Special Issue 2012

Harald Hammarström & Wilco van den Heuvel (eds.)
History, contact and classification of Papuan languages

Part One

The historical relation of the Papuan languages of Timor and Kisar

Antoinette Schapper [a], Juliette Huber [b] & Aone van Engelenhoven [a]
[a] LEIDEN UNIVERSITY & [b] LUND UNIVERSITY

The Papuan languages on the islands of Timor, Alor, Pantar and Kisar have typically been supposed to be related to one another. Based on their geographical proximity, this assumption of relatedness has been supported by impressionistic assessments of lexical and phonological similarity and the results of lexicostatistical measurements. Whilst recent historical work has shown definitively that the languages of Alor and Pantar are related to one another (Holton et al. 2012), the relation of the Papuan languages on Timor and Kisar has not yet been established with the proper methodological rigour. In this paper we present the results of a study applying the comparative method to the Timor and Kisar languages. We demonstrate the inter-relatedness of the Timor-Kisar languages through the identification of abundant cognate sets with regular sound correspondences between languages.

1.   Introduction[1]

Due to their proximity, the Papuan languages on the islands of Timor, Alor, Pantar and Kisar are typically seen to be related to one another (see section 2.2). The assumptions of geography have been reinforced by impressionistic assessments of lexical and phonological similarity and by the results of lexicostatistical measurements, but have for the most part not been tested through rigorous application of the comparative method. Whilst recent work has shown the Papuan languages of Alor and Pantar to be related to one another (Holton et al. 2012), the inter-relation of the Papuan languages on Timor and Kisar (henceforth, simply 'Timor-Kisar languages') has not yet been demonstrated. In this paper we investigate the historical relations of the Timor-Kisar languages. Our aims are two-fold: (i) to establish that the Timor-Kisar languages are related to one another, and (ii) to identify the changes defining sub-groups amongst the Timor-Kisar languages. We show that the regular sound correspondences which can be established on the basis of

form-meaning pairings support seeing the Timor-Kisar languages as inter-related. We further identify sound and lexical changes that define low-level sub-groupings amongst the Timor-Kisar languages.

Timor-Kisar languages are typically discussed together with the Alor-Pantar languages, and indeed it is almost certain that they are related (cf. section 7). Nevertheless, in this paper, we will limit ourselves to the consideration of the Timor-Kisar languages exclusively. Whilst the integrity of the Alor-Pantar languages as a genetic group has been established, the reconstruction of Proto-Alor-Pantar (PAP) is only preliminary, with fewer than one hundred proto-lexemes reconstructed. Given the smallness of the proto-vocabulary available at this time, the possibility of finding cognates in Timor-Kisar languages is limited. It is for the moment thus more useful to establish the existence of the relationship between the Timor-Kisar languages, for which the authors have their own detailed lexical materials. Once this is established and the reconstruction of the PAP lexicon further advanced, it will be possible to meaningfully proceed with determining the inter-relationship of the two geographical groups of Papuan languages in the region.

The paper is structured as follows. In section 2 we overview the history of descriptive work on and historical classification of the Timor-Kisar languages. In section 3 we outline the problem of relexification in Timor-Kisar languages due to contact with Austronesian languages. In section 4, we turn to the examination of form-meaning pairings in cognate sets to establish regular sound correspondences that attest to a genetic relationship between the languages. Section 5 looks briefly at issues complicating the reconstruction, namely word-edge erosion and morpheme fossilisation. In section 6 we present arguments for subgroups based on shared innovations in sounds and the lexicon. An appendix is included with all the supporting data from each of the five languages for easy reference.

2.   Previous work on the Timor-Kisar languages

2.1.   Descriptive work

There are four Papuan languages on Timor. Bunaq ([bfn]) is spoken by some 80,000 people straddling the border between Indonesian West Timor and independent East Timor. Makasae ([mkz]), with about 70,000 speakers, Makalero (no ISO 639-3 code), spoken by about 6,000, and Fataluku ([ddg]), with a population of about 30,000, occupy a contiguous region on the eastern tip of the island. Finally, Oirata ([oia]), is a Papuan language spoken by some 1,200 people in the southern half of Kisar Island (Lewis 2009). See Map 1.

Map 1: The Papuan languages of Timor and Kisar †
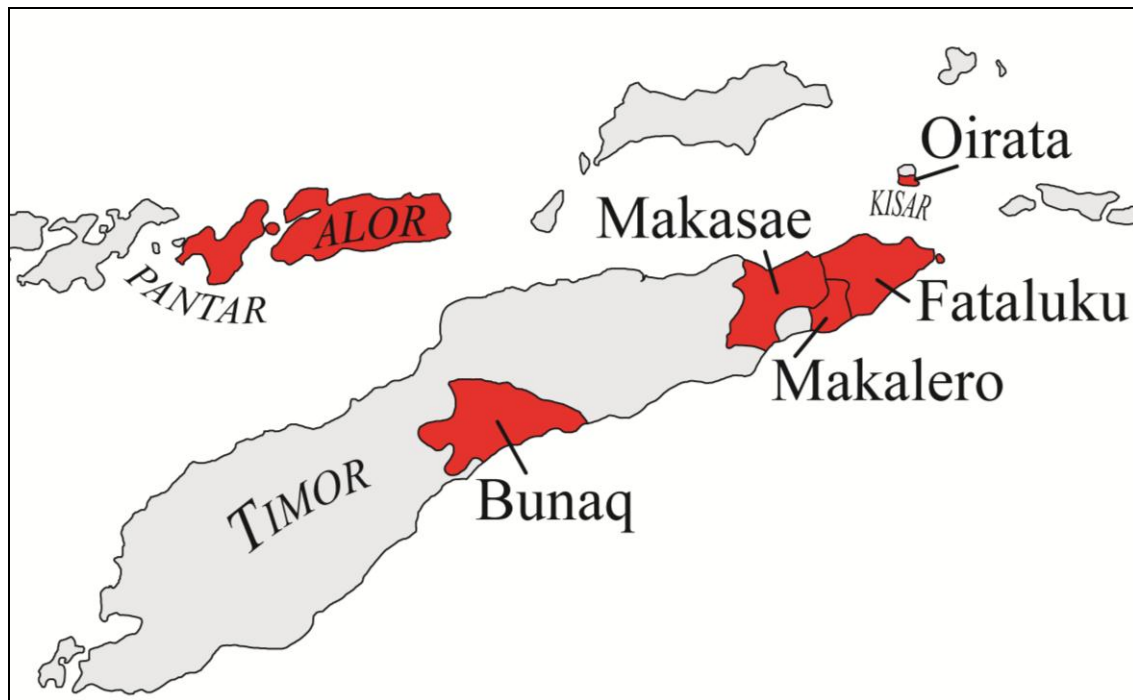
† Dark shading marks areas where Papuan languages are found; light shading where Austronesian languages are found. Only Timor-Kisar languages are marked on the map by name.

The Timor-Kisar languages are among the western-most Papuan languages, occupying a small corner in a large region which is dominated by Austronesian languages. The presence of Papuan languages at the eastern end of the Minor Sundic Island chain, where Timor and Kisar are located, was first brought to light by Josselin de Jong's (1937)

description of the Oirata language on Kisar. Most subsequent treatments and discussions of Oirata (Cowan 1965, Donohue and Brown 1999, and Faust 2005) have been based on Josselin de Jong's monumental monograph. According to Dutch historical records, the Oirata arrived in Kisar in 1721, having fled from Loikera in the Fataluku-speaking area (Riedel 1886: 403). Still today the northern Fataluku dialect is claimed to be, at least partly, mutually intelligible with Oirata (Katrina Langford pers. comm.), though other Fataluku dialects appear to be too divergent.

Capell (1943a, 1943b, 1944, 1972) identified four further Papuan languages on Timor, Bunaq and Makasae, for which he provides grammar sketches, and Fataluku and Makuva/Lovaia[2]. Further details of Bunaq were revealed in the pioneering work of Berthe (some early results were published as Berthe 1959, 1963). Berthe returned to Timor in the mid-1960s at the head of a French-Portuguese Ethnological Expedition. Out of this fieldwork came more linguistic materials on Bunaq (notably Berthe 1972, 1978 and Friedberg 1990) and further materials on Fataluku (Campagnolo 1973).

Between 1953 and 1975, António de Almeida, head of the Portuguese Missão Antropológica de Timor, collected word lists and elicited sentences on Bunaq, Makalero, Makasae, and Fataluku.[3] Today these materials are held at the Instituto de Investigação Científica Tropical (IICT) in Lisbon and are partly published in Almeida (1994). Between 1955 and 1968, Alfonso Nacher, a priest at the Missão Salesiana in Fuiloro, compiled a sizeable Fataluku dictionary which was later published as Nacher (2003, 2004).

A long hiatus in research ensued during the era of Indonesian occupation of East Timor from 1975 till 1999.[4] Towards the end of this quarter century, there appeared a Makasae language course in Portuguese (Marques 1990), and some error-ridden word lists and sketches from the Indonesian *Pusat Pembinaan dan Pengembangan Bahasa* (Sawardo et al. 1996, Sudiartha et al. 1994, and Sudiartha et al. 1998). It was not until the beginning

of the 21st century that new materials on the Papuan languages of Timor really began to emerge. Initially, basic materials on a number of the languages appeared in Hajek and Tilman (2001) and Hull (2004b), but more comprehensive documentation followed soon after for all languages.

Makasae garnered attention from several corners. Following a field methods course on the Ossu dialect, two descriptive theses on specific language domains came out of the Australian National University, one on Makasae deixis and space (Brotherson 2003) and one on Makasae speech registers (Carr 2004). Around 2005, two dialects of Makasae were studied at the Language Documentation Training Centre at the University of Hawaii, resulting in mini-sketches, word lists and some text material appearing online.[5] The language has also been the subject of a detailed phonological study (Fogaça 2011), a short dictionary (Hull and Correira 2006), and two grammar sketches, one on the Baucau dialect (Hull 2005a) and another on the Ossu dialect (Huber 2005 later published as Huber 2008a).

Makalero has also been astoundingly well-documented for a language which has only been recognised since the independence of East Timor (see, e.g., Hull 2004a). Academic documentation comprises a reference grammar (Huber 2011) and a short paper on agreement morphology (Huber 2008b). A monolingual Makalero dictionary (Pinto 2004) and a Makalero language course with Tetum and English translations (Pinto 2007) have also appeared with the support of the Timor Loro Sa'e Nippon Culture Center.

Fataluku is the subject of extensive and ongoing development and documentation driven by an engaged community of native speakers (cf. Langford 2011). Local materials produced for and by the Fataluku community include a monolingual dictionary (Valentim 2002) and several school readers (e.g., Langford 2012). A grammar sketch came out of the National Institute of Linguistics of the University of East Timor (Hull 2005b). Between 2005 and 2008, an academic documentation project on Fataluku funded by the

Endangered Languages Programme of *De Nederlandse Organisatie voor Wetenschappelijk Onderzoek* was run out of the University of Leiden. This project resulted in several descriptive papers on the central dialect (e.g., Stoel 2007a, 2007b, Engelenhoven 2009, 2010a, 2010b). Some documentation of Fataluku has also been conducted at the University of Hawaii's Language Documentation Training Centre.[6]

Finally, Bunaq as spoken in the west Timorese area of Lamaknen has been treated in a reference grammar (Schapper 2010) and a suite of academic papers (e.g., Klamer and Schapper forthcoming, Schapper and San Roque 2011, Schapper 2011a, 2011b, 2011c). However, much descriptive work remains to be done on the many, often highly divergent, dialects of Bunaq spoken in East Timor. Currently, a range of school readers are being produced to support Bunaq literacy (e.g., Fundasaun Alola 2011), but there is as yet no program to describe the dialects in East Timor.

