Documentation, Analysis, and Writing of TAM markers in Vera’a

Stefan Schnell
stefan.schnell@mail.ru

Kiel University

1st International Conference on Language Documentation and Conservation
Language and speakers

genetic affiliation: Austronesian, Oceanic, Northern Vanuatu, Banks Is. (Lynch et al. 2001)

location: West / North-West Vanua Lava, Banks Is., Northern Vanuatu

number of speakers: approx. 300

neighbouring languages: closely related Vurës (approx. 1,000 speakers), Lemerig, Mwesen (both moribund)
VolkswagenStiftung-funded DoBeS Project: "Documentation of Vurës and Vera’a, the two surviving endangered languages of Vanua Lava, Vanuatu"

duration: Nov 2006 - Oct 2009

linguistics: Catriona Malau (Vurës);
Stefan Schnell (Vera’a)

other disciplines: anthropology, ethno-musicology,
marine-biology, botany

aim: thorough documentation of the languages in cultural context

academic output: description of the language (sketch grammar)
multi-media lexical database
doctoral thesis

output for community: story books
printed dictionary
Data collection and processing in the field

Types of data I

**Recordings:**

**format:** audio and video recordings

**interaction:** (mostly) non-interactive (monologues)

**text genres:**
- traditional and contemporary stories
- memories of local history (e.g. hurricane, landslide, WWII)
- procedural texts (e.g. house building, cooking)
- descriptions / explanations (e.g. plants and their use as food, medicine, etc.)
Data collection and processing in the field
Transcription, translation, editing

**Transcriptions:**
by two, later only one native speaker
later checked by myself

**Translations:**
Bislama: by native speaker
English: by myself with assistance
of language teachers

**Editing of stories:**
by one native speaker
(who also transcribes)
Data collection and processing in the field

Types of data II

**Written text:**

transcriptions of recordings

edited texts (stories)
Vera’a has SVO word order and the first element of the verbal predicate is a tense-aspect-mood (TAM) marker:

\[
\text{SubjNP TAM Verb (ObjNP)}
\]

(1) nak susuõ ga tik
    boat paddling TAM small

'The canoe is (too) small.'

HHAK.010
Basic clause structure and position of TAM markers

TAM marker occurs in clause-initial position:

e.g. where subject is omitted under co-reference

(2) ‘ama-ruô ne kalraka [ne van lē =n gamal]
father-3DL TAM get.up TAM go LOC =ART mens’.house
’Then their (two) father got up [(and) went to the mens’ house].’ MVBW.107
The problematic cases
Formal status of TAM markers

(3) no m le n gunu-k
    1SG TAM take ART spouse-1SG
    'I took a wife.' 1.JS.094

(4) no k van ma
    1SG TAM go hither
    'Then I came here..' JJQ.403

Problem of segmentation / word boundaries (cf. Himmelmann 2006):

1. In transcriptions of the recordings
2. For linguistic analysis
The problematic cases
Formal status of TAM markers

The morph k is one allomorph of a more complex morpheme (called 'AORIST in Francois (2007)), the allomorphy of which is conditioned by Pers / Num of Subj (examples meaning: 'I will go.'; 'You will go.'; ...):

<table>
<thead>
<tr>
<th>AORIST:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pers/Num</strong></td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td><strong>SINGULAR:</strong></td>
</tr>
<tr>
<td>1sg</td>
</tr>
<tr>
<td>2sg</td>
</tr>
<tr>
<td>3sg</td>
</tr>
<tr>
<td><strong>NON-SINGULAR:</strong></td>
</tr>
<tr>
<td>e.g.</td>
</tr>
<tr>
<td>3pl</td>
</tr>
<tr>
<td>2dl</td>
</tr>
<tr>
<td>3tl</td>
</tr>
</tbody>
</table>
The problematic case
Formal status of TAM markers

The morph $m$ is the only allomorph of a morpheme $M$.

Comparison with neighbouring languages and historical reconstructions suggest that both forms once consisted of a CV syllable, probably a prefix or proclitic to the verb. Subsequently, vowel deletion occurred, a frequent development in the languages of the area (cf. Francois 2007).
Phonological word = preceding PRO + TAM marker

(3') a. no mlen [...]  
   b. |unm|len|

(4') a. no kvanma  
   b. |nɔk'|van|ma|

syllable structure: CCV(C) or (C)VCC are not attested elsewhere

allophones of /k/: word-final: [k'], else: [k]
Fusional tendencies with preceding elements, either nouns or pronouns:

(5)  

a. /mɛʔ/ - 'reef'

me m löl
reef TAM be.flooded

'The reef got flooded'

b. |mɛm|lɔl|

(6)  

a. /kamam/ - '1PL.EXCL'

kama k mul kal ma
1PL.EXCL TAM go upwards hither

'We came up here.'

b. |kamak'|mulɔ|
The category of the word preceding the TAM markers is open, usually the last or only constituent of the subject NP:

(7) a. e raga ’anē m rem-rem lē n ’erē dara
   ART people DEM TAM RED:climb LOC ART PL tree.sp.
   'They climb on the dara trees.' JJQ.143
b. |ʔanēm|remrem|

(8) a. e raga ’anē k van ma
   ART people DEM TAM go hither
   'Then these people came here.' HHAK.028
b. |ʔanēk|van|

→While the word preceding the TAM marker varies, the word following it appears to be invariably a verb.
TAM markers occur in clause-initial position under subject deletion:

(9) a. tētē m rara [m rara] ...
    baby TAM cry TAM cry
    'The baby was crying [and crying] ...'
    ASB.053

   b. ... |raram|rara|...

(10) a. dirē k van 'anē [k van rōw ma 'anē]
    3PL TAM go DEM TAM go down hither DEM
    'Then they went [and came down].'
    ASMS.163

   b. ... |ʔanék'|van|

→ While the subject may be deleted if the context provides the necessary information, this is not true for the verb.
Other morphemes that are not part of the subject NP may intervene between the latter and the TAM marker:

(11) nik wo m ma’ ba no k ma’ rek
1SG COND TAM dead CONJ 1SG TAM dead too

'If you die, then I want to die, too.'  
	ANV.079

(12) no so k mi’ir
1SG PROSP TAM sleep

'I’d like to sleep.' 
	Francois 2007
However, one example in my corpus shows the prospective marker intervening between TAM marker and verb:

(13) ‘alē duru k van ‘anē [k so ‘ara bē]
     alright 3DL TAM go DEM TAM PROSP scoop water
     ‘Alright, then the two went in order to get water.’

     ASMW.183
Preliminary conclusions

**Phonological word:** preceding word + TAM

**Grammatical word:** TAM marker + verb(al root)

Mismatch between phonological and grammatical words (cf. Aikhenvald & Dixon 2002:29)

Confirms Francois’ (2007) analysis of k as a prefix in Vera’a rather than part of a protmonteau morph formed with preceding pronoun, as in neighbouring languages.
How to treat pauses?
Pausing before TAM marker

Pauses occur before TAM markers only if no subject NP is present in the clause:

(14) ... lē n qila ’anē’e (0.6) m kal ba ’enteg sar
     LOC ART bay DEM PAUSE TAM enter inside well inside
     ’... into the bay, went well into it.’ ISAM.039

(15) okay (1.9) k van ma mē diē
     okay (FILLED PAUSE) TAM go hither DAT 3SG
     ’Okay, came to him.’ JEM.031

These examples are extremely rare in the corpus. For k, it is the only one.
How to treat pauses?

Pausing after TAM marker

Pauses frequently occur after subject NP + TAM marker and also after sole TAM markers:

(16)  
dirē k  (0.7) sag  gōr  die
3PL TAM PAUSE sit  protect 3SG

'They were sitting guarding after her.'  ASMS.058

(17)  
no k  vanē k  (1.5)  bu’uñ  ēn  nōk  ēn  nōl  qe’ē
1SG TAM go  ?? TAM FILLED.PAUSE plant  ART my.thing  ART top  taro

'I'll go and plant my taro tops.'  ASMS.094

Does Vera’a allow for pausing within a grammatical word, between affix and lexical root?

Not with e.g. possessive suffixes attached to nouns, clear cases of affixes.
False starts and corrections
Restarts after TAM marker

If the correction triggers the subject NP, then the cut-off point is always after subject NP + TAM marker:

(18) no k (CORR) nikē da n mogi
1SG TAM 2SG.AOR make ART his.drink
'I ... you’re gonna prepare his (kava).’

(19) duru m (CORR) dirē m van ’ō kal ...
3DL TAM 3PL M go carry upwards
'The two ... they brought (us) up here …’
If the correction triggers the TAM marker, the pronoun is repeated together with the latter:

(20) dirē m (1.7) dirē k van
    3PL TAM (PAUSE) 3PL TAM go

'They ?? ... then they went.'

These examples seem to suggest that we do not deal simply with a 'mismatch' of phonological and grammatical word.
Writing of TAM markers in native speaker's transcriptions

The 'unproblematic' cases

In structures \([\text{Subj NP TAM Verb}]\), TAM markers \(m\) and \(k\) are written in a unit with only or rightmost NP constituent, i.e. the phonological word:

(3')

a. no \(m\) len gunuk
b. nom len gunuk

(4')

a. no \(k\) van ma
b. nok van ma
Writing of TAM markers

Corrections

With corrections, TAM markers are written in unit with e.g. the preceding pronoun, dots seem to indicate that something is 'wrong' here:

20’ a. \( \text{dirē m (1.7) dirē k \ } \text{van ...} \)  
   \( \text{3PL \ TAM \ PAUSE \ 3PL \ TAM \ go} \)  
   'They ?? ... then they went ...’  
   HHAK.013  

b. \( \text{dirēm...dirēk \ van} \)
TAM markers in clause-initial position are written in one orthographical word with the preceding word, though this is part of the preceding clause:

10’ a. dirēk van ’anē [k van rōw ma ’anē]
    3PL TAM go DEM TAM go down hither DEM
    'Then they went [and came down].'

b. dirēk van ’anēk van rōw ma
If TAM markers occur after pauses at the beginning of a clause, they are edited out:

15' a. okey (1.9) k van ma mē diē
   okay (FILLED PAUSE) TAM go hither DAT 3SG
   'Okay, came to him.' JEM.031

b. okey, van ma mē diē
Conclusions
Linguistic analysis

A lot of evidence for a prefix analysis (cf. Francois 2007)

But: An enclitic analysis might be available, as well.

The formal characteristics of the TAM markers in questions make it hard to accommodate these morphophonologically in some contexts.

Is a language change underway, possibly resolving the 'unstable' situation?

Will Vera’a follow the tendency in neighbouring languages (cf. Francois 2007) to develop fused forms of PRO + TAM?

Is probable also from typological perspective (cf. Mushin 2006): tendency for PERS / NUM information to cluster with TAM information at the beginning of a clause or in second position.

Still, a prefix analysis (following Francois (2007)) may be the most 'straightforward', as it covers most of the examples, and may be the one opted for in the sketch grammar - with comments.
The writing of TAM markers should be kept as it is, also in edited texts.

Hand-written transcriptions should be documented, e.g. as scanned texts which can be archived alongside with the recordings and annotations.

Though they provide valuable information on (mainly one!) native speaker's intuition, they should not drive the linguistic analysis.

These transcriptions need to be further analyzed.
Conclusions
Documentation

The documentation should capture all the linguistic facts, as 'untidy' as they may be in some instances.

It should be a source for further research rather than a 'convincing', final analysis.

The linguistic analysis should not be imposed on all the outputs involved in a language documentation.

Conflicts between writing and linguistic analysis should be commented on in the annotations of the recordings, and these should be linked to the linguistic analysis, i.e. the sketch grammar (cf. Schultze-Berndt 2006).
Conclusions

Interlinking (Schultze-Berndt 2006)

Vice versa, the respective chapter of the sketch grammar contains information on how to search the corpus for relevant data, e.g. via key words (preliminarily).