *u* = ˈmule	>	næk Ø-mʊl
'let him go back'	*(ˈnia) *ni* = ˈmule	>	(ti) **n**-mʊl
'let's go back'	*ˈkida *a* = ˈmule	>	ɣæt Ø-mʊl

This vowel reduction process explains why the former clitics generally became prefixes. It also accounts for the formal convergence between 2sg *u* and the non-singular clitic *a* – variant of *(k)a* – in the form of *zero*. As for Mwesen and Mota (*Table 2*), this convergence rather results from the spread of the non-singular clitic *a* to 2sg (MTA *ka* <*ko a*); see also below for the case of 1sg.

Among the four clitics {*gu*, *u*, *ni*, *(k)a*}, only 3sg *ni* is preserved in the eleven languages endowed with a genuine "aorist", from Hiw all the way down to Mota. On the other hand, *(k)a* seems to show the widest historical extension, as its reflexes are scattered in various places from Lo-Toga *ɣə* down to Lakon/Olrat *(ɣ)a* (*Table 1*).

4.2.2   The intricate destiny of 1sg *gu

As for 1sg **gu*, it has survived in the eight (or nine: see below for Vurës) northernmost languages of the area, but has left little trace in any other language further south. To the best of my knowledge, the only other NCV language with a reflex of the subject marker **gu* is Tamabo, with *ku* '1sg' (Jauncey 2002:610).

The detailed evolution of this form **gu* in north Vanuatu languages is intricate. The former proclitic **gu=* is regularly reflected as a proclitic *kə=* in the Torres languages, or a prefix *k-* in northwest Banks languages (*Table 3*). The 1sg pronoun alternation (*nɔ/nɔk*) of Mwotlap – the starting point of this study – results from a process of reanalysis: the sequence /*nɔ* + *k-*/, with no intervening element, was so overwhelmingly frequent in speech, that it was eventually resegmented as *nɔk*; the former verbal clitic became accreted to the preceding pronoun, while the verb itself appeared in its bare form. The zero form taken by the Aorist with other persons (*Table 3*) probably added to the pressure towards morphological leveling. Volow followed a similar path, leading to an allomorphic variation between *nɛ* and *nɛŋ*; as well as Löyöp, with the forms *nɔ* and *nʊk*.

So far, the accreted form *nɔk* in Mwotlap has remained restricted to its original function – namely, 1sg subject of an Aorist predicate – without much affecting the regular form of the 1sg pronoun *nɔ* in other contexts. However, a slight tendency towards the expansion of *nɔk* is perceptible with other TAM categories. Besides the six markers that are formally derived from the Aorist (2.1), the use of *nɔk* is increasingly frequent, albeit optional, for as many as ten TAM markers, historically unrelated to the Aorist – for example, the negative markers.

This tendency for the accreted form to gain ground over the original 1sg pronoun is only incipient in Mwotlap, but has reached its final stage in the two Torres languages. Due to their high frequency in discourse, the augmented forms – *nɔkə* in Hiw and Lo, *nɛkə* in Toga – have now become the ordinary 1sg pronoun regardless of the predicate's TAM-marking, and indeed for all functions (subject, object, etc.).[13] In other words, the pronoun resulting from the coalescence of **nau gu* has replaced **nau* in all positions.

Unlike Mwotlap, the coalescence of **nau* and **gu* in the Torres languages did not eliminate **gu* as an independent morpheme (*kə*) in the system. But the fact it was historically incorporated into the free pronoun explains why the two forms are incompatible (**nɔkə kə* is ungrammatical). Crucially, this is the key to many of the so-called "deletion rules" which are required in the synchronic description of the Torres languages (3.4.2). A similar process of accretion is the origin of the form *ṭəɣə* used in the Lo dialect, the only plural form in Lo-Toga to be incompatible with *ɣə* (see fn.10 p.2): *ṭəɣə < *(ɣi)ṭə ɣə < *kida ka*. Likewise, in Hiw, *ninə* has incorporated the clitic *nə* (*ninə < *ni(ə) nə < *nia ni*); *titə* has incorporated *tə*; *sisə* has incorporated *sə*; and so on. Hiw is the language where the contamination of the pronoun system with Aorist

---

[13] The earlier forms **nɔ* and **nɛ* only survive vestigially, in the 1sg possessive of Lo (*mi-nɔ*) and Toga (*mi-nɛ*) – etymologically 'with me'.

clitics has been maximal.

Interestingly, the 1sg Aorist clitic *gu* has also undergone another reanalysis of a completely different kind. In modern Vera'a, the prefix *k-* encodes the Aorist not only for 1sg, but for all non-singular persons as well (*Table 3*). It looks as if this prefix were in the process of being reanalyzed as the general marker for Aorist, thereby losing its original connection with the first person; ironically, this is the exact opposite of the Torres evolution. Even 2sg and 3sg seem threatened by the expansion of VRA *k-* to all persons, judging by the attestation of *nik sɔ **k**-van* and *di sɔ **k**-van* as alternatives to the regular *nik sɔ **Ø**-van* and *di sɔ **nɛ**-van* (respectively 'you/he should go').

Finally, this generalization of /k/ to non-singular persons in Vera'a possibly provides the explanation for a non-etymological /k/ found in several non-singular pronouns in Vera'a's closest neighbor, Vurës: *kɛmɛ**k*** '1exc.pl'; *kumuru**k*** '1exc.du'; *duru**k*** '1inc.du'. If this hypothesis is right, then Vurës would combine the evolution patterns of Vera'a (spreading of *gu* to non-singular persons) and of the Torres languages (accretion of *gu* to the preceding pronoun, and generalization of the augmented form to all functions, regardless of the verb's aspect). If this hypothesis is true, then these three non-etymological /k/ are the only trace left by *gu* in Vurës. Indeed, this language – like Mota and Mwesen – encodes its 1sg Aorist subjects with an innovative form *na* (*Table 2*), surely the result of the coalescence of *nɔ* <*nau* and *a* <*(k)a*.

5   Conclusion

Out of the seventeen languages of north Vanuatu, eleven share an aspect category labeled "Aorist", whose function is to represent a new event regardless of its deictic coordinates in terms of tense or modality. Despite their differences, the various morphosyntactic patterns that encode this Aorist can be shown to derive ultimately from a single protosystem: a set of four portmanteau proclitics {*gu–*u–*ni–*(k)a} combining aspect marking and subject agreement.

The reconstruction proposed in this paper not only helps to explain the development of these languages' TAM systems, but also proves indispensable in unravelling the historical morphology of personal pronouns in this part of Oceania. The next challenge is now to define the precise link – both formal and functional – that connects these four clitics to the set of personal pronouns which Robert Blust reconstructed for the remote ancestors of these languages.

17

References

Blust, Robert. 1977. The Proto-Austronesian pronouns and Austronesian subgrouping: a preliminary report. *University of Hawai'i Working Papers in Linguistics*, 9/2:1-15.

——. 2003. Three notes on early Austronesian morphology. *Oceanic Linguistics* 42-2, 438-478.

Benveniste, Émile. 1966 [1959]. Les relations de temps dans le verbe français. In *Problèmes de Linguistique Générale*, vol.1. Tel. Paris: Gallimard. Pp.237-250.

Clark, Ross. 1985. Languages of North and Central Vanuatu: Groups, chains, clusters and waves. In Andrew Pawley & Lois Carrington (eds), *Austronesian linguistics at the 15th Pacific Science Congress*. C-88. Canberra: Pacific Linguistics. Pp.199-236.

Comrie, Bernard. 1976. *Aspect. An introduction to the study of verbal aspects and related problems*. Cambridge: Cambridge University Press.

Culioli, Antoine. 1978. Valeurs aspectuelles et opérations énonciatives : l'aoristique. In J. David & Robert Martin (eds), *Actes du colloque sur la notion d'aspect*. Metz: Klincksieck. Pp.182-193.

Depuydt, Leo. 1993. *Conjunction, Contiguity, Contingency: On Relationships between Events in the Egyptian and Coptic Verbal Systems*. New York: Oxford University Press.

François, Alexandre. 2003. *La sémantique du prédicat en mwotlap (Vanuatu)*. Collection Linguistique de la Société de Linguistique de Paris, 84. Paris, Louvain: Peeters.

——. 2004. La réduplication en mwotlap : les paradoxes du fractionnement. In E. Zeitoun (ed.), *Les langues austronésiennes*. *Faits de langues* n°23-24 : 177-194.

——. 2005. Unraveling the history of the vowels of seventeen northern Vanuatu languages. *Oceanic Linguistics* 44-2:443-504.

——. 2007. "Verbal aspect and personal pronouns: The history of aorist markers in north Vanuatu". Paper presented at 7th *International Congress on Oceanic Linguistics* (COOL7), Université de Nouvelle-Calédonie, Nouméa, July 2007.

Galand, Lionel. 2003. L'aoriste berbère, l'aspect et les valeurs modales. In Jérôme Lentin & Antoine Lonnet (eds), *Mélanges David Cohen*. Paris: Maisonneuve & Larose. Pp.235-246.

Jauncey, Dorothy. 2002. Tamabo. In Lynch, Ross & Crowley (eds). Pp.608-625.

Kikusawa, Ritsuko. 2003. A new view of the Proto Oceanic pronominal system. *Oceanic Linguistics*. 42-1:161-186.

Lynch, John & Ozanne-Rivierre, Françoise. 2001. Some shared developments in pronouns in languages of Southern Oceania. *Oceanic Linguistics* 40-1, 33-66.

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

Robert, Stéphane. 1996. Aspect zéro et dépendance situationnelle: L'exemple du Wolof. In C. Müller (ed.), *Dépendance et intégration syntaxique (Subordination, coordination, connexion)*. Linguistische Arbeiten. Tübingen: Niemeyer. Pp.153-161.