In short, the last decade has seen dramatic advances in our knowledge and understanding of the character and composition of the Papuan languages on Timor. While much work remains to be done in terms of documentation of all speech domains and dialects, we are now much better placed to investigate the historical relations of the languages. The data used in this study have been collected from a variety of sources. Bunaq data are from the Lamaknen dialect and are drawn from a lexical corpus of approximately 3000 items collected by Schapper on the basis of firsthand fieldwork. Makasae data come from a range of dialects treated in the following sources: Huber's own field notes, Brotherson (2003), Carr (2004), and the Makasae materials at the Language Documentation Training Centre. Makalero data used here were collected by Huber in the area of Iliomar and comprise a corpus of approximately 2000 lexical items. Fataluku data are drawn from fieldwork notes by Engelenhoven and from the two Fataluku dictionaries (Nacher 2003, 2004, and Valentim 2002). Finally, the data for Oirata comes from Josselin de Jong (1937).

2.2.   Historical work

Despite the scarcity of materials, early discussions of the Timor-Kisar languages were dominated by speculation about their historical relations.

The conjecture began with Cowan (1963), who argued that Bunaq and the other Timor-Kisar languages were related to the languages of the Bird's Head of New Guinea. Similarly, Capell (1975) suggested that the Timor-Kisar languages should be grouped with the West Papuan phylum together with the Papuan languages of the Bird's Head and Bomberai peninsula. This classification was mainly based on typological similarities; the absence of identifiable lexical correspondences, however, forced Capell to propose a major split between the "Alor-Timor languages" and the rest of the West Papuan Phylum. Capell's article is marked with a hasty note from the editor, Stephen Wurm, pointing the reader to Voorhoeve's (1975) contribution in the same volume. This argued that the languages of the southern Bird's Head and Bomberai peninsula belonged to the Trans-New-Guinea (TNG) phylum, with the implication being that all Papuan languages in the Timor area must also be TNG. This was made explicit by Wurm, Voorhoeve and McElhanon (1975), albeit with the caveat that 'whichever way they are classified, they contain strong substratum elements of the other … phyla involved' (1975: 318). This cautionary note is repeated by Pawley (2005: 73).

In recent years, several authors have supported a link with Bomberai languages. Hull (2004b) suggested a genetic relationship between the Timor-Alor-Pantar (TAP) languages and the languages of the Bomberai peninsula on the basis of some rather tenuous lexical similarities. Ross (2005) used a comparison of independent pronoun forms as evidence to propose a classification of the Timor-Kisar languages with Bomberai languages, making them part of a large 'western linkage' within TNG.

While the high-level classification of the TAP languages has been the subject of much speculation and hypothesis, their internal affiliations are still underexplored. Stokhof

(1975: 12-16, 22-24) tentatively made a case for a genetic relationship between AP languages and the Timor-Kisar languages. Furthermore, based on lexicostatistical evidence, he argued for a closer affinity between Makasae and the Alor-Pantar group (1975: 24). Ross (2005), based on the aforementioned pronominal evidence, contradicts this by grouping Bunaq with Alor-Pantar, as opposed to an East Timor group. On the basis of morphological evidence, Schapper (2010) suggests that the Papuan languages of Timor and Kisar may be more closely related to one another than to the Alor-Pantar languages (Schapper 2010: 21, 346). Mandala (2010), expanding on initial work presented in Mandala (2003), examines specifically the relationship between Oirata and the Eastern Timor languages (Fataluku and Makasae, to the exclusion of Makalero). Based on lexical data gathered by himself in Kisar and Timor, the author argues for a Fataluku-Oirata subgroup and a Makasae subgroup in Eastern Timor and reconstructs 172 proto-forms (in which, however, are included numerous loan words). These findings are summarized in English in Mandala et al. (2011).

This paper differs from the research described above in that it focuses on establishing relatedness through the application of the comparative method on all of the Timor-Kisar languages, that is, a close comparison of cognate lexical forms to establish regular sound correspondences. It thus ties in with Holton et al. (2012), which presents the results of the same method applied to the Alor-Pantar languages.

3.   Note on the socio-linguistic context in Timor

As in much of the Philippines, western Indonesia and the Pacific, languages of the Austronesian family dominate the east Nusantara region, where Timor is located (Adelaar 2005). The Alor archipelago (Alor, Pantar, and the intervening islands of the Pantar strait, Pura, Ternate, and Tereweng), however, is almost exclusively home to Papuan languages.[7] Such linguistically homogenous patterns are starkly contrasted by the picture presented by the island of Timor, which is home to multiple Papuan languages as well as numerous Austronesian languages. Speakers of these Papuan languages live in close

contact with speakers of Austronesian languages, and they show widespread borrowing and adaptation from Austronesian language and society.

McWilliam (2007) points out that Fataluku exhibits so much which is typical of Austronesian cultural categories and forms that, whilst the language is grammatically "Papuan", Fataluku communities have more in common with their Austronesian speaking regional neighbours than they have differences. He notes the extensive presence of lexemes with clearly Austronesian etymologies in key areas of Fataluku social life, such that they must be considered Austronesians merely "in linguistic disguise". Most of the Austronesian features identified by McWilliam in the Fataluku language and culture are also found in Makalero and Makasae. Furthermore, based on common taboos, Makalero clans or houses readily identify themselves as being the same as clans belonging to other linguistic groups of East Timor, suggesting that the cultural frameworks of Papuan-speaking and Austronesian-speaking communities are essentially analogous. There is also speculation in the literature that the Fataluku, Makasae and Makalero have expanded, taking over earlier Austronesian-held lands (O'Connor and Oliveira 2007). This has clearly happened in the case of the moribund Austronesian language Makuva, which has been absorbed by the expansion of Fataluku (Engelenhoven 2010c). In this light, it is perhaps then not surprising to note that Austronesian loan words permeate even the core vocabulary of these languages.

Berthe (1963) was similarly impressed by the sheer quantity of Austronesianisms present in Bunaq, so that he described the language as being of 'mixed' Papuan-Austronesian stock. Schapper (2010: 22-24, 2011b, 2011c) sets out the various strata of Austronesian influence that pervade every aspect of the Bunaq language. In particular, she notes that lexemes with traceable Austronesian sources dominate in the domains of society and governance, of male kinship and marriage, and of material culture and trade.

The extensive Austronesian influence found in Timor-Kisar languages is hardly surprising given their long shared history on Timor. However, it does create problems for determining the genetic relatedness of the Timor-Kisar languages in that the search for cognates is significantly impeded by relexification from surrounding Austronesian languages. Nevertheless, in this preliminary investigation we do still find enough cognate forms showing predominantly regular sound correspondences to establish the relatedness of the languages. There is, of course, the issue of borrowed lexeme identification. Where clear Austronesian etymologies are not established, there is the chance that parallel Austronesian borrowings into Timor-Kisar languages be falsely deemed cognate (as, for instance, occurs in Naerssen 2007). As such, we have taken great care to ensure that the cognates we identify are not also present in the neighbouring Austronesian languages, as far as the materials we have available allow.

4. Sound correspondences

In this section, we describe the consonant correspondences that we have identified in the Timor-Kisar languages. We do draw on vowel correspondences where they condition particular sound changes in consonants, but otherwise we do not deal with vowels in this preliminary reconstruction. In many cases, sound correspondences depend on the environment in which the consonant occurs. Thus, we provide examples of the correspondences in all environments (i.e., word-initial, word-medial, and word-final position) where possible.

In Table 1, we provide an overview of the consonant correspondences we observe. In addition, we give the reconstructed Proto-Timor (PTIM) phoneme for each correspondence set.[8] In the "Environment" column, we indicate whether the correspondence applies in initial (#__), medial (V__V), or final (__#) position. Wavy brackets are used to indicate the environment of specific phonemes. An empty slot means that there is no particular conditioning environment for the correspondence. The symbol

'Ø' in a column indicates that the PTIM sound in question is lost in that language. The symbol '~' indicates that more than one reflex is found in that language.

In the subsections that follow, transcription of language data adheres to IPA conventions. Long vowels are indicated with a length mark (:). The languages are arranged in the tables in order from west to east, with the western-most language, Bunaq, on the left-hand side and the eastern-most language, Oirata, on the right-hand side. In the tables in the paper, where a language has irregularly lost the particular segment in question, say, due to apocope (cf. section 5.1), we do not give that item in the correspondence tables, although it may still be deemed cognate. In the text of the paper itself, for reasons of compactness, we only give simple one word glosses which reflect the presumed meaning of the proto-lexeme. Bracketed segments are those deemed to be non-etymological, that is, typically reflecting some morpheme which has fossilised on a root (cf. section 5.2).

Table 1: Sound correspondences

		Bunaq	Makasae	Makalero	Fataluku	Oirata	Environment
1	*p	p	f	f	p	h	
2	*f	w	f	f	f	p	V_V
3	*f	p	f	f	f	p	elsewhere
4	*b	b	b	p	p	h	
5	*t	tʃ (t)	t	t	t	t	#_i
6	*t	t	t	t	t	t	#_V$_{\neq i}$
7	*t	t	t	t	c~t	t	
8	*t	z	t	t	t	t	V_{a, o}
9	*t	t	t	t	t	t	V_V$_{\neq a, o}$
10	*T	tʃ	t	t	c	t	#_i
11	*T	t	t	t	t~c	t	#_V$_{\neq i}$
12	*T	t	t	t	c	t	V_V
13	*d	t	d	d	c	ʈ	
14	*D	z	d	s	Ø	Ø	#_V$_{\text{FRONT}}$
15	*D	z	d	s	c	ʈ	#_V$_{\text{NON-FRONT}}$
16	*D	z (s)	s	s	c	s	V_V
17	*s	s	s	s~h	h	Ø	#_
18	*s	s	s(~h)	s(~h)	h~ʔ	Ø	V_V$_{\text{NON-LOW}}$
19	*s	s	s(~h)	s(~h)	s	s	V_V$_{\text{LOW}}$

20	*s	h~s~g (~t)	s(~d~h)	h~s	h~s	h	#_ (verbs)
24	*k	k	k	k	k	k	#_
25	*k	g~k	ʔ	ʔ	ʔ	ʔ	V_VC#
26	*k	g~k	ʔ	ʔ	ʔ	Ø	V_V#
27	*g	g	g	k	Ø	Ø	#_
28	*g	g~k	g	k	ʔ~k	ʔ~k	V_V
29	*h	h	h	h	Ø	Ø	_i
30	*h	h	Ø	Ø	Ø	Ø	_V_{NON-FRONT}
31	*w	w	w	w	v	w	_V_{FRONT}
32	*w	h	w	w	v	w	_V_{NON-FRONT}
33	*w	w~h	h	Ø	Ø	Ø	_V_{HIGH}
34	*l	l	l	l	l	l	
35	*r	l	r	r	r	r	
36	*m	m	m	m	m	m	
37	*n	n	n	n	n	n	

So that the reader can still easily view and compare cognates together with one another, we include an appendix with all the data we examine along with the details of any irregularities they show. In the appendix we also give details about the semantics of lexemes where relevant, such as when a semantic change has taken place. The full forms of reconstructed lexemes are also given in the appendix.

We now turn to the examination of cognates that illustrate the sound correspondences we identify in the Timor-Kisar languages.

4.1.   Reconstruction of labial obstruents

We reconstruct three bilabial obstruents: *p, *f and *b. Although no single modern language has three distinct labial obstruent phonemes, the reconstruction is secure, being based on three very clearly distinct correspondence sets. The three sets are set out in Table 2.

Table 2: Correspondent sets for *p, *f and *b

meaning	Bunaq	Makasae	Makalero	Fataluku	Oirata
*p	p	f	f	p	h
carry	––	(g)afu, afu	afu ~ kafu	apu	ahau(re)
fish	––	afi	afi	api	ahi
dog	zap	defa	sefar	ipar(u)	ihar(a)
owl	tupi	–– †	–– †	tupu(kuru)	––
tongue	–up	ifi	ifil	epul(u)	uhul(u)
sugarcane	up	ufa	ufa	upa	uha
mountain ₁	op	afa	(war)afar	apa	––
*f	p ~ w	f	f	f	p
swell	pe	fatu	fatu	fatu	patu
girl	pana	fana(rae)	fana(r)	fana(r)	pana(rai)
new	tip	sufa	hofa(r)	tufa(tufa)	––
run	t͡ʃiwal	–– †	–– †	tifar	tipar(e)
face	–ewen	fanu	fanu	fanu	panu
*b	b	b	p	p	h
pig	––	baj	paj	paj	haj
price	bol	bura	pura	pura	hura
star	bi	(ifi)bere	–– ‡	–– ‡	–– ‡
wind	bel	––	pare	pari	hari

† Cognates in these languages show consonant harmony so that reflexes of *p are lost; see item 83 in the Appendix.

‡ Lexemes in these languages reflect the other half of the PTAP lexical doublet for 'star'; see item 98 in the Appendix.

In Bunaq reflexes of *f merge with /p/ initially and finally, and with /w/ medially; reflexes of *p and *b are /p/ and /b/ respectively. Makasae has merged reflexes of *p and *f into a single phoneme, /f/.[9] Makalero has merged *p and *f into /f/ in all positions, and has recreated /p/ from reflexes of *b. Fataluku and Oirata have merged *p and *b in all positions. Reflexes of both these come out as /p/ in Fataluku and /h/ in Oirata.

4.2.   Reconstruction of coronal stops

We reconstruct four coronal stops: *t, *T, *d, and *D. Of these, *t, and *d are thought to have corresponded more or less with the IPA values of the characters. There is no conclusive evidence regarding the phonetic value of *T and *D; hence the use of the capital letters here. One possibility is that they were affricates, t͡ʃ and d͡ʒ respectively. This

would account for the diversity of their reflexes in the modern languages, which include plosives, affricates and fricatives. Alternatively, we could take values from segments found in the modern languages. For instance, *T could be seen as a voiceless palatal stop such as found as in Fataluku, and *D could be seen as a voiceless retroflex stop such as found in Oirata, or a voiced alveolar fricative such as in Bunaq. However, until more evidence emerges, we choose to remain agnostic about the exact value of these proto-phonemes.

Table 3: Correspondence sets for *t

meaning	Bunaq	Makasae	Makalero	Fataluku	Oirata
initial	**t ~ tʃ**	**t**	**t**	**t**	**t**
hand	ton	tana	tana	tana	tana
where	teo	––	tau–	te	te(nai)
owl	tupi	tutu(guru)	toutou	tupu(kuru)	––
RECP	tV–	ta	ta	(ni)ta	ta
pierce	––	––	tafa ~ dafa	tafa	––
cook	tinik ~ ginik	tina	tina ~ dina	––	––
run	tʃiwal	ditar	titar	tifar(e)	tipar(e)
medial	**t ~ z**	**t**	**t**	**t**	**t**
tree	hote(l)	ate	ate	ete	ete
cover	–utu	(g)utu	kutu– ~ utu–	utu	utu
clew	mot	––	––	matar(u)	matar(a)
sit	mit	mit ~ mi	mit	(i)mir(e)	mir(e)
stand	net	––†	nat	(a)nat(e)	nat(e)
P. indicus	mazoʔ	mater	mater	matar(ia)	––
coconut	hoza	wata	wata	vata	wata
sea	––†	––	––	mata	mata
swell	––†	fatu	fatu	fatu	patu

† The relevant segment is not reflected in these cognates due to apocope; see section 5.1.

Table 3 sets out the correspondences for the PTɪᴍ *t. The correspondence of *t is for the most part steady and unchanging as /t/ in the four eastern languages, Makasae, Makalero, Fataluku and Oirata, in both initial and medial positions. In initial position, the only exception we find is Makasae *ditar* 'run' which exhibits an irregular change to /d/, possibly the result of historical prefixation with *n- (as suggested by Schapper 2010: 21,

346). In medial position, *t initially appears to be an irregular change to /r/ in Fataluku and Oirata 'sit', but it seems also possible to explain the irregularity as the result of historical suffixation. That is, PTIM *mit 'sit' appears to have been suffixed with a morpheme like *-*re* as part of an innovation of singular-plural subject distinction on verbs. As a result of this suffixation, the final *t was lost. Similarly, we find irregularities in Makasae posture verbs with singular-plural subject distinction: Makasae has *na* 'stand' instead of expected *nat, and *mit* 'sit' is often also realised as *mi*.

In initial position, Bunaq reflects *t consistently as /t/ in all environments except before /i/. In this environment, *t becomes /ʧ/ except where the initial segment represents a 3rd person inanimate prefix *t-*, in which case it remains a plosive (see Schapper 2010: 43-45 for more details on this). So, the initial /t/ in Bunaq *tinik* 'cook' does not affricate as it is a prefix which alternates with the 3rd person animate prefix *g-* as seen in the form *ginik.* In medial position, Bunaq reflects *t as either /t/ or /z/. It is not entirely clear at this stage why there is this variation in reflexes.

Table 4 presents the small set of cognates in which we find an aberrant correspondence between Makasae final /r/ and Makalero final /t/. Comparison with the other Timor-Kisar languages suggests the reason for this deviant pattern. Notice that Fataluku and Oirata have a corresponding /t/ as in Makalero, but this is always followed by the verbalising suffix –*e* (discussed in section 5.2). By contrast, Bunaq has a final liquid in cognates, that is, following the Makasae pattern and indicating that this is the earlier final consonant (since Bunaq is an earlier break-off in the family; see section 6). We hypothesise that the /t/ in Makalero, Fataluku and Oirata was originally the result of a morphophonemic change in final *r (or perhaps *l) brought about by the suffixation of a morpheme –*e*; Makalero subsequently lost the final /e/, but retained the /t/ that was conditioned by it. The result is the deviant correspondence we find today.

Table 4: r ~ t correspondences in Makasae-Makalero

	Bunaq	Makasae	Makalero	Fataluku	Oirata
long (time)	––	muʔir	muʔit	muʔit(e)	––
old	legul	laʔir	laʔit	laʔit(e)	lait(e)
blow	––	gur	kut	––	––
sick	heser	sisir	hisit ~ sisit	––	––
dry	–– †	sahar	haʔat ~ saʔat	sasak(e)	hatat(e)

† The cognate in these languages shows apocope and does not reflect the segment in question.

In initial position, *T is reflected as /t/ in Bunaq (except before /i/, as above), Makasae, Makalero and Oirata. Fataluku, on the other hand, shows a morphophonologically conditioned variation between /t/ and /c/, although it appears before /i/, only /c/ is found, as in *ciʔir(e)* 'heavy'. In the medial position, *T also becomes /t/ in Bunaq, Makasae and Makalero, but changes to Fataluku /c/ and Oirata /ʈ/. /t/ is found as a reflex of medial *T in Fataluku only in the case of *uta* 'beat', which is used with multiple patients. The corresponding form for a single patient, *uca*, is regular.

Table 5: Correspondence sets for *T

meaning	Bunaq	Makasae	Makalero	Fataluku	Oirata
initial *T	**t**	**t**	**t**	**t ~ c**	**t**
wake	otin	tane	tane	tani ~ cani	––
heavy	––	tiʔir	tiʔir	ciʔir(e)	tu:r(e)
sleep	tʃier	taʔe	tia	taja ~ caja	taja
medial *T	**t**	**t**	**t**	**c**	**ʈ**
contact	ata ~ gita	(g)ata	kata– ~ ata–	aca	aʈa
fire	hoto	ata	ata	aca	aʈa
tick	kata	––	––	kocu	ko:ʈo
sun	hot	watu	watu	vacu	waʈu
kill	–ita	(g)uta	kuta ~ uta	uca ~ uta	uʈa
egg	–ut	––	uta	ucu	uʈu
child	––	mata	mata	moco	moʈo
bat	––	––	––	maca	maʈa

The evidence for PTIM *d is at this stage extremely limited. Table 6 presents the three cognate sets. This proto-phoneme is the only coronal sound to show such defective

distribution, being found only in a handful of cases in initial position. It is reflected as /d/ in Makasae and Makalero, /ʈ/ in Oirata, but as /c/ in Fataluku and /t/ in Bunaq. However, much more data is needed to confirm this reconstruction.

Table 6: Correspondence sets for *d

meaning	Bunaq	Makasae	Makalero	Fataluku	Oirata
*d	**t**	**d**	**d**	**c**	**ʈ**
eucalypt	tal	dara	daraʔ	cara(nu)	––
wake	––	dur	dur	cur(e)	ʈu:r(e)
sit.PL	––	diar	diar	cuar(e)	––†

  † The Oirata reflex of this item has undergone metathesis and does not have a reflex of *d; see item 91 in
    the Appendix.

The reconstruction of *D is more secure than that of *d. We have clear cognates across the full range of Timor-Kisar languages, but the number remains relatively small (Table 7). *D regularly becomes Bunaq /z/; the only exception is where the reflex of *D is word-final, in which case it devoices in accordance with Bunaq phonotactic rules (as in *hos* 'bird'). In Makasae medial *D and initial *D are reflected as /d/ and /s/ respectively. In Makalero *D is reflected as /s/ in all positions. In Fataluku *D is reflected as /c/ in all positions, and in Oirata as /ʈ/ initially and /s/ medially. We find one irregularity in the reflexes in Fataluku and Oirata, namely, that a reflex of *D is lost in the lexeme 'dog'. It may be that this was a conditioned loss of *D, say, before the high front vowel, but more data is required for a conclusion to be drawn.

Table 7: Correspondence sets for *D

meaning	Bunaq	Makasae	Makalero	Fataluku	Oirata
initial *D	**z**	**d**	**s**	**c ~ Ø**	**ʈ ~ Ø**
rat	zul	dura	sura	cura	ʈura
wife	––	da	sa	––	––
cuscus	zulo	––	––	acur(u)	aʈur(u)
head	––	daʔe	saʔe	caʔu	ʈau
dog	zap	depa	sefar	ipar(u)	ihar(a)

medial *D	z (s)	s	s	c	s
beeswax	wezun	(badu)husu	usu(pada)	ucu	usu
see	haza(l)	––	––	aci	asi
bird	hos	asa	asa	aca	asa
together	––	––	––	acu	asu

4.3.   Reconstruction of *s

Correspondences for *s present a somewhat chaotic picture, with several reflexes of the apparent proto-phoneme being found in each language. This appears to be, at least in part, the result of fossilised verbal prefix morphology.

Table 8: Correspondences of *s

meaning	Bunaq	Makasae	Makalero	Fataluku	Oirata
initial *s	s	s	h ~ s	h	Ø (?)
meat	sael	seu	seur	––	––
branch	esu †	sika	saka	––	––
war	asu †	sala	hala	hal(u)	al(a)
wall	sirin	sidi	hir	––	––
spoon	sulu	sulu	hulu	hula	––
bone	––	safa	hafa	hafa	opo
banyan	––	sama	hama	hama	––
medial *s	s	s (~ h)	s (~ h)	s ~ h ~ ʔ	s ~ Ø
for	hos ~ gos	(g)asu	kasu ~ asu	ahu	––
sick	heser	sisir	hisit ~ –sisit	––	––
rare	lesa	––	lesa	––	––
tooth	––	wasi	wasi	vahi(nu)	wai(ni)
cut	––	lasi	lasi	laʔi	lai
hit	––	base	pase	paha	––
leaf	––	asa	hasa	asa	asah(a)
stand.PL	––	nahar	naser	neher(e)	––
scared	––	––	masan	mahan(e)	man(e)
urinate	sele ‡	––	irih	iris(e)	iris(e)

† The final CV has been lost in these Bunaq items. See section 5.1 on Bunaq apocope.
‡ This item shows historical metathesis. See this item in the Appendix for reconstruction of the
  stepwise development of the lexeme.

Table 8 presents cognate sets with initial *s correspondences on non-verbs and with medial correspondences (on both verbs and non-verbs). There is by and large regularity.

Initially, Bunaq and Makasae consistently show /s/, and Fataluku /h/. Makalero is messy, with both /s/ and /h/ reflexes being found without any apparent environment conditioning. We do not at this stage have enough data for Oirata to make a judgement about its initial reflexes. Medially for *s, the reflex /s/ is consistently found in the few Bunaq cognates we identify. In Makasae and Makalero, medial *s is reflected as /s/ in all but one case each where *h is found. In Fataluku, *s is medially subject to lenition in front of *u, *i and *e, and is reflected in these environments as either /h/ or /ʔ/, but retained as /s/ elsewhere. Similarly, in Oirata, *s is lost medially in front of /i/, and reflects otherwise as /s/.

When we turn to verbs, we find that initial *s shows a much greater degree of variation across the languages. Table 9 presents a sample of the unusual correspondences we find. It appears that various forms of morphological alternations involving the initial segments *s, *h and *g have been preserved in different ways. In Bunaq, we find productive alternations where an initial /h/ or /s/ marking a 3rd person inanimate P is replaced by /g/ when an animate 3rd person is P. In Makasae, mostly the /s/ form of the verb has been preserved, except in one case (*haga* 'seek') where we find an /h/ initial form. Makalero also shows an /h ~ s/ alternation with many verbs; the /s/ initial form is conditioned by a locative element combining with the verb. On other verbs either the /s/ or the /h/ initial form is preserved. In Fataluku and Oirata, the /h/ initial form of the verb is typically preserved. In other contexts, /s/ appears. It is not yet clear what underlies the observed alternations on verbs with original initial *s, and more investigation of these unusual forms is necessary.

Table 9: Correspondences of initial /s/ and /h/ on verbs

meaning	Bunaq	Makasae	Makalero	Fataluku	Oirata
sick	heser	sisir	hisit ~ –sisit	––	––
finish	haʔa(l) ~ gaʔa(l)	saʔi	haiʔ ~ –saiʔ	sai	––
shoot	holi ~ goli	suri ~ duri	huri ~ –suri	suri(te)	––
hold	––	sifaʔ ~ difaʔ	hifaʔ ~ –sifaʔ	––	––
clean	––	sare	hareʔ ~ –sareʔ	(pa)hal(e)	ha:l(e)
seek	sagal ~ gagal	haga	haka ~ –saka	haʔa	(aṭa)hoko
weave	sien †	sina	hina	hina	––
sweat	sil	––	––	her(u)	––
tie	ti	sil	sil	(i)sil(e)	i:l(e)
new	tip	sufa	hofar	tufa(tufa)	––

† This item shows historical metathesis. See this item in the Appendix for reconstruction of the stepwise development of the lexeme.

4.4. Reconstruction of velar stops

We reconstruct two velar plosives: *k and *g. We have evidence for *k and *g in initial and medial positions, but not in word-final position. The correspondence sets are laid out in Table 10 and Table 11 for *k and *g respectively.

Correspondences for *k are relatively straight-forward. The correspondence of *k in initial position is constant and unchanging in all languages. We do however note that only a small number of cognates could be identified. Reflexes of medial *k are more abundant. In Bunaq, medial *k becomes /g/ where the onset it occurs on is still medial. However, where the final vowel of the proto-lexeme has been lost, *g devoices to k as is required by Bunaq phonotactics (Schapper 2010: 56-57). Medially *k changes regularly to glottal stop in Makasae, Makalero and Fataluku. In Oirata, the change of *k to glottal stop has also taken place, but with an extra step forward, that is, the loss of the glottal stop in some items. This loss is conditioned by the openness of the syllable in which medial *k occurred. That is, where the proto-lexeme had the shape *(C)VkVC, *k is retained as glottal stop, as in the correspondence sets for 'green' and 'mouth'; where the proto-lexeme had the shape *(C)VkV, no reflex of *k is retained in Oirata.

Table 10: Correspondence sets for *k

meaning	Bunaq	Makasae	Makalero	Fataluku	Oirata
initial *k	**k**	**k**	**k**	**k**	**k**
P. buceroides	koak	kaukua	––	(folo)kua	––
tick	kata	––	––	kocu	ko:ṭo
small	––	kaʔu(lai)	kaʔu	kaʔu(sila)	––
sing	––	kaul	kaul	kol(e)	––
medial *k	**g ~ k**	**ʔ**	**ʔ**	**ʔ**	**ʔ ~ Ø**
green	ugar	uʔur	huʔur ~ uʔur †	uʔur(eke)	uʔul(e)
laugh	higa(l)	hiʔa	hiʔe	––	––
mouth	–agar	aʔa	haʔa ~ aʔa †	oʔo	oʔo
bite	gagil	gaʔel	kaʔel	––	––
banana	mok	muʔu	muʔu	muʔu	mu:
earth	muk	muʔa	muʔa	muʔa	mua
head	––	daʔe	saʔe	caʔu	ṭau
hear	mak	maʔen	maʔen	––	––
path	hik	hiʔa	hiʔa	iʔa	ia(ra)

† /h/ is very weak in Makalero and often dropped completely (see Huber 2011: 51).

The reconstruction of *g is supported by a relatively small set of correspondences. Its reconstruction in initial position hinges entirely on two 3rd person markers. This will be discussed below. The evidence for medial *g is rather tenuous at this stage, as not many cognates could be located outside of Makasae and Makalero.

Table 11: Correspondence sets for *g

meaning	Bunaq	Makasae	Makalero	Fataluku	Oirata
initial *g	**g**	**g**	**k**	**Ø**	**Ø**
3 PREFIX	g–	g–	k–	––	––
3 POSSESSOR	gie	gi	ki	i	ue
medial *g	**g ~ k**	**g**	**k**	**ʔ ~ k**	**ʔ ~ k**
difficult	––	sege	heke ~ –seke	heʔe	heʔe
seek	saga(l)	haga	haka ~ –saka	haʔa	(aṭa)hoko
tell a lie	––	logo	loko	––	––
fantail	––	sigilai	(taru)sikiloti	––	––
owl	––	(tutu)guru	––	(tupu)kuru	––
cold †	huruk	gawa	kauar	––	––

† This set shows unusual metathesis patterns. See 16 in the appendix for details.

Initial *g is reflected in Bunaq as a productive agreement prefix marking 3rd person animates. It is an inflection which occurs in a paradigmatic relationship with two other prefixes, *n-* '1EXCL' and *V-* '1INCL/2'. Thus, the following examples:

(1)   a.   INFLECTIONS ON POSSESSIVE CLASSIFIER:
                n-ie '1EXCL-POSS', *Ø-ie* '1INCL/2-POSS', *g-ie* '3AN-POSS'.
        b.   INFLECTIONS ON VERB, e.g.:
                n-ube '1EXCL-block, *Ø-ube* '1INCL/2-block', *g-ube* '3AN-block'.

In Fataluku and Oirata, all such person marking has been stripped from roots, but in Makasae and Makalero some remnants of *g-, the 3rd person inflection, are found. Consider the data in Table 12. Makalero has a small set of verbs that inflect for *k-* marking a 3rd person P argument (for details and subtypes, see Huber 2011: 349ff); this prefix is cognate with Bunaq productive *g-* '3AN'. Makasae retains a reflex of the same prefix, although it is no longer an inflectional morpheme. We see that the Makasae cognates of the Makalero inflecting verbs for the most part have fossilised the prefix and thus always have initial /g/. There are two exceptions to this fossilisation pattern: (i) *gapu* 'with' and *apu* 'carry', where Makasae retains both prefixed and unprefixed forms, albeit with differing semantics, and; (ii) *umu* where no prefix has fossilised.

Table 12: Reflexes of 3rd person *g- on roots

meaning	Bunaq		Makalero		Makasae		Fataluku	Oirata
	inanimate 3rd person	animate 3rd person †	vowel form	k-prefixed form	vowel form	g-form		
for	hos	g-os	asu	k-asu	--	gau	ahu	--
kill	--	g-it	uta	k-uta	--	guta	uta	uṭa
impact, touch	--	gene	ene	k-ene	--	gene	--	--
cover, wear	--	g-utu	utu	k-utu	--	gutu	utu	utu
contact, (on/in)to	ata	g-ita	ata-	k-ata-	--	gata	aca	aṭa
die	--	g-ume	umu-	k-umu-	umu	--	umu	umu
carry	--	--	afu	k-afu	afu	gafu	apu	ahau(re)
give	hini	g-ini	ini	k-ini	--	gini	ina	ina
on top	wa	ga-wa	ua-	k-ua-	--	gua	--	--

† The hyphen separating the Bunaq *g-* '3AN' prefix from its root is included here to make clear that this is an inflectional morpheme, that can be replaced by other items in the paradigm.

4.5.   Reconstruction of *h and *w

We reconstruct both *h and *w phonemes to PTIM on the basis of two clear and distinct correspondence sets. These are set out in Table 13. However, we do note that the reconstruction of *w and *h can only be made for word-initial position; there is no evidence for the reconstructed phonemes being present in medial or final positions.

Table 13: Correspondence sets for *h and *w

meaning	Bunaq	Makasae	Makalero	Fataluku	Oirata
*h	h	Ø ~ h	Ø ~ h	Ø	Ø
bird	hos	asa	asa	aca	asa
fire	hoto	ata	ata	aca	aṭa
tree	hotel	ate	ate	ete	ete
moon	hul	uru	uru	uru	uru
laugh	higa(l)	hiʔa	hiʔe	––	––
path	hik	hiʔa	hiʔa	iʔa	ia(ra)
*w	w ~ h	w (~ h)	w (~ Ø)	v (~ Ø)	w (~ Ø)
bathe	wer	waruʔ	waroʔ	vahu	wau
tooth	–ewe	wasi	wasi	vahi(nu)	wai(ni)
top, above	wa	gua	wa– ~ kua–†	––	––
beeswax	wezun	(badu)husu	usu(pada)	ucu	usu
blood	ho	waj	wej	vehe	we
coconut	hoza	wata	wata	vaca	wata
pea, bean	ho	wa:	wa	––	wa
stone	hol	––	war	––	war(aha)

† When *k-* is prefixed the initial /w/ vocalises.

The phoneme *h is invariably retained in Bunaq as /h/, while in Fataluku and Oirata it is always lost. In Makasae and Makalero, the retention of *h depends on the vocalic environment. Before the high front vowel /i/, Makasae and Makalero retain *h as /h/; before all other vowels, *h is lost.

Reflexes of *w are retained in all Timor-Kisar languages in at least some vocalic environments. In Bunaq, *w is retained as /w/ before non-back vowels, but has become /h/ before back vowels. The phoneme *w is typically retained as /w/ in Makasae, Makalero and Oirata, and as /v/ in Fataluku. However, there is some evidence to suggest

that *w was not retained as /w/ before high vowels in eastern Timor and Kisar. Firstly, in these languages, it is synchronically rare to find /w/. Secondly, we have evidence from one cognate set (PTIM *wiDu(n) 'beeswax') that *w in fact went to *h before the high vowel *u, since it is retained as /h/ in Makasae; this *h was then lost in the other three languages. More cognate sets need to be identified to confirm this hypothesis.

4.6.   Reconstruction of liquids

We reconstruct two liquids *l and *r to PTIM. The correspondence sets for these are set out in Table 14 and Table 15 respectively.

The proto-phoneme *l is evidenced in all positions (Table 14). Whilst not abundant in number, the correspondences are for the most part steady and unchanging. However, the number of sets that have been identified is small and the reconstruction of *l can not yet be said to be secure.

Table 14: Correspondence sets for *l

meaning	Bunaq	Makasae	Makalero	Fataluku	Oirata
*l, initial	l	l	l	l	l
walk	lagor	laʔa	laʔa	laʔa	lara
live	––	lafuʔ	lafuʔ	lau(he)	lau(se)
rare	lesa	––	lesa	––	––
fly	lore	lor	lor	––	––
*l, medial	l	l	l	l	l
tail	–uloʔ	ula	ula	ula(fuka)	ula(puʔa)
narrow	moel	––	––	male(te)	male(te)
ear	––	wala(ku:)	wali	vali	wali
*l, final	l	l	l	l	l
tongue	––	––	ipil	epul(u)	uhul(u)
bite	gagil	gaʔel	kaʔel	––	––
tie	––	sil	hil ~ –sil	acil(e) ~ isil(e)	i:l(e)

Note that Bunaq in many cases shows an unetymological final /l/. This is discussed as a fossilised suffixal morpheme in section 5.2.

The proto-phoneme *r is robustly supported by cognate sets, albeit only in word-medial position. We see in Table 15 that there are only two instances of final *r reconstructable on the basis of the cognates we have identified. There is no evidence for *r having been present word-initially in PTIM. In the sets we have, *r is regularly reflected as /l/ in Bunaq, but retained as /r/ in all other languages. See section 5.1 for discussion of the loss of final vowels in Bunaq.

Table 15: Correspondence of *r

meaning	Bunaq	Makasae	Makalero	Fataluku	Oirata
*r	l	r	r	r	r
eucalypt	tal	dara	dara?	cara(nu)	––
go	mal	mara	mara	mara	mara
moon	hul	uru	uru	uru	uru
price	bol	bura	pura	pura	hura
rat	zul	dura	sura	cura	ʈura
water	il	ira	ira	ira	ira
wind	bel	––	pare	pari	hari
vagina	––	aru	aru	aru	aru
wake	––	dur	dur	cur(e)	ʈur(e)
creep	el	––	––	er(eke)	––
cuscus	zulo	––	––	acur(u)	aʈur(u)
shoot	holi ~ goli	suri ~ duri	huri ~ –suri	surit(e)	––
urinate	sele	iri	iri?	iris(e)	iris(e)
stone	hol	––	war	––	war(aha)
run	tʃiwal	ditar	titar	tifar(e)	tipar(e)

Irregularity is a major issue in the reconstruction of liquids, in particular final liquids. In Table 16 we present a variety of cognate sets showing irregularity in final liquid reflexes. There are several trends that can be recognised in the data. Firstly, in Bunaq apocope of final syllables of the shape VL (where 'L' stands for any liquid) is common in disyllabic proto-lexemes with the shape (C)VCVL, but not predictable (see section 5.1). Similarly, in Makasae we repeatedly observe the loss of final liquids in disyllables (see section 5.1). Secondly, in other cases we find that Bunaq has a final liquid /r/ where other languages do not. Thirdly, Oirata occasionally has /l/ instead of the expected *r before a suffix *-e*

which has fossilised on the item. Finally, we note the existence of a few cognates where Makalero and Makasae /r/ correspond to Bunaq /r/ instead of the expected *l reflex.

Table 16: Examples of cognate sets with irregular reflexes of liquids †

meaning	Bunaq	Makasae	Makalero	Fataluku	Oirata
new	tip Ø	sufa Ø	hofar	––	––
dog	zap Ø	defa Ø	sefar	ipar(u)	ihar(a)
stone circle	mot Ø	––	––	matar(u)	matar(a)
tongue	–up Ø	ifi Ø	ifil	epul(u)	uhul(u)
meat	sael	seu Ø	seur	leur(a)	leur(a)
garden	mar	ama Ø	ama Ø	––	uma Ø
mouth	–agar	aʔa Ø	haʔa Ø	oʔo Ø	oʔo Ø
sleep	tʃier	tiʔe Ø	tia Ø	taja Ø	taja Ø
green	ugar	uʔur	huʔur	uʔur(eke)	uʔul(e) *r
clean	––	sare	hareʔ ~ –sareʔ	(pa)hal(e) *r	ha:l(e) *r
bathe	wer *l	waruʔ	waroʔ	vahu Ø	wau Ø
fly	lore *l	lor	lor	––	––
sick	heser *l	sisir	hisit ~ sisit ‡	––	––

† Ø represents the absence of a liquid where one would be expected. A starred segment represents the expected, but not occurring liquid segment.

‡ The issue of final r ~ t correspondences between Makasae and Makalero was already discussed in section 4.2.

In short, more work needs to be done to clarify the history of liquids in Timor-Kisar languages.

4.7.   Reconstruction of nasals

Two nasals *m and *n can be unproblematically reconstructed to PTIM. The correspondence sets for these are set out in Table 17 and Table 18 respectively. We see that the correspondences are for the most part steady and unchanging. However, there is for both proto-phonemes little clear evidence of nasals in final positions.

The reconstruction of initial *m is well supported by a multitude of correspondence sets (Table 17). We have, however, only been able to identify one instance of medial *m, namely in PTIM *-umV 'die'. We also have no evidence for *m in word-final position,

and this restriction persists in the modern languages; none of the Timor-Kisar languages allow final /m/.

Table 17: Correspondence sets for *m

meaning	Bunaq	Makasae	Makalero	Fataluku	Oirata
*m	**m**	**m**	**m**	**m**	**m**
bamboo	ma	maeri	mar	––	––
banana	mok	muʔu	muʔu	muʔu	mu:
bat	––	––	––	maca	maʈa
earth	muk	muʔa	muʔa	muʔa	mua
garden	mar	ama	ama	––	uma
hear	mak	maʔen	maʔen	––	––
inside	mi(l)	mu(tu)	mu(tu–)	mu(cu)	mu(ʈu)
narrow	muel	––	––	male(te)	male(te)
P. indicus	mazoʔ	mater	mater	matar(ia)	––
sea	mo	––	––	mata	mata
clew	mot	––	––	matar(u)	matar(a)
sit	mit	mit ~ mi	mit	(i)mir(e)	mir(e)
take	––	ma	mej	me	me
die	–ume	umu	(k)umu	umu	umu

The reconstruction of *n in initial and medial positions is also well supported by a multitude of correspondence sets (Table 18). At this stage, however, we don't have any correspondence sets which clearly involve final *n. This will be discussed briefly below.

Table 18: Correspondence sets for *n

meaning	Bunaq	Makasae	Makalero	Fataluku	Oirata
*n	**n**	**n**	**n**	**n**	**n**
stand	net	na	nat	(a)nat(e)	nat(e)
stand.PL	––	nahar	naser	(i)neher(e)	––
name	–ini(l)	naj	nej	ne	ne:(ne)
cook	tinik	tina	tina ~ dina	––	––
face	–ewen	fanu	fanu	fanu	panu
girl	pana	fana(rae)	fana(r)	fana(ru)	pana(rai)
give	–ini	(g)ini	–ini ~ kini	ina	ina
hand	ton	tana	tana	tana	tana
impact	(g)ene	(g)ene	kene ~ ene	––	––

| wake | otin | tane | tane | tani ~ cani | –– |
| weave | sien | sina | hina | hina | –– |

At this stage, it is difficult to reconstruct final *n because we do not find sufficiently regular and stable correspondences across the Timor-Kisar languages. This is clear from a consideration of the data in Table 19. For instance, we observe that in Bunaq final /(V)n/ is frequently lost (as in the sets 'holy', 'hear' and 'mountain'). However, in some cases it is retained (as in the set 'long'), and in other cases Bunaq even appears to add a final /n/ to correspondences where other languages don't have it (as in the set 'beeswax'). So, in short, the identification of more correspondence sets is necessary before final *n can be reconstructed.

Table 19: Appearance of final /n/ in Timor-Kisar languages

meaning	Bunaq	Makasae	Makalero	Fataluku	Oirata
holy	por	falun	falun	––	––
hear	mak	maʔen	maʔen	––	––
mountain $_2$	lolo	larin	larin	––	––
right–hand	––	tane	tanen	tenen(u)	tenen(e)
throw	––	liʔan	liʔan	(li)liʔen(e)	len(e)
long	esen	asan	asan	––	––
beeswax	wezun	(badu)husu	usu(pada)	ucu	usu

5.   The problem of final word-edges in the reconstruction

In the previous section, we have occasionally mentioned the existence of fossilised morphemes whose identification has been necessary to clarify the correspondences and reconstruction. In this section, we expand on this and related issues looking at irregularities that occur at the word-edges of Timor-Kisar languages and how they complicate the reconstruction.

5.1.　　Apocope

The reader may already have noticed that Bunaq lexemes are often "shorter" than their cognates in the other Timor-Kisar languages. Table 20 shows the instances of apocope in Bunaq and summarises the patterns found. The largest set of examples involves the loss of the final vowel of a proto-lexeme of the form (C)VCV. Less frequent patterns include the loss of a full CV syllable, the loss of a VC and of a V(C) sequence. The majority of these patterns yield monosyllabic Bunaq lexemes of the form (C)VC. Table 19 shows only five exceptions to this pattern, namely those items which drop a CV syllable. We are at present unable to explain when this exceptional pattern occurs.

Table 20: Instances of apocope in Bunaq

Bunaq	PTIM	meaning	Bunaq	PTIM	meaning
V			**CV**		
mok	*muku	banana	ma	*mari	bamboo
hos	*haTa	bird	mo	*mata	sea
muk	*muka	earth	bi	*(ipi–)bere	star
–ut	*–uTa	egg	pe	*patu	swell
–ewen	*–panu	face	–ewe	*–wasi	tooth
mal	*mara	go	**VC**		
ton	*tana	hand	el	*erek	crawl
hul	*huru	moon	mot	*matar	stone circle
hik	*hika	path	zap	*Depar	dog
bol	*bura	price	**V(C)**		
zul	*Dura	rat	wer	*weru(ʔ)	wash
net	*nate	stand	tal	*Dara(ʔ)	eucalypt
up	*upa	sugarcane	mak	*make(n)	hear
hot	*watu	sun	por	*falu(n)	holy
il	*ira	water	tip	*(s, t)ipa(r)	new
bel	*bari	wind	lolo	*laru(n)	mountain [2]

Apocope is also found in a variety of Makasae lexemes, which lose a final consonant. Table 21 gives a list of these items, with, for comparison, their consonant-final Makalero counterparts. The consonant in question in the reconstructed proto-forms are the liquids, *l and *r, *n and *t. Most of these items involve some irregular changes in the other

Timorese languages, too, and we suggest in section 5.2 that this might be due to fossilised morphology from the proto-language.

Table 21: Apocope in Makasae

	Makasae	Makalero	PTɪᴍ
new	sufa	hofar	*(s, t)ifa(r)
meat	seu	seur	*seor
cold	gawa	kawar	*hugur †
dog	defa	sefar	*Depar
tongue	ifi	ifil	*ipul
right–hand	tane	tanen	*tanen
stand	na	nat	*nat
sit	mi(t)	mit	*mit

† This item has undergone a complex process of metathesis.

5.2.   Fossilised suffixes

Throughout the previous sections we have mentioned various instances of fossilized morphology in the Timor-Kisar languages obscuring regular sound correspondences. In this section, we give a brief overview of the morphemes that we have thus far identified. Throughout this section we mark the additional morphemes under discussion in the tables in bold.

One of the most readily apparent irregularities is the presence of a final /l/ on a dozen or so items in Bunaq, with no hint of it ever having been present in the other languages.[10] Table 22 presents a few examples from our corpus. This unetymological final /l/ in Bunaq is likely to have been a suffix of some kind which fossilised, but it is, at this stage, not yet known what the exact nature of the morpheme was.

Table 22: Examples of unetymological of final /l/ in Bunaq

meaning	Bunaq	Makasae	Makalero	Fataluku	Oirata
name	–inil	nai	nei	ne	ne:(ne)
tree	hotel	ate	ate	ete	ete
seek	sagal	haga	saka ~haka	haʔa	(aṭa)hoko
inside	mil	mu(tu)	mu(tu–)	mu(cu)	mu(ṭu)
laugh	higal	hiʔa	hiʔe	––	––

Fataluku has some derivational morphology which is found on many lexical items: a verbal suffix *–e* and a nominal suffix *–u*. As a consequence, many Fataluku items are longer by one syllable than their Makasae and Makalero cognates. In the case of consonant-final lexemes, this is usually mirrored in Oirata by a verbal suffix *–e* and a nominal suffix *–V* whose vocalic value is specified through harmonisation with the final vowel of the root.[11] Table 23 gives a few examples from our data set.

Table 23: Fataluku & Oirata derivational morphology

		Bunaq	Makasae	Makalero	Fataluku	Oirata
Verbal	wake	––	dur	dur	cure	ṭu:re
	heavy	––	tiʔir	tiʔir	ciʔire	tu:re
	scared	––	––	masan	mahane	mane
Nominal	dog	––	defa	sefar	iparu	ihara
	clew	mot	––	––	mataru	matara
	tongue	–up	ifi	ifil	epulu	uhulu

Fataluku items with these suffixes often coincide with cognates exhibiting irregular changes in the other languages. This suggests that this morphology may have affected the working of the regular sound changes identified in section 4. Table 24 gives some examples where the Fataluku derivational affixes coincide with apocope of the final syllable in Bunaq and loss or change of the final consonant in Makasae. In one case, *titlene* 'dry', Oirata displays the addition of a further syllable at the right edge.

Table 24: Irregularities I

	Bunaq	Makasae	Makalero	Fataluku	Oirata	PTIM
stone circle	mot	--	--	mataru	matar	*matar
dog	zap	defa	sefar	iparu	ihara	*Depar
tongue	-up	ifi	ifil	epulu	uhulu	*ipul
dry	--	titi	titil	titile	titlene	*titil
right-hand	--	tane	tanen	tenenu	tenene	*tanen
old	--	la?ir	la?it	la?ite	laite	*lakir
long (time)	--	mu?ir	mu?it	mu?ite	--	*mukir

In some cases, Fataluku not only adds a vocalic derivational suffix, but one of either the form *-Cu* or *-Ce*. The unspecified consonant ('C') is usually also found in Oirata. Table 25 lists a variety of examples. Again Bunaq cognates generally involve apocope, while the Makasae ones in some cases show loss or change of the final consonant.

Table 25: Irregularities II

	Bunaq	Makasae	Makalero	Fataluku	Oirata
eucalypt	tal	dara	dara?	caranu	--
tooth	-ewe	wasi	wasi	vahinu	waini
edge	--	wali	walir	valiku	--
live	--	lafu?	lafu?	lauhe	lause

In a variety of other cases, laid out in Table 26, Fataluku and/or Oirata lexemes seem to be originally compounds, or lexicalised parallel expressions. We refer to these as "lexical doublets".[12] It is noticeable that the operating of the regular sound changes seems to have been unaffected, as, for instance, in *ukar 'green', which shows the regular reflexes of *k as /g/ in Bunaq, but as /?/ in all the other languages. Also, *r is retained in all daughter languages.

Table 26: Fataluku and Oirata compounds

	Bunaq	Makasae	Makalero	Fataluku	Oirata
stone	hol	--	war	--	waraha
tail	–uloʔ	ula	ula	ulafuka	ulapua
small	--	kaʔu	kaʔu	kaʔusila	--
green	ugar	uʔur	huʔur	uʔureke †	uʔule

† *-eke* is often found on colour terms in Fataluku

Finally, in some instances, Fataluku items also have an unetymological prefixal element which has fossilised. The absence of this additional vowel is conspicuous in other Timor-Kisar languages. The list in Table 27 suggests that this is an old locative element associated with positional verbs. These forms are also marked by the already mentioned verbaliser *–e* in Fataluku and Oirata.

Table 27: Fataluku prefixes

	Bunaq	Makasae	Makalero	Fataluku	Oirata
sit	mit	mi ~ mit	mit	imire	mire
sit.PL	--	diar	diar	icuare	--
stand	net	na	nat	anate	nate
stand.PL	--	nahar	naser	inehere	--

6.   Subgrouping

Now that we have established the relatedness of the Timor-Kisar languages, we turn to the identification of subgroup-defining sound changes.

We propose that there are two primary subgroups, with Bunaq in one, and the remaining four languages in the other. The eastern languages, Makasae, Makalero, Fataluku and Oirata form a subgroup, which we call Proto-Eastern-Timor (PETIM). PETIM is characterised by the following sound changes from PTIM:

(i)   *k > ʔ /V_V

(ii)   *h > Ø /#_V$_{NON-FRONT}$

(iii)   *w > Ø / _u

The other primary subgroup– comprised of Bunaq alone– also has several exclusive sound changes from PTᴵᴹ that set it apart from PETᴵᴹ. They are:

(i)   Merger of *p and *f in initial and final positions as p
(ii)   *w > h / #_V_NON-FRONT
(iii)   *t > ʧ /#_i
(iv)   *r > l /V_V

Within PETᴵᴹ, we can further distinguish two subgroups, Proto-Makasae-Makalero (PMᴀᴋᴀ) and Proto-Fataluku-Oirata (PFʀᴀᴛᴀ). PFʀᴀᴛᴀ is identified as a subgroup within PETᴵᴹ on the basis of the following sound changes:

(i)   Loss of 3rd person prefix *g-
(ii)   PETᴵᴹ *h > Ø / #_V_FRONT
(iii)   Merger of PETᴵᴹ *p and *b as PFʀᴀᴛᴀ *p

Makalero and Makasae are very closely related. PMᴀᴋᴀ is defined by the following sound changes:

(i)   Merger of PETᴵᴹ *p and PETᴵᴹ *f as PMᴀᴋᴀ *f
(ii)   Morphophonologically conditioned variation of the initial consonants in verbs with PETᴵᴹ initial *s

In summary, the phonological evidence suggests a subgrouping as shown in Figure 1.

Figure 1: Subgrouping of the Timor-Kisar languages
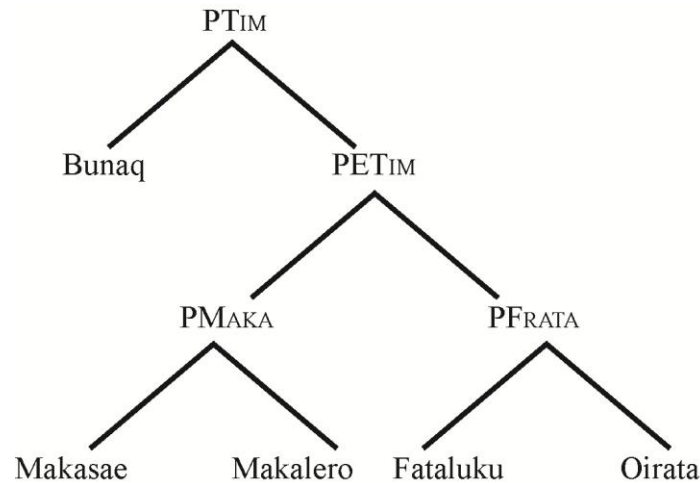
7.   Concluding discussion

In this paper, we have shown that the regularity of sound correspondences in cognate vocabulary supports the view that the Papuan languages of Timor and Kisar languages are genetically related to one another. We have further adduced sound changes that define subgroups for the hypothesised Proto-Timor language.

It was mentioned at the beginning of this paper that the existence of a Proto-Timor node which encompasses Timor-Kisar languages to the exclusion of Alor-Pantar languages remains speculative. That is, there are two significant points that remain to be clarified by the application of the comparative method. They are:

(i)   Are the Alor-Pantar languages related to the Timor-Kisar languages?

(ii)   If they are related, what are the primary subgroups of Proto-Timor-Alor-Pantar (PTAP)?

We are confident that the answer to (i) is yes. Even upon cursory inspection, cognates are immediately obvious to the linguist with experience in the TAP languages; Schapper et al. (2012) presented initial consonant correspondence sets between Timor-Kisar languages and PAP. A more extensive comparison is still to be done, but will likely result

in the confirmation of a historical relationship. The greater question is (ii). Figure 2 presents the two *prima facie* possibilities for the primary subgrouping of the family. In the one scenario, the TAP family has two primary subgroups, PAP on the one hand and PTIM on the other (as assumed in this paper for the purposes of simplicity). In the other scenario, the TAP family has three primary subgroups, namely, PAP, PETIM and Bunaq.

Figure 2: Primary sub-grouping possibilities for the Timor-Alor-Pantar languages
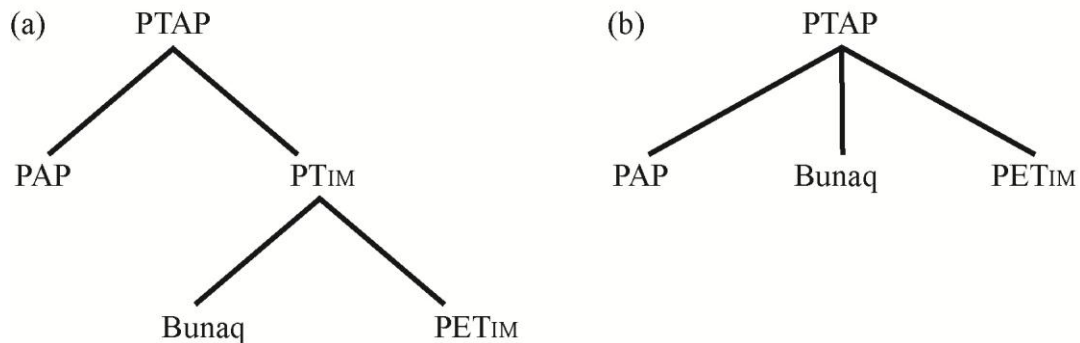
Work is already underway to determine the exact level of the relationship between Timor-Kisar and Alor-Pantar languages. The solving of this question (to which this paper represents an intermediate step) will take us one step closer to a better understanding of the linguistic prehistory of Eastern Indonesia. The received wisdom holds that pre-existing Papuan language communities were overwhelmed by the relentless spread of the Austronesians across insular South-East Asia (Bellwood, Fox and Tryon 1995). This view has been challenged by suggestions that TAP languages were a late arrival in the Timor region from Bomberai post-dating the Austronesian arrival. The finalisation of a reconstruction of PTAP will permit the proper comparison with the languages of the Bomberai peninsula in order to test this notion, which is amongst the most intriguing and most speculative claims in Papuan studies.

References

ADELAAR, ALEXANDER. 2005. The Austronesian languages of Asia and Madagascar: a historical perspective. In: Alexander Adelaar, and Nikolaus P. Himmelmann (eds.).

The Austronesian languages of Asia and Madagascar, 1-42. London and New York: Routledge.

ALMEIDA, ANTÓNIO DE. 1994. *O Oriente de Expressão Portuguesa*. Lisbon: Fundação Oriente and Centro de Estudos Orientais.

BELLWOOD, PETER; JAMES FOX; and DARRELL TRYON (eds). 1995. *The Austronesians: Historical and Comparative Perspectives*, Canberra: Department of Anthropology, Research School of Pacific and Asian Studies, Australian National University.

BERTHE, LOUIS. 1959. Sur quelques distiques bunaq (Timor central). *Bijdragen tot de taal, land en volkenkunde* 115: 336-371.

BERTHE, LOUIS. 1963. Morpho-syntaxe du Bunaq (Timor central). *L'Homme* Jan-Apr: 106-116.

BERTHE, LOUIS. 1972. *Bei Gua, itinéraire des ancétres, mythes des Bunaq de Timor*. Paris: CNRS.

BERTHE, LOUIS. 1978. *Comment fut tranchée la liane céleste et autres textes de littérature orale bunaq (Timor, Indonésie)*. Paris: SELAF.

BROTHERSON, ANNA. 2003. *A Spatial Odyssey: Referring to Space in Makasai*. Honours thesis, Australian National University, Canberra.

CAMPAGNOLO, HENRI. 1973. *La Langue des Fataluku de Lórehe (Timor Portugais)*. Paris: Université René Descartes PhD thesis.

CAPELL, ARTHUR. 1943a. Peoples and Languages of Timor. *Oceania* 14.3: 191-219.

CAPELL, ARTHUR. 1943b. Peoples and Languages of Timor. *Oceania* 14.4: 311-37.

CAPELL, ARTHUR. 1944. Peoples and Languages of Timor. *Oceania* 15.3: 19-48.

CAPELL, ARTHUR. 1972. Portuguese Timor: Two More Non-Austronesian Languages. *Oceania Linguistic Monographs* 15: 95-104.

CAPELL, ARTHUR. 1975. The West Papuan Phylum: General, and Timor and areas further west. In: S.A. Wurm (ed.). *New Guinea area languages and language study*, 667-716. Canberra: Pacific Linguistics.

CARR, FELICITA M. 2004. *Pupuuk, Masu Moru and Data: Three genres of Makasai*. Canberra: Australian National University Masters thesis.

COWAN, H.K.J. 1963. Le Buna' de Timor: une langue "ouest-papoue". *Bijdragen tot de taal-, land- en volkenkunde* 119: 387-400.

COWAN, H.K.J. 1965. The Oirata language. *Lingua* 14: 360-370.

DONOHUE, MARK, and LEA BROWN. 1999. Ergativity: Some additions from Indonesia. *Australian Journal of Linguistics* 19: 57-76.

ENGELENHOVEN, AONE VAN. 2009. On derivational processes in Fataluku, a non-Austronesian language in East Timor. In: W.L. Wetzels (ed.). *The Linguistics of Endangered Languages, Contributions to Morphology and Morphosyntax,* 331-362. Utrecht: LOT.

ENGELENHOVEN, AONE VAN. 2010a. Verb serialization in Fataluku: The case of 'take'. In: Azeb Amha, Christian J. Rapold, Sascha Völlmin and Silvia Zaugg-Coretti (eds.). *Converbs, medial verbs, clause chaining and related issues: Proceedings of the workshop held in December 2006 at Leiden University,* 185-211. Frankfurter Afrikanistische Blätter. Köln: Köppe.

ENGELENHOVEN, AONE VAN. 2010b. The War of the Words. Lexical parallelism in Fataluku ritual discourse. In: Clara Sarmento (ed.). *From Here to Diversity:*

Globalization and Intercultural Dialogues, 241-252. Newcastle Upon Tyne: Cambridge Scholars Publishing.

ENGELENHOVEN, AONE VAN. 2010c. The Makuva Enigma: Locating a hidden language in East-Timor. *Revue Roumaine de linguistique* 80 (2), 161-181.

FAUST, VERA. 2005. *Oirata, a language of Kisar. 1. Grammatical sketch, 2. Myth of Creation (glossed text)*. Leiden: Leiden University Masters thesis.

FOGAÇA, JESSÉ SILVEIRA. 2011. *Fonética e fonologia da língua Makasae*. Brasilia: Universidade de Brasília Masters thesis.

FRIEDBERG, CLAUDINE. 1990. *Le savoir botanique des Bunaq: Percevoir et classer dans le Haut Lamaknen (Timor, Indonésie)*. Paris: Éditions du Muséum.

FUNDASAUN ALOLA. 2011. Man i assa [Bunak numeracy (numerals 1-10) tool]. Dili. MS. Online: http://www.scribd.com/doc/54699657/Man-i-assa.

GUNN, GEOFFREY. 2007. The state of East Timor studies since 1999. *Journal of Contemporary Asia* 37.1: 95-114.

HAJEK, JOHN, and ALEX V. TILMAN. 2001. *East Timor Phrasebook*. Hong Kong: Lonely Planet Publications.

HOLTON, GARY; MARIAN KLAMER; FRANTIŠEK KRATOCHVÍL; LAURA C. ROBINSON; and ANTOINETTE SCHAPPER. 2012. The historical relations of the Papuan languages of Alor and Pantar. *Oceanic Linguistics* 51.1: 86-122.

HUBER, JULIETTE. 2005. *First steps towards a grammar of Makasae*. Zurich: Zurich University MA thesis.

HUBER, JULIETTE. 2008a. *First steps towards a reference grammar of Makasae*. München: Lincom.

HUBER, JULIETTE. 2008b. Preliminary Remarks on Object-Marking in Makalero. *Linguistics in the Netherlands* 25: 87-96.

HUBER, JULIETTE. 2011. *A grammar of Makalero: A Papuan language of East Timor*. LOT dissertation series 279. Utrecht: LOT.

HULL, GEOFFREY. 2004a. The Languages of East Timor: Some Basic Facts (revised 24.8.2004). Online: http://www.ramelau.com/news/facts.html.

HULL, GEOFFREY. 2004b. The Papuan Languages of East Timor. *Studies in Languages and Cultures of East Timor* 6: 23-99.

HULL, GEOFFREY. 2005a. *Makasai*. East Timor language profile no. 6. Dili: Instituto Nacional de Linguística.

HULL, GEOFFREY. 2005b. *Fataluku*. East Timor language profile no. 7. Dili: Instituto Nacional de Linguística.

HULL, GEOFFREY, and ADÉRITO JOSÉ GUETERRES CORREIRA. 2006. *Disionáriu badak Makasae-Tetun, Tetun-Makasae = Disionari gi digara Makasae-Tetun, Tetun-Makasae*. Dili: Instituto Nacional de Linguística.

JOSSELIN DE JONG, J. P. B. 1937. *Oirata, a Timorese settlement on Kisar*. Studies in Indonesian culture 1. Amsterdam: N.V. Noord-Hollandsche Uitgevers-Maatschappij.

KLAMER, MARIAN, and ANTOINETTE SCHAPPER. Forthcoming. The history of 'give' constructions in the Papuan languages of Timor-Alor-Pantar. *Linguistic Discovery*.

LANGFORD, KATRINA. 2011. Local Language Council- Fataluku: Summary Report of workshop on orthography, materials development, and dictionary development. Melbourne: Monash University, MS.

LANGFORD, KATRINA. 2012. Fataluku readers. Melbourne: Monash University, MS. Online: http://www.scribd.com/livrutimor.

LEWIS, M. PAUL (ed.), 2009. *Ethnologue: Languages of the World*, Sixteenth edition. Dallas, Tex.: SIL International. Online: http://www.ethnologue.com/.

MANDALA, HALUS. 2003. Bahasa Oirata dan Relasinya Dengan Bahasa-Bahasa non-Austronesian di Timor Leste. In: I Gde Semadi Astra; Aron Meko Mbete; Ida Bagus Puja Astawa; and I Nyoman Darma Putra (eds.). *Guratan budaya dalam perspektif multikultural: Katurang ri kalaning purnabakti Prof. Dr. I Gusti Ngurah Bagus*, 662-697. Denpasar: Fakultas Sastra dan Budaya, Universitas Udayana.

MANDALA, HALUS. 2010. *Evolusi fonologis bahasa Oirata dan kekerabatannya dengan bahahasa-bahasa nonaustronesia di Timor Leste.* Denpasar: Universitas Udayana PhD thesis.

MANDALA, HALUS; ARON MEKO MBETE; NI MADE DHANAWATY; and INYO YOS FERNANDEZ. 2011. Phonological Evolution of Oirata and its Genetic Relationship with Non-Austronesian Languages in Timor-Leste. Denpasar: Ejournal Universitas Udayana. Online: http://ejournal.unud.ac.id/abstrak/halus%20mandala%20e_journal.pdf

MARQUES, TITO LÍVIO NUNES. 1990. *Metodo pratico para apreender o Makasae*. Lisbon: Self-published.

MCWILLIAM, ANDREW. 2007. Austronesians in Linguistic Disguise: Fataluku Cultural Fusion in East Timor. *Journal of Southeast Asian Studies* 38.2: 355-375.

NAERSSEN, MAAIKE VAN. 2007. *Klankverschuivingen van drie non-Austronesische talen op Oost-Timor: Fataluku, Oirata en Makasai.* Leiden: Leiden University Bachelor thesis.

NACHER, ALFONSO. 2003. Léxico Fataluco-Português: Primeira Parte. *Studies in Languages and Cultures of East Timor* 5: 138-196.

NACHER, ALFONSO. 2004. Léxico Fataluco-Português: Parte Segunda. *Studies in Languages and Cultures of East Timor* 6: 119-177.

O'CONNOR, SUE, and NUNO VASCO OLIVEIRA. 2007. Inter- and intraregional variation in the Austronesian painting tradition: A view from East Timor. *Asian Perspectives* 46.2: 389-403.

PINTO, TOMAS M. 2004. *Makalero-Makalero Disionariu*. Dili: Timor Lorosa'e Nippon Culture Center.

PINTO, TOMAS M. 2007. *Ma'u Fi Makalero Lolo*. Dili: Timor Lorosa'e Nippon Culture Center.

RIEDEL, JOHANN FRIEDRICH. 1886. *De sluik- en kroesharige rassen tussen Selebes en Papoea.* 's Gravenhage: Martinus Nijhoff.

ROSS, MALCOLM. 2005. Pronouns as a preliminary diagnostic for grouping Papuan languages. In: Andrew Pawley; Robert Attenborough; Robin Hide; and Jack Golson (eds.). *Papuan pasts: cultural, linguistic and biological histories of Papuan-speaking peoples*, 15-66. Canberra: Pacific Linguistics.

SAWARDO, P.; F. SANDA; H. JEHANE; S. H. NITBANI; and S. KUSHARYANTO. 1996. *Fonologi, morfologi, dan sintaksis Bahasa Buna*. Jakarta: Departemen Pendidikan dan Kebudayaan.

SCHAPPER, ANTOINETTE. 2010. *Bunaq, a Papuan language of central Timor*. Canberra: Australian National University PhD thesis.

SCHAPPER, ANTOINETTE. 2011a. Iconicity of sequence in the coding of source and goal in two Papuan languages. *Linguistics in the Netherlands* 2011: 101-113.

SCHAPPER, ANTOINETTE. 2011b. Crossing the border: Historical and linguistic divides among the Bunaq in central Timor. *Wacana, Journal of the Humanities of Indonesia* 13.1: 29-49.

SCHAPPER, ANTOINETTE. 2011c. Finding Bunaq: The homeland and expansion of the Bunaq in central Timor. In: Andrew McWilliam, and Elizabeth G. Traube (eds.). *Life and Land in Timor: Ethnographic Papers*, 163-186. Canberra: ANU E Press.

SCHAPPER, ANTOINETTE, and LILA SAN ROQUE. 2011. Demonstratives and non-embedded nominalisations in the Papuan languages of south-east Indonesia. *Studies in Language* 35.2: 378-406.

SCHAPPER, ANTOINETTE; AONE VAN ENGELENHOVEN; and JULIETTE HUBER. 2012. The historical relations of the Papuan languages of Timor, Kisar, Alor and Pantar. Talk given at the Conference on the history, contact and classification of Papuan languages, February 2-3, Vrije Universiteit, Amsterdam. Online: http://leidenuniv.academia.edu/AntoinetteSchapper/Talks/87127/The_historical_relations_of_the_Papuan_languages_of_Timor_and_Kisar.

STOEL, RUBEN. 2007a. The prosody of Fataluku. Paper presented at the Workshop on the Languages of Papua, August 8-10, Universitas Negeri Papua, Manokwari. Online: http://www.fataluku.com/staff/papers/Stoel_Fataluku_Papua.pdf.

STOEL, RUBEN. 2007b. Question Intonation in Fataluku. Paper presented at the Fifth East Nusantara Conference, August 2-5, Kupang, Indonesia. Online: http://www.fataluku.com/staff/papers/Stoel_Fataluku_ENUS.pdf

STOKHOF, W.A.L. 1975. *Preliminary notes on the Alor and Pantar languages (East Indonesia)*. Canberra: Pacific Linguistics Series B-43.

SUDIARTHA, I WAYAN; MADE DENES; I WAYAN TAMA; and R.B. SUPRIHANTO. 1994. *Survei bahasa dan sastra di Timor Timur*. Jakarta: Departemen Pendidikan dan Kebudayaan.

SUDIARTHA, I WAYAN; I NENGAH BUDIASA; NI LUH PARTAMI; and ANAK AGUNG PUTRA. 1998. *Struktur bahasa Makasai*. Jakarta: Departemen Pendidikan dan Kebudayaan.

VALENTIM, JUSTINO. 2001. Lia Fataluku iha Lautem. *Hakerek Kona Ba Timor Loro Sa'e* 3. Online: http://www.geocities.jp/hkbtls/hakerek/Hakerek03/lia_tetun.htm

VALENTIM, JUSTINO. 2002. *Fata-Lukunu i Disionariu*. Dili: Timor Loro Sa'e- Nippon Culture Center.

VOORHOEVE, C.L. 1975. Central and Western Trans-New Guinea Phylum Languages. In: S.A. Wurm (ed.). *New Guinea area languages and language study*, 345-460. Canberra: Pacific Linguistics.

WURM, S.A., C.L. VOORHOEVE, and K. MCELHANON. 1975. The Trans New Guinea Phylum in general. In: S.A. Wurm (ed.). *New Guinea area languages and language study,* vol. I: *Papuan languages and the New Guinea linguistic scene*, 299-322. Canberra: Pacific Linguistics.

Appendix: Timor-Kisar Cognates

The following tables give the supporting data for all the cognate sets found in the data sample. Cognates are presented from 122 sets. Fossilised morphology is indicated in brackets '( )'. Morphological variants of the same lexeme are joined with '~'. Notes on semantic innovations, notably sound changes and so forth are provided in the notes at the bottom of each table.

We also present whole reconstructed lexemes. If cognates occur in both primary subgroups (i.e., Bunaq and at least one of the other languages), then we reconstruct the lexeme to Proto-Timor. If cognates are found in both subgroups of the Eastern Timor group (i.e., Proto-Maka and Proto-Frata) but not Bunaq, then the reconstructions are bracketed indicating that they are at this stage only reconstructed to Proto-Eastern Timor. If cognates are found only within the Maka or the Frata subgroups, then we do not present a reconstruction.

	meaning	Bunaq	Makasae	Makalero	Fataluku	Oirata	Pᴛɪᴍ
1	bamboo	ma	maeri	mar	––	––	*mari
2	banana	mok	muʔu	muʔu	muʔu	mu:	*muku
3	banyan	––	sama	hama	hama	––	(*sama)
4	bat	––	––	––	maca	maṭa	––
5	bathe	wer	waruʔ	waroʔ	vahu	wau	*weru
6	beeswax	wezun a	(badu)husu b	usu(pada) b	ucu	usu	*wiDu(n)
7	bird	hos	asa	asa	aca	asa	*haDa
8	bite	gagil	gaʔel	kaʔel	––	––	*gakel
9	blood	ho	waj	wej	vehe	we	*waj
10	blow	––	gur	kut	––	––	––
11	bone	––	safa	hafa	hafa	opo	(*safa)
12	branch	esu c	sika	saka	––	––	*sika
13	carry	––	(g)afu, afu d	kafu ~ afu	apu e	ahau(re) e	(*–apu)
14	child	––	mata	mata	moco	moṭo	(*moTo)
15	clean	––	sare	hareʔ ~ –sareʔ	(pa)hal(e)	ha:l(e)	(*sale(ʔ))
16	coconut	hoza	wata	wata	vaca	wata	*wa(t, D)a f
17	cold	huruk g	gawa h	kawar h	––	––	*wogor
18	contact, (on/in)to	ata ~ gita i	(g)ata j	kata– ~ ata– j	aca j	aṭa	*–aTa k
19	cover, wear	–utu l	(g)utu	kutu– ~ utu–	utu	utu	*–utu
20	cook	tinik ~ ginik	tina	tina ~ dina	––	––	*tina(k)

a Can be used to mean either 'beeswax' or 'candle'. b Reflexes are only found in compounds meaning 'candle'. c Item shows final CV apocope. d There has been a semantic split between prefixed and unprefixed reflexes of this item: *gafu* means 'carry' and *afu* means 'with'. e Semantically bleached to mean 'with'. f The variation in the reconstructed medial consonant is because the Oirata form is irregular (expected *waṭa 'coconut'). g This form results from medial–final consonant metathesis: *wogor > *hogor > *horok (devoicing phonotactically required) > huruk. h These items have been semantically extended to mean not just "cold" but also "wind"; see also item 125; in Makasae and Makalero, the forms result from medial–final consonant metathesis: *wogor > *wagar > *gawar > Mks *gawa*, Mkl *kawar*. i Used synchronically to introduce a goal argument. j Used synchronically to denote a location with which a participant is or comes into contact. k This item appears to have had very general semantics coding goals and contact relations. l This item has semantically been semantically bleached and is now used as a medial verb denoting an "accompanier".

	meaning	Bunaq	Makasae	Makalero	Fataluku	Oirata	Pᴛɪᴍ
21	cuscus	zulo [a]	––	––	acur(u)	aʈur(u)	*(a)Dura
22	cut	––	lasi	lasi	laʔi	lai	(*lasi)
23	crawl, creep	el [b]	––	––	er(eke)	––	*er(ek)
24	clew, stone circle	mot	––	––	matar(u) [c]	matar(a)	*matar
25	die	–ume [d]	umu	umu	umu	umu	*–umV
26	difficult	––	sege	heke ~ –seke	heʔe	heʔe [e]	(*sege)
27	dog	zap	defa	sefar	ipar(u) [f]	ihar(a) [f]	*Depar
28	dry	saʔ	sahar	haʔat ~ saʔat	sasa(ke)	hatat(e)	*sasar
29	ear	––	wala(ku:)	wali	vali	wali	(*wali)
30	earth	muk	muʔa	muʔa	muʔa	mua	*muka
31	edge	––	wali	walir	vali(ku)	––	(*walir)
32	eucalypt	tal [g]	dara	daraʔ	cara(nu)	––	*dara(ʔ)
33	egg	–ut	––	uta [h]	ucu	uʈu	*uTa
34	face	–ewen	fanu	fanu	fanu	panu	*fenu
35	fantail	––	sigilai	(taru)sikiloti	––	––	––
36	finish	haʔa(l) ~ gaʔa(l)	saʔi	haiʔ ~ –saiʔ	sai	––	*sa(ʔ)i [f]
37	fire	hoto	ata	ata	aca	aʈa	*haTa
38	fish	––	afi	afi	api	ahi	*api
39	fly	lore	lor	lor	––	––	*lore
40	for	hos ~ gos [i]	(g)au [j]	kasu ~ asu	ahu ~hau [k]	––	*–asu

[a] Semantic shift to mean 'civet'; synchronically, the related form *zulet* means 'cuscus'. [b] Final VC lost through apocope. [c] Semantic shift to mean 'stone'. [d] Syntactic change to transitive verb meaning 'kill, murder'; see Bunaq *heser* 'dead' in 89. [e] Semantic shift to mean 'oppressed, obstructed'. [f] There has been irregular loss of *D before the high vowel here. [g] Refers to *E. alba*, whereas the cognate items in other languages refer to *E. urophylla*. [h] Semantic flip to mean 'pea, bean'; see 74 for 'egg'. [i] A glottal stop can in general not be reconstructed amongst Timor–Kisar languages; hence its tentative status here. [j] As a full verb means to 'wait for', but medially to introduce a participant for whom something is done in their absence. [k] Irregular loss of medial *s.

	meaning	Bunaq	Makasae	Makalero	Fataluku	Oirata	PTIM
41	garden	mar	ama	ama	––	uma	*(u)mar
42	girl	pana [a]	fana(rae)	fana(r)	fana(ru)	pana(rai)	*fana
43	give	–ini [b]	gini [c]	–ini ~ kini [d]	ina	ina	*–inV
44	go	mal	mara	mara	mara	mara	*mara
45	green	ugar	uʔur	huʔur ~ uʔur	uʔur(eke)	uʔul(e)	*ugar
46	hand	ton [e]	tana	tana	tana	tana	*tana
47	head	––	daʔe	saʔe	caʔu	ʈau	(*Dake)
48	hear, understand	mak	maʔen [f]	maʔen [f]	––	––	*make(n)
49	heavy	––	tiʔir	tiʔir	ciʔir(e)	tuːr(e)	(*Tiʔir)
50	_P. buceroides_	koak	kaukua	fula	folo kua	––	*fula–kaukoa [c]
51	hit	––	nehe	nese–	––	––	––
52	hold	––	sifaʔ ~ difaʔ	hifaʔ ~ –sifaʔ	––	––	––
58	holy, taboo	por	falun	falun	––	––	*falu(n)
53	impact, contact with	(g)ene [g]	(g)ene	kene ~ ene	––	––	*–ene [e]
54	inside, (with)in	mi(l)	mu(tu)	mu(tu–)	mu(cu)	mu(ʈu)	*mi
55	kill	–ita [h]	(g)uta	kuta ~ uta	uca ~ uta [i]	uʈa	*–uTa
56	laugh	higa(l)	hiʔa	hiʔe	––	––	*hika
57	leaf	––	asa	hasa	asa	asah(a)	(*hasa)
59	lie	––	logo	loko	––	––	––
60	live	––	lafuʔ	lafuʔ	lau(he) [j]	lau(se) [j]	(*lapuʔ)

[a] Semantic shift to 'woman'. [b] Grammaticalised into a causative verb 'make'. [c] Used as both 'give' and a causative verb 'