Journal of the Linguistic Society of Papua New Guinea

ISSN: 0023-1959

Vol. 35, 2017

Writing Lamap: the representation of person markers

Julie Barbour & Claudia Williams
Malekula Languages Project, University of Waikato
julie.barbour@waikato.ac.nz; claudiamaria.williams@gmail.com

Abstract

Primary school teachers have faced a number of representational issues when developing a standard written form of the Lamap language of Malekula Island, Vanuatu. The issue discussed in this paper concerns the treatment of subject person markers, which teachers variably represent as independent word forms surrounded by space, or as prefixes attached to verbs. We employ linguistic theory in an attempt to resolve the representational issue, applying a selection of word segmentation criteria compiled by Haspelmath (2011) to the Lamap data. The criteria of non-selectivity, free occurrence, and non-coordinatability prove to be of relevance, while the search for morphophonological rules and idiosyncracies results in an interesting but separate discovery. Our analysis indicates that the Lamap person markers display properties of bound forms rather than of independent word forms. There is some evidence for their status as affixes as compared with free-form grammatical particles. While our findings help us to better understand the variation that we have observed in Lamap, ultimately it is the community of emerging Lamap writers who will determine how the language is represented.

Key words: Malekula, Lamap, person marking, vernacular literacy, orthography.

1. Introduction

1.1 The Lamap language of Port Sandwich

The language of Lamap is spoken in the southeast of Malekula Island, Vanuatu, on the southern side of Port Sandwich. The broader community population, including second language speakers of Lamap, has been estimated to comprise between 750 speakers (Crowley 2002: 650) and 1200 speakers (Lynch and Crowley 2001: 76). The 2009 census population statistics indicate a population of 783 residing in the villages within the immediate Lamap-speaking region (Vanuatu National Statistics Office 2009: unpublished census data). [1]

The language has recently been classified as belonging to the Southeastern subgroup of the Eastern Malekula Linkage (Lynch 2016: 418). Lynch (2016: 399) comments that all the languages of Malekula "probably belong to a single subgroup of the Central Vanuatu subgroup of Southern Oceanic". Lamap is one of 24 languages thought to be actively spoken on Malekula, with a further 15 known to be moribund or recently extinct (Lynch and Crowley 2001: 68, 85). For a population of some 37,000 (Vanuatu National Statistics Office 2009: 13), this gives an extremely high number of languages per capita. In the broader context of

---

[1] A Bislama summary of the paper is provided for ni-Vanuatu readers in Appendix I.

Vanuatu, Lamap is one of more than 100 indigenous languages (Lynch and Crowley 2001: 1, Simons and Fennig 2017).

Lamap is better known as Port Sandwich (ISO psw) for the European name designated to the trading port near Lamap's location. Lamap was the focus of research by French linguist Jean-Michel Charpentier (1979, 1982, 1995). Charpentier's work informed Crowley's (2002: 650-659) brief English grammar sketch of the language. Charpentier's research activities did not support the development of a written tradition within the community,[2] and as per government policy, Lamap children have been educated in the medium of French.

1.2 Education in Lamap

When Vanuatu gained independence in 1980, English and French were named as the languages of education (Republic of Vanuatu 2006 [1980]). In Francophone regions like Lamap, where a Roman Catholic Mission was established in the late 1800s (Pacific Manuscripts Bureau 2015), French became the language of formal schooling. This situation has only recently been reconsidered through the Vanuatu National Language Policy (Vanuatu Ministry of Education 2012). In the new policy, vernacular languages are identified as suitable for the early years of learning, to meet the educational and cultural needs of children (Vanuatu Ministry of Education 2012: 2). According to the policy, Years 1 and 2 of formal education are to be conducted in the medium of the local vernacular language, or Bislama. Children then transition into the official languages of education, these being French and English, in Year 3. Provision is made for the continued use of the vernacular through primary education to facilitate learning.

The Vanuatu National Language Policy will see a shift in the medium of early primary education from French to Lamap. To implement this shift, an intensive programme of materials writing has taken place, generating materials for Years 1 to 3 learning. Year 1 materials were translated into Lamap at a Vanuatu Education Support Program [VESP] workshop on Malekula in late 2015. Year 2 and 3 materials were translated at a second VESP workshop on Malekula in late 2016. During the workshops, two Lamap primary school teachers produced translations of 50 graded readers, an additional 24 fluency readers, 9 posters for classroom walls, and alphabet and number charts. The Lamap teachers were supported by University of Waikato student linguist Claudia Williams, under the supervision of Dr. Julie Barbour.

1.3 Writing Lamap

Translation work presupposes a written tradition. While some of Vanuatu's languages have been fully analysed, and orthographies have been proposed, Lamap does not have an established orthographic tradition that is widely known within the community. Because of this, the Lamap primary school teachers at the 2015 and 2016 VESP workshops on Malekula

---

[2] Charpentier (1997: 223, 227) expressed the view that the development of literacy skills in either an indigenous language or a pidgin language such as Bislama, represented an unwelcome intrusion into traditional cultural practices.

faced many practical issues concerning the representation of their language on paper. At the time of writing, aspects of the Lamap orthographic system employed during the VESP workshops remain fluid.

The representation of contrasting segments is a core concern for Lamap writers. Malekula languages are analysed as having complex segments, and there is considerable variation in the representation of these. The prenasalised plosive series, commonly /b^w, b, d, g/ in Malekula languages, can be written as variously as 'mbw, mb, nd, ngg', 'bw b, d, g', or 'mpw, mp, nt, ngk'.[3] In lieu of a full phonological analysis of Lamap, the potential complex trill phoneme [ndr] was represented with the digraph 'dr' or 'ndr' by teachers in 2015. The linguistically recommended representation 'dr' followed conventions adopted elsewhere on Malekula. On the teachers' return to the Lamap community, they consulted with other community members, and 'dr' was rejected. Instead, 'dd' was proposed as an alternative, and in 2016 this preferred representation was employed in writing new literacy materials. The locally preferred spelling 'dd' is analogous to the use of 'bb', employed to represent the prenasalised bilabial trill [^{m}ʙ] in Lamap.

In a second example of emerging orthographic practices, in 2015, the digraph 'ch' was used to represent the voiceless affricate [ʧ]. This was revised to 'j' in 2016 materials. The character 'j' is used elsewhere on Malekula for the voiceless affricate segment (see e.g. Uripiv [McKerrras 2001]), although there are also Malekula languages with a prenasalised voiced affricate [^{n}ʤ], represented as 'j' (see e.g. Neverver [Barbour 2012: 36]), and with contrasting plain [ʧ] and prenasalised [^{n}ʤ] affricates, represented as 'ch' and 'j' respectively (see e.g. Malua Bay [Wessels 2013: 42]).

Beyond phonology, a representational issue that the Lamap teachers have struggled with has been the systematic placement of word boundaries, indicated by the presence or absence of spaces around written elements. This is particularly variable in the representation of person markers which are associated with verbs. These person markers express person and number properties of the subject of the clause, and are employed for both intransitive and transitive subjects. On some occasions, the teachers wrote the person markers as independent words, surrounded by space. At other times, the markers were written as prefixes attached to their associated verb, or associated negative particle. Variation was observed within paradigms, where some person markers were more consistently written in one way or the other. Variation also occurred with the representation of individual markers, where the same marker would sometimes be written as an independent 'word', and sometimes as a prefix. Example (1a) shows the word representation, while (1b) displays the prefix representation.[4]

---

[3] Word-finally, voiced segments can undergo a process of devoicing, thus accounting for the orthographic representation of /b/ as 'mp'. Word-final devoicing is reported for example in Neverver (Barbour 2012: 31).

[4] The linguistic glosses used in this paper include: 1 first person, 2 second person, 3 third person, ADJ adjective, CONJ conjunction, DU dual, EXCL exclusive, INCL inclusive, IRR irrealis mood, LOC locative, NEG negator, POSS possessive, PL plural, QUANT quantifier, REAL realis mood, and SG singular.

(1)   a.   *No*   *rox*   *lä*   *näräs.*
              1SG.REAL   live   LOC   sea

        b.   *Norox*   *lä*   *näräs.*
              1SG.REAL-live LOC   sea

        'I live in the sea.' [psw_icr19.005]

While both the word and prefix options are possible representational choices, a standard writing system would choose one means of representation, and use this systematically. Such a written standard is needed to create literacy materials and deliver literacy skills to children, and it is needed to assess children's acquisition of literacy skills through their written work. Over time, the current fluidity in the written form of Lamap will likely need to transition towards a more standard representation. This means choosing either (1a) or (1b), and using the chosen writing strategy consistently.

In order to support the local Lamap teachers in their writing efforts, we have undertaken a linguistic analysis of Lamap's person markers. The purpose of our analysis is to establish whether the markers behave more like free morphemes or bound morphemes. If they are found to behave more like bound morphemes, we will consider whether there is any evidence that they should be represented as prefixes, as in (1b), or whether they are more appropriately represented as free-form grammatical particles. Primarily, our analysis draws on Martin Haspelmath's (2011) paper on word segmentation. In this paper, Haspelmath collates ten general criteria for wordhood. We explore a selection of these criteria, in an attempt to establish the language-specific status of Lamap's person markers.

The data for our analysis is drawn from the graded readers translated at the VESP workshops by the Lamap teachers, along with further translations of locally produced fluency readers. The readers have been collated with permission as research data in the Lamap Reader Corpus.[5] The books range from one sentence per page, to more extended narratives, and there are currently over 70 readers in total.[6]

To set the scene for the discussion of person markers in Lamap, we briefly review understandings of person markers in the Oceanic language family in Section 2. We then introduce two paradigms of person morphology in Lamap, drawing on evidence from the translated school materials, as well as earlier linguistic research, in Section 3. In Section 4 of the paper we apply a selection of Haspelmath's (2011) listed criteria for wordhood in an

---

[5] Permission to use workshop material for research purposes prior to publication was granted in writing by the teachers involved in translation. The project was approved by the University of Waikato's devolved Human Research Ethics committee for the Faculty of Arts and Social Sciences under approvals FS2012-04 and FS2016-58, and data was collected with the permission of the Vanuatu National Cultural Council (Research Permit granted 2012-2015, extended 2016).

[6] While we acknowledge that translated materials are a less-than-ideal data source, at the same time it is important to acknowledge the functional multilingualism of the Lamap teachers. The teachers were working from Bislama texts into Lamap. At several points during the workshop, the Lamap teachers, and indeed all teachers at the Malekula VESP workshops, overtly expressed their understanding that an effective translation is one of meaning rather than one of form. We felt that the focus on translation of meaning rather than form was enhanced by the whole texts that teachers were working with. This distinguishes text translation from clause by clause grammatical elicitation, which has been critiqued as a method of linguistic data collection (see e.g. Samarin 1967, Gil 2001, Bowern 2008, Lüpke 2010), although elicitation retains an important place in linguistic fieldwork.

attempt to support the analysis of Lamap's person morphology as being free or bound, and to identify properties of the morphology that might suggest affixation. We finish in Section 5 with comments on the status of the Lamap person markers, and the application of our findings to the issue at hand: "Are person markers in Lamap more appropriately represented as separate words or as prefixes?"

2.   Person Markers in the Oceanic languages

A number of observations have been made about person markers in the Oceanic languages. Lynch, Ross and Crowley (2002: 35-36) and Ross (2004: 498) comment that canonical Oceanic languages usually present four separate paradigms of "pronominal" forms. The four paradigms identified by Lynch, Ross and Crowley (2002), and later Ross (2004) are listed in (2).

(2)   a.   independent pronouns;
         b.   possessor suffixes which attach to possessed nouns, possessive classifiers, and in some cases prepositions and adjectives;
         c.   preverbal subject markers, which may present as clitics or prefixes;
         d.   postverbal object markers, which may present as clitics or suffixes.

Of interest in this paper are category (2a) forms, comprising independent person markers, and their relationship to category (2c) forms, comprising preverbal subject markers. Person forms in these two categories may co-occur in Oceanic languages, although Ross (2004: 499) comments that "in most canonic languages... the subject proclitic [subject marker] occurs whether or not there is a subject noun phrase". In this way it is possible for the category (2c) subject marker to be the only overt representation of the entity encoded as subject. Alternatively, the properties of the entity functioning as subject may be distributed between a separate subject noun phrase (including one comprising an independent pronominal), and the preverbal subject morpheme.

Category (2c) subject markers associated with verbs are observed to be either free forms, or prefixes, with both types found in Melanesia (Lynch 1998: 103; Lynch, Ross and Crowley 2002: 45). In a given language, subject morphology may combine with Tense/Aspect/Mood [TAM] morphology before the verb through concatenation, or the subject markers may form portmanteau morphemes including TAM meanings (Lynch 1998: 138; Lynch, Ross and Crowley 2002: 45, Ross 2004: 500-501). Ross (2004: 500-501) observes that mood contrasts in particular are a major feature of verbal morphology, where "there is usually a distinction between realis [REAL] and irrealis [IRR] mood", and where such morphemes "usually come immediately before or immediately after the subject marker". Negation may also be marked in a pre-verbal position (Lynch 1998: 159; Lynch, Ross and Crowley 2002: 45).

These observations of Oceanic languages direct our attention to the possibility of a number of structural positions before the verb, filled by morphology which is likely to express information concerning the grammatical subject, mood categories and negation in particular.

3. Person Markers and the Verb Complex in Lamap

Previous research on Lamap, combined with new evidence from the Lamap Reader Corpus, indicates that there are at least two paradigms of person markers that are relevant to this discussion.[7] The first paradigm is presented in Table 1, and all forms are attested in the Lamap Reader Corpus. The members of Paradigm I are described by Crowley (2002: 652) as "pronouns".

	Singular		Dual	Plural
1	*inao*	INCL	*ddäu*	*ddate/o*
		EXCL	*namu*	*nämit(e)*
2	*xaing*		*xamu*	*xamite*
3	*nai*		*xau*	*xate*

Table 1. Paradigm I Person Markers in Lamap

The second paradigm of person markers is presented in Table 2. The members of Paradigm II are described by Crowley (2002: 655) as "subject prefixes" which attach to verbs. All Paradigm II person markers are attested in the corpus of Lamap literacy materials, with the exception of the second person plural marker. Crowley (2002: 655) lists *xati* as the second person dual form; however, it is possible that *xati* belongs to a separate paradigm. As such, we have left that cell empty. Further comments on forms with final-*i* vowels are offered in section 4.4.

The person markers shown in Table 1 and Table 2 contrast singular, dual and plural number. Non-singular first person forms also contrast in terms of clusivity. Lamap makes no gender distinctions in either paradigm.[8]

---

[7] Earlier work on Lamap (Crowley 2002: 654-655), as well as evidence from the Lamap Reader Corpus, indicates that there is a paradigm of possessive suffixes in the language, and an additional set of object suffixes which apparently indicate properties other than person and number. Identification and analysis of the properties of these morphemes is part of a 2017-2018 masters thesis project being conducted by Claudia Williams under the supervision of Julie Barbour. Regarding the Paradigm II forms shown in Table 2, Crowley suspected patterns of vowel harmony, but was unable to establish these with any certainty (Crowley 2002: 655). Our data indicates that vowel variation patterns according to mood. In Table 2, we present the paradigm established for realis contexts.

[8] The vowels in Lamap require further investigation. Teachers vary between writing 'a' [a] and 'ä' [æ] in the same words. Likewise, there is variation in writing between 'o' [o] and 'ö' [ɵ, ø], and between 'u' [u] and 'ü' [ʉ]. There is an audible distinction in the fronting of these pairs, but until audio data is collected from a large number of speakers, it will not be possible to establish the range of allophonic variation for each contrasting phoneme. As such, vowels with diaeresis appear somewhat sporadically in this paper as they were written by teachers. They are not necessarily in contrast with plain vowels. Given the Francophone nature of the Port Sandwich area, there may well be a French influence in the speech of some language users.

	Singular		Dual	Plural
1	*no*	INCL	*ddu*	*ddato*
		EXCL	*namu*	*nato*
2	*ko*		*x(a)mu*	---
3	*e*		*ku*	*(xa)to*

Table 2. Paradigm II Person Markers in Lamap

Evidence for the difference between the person markers in Paradigm I and Paradigm II is presented in examples (3) to (8) below. Looking at the behaviour of the elements in the tables, we can observe that Paradigm I person markers encode intransitive S-function subjects (3) and transitive A-function subjects (4). In these functions, they are positioned before the verb, and the Paradigm I person markers co-occur with Paradigm II person markers.

(3)   Paradigm I form as S subject/Paradigm II form as S subject
         "Inao$_S$ *no*   *bao."*
         1SG   1SG.REAL   be.big
         'I am big.' [psw_icr19.004]

(4)   Paradigm I form as A subject/Paradigm I form as A subject
         Xate$_A$ *to*   *sba*   *kakadd*   *wil.*
         3PL   3PL.REAL   NEG   have   wheel
         'They do not have wheels.' [psw_icr20.014]

Paradigm I person markers also encode O-function objects (5), and serve as the objects of prepositions (6). In these functions, they are positioned after the verb or preposition respectively. There is no accompanying Paradigm II morpheme associated with these functions. It may transpire that there is person or transitivity morphology on some verbs or prepositions, but it is too early in our analysis to comment on such morphology.

(5)   Paradigm I form as O object
         Räbä   *nisävä*   *nanam*   *e*   *raja*   *ddate*$_O$
         because   what   mosquito   3SG.REAL   sting   1PL.INCL
         'Why does the mosquito sting us?' [psw_Yia3Rida09.001]

(6)   Paradigm I form as object of preposition
         E   *bravüx*   *kanä*   *xate*$_O$
         3SG.REAL   join.in   with   3PL
         'He joined in with them.' [psw_Yia3Rida05.012]

Finally, Paradigm I markers are attested in our data functioning as the subject (7) of a non-verbal clause, and as an elliptical utterance (8). Again, there are no accompanying Paradigm II markers with the Paradigm I forms in these examples.

(7)   Paradigm I form as subject of non-verbal predicate
        "Nai   äbi   amo."
        3SG   where   here
        "He is here/Here he is" [psw_icr23.013]

(8)   Paradigm I form as elliptical utterance
        "Xaing   miji."
        2SG   immediate
        "(It is) you now/It is your turn now." [psw_van03.013]

Based on the distribution of Paradigm I person markers in examples (3) to (8), the analysis of these morphemes as independent pronouns is appropriate. Such an analysis is reflected in the written representation of Lamap by teachers, who consistently treat the Paradigm I person markers as separate words, surrounded by space. Paradigm II person markers, on the other hand, are restricted in their distribution. In examples (3), (5) and (6) show the person marker positioned immediately before the verb root, while (4) shows the person marker positioned before the negator, which is followed by the verb root. Paradigm II person markers are thus more positionally bound than Paradigm I person markers. As grammatical morphemes, they are also more functionally restricted than the Paradigm I pronouns, and can only encode properties of the S/A subject of a verbal clause.

In Lamap, the morpheme expressing negation is positioned between the Paradigm II person marker and the verb root, as shown in (9) and (10). While negation in Melanesian languages is sometimes expressed in a distributed fashion (Lynch 1998: 160, Lynch, Ross & Crowley 2002: 51), Lamap is one of a number of Malekula languages that has only one negative marker (see e.g. Barbour 2015: 436-438). Lamap's negator takes the form *sba*.

(9)   *Nürükav   e   sba   bao*
        ant   3SG.REAL   NEG   be.big
        'The ant is not big.' [psw_mlp_y1-6a.07]

(10)   *Eme   kü   sba   tädd   aim.*
        No   3DU.REAL   NEG   be/sit/stay house
        'No, they (two) are not at home.' [psw_icr23.004]

As noted in section 2, the morphological contrast between realis and irrealis mood is widespread in the Oceanic languages. Crowley (2002: 655), working with Lamap data from Charpentier, observed that where one of the Paradigm II subject markers is present, but where no other morphology occurs, the default interpretation is present/past time. This is consistent with a realis mood interpretation. Whether Lamap overtly encodes the contrast between realis and irrealis mood is of considerable interest, and we discuss this matter in some detail in section 4.4.

In the Lamap Reader Corpus, the verb complex involves at least two structural positions before the verb root. Firstly, there is a position for the person marker representing person and number properties of the entity functioning as the S/A subject of the clause, as well as the mood of the clause. Secondly, there is a position for the negative morpheme, when the situation expressed in the clause has negative polarity. In addition, according to

Charpentier (1979:175-176), there are morphemes that express inchoative, dubitative, and continuous aspect that co-occur with the first and second order morphemes and precede the verb root. Only one of these morphemes is potentially attested in the Lamap Reader Corpus, and it is unclear to what extent the morphology is productive in the language. As such, only the two well-attested positions are presented in this study. The verb root follows these two positions, as shown in (11).

(11)   1ST ORDER   2ND ORDER
            SUBJECT MARKER   (NEGATION)   VERB ROOT

While the verb complex hypothesised in (11) is not necessarily maximal, it provides a useful template for the discussion of wordhood that follows. The first and second order positions are evidenced throughout the data in the Lamap Reader Corpus. The question that we are hoping to resolve is whether the verb complex comprises a series of morphemes associated with the verb root in a phrasal structure, indicating a separate representation of each morpheme, or series of prefixes attached to the verb root, indicating a 'word' representation where the morphemes are presented with no intervening spaces between them. The relationships between the subject marker, the negator, and the verb root, and between the subject marker and other elements in the clause, are thus the focus of the remainder of this paper.

4. The status of person markers in Lamap

Evidence presented in Section 3 indicates that Paradigm I person markers are free forms. In contrast, Paradigm II person markers are bound, at least in terms of being positionally restricted. While Lamap teachers represent Paradigm I person markers as words surrounded by space, Paradigm II person markers are represented variably as free forms or as prefixes. In this section, the status of Paradigm II person markers will be examined closely. Primarily, we employ Martin Haspelmath's (2011: 38-59) criteria for wordhood,[9] investigating wordhood four criteria in detail. The criterion of non-selectivity is considered in section 4.1, free occurrence is considered in section 4.2, non-coordinatability is discussed in section 4.3, and morphophonological rules and idiosyncracies are investigated in section 4.4. The remaining criteria collated by Haspelmath (2011) are either irrelevant to Lamap, or cannot be applied due to the nature of the corpus we are currently working with.

---

[9] An important purpose of Haspelmath's paper is to make the point that attempts to define "a cross-linguistically valid concept of '(morphosyntactic) word'" have failed (Haspelmath 2011: 31). None of the ten criteria is sufficient or necessary to define 'word' in a cross-linguistic application. As such, the traditional distinction between morphology and syntax is called into question, and Haspelmath (2011: 72) concludes that "the part of (the study of) language structure that deals with sign combinations can be called morphosyntax, and for theoretical purposes this is currently best viewed as a unitary domain". While Haspelmath's conclusion resonates with us as field linguists and language analysts, we see the value in applying wordhood criteria to data from an individual language, in an attempt to produce a language-specific understanding that may be employed to inform the development of a written standard for the language. We acknowledge that not all of the ten criteria will be of relevance to Lamap data.

4.1.Non-Selectivity

The criterion of non-selectivity directs our attention to the combinatorial possibilities for person markers. Haspelmath (2011: 45, see also Haspelmath & Sims 2010: 198) makes the observation that, "while an affix tends to be highly selective with respect to the kinds of hosts it can combine with, (function) words are often able to combine with a wide range of hosts".

Evidence of the subject markers and their neighbouring elements was provided in examples (3) to (6), and (9) to (10) in section 3. Further examples are presented in (12a, b, c) for a number of paradigm members. While a full analysis of morphological categories in Lamap remains to be conducted, the examples indicate that in each case the person marker occurs before a verb.

(12)   a.   1SG person markers
no vän   1SG.REAL go   [psw_icr15.015]
no bao   1SG.REAL be.big   [psw_icr19.004]
no lümi   1SG.REAL plant something [psw_van04.002]

        b.   2SG person markers
ko män   2SG.REAL laugh   [psw_Yia3Rida.04.024]
ko rox   2SG.REAL live/stay/exist [psw_Yia2Rida.06.007]
ko lngonä   2SG.REAL hear/feel something [psw_Yia3Rida.09.047]

        c.   3SG person markers
e ov   3SG.REAL run   [psw_icr10.010]
e ddang   3SG.REAL be.strong   [psw_Yia3Rida.03.008]
e risä   3SG.REAL see something   [psw_Yia2Rida.01.007]

As observed in section 3, when a verbal clause has negative polarity, the negator occurs in a fixed position between the person marker and the verb. Examples in (13) show this pattern.

(13)   *no*   *sba*   *tadd*   1SG.REAL   NEG   sit/be/stay   [psw_Yia3Rida.05.007]
        e   *sba*   *jaxin*   3SG.REAL   NEG   catch (of fish) [psw_Yia2Rida.07.009]
        to   *sba*   *tas*   3PL.REAL   NEG   be.tired   [pws_Yia3Rida.05.012]

There is orthographic variation in the representation of person markers depending on whether the negator is present or not. When clauses are affirmative, teachers sometimes affix the person marker directly to the verb root. When negation is present, the person marker is more consistently affixed to its new neighbour, and this combined orthographic unit is written separately from the verb root. This orthographic pattern is demonstrated in (14).

(14)   a.   *Näim   pajipaj   ebao,*
              Näim   pajipaj   e-bao
              house   (sleep)   3SG.REAL-be.big
              'The sleeping house is big,'

        b.   *ä   näim   mokekë   esba   bao.*
              ä   näim   mo-kekë   e-sba   bao.
              but   house   ADJ-be.small 3SG.REAL-NEG be.big
              'but the bathroom is not big.'

c.   *Ekekë.*
       E-kekë.
       3SG.REAL-be.small
       'It is small.' (psw_mlp_y1-6a.03)

The representational contrast shown in (14) may derive from the dual function of the form *sba,* both as a verbal modifier and as a negative existential verb root meaning 'be not'.[10] Example (15) demonstrates the negative existential function. Note that in (15), the orthographic unit of the person marker followed by negative existential is separated by a space from the noun which functions as its object.

(15)   *Esba*   *brevuj.*
        E-sba   *brevuj.*
        3SG.REAL-be.not   banana
        'It is not a banana.' [psw_icr22.005]

The data reveal that subject person markers are highly selective of their hosts, occurring in the first position of the verb complex, either followed directly by the verb, or followed by the negator and then the verb. The selectivity of person markers points towards their analysis as affixes. In morphological texts, the corresponding property of non-selectivity, or freedom of host selection, is widely cited as a property of clitics as opposed to affixes (see e.g. Bickle & Nichols 2007: 174-175; Haspelmath 2011: 45; McGregor 2015: 64; and Payne 1997: 22). In contrast to the individual elements that comprise the verb complex, the verb complex as a whole is far less selective, combining with preceding and following entities from a variety of categories, or no entity at all. Example (16) shows the inflected verb in square brackets as a minimal complete clause, this time with an irrealis inflection.

(16)   *[Ni   kübax.]*
         1SG.IRR   jump
        'I will jump.' [psw_icr01.003]

Example (17) shows the verb complex preceded by a modified nominal subject, and followed by a nominal object.

(17)   *Xaritav   märäx   [to   ringa]   medde   sa   xate.*
        butterfly   female   3PL.REAL spread   egg   POSS   3PL
        'The female butterflies lay their eggs.' [psw_Yia2Rida01.008]

Example (18) shows the verb complex preceded by an adverbial morpheme *mijä* 'then', while (19) shows the verb complex followed by an adverbial morpheme *räbä* 'because'.

---

[10] Although *sba* can function as a lexical negative (a negative existential verb), there is no evidence that the combination of the negator *sba* followed by a lexical verb, illustrated in examples (13) and (14a), involves anything other than standard negation of a verb. The negative morpheme and the following verb do not carry separate person markers, and they form an uninterruptible unit.

(18)   *Mijä   [to   lavüx]   näran.*
         Then   3PL.REAL   drop   ground
         'Then they fall to the ground.' [psw_Yia2Rida03.012]

(19)   *"[Ko   män]   räbä   nisä?"*
         2SG.REAL   laugh   because   what
         "'Why are you laughing?'" [psw_Yia3Rida04.024]

The selectivity of person markers which consistently co-occur with verb roots, and with the negator when present, points towards their analysis as affixes, while the larger verb complex is non-selective, and is found with a variety of preceding and following material. The verb complex behaves as an independent morphosyntactic word, and as a whole unit displays the further property of Free Occurrence.

4.2. Free Occurrence

The free occurrence criterion is based on the understanding that a word is a "*minimum free form*" (Bloomfield 1935: 178) that can occur on its own. Haspelmath (2011: 40) claims that "if an element can occur independently (i.e. as a complete utterance all by itself), it clearly cannot be an affix but must be minimally a morphosyntactic word". An affix, by definition, cannot stand on its own. It has to be attached to another element.

In the data for Lamap, there are no instances of Paradigm II person markers "occurring as a well-formed complete (but possibly elliptical) utterance" (Haspelmath 2011: 39). Paradigm II person markers always precede a verb. When the negator is present, the person markers precede the negator and the verb in that order. This means that there is no free occurrence of person markers, of negators, or of verb roots. Free occurrence is only possible when these morphemes combine to form the verb complex, as illustrated in section 4.1, example (16).

4.3. Non-coordinatability

Another criterion that is relevant to our Lamap data is that of non-coordinatability. According to the criteria of non-coordinatability, only words can be ellipted, having a wide scope of coordination. If an element cannot undergo coordination ellipsis but has to occur in each coordinated clause, it may be considered an affix (Haspelmath 2011: 47-48).

Evidence presented in (20), (21) and (22) shows that person markers behave differently from subject nominals under coordination. In example (20a) the subject noun phrase *rüare* 'child(ren)' is expressed before the verb. The person and number properties of the subject are indexed by the person marker *to* '3PL', immediately afterwards. The plurality of the subject entity 'children' is only coded by the subject marker. In clauses (20b) and (20c), the subject nominal *rüare* 'child(ren)' is ellipted, but the subject marker *to* '3PL' is repeated with each verb.

(20)   a.   *Rüare   [to   jibü]   ä   rieväj*
            child   3PL.REAL   get-out   LOC.of/from   truck
            'The children got off the truck,'

       b.   *[to   juäni]*
            3PL.REAL   push
            '(and) they pushed,'

       c.   *[to   juäni...]*
            3PL.REAL   push
            '(and) they pushed…' [psw_icr10.004]

The same pattern is seen in (21), where *nürükäv älinä* 'ants' occurs only in (21a), while in (21b) it is ellipted. The subject marker *to* '3PL' occurs with each subsequent verb.

(21)   a.   *Mijä,   nürükäv älinä   [to   vän]   penaxerä   märu   jika,*
            Then   ant   3PL.REAL   go   LOC.under   coconut   QUANT.one
            'Then the ants went under a coconut,'

       b.   *mononginong   [to   vän]   krövä   nabü   evis.*
            then/following 3PL.REAL go   LOC.over   bamboo   QUANT.some
            'and after that they went over some bamboo.' [psw_icr18.004]

Example (22) repeats the pattern, but in this case, the subject nominal is expressed in clause (22a) as a Paradigm I person marker, modified by a quantifier *xate piji* 'they all'. This subject is indexed on the verb with *to,* and only *to* occurs in the second clause (22b). The subject pronominal is ellipted, but the person marker remains with the verb.

(22)   a.   *Xate   piji   [to   vänä-karäv]   sur   naür*
            3PL   all/every 3PL.REAL   go-look   around   place
            'They all went looking around the place,'

       b.   *rä   [to   bärinä]   betlivär   mo   paj*
            CONJ   3PL.REAL   find   big.stone   ADJ   be.lying
            'and they found a big stone that was lying down...' [psw_Yia3Rida06.010.]

Haspelmath (2011: 48-49) acknowledges that in some languages, affixal-type elements can be ellipted under coordination, weakening the criterion in cross-linguistic comparison. In Lamap however, ellipsis of person markers is prohibited. The criterion thus proves helpful in establishing the status of the Lamap person markers as affixes.

4.4.Morphophonological rules and idiosyncrasies

Morphophonological rules are considered to be a predictor of affix-stem combinations, with Haspelmath and Sims (2010: 198) observing that, "morphophonological rules are less likely to operate across the boundary between a host and a clitic than across the boundary between a stem and an affix". A morphophonological rule of particular interest in Lamap involves vowel harmony. Crowley (2002: 655), in his summary of Charpentier's work on Lamap, observed that Lamap person markers display variation, where "the variation in the shape of some prefixes involves partly vowel harmony and partly free variation". Crowley (2002: 655) went on to say that, "details of the conditioning factors for this allomorphic variation are not clear". If patterns of vowel harmony between the subject marker, (negator), and verb root could be established in Lamap, we could have evidence of morphophonological rules that support the analysis of person markers as affixes.

One point of vowel variation concerns the third person singular marker, which is reported by Crowley as *e* and *i*. In our data it also appears as *bi*. It is plausible that forms with the high front vowel *i* are triggered by roots with a high front vowel; however, example (23a) shows *e* with the verb *kis* 'be nice'. The person marker and verb root also remain stable under negation, as shown in (23b).

(23)   a.   *Näxöm*   *e*   *kis*   *soxor.*
                face.2SG.POSS   3SG.REAL   be.nice   very
                'Your face is very nice.' [Yia3Rida08.081]

        b.   *E*   *sba*   *kis*   *rai.*
                3SG.REAL   NEG   be.nice   at.all
                'It wasn't nice at all.' [Yia3Rida01.003]

Data in the Lamap Reader Corpus indicate that the vowels in person forms do indeed vary; however, the variation cannot be explained satisfactorily as involving vowel harmony. Examples (24a) and (24b) show two forms of the person marker for the first person plural inclusive 'we'. In (24a), the question is formed in response to a picture, where the answer is clearly visible and the action is in progress. In (24b), the question is formed to inquire about a future activity.

(24)   a.   *Ddäte*   *ddäto*   *maxa*   *nisä?*
                1PL.INCL   1PL.INCL.REAL   do/make   what
                'What are we doing (in this picture)? [psw_icr16.014]

        b.   *Lönixa*   *ddäti*   *maxa*   *nisa*   *va-miji?*
                now   1PL.INCL.IRR   do/make   what   immediate
                'Now what are we going to do? [Yia3Rida01.036]

The variation in person markers appears to represent two distinct person marking paradigms. Paradigm II person markers occur in realis mood contexts, while Paradigm III person markers occur in irrealis mood contexts. Paradigm III is as yet incomplete, although the forms that are attested in the Lamap Reader Corpus are shown in Table 3.

	Singular		Dual	Plural
1	*ni*	INCL	*ddi*	*ddati*
		EXCL	---	*nati*
2	*ki*		---	---
3	*(b)i*		---	*(xa)ti*

Table 3. Paradigm III Person Markers in Lamap[11]

Examples (25a) and (25b) show two forms of the person marker for the first person singular associated with verbs of the shape [Ca]. The larger structure involves a desiderative verb, followed by its sentential complement. The desiderative verb takes the realis *o*-marker, while the verb in the sentential complement takes an *i*-marker.[12]

(25)　a.　*Vavü　　　no　　　ka...*
　　　　　grandparent　1SG.REAL　say
　　　　　'Grandparent, I want...'

　　　b.　*ni　　　　van　　　vasav*
　　　　　1SG.IRR　　go　　　　dance
　　　　　'to go dance.' [Yia3Rida05.006]

The vowel contrast shown in (25) is repeated in (26) with the verb form *vänä* 'go', combined with the first person plural exclusive person marker *nato* for realis in (26a) and *nati* for irrealis in (26b).

(26)　a.　*Nabong　jika　nato　　　　vänä　rieväj*
　　　　　time/day　one　1PL.EXCL.REAL　go　　vehicle
　　　　　'One day, we went on the truck...' [psw_icr10.002]

　　　b.　*Nato　　　　ka　　nati　　　　vänä　jajax.*
　　　　　1PL.EXCL.REAL　say　1PL.EXCL.IRR　go　　fishing
　　　　　'We decided to go fishing (when the tide went out later in the day).' [Yia3Rida01.003]

---

[11] As an anonymous reviewer has pointed out, it would be ideal to present a complete irrealis paradigm. We agree, and this is a goal of future fieldwork; however, at this point we wanted to include the preliminary evidence for the paradigm as we were able, rather than limit the paper to a discussion of Paradigm II realis mood forms which are more fully attested in the Lamap Reader Corpus.

[12] The vowel-based mood distinction that we have identified in Lamap corresponds with a vowel-based mood distinction reported in the Uluveu language of the Maskelynes Islands. Like Lamap, Uluveu is in the same Southeastern subgroup of the Eastern Malekula Linkage (Lynch 2016: 418), and it is spoken immediately to the south of Lamap. Healey (2013: 188-189) analyses Uluveu as encoding realis mood with the morphophoneme *U,* and irrealis mood with the morphophoneme *I.* Both morphophonemes undergo dissimilation processes in specified contexts. Lamap's Paradigm II *o*-form realis person markers, and hypothesised Paradigm III *i*-form irrealis person markers follow the basic pattern identified by Healey, although it is too early to speculate about paradigm-internal harmony or dissimilation processes in Lamap. The vowel-based mood distinction may transpire to be an important shared innovation of the Southeastern subgroup of the Eastern Malekula Linkage. Work in progress on the Axamb language by University of Waikato doctoral student Tihomir Rangelov, under the supervision of Julie Barbour, will contribute to our understanding of this innovation. Axamb is also classified as a Southeastern Malekula language.

In investigating the occurrence of morphophonological rules in Lamap, we expected to find evidence of vowel harmony patterns across the morpheme boundary between person marker and verb root. We did not find such patterns. Rather than being phonologically conditioned, the variation in person markers appears to be grammatically conditioned.

Morphophonological idiosyncrasies are proposed to be more indicative of affix-root combinations than clitic and host combinations (Haspelmath 2011: 52); however, there is no evidence of idiosyncratic combinations, or meanings of combinations, in the Lamap data thus far. The forms, combinatorial possibilities, and resultant meanings are highly predictable.

4.5. Summary

We have applied a selection of Haspelmath's (2011) criteria for defining the morphosyntactic word to data from the Lamap language. Not all criteria could be tested in this study due to the nature and quantity of available data, and just four criteria have been investigated in detail. Paradigm II (and III) person markers are highly selective, co-occurring only with a verb root or negator followed by verb root in negative polarity contexts (section 4.1). They do not display free occurrence in the corpus, but are required to occur with their associated verb (section 4.2). They do not demonstrate ellipsis under coordination, but again are required to occur with their associated verbs in a multi-clause structure (section 4.3). Following Haspelmath's description of these three criteria, the patterning of the Lamap person markers points towards their analysis as affixes.

Further evidence for the analysis of person markers as affixes would have come from the identification of morphophonological rules and idiosyncrasies (section 4.4). In the search for such patterns, we uncovered a system of mood marking. Vowel choices appear to reflect the grammatical properties of the situation expressed in a clause, rather than the phonological properties of the verb root, or negator where present. The discovery of this system is fundamental to our understanding of the Lamap language. It fails, however, to furnish us with evidence of the phonological integration of the person markers, negation, and verb root, which might further inform our representational recommendations. We acknowledge that evidence of phonological integration may emerge in our future research.

5. Applying the results of linguistic research

Our search for a linguistic response to the issue of writing person markers in Lamap has greatly enhanced our understanding of the language. The Lamap Reader Corpus provides evidence that points towards the linguistic analysis of person markers as affixes, and it is thus tempting for us to propose a contiguous representation of person markers and verb stems on the basis of Haspelmath's (2011) relevant wordhood criteria. It is conventional in English and French to write many affixes as contiguous to their associated stems. It does not follow, however, that Lamap writers must represent person markers as affixes. The lack of evidence to date for phonological integration between affix and stem makes the representation of person markers as free form particles a plausible alternative, and one that is already being used by writers. We therefore recognise that there is a genuine representational decision to be made by writers of Lamap.

Lamap community members have demonstrated that they have preferences in how their language is represented, and we recall their instructions to us to represent the phonemes pronounced as [ndr] and [ʧ] as 'dd' and 'j' respectively (as described in section 1.3). We can draw their attention to the representation of person markers, and encourage writers to consider the options available to them, with a view towards selecting one form or another as the written standard.

An important long-term goal of a community literacy programme such as that unfolding in the Lamap community is to foster a transfer of knowledge and skill from the linguist to the community. As Mary Raymond (2007: 194) comments, "the ultimate goal of any literacy programme... is that the programme will eventually gain sufficient momentum in the hands of the community that it can continue long term without outside assistance". Linguistic research can certainly support literacy work, but ultimately, the community of writers must take ownership of the written form of their language. We look forward to sharing more of our research findings with the Lamap community, and to working alongside them as they make decisions about the written representation of their language.

Appendix I: Yumi Raetem Lamap - Bislama Samari

Gayleen Tarosa, PhD Candidate, University blong Waikato, hemi bin mekem translesen

Long yia 2015 mo 2016, Dr. Julie Barbour wetem tim blong Yuniveseti blong Waikato, New Zealand, oli bin go long Malekula Aelan long Vanuatu blong tekempat long MOET (Ministri blong Edukesen mo Trening) woksop long saed blong translesen. Olgeta evriwan sikisfala, mo olgeta oli mekem stadi long saed blong lingwistik o lanwis.

Tim blong Waikato ia i bin givhan long 7 lanwis komuniti blong transletem fulap smol buk blong Yia 1, Yia 2 mo Yia 3. Wan lanwis ia, hemi *Lamap*, blong Port Sandwich, Malekula, mo Claudia Williams blong Waikato, hemi bin givhan long lanwis ia.
Wan samting we i had wok lelebet, hemi, "Hao nao blong raetem wod – yumi raetem wan wan, or yumi joinem tugeta?"

Blong givhan long Lamap tija, mitufala blong Waikato (Claudia mo Julie), mitufala lukluk gud long ol smol pat blong lanwis. Mifala i faenemaot sam samting:

(1) Ol fom olsem *no* 'mi', *ko* 'yu', *e* 'hem', oli save stap klosap long aksen wod ('verb' o 'la verbe').

        Inao no-xan   'Mi kakae.'

(2) Ol fom olsem *no* 'mi', *ko* 'yu', *e* 'hem', oli save stap klosap wetem *sba* 'no'.

        Inao no-sba-bao.   'Mi no bigfala.'

(3) Ol fom olsem *no* 'mi', *ko* 'yu', *e* 'hem', oli no save stap hem wan long lanwis. Spos yu wantem talem 'mi' o 'yu', yu no save talem *no* or *ko*. Yu mas talem wan nara samting:

        Inao 'mi'   *Xaing* 'yu'

(4) Mo evri taem i gat wan aksen wod we i stap festaem lo lanwis, i mas gat wan fom olsem *no* 'mi', *ko* 'yu', *e* 'hem' we i stap wetem. Yu no save talem *rox* 'stap' nomo. Yu mas talem olsem:

        No-rox. 'Mi stap.'   *Ko-rox.* 'Yu stap.'

Nao ia, kwesten we mifala i askem, i go olsem: Hao nao blong raetem ol samting ia? Olsem taem yumi raetem "Mi (no) stap long solwota", yumi raetem olsem (a) o olsem (b):

        (a) ***No rox** lä näräs.*   ***No sba rox** lä näräs.*
        (b) ***Norox** lä näräs.*   ***Nosbarox** lä näräs.*   ***Nosba rox** lä näräs.*

Blong mekem lanwis i no hevi tumas blong pikinini, i gud yufala jusemaot wan rod blong raetem. Disesen ia, i go long yufala long Lamap komuniti, mo bae mifala implementem disesen blong yufala.

References

Barbour, Julie. 2012. *A grammar of Neverver*. Berlin: De Gruyter Mouton.

Barbour, Julie. 2015. Jespersen Cycles on Malekula. *Linguistic Typology 19*(3). 425-462.

Bickle, Balthasar & Johanna Nichols. 2007. Inflectional morphology. In Shopen, Timothy (ed.), 169-240. *Language typology and syntactic description. Volume III: Grammatical categories and the lexicon.* 2nd edn.

Bloomfield, Leonard. 1935. *Language*. London [England]: Allen & Unwin.

Bowern, Claire. 2008. *Linguistic fieldwork: a practical guide.* Basingstoke England; New York: Palgrave Macmillan.

Charpentier, Jean-Michel. 1979. *La Langue de Port-Sandwich (Nouvelles Hébrides : introduction phonologique et grammaire*. Langues et Civilisations à Tradition Orale 34. Paris: Société d'Études Linguistiques et Anthropologiques de France (SELAF).

Charpentier, Jean-Michel. 1982. *Atlas linguistique du Sud-Malakula – Lingusitc atlas of South Malekula (Vanuatu)*, vol.2 : Langues et cultures du Pacifique 2. Paris : Centre National de la Recherche Scientifique/L'Agence de Coopération Culturelle et Technique.

Charpentier, Jean-Michel. 1995. Port Sandwich. In Tryon, Darrell T., and Shigeru Tsuchida (eds.), 829-836. *Comparative Austronesian Dictionary : An Introduction to Austronesian Studies*. Trends in Linguistics. Documentation; 10. Berlin ; New York: Mouton De Gruyter.

Charpentier, Jean-Michel. 1997. Literacy in a Pidgin Vernacular. In Tabouret-Keller, Andree Robert B. Le Page, Penelope Gardner-Chloros, and Gabrielle Varro (eds.). *Vernacular Literacy: A re-evaluation.* Oxford; New York: Clarendon Press.

Crowley, Terry. 2002. Port Sandwich. In Lynch, John, Malcolm Ross, and Terry Crowley (eds). *The Oceanic languages*. Richmond, England: Curzon. 650-659.

Gil, David. 2001. Escaping Eurocentrism: fieldwork as a process of unlearning. In Paul Newman & Martha Ratliff (eds.), 102-132. *Linguistic Fieldwork.* Cambridge, New York: Cambridge University Press.

Haspelmath, Martin. 2011. The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica* 45 (1): 31-80.

Haspelmath, Martin & Andrea Sims. 2010. *Understanding Morphology, 2nd edn.* London: Hodder Education.

Healey, David. 2013. *A grammar of Maskelynes: the language of Uluveu Island, Vanuatu.* Port Vila, Vanuatu: University of the South Pacific doctoral dissertation.

Lüpke, Friederike. 2010. Research methods in language documentation. *Language Documentation and Description 7.* 55-104.

Lynch, John. 1998. *Pacific languages: an introduction*. Honolulu: University of Hawai'i Press.

Lynch, John. 2016. Malakula internal subgrouping: phonological evidence. *Oceanic Linguistics* 55(2). 399-431.

Lynch, John & Terry Crowley. 2001. *Languages of Vanuatu: A new survey and bibliography.* Canberra: Pacific Linguistics, Research School of Pacific and Asian Studies, Australian National University.

Lynch, John, Malcolm Ross, & Terry Crowley. 2002. *The Oceanic languages*. Richmond [England]: Curzon.

McGregor, William B. 2015. *Linguistics: An introduction.* 2nd edn. London, New York: Bloomsbury.

McKerras, Ross. 2001. *Uripiv*. Unpublished manuscript. Summer Institute of Linguistics, Port Vila, Vanuatu.

Pacific Manuscripts Bureau. 2015. *Pacific Manuscripts Bureau titles documenting Vanuatu/New Hebrides (Church Records): AU PMB MS53.* Canberra: Australian National University. Retrieved from: http://asiapacific.anu.edu.au/pambu/catalogue/index.php/papers-4.

Payne, Thomas E. 1997. Describing morphosyntax: a guide for field linguists. Cambridge, New York: Cambridge University Press.

Raymond, Mary. 2007. Literacy work in Papua New Guinea: the accidental and the planned. *Language Documentation and Description, Volume 4.* 174-194.

Republic of Vanuatu. 2006 [1980]. *Constitution of the Republic of Vanuatu: Laws of the Republic of Vanuatu Consolidated Edition 2006.* Retrieved from: http://www.paclii.org/vu/legis/consol_act/cotrov406/.

Ross, Malcolm. 2004. The morphosyntactic typology of Oceanic languages. *Language and Linguistics 5*(2). 491-541.

Samarin, William J. 1967. *Field Linguistics; a Guide to Linguistic Field Work*. New York: Holt, Rinehart and Winston,

Simons, Gary F. & Charles D. Fennig. 2017. Ethnologue: Languages of Vanuatu. In Simons, Gary F. & Charles D. Fennig (eds.). *Ethnologue: Languages of the World,* 20th edn. Dallas, Texas: SIL International.

Vanuatu Ministry of Education. 2012. *Vanuatu National Language Policy*. Port Vila: Vanuatu Ministry of Education & Training. Retrieved from: https://education.gov.vu/docs/policies/Vanuatu%20National%20Language%20Policy%20(English)_2012.pdf.

Vanuatu National Statistics Office. 2009. *2009 National population and housing census: Basic tables report Volume I.* Port Vila: Vanuatu National Statistics Office.

Wessels, Kanauhea. 2013. *Malua Bay: A description of the Malua Bay language (Malekula, Vanuatu).* Hamilton, New Zealand: University of Waikato Master's thesis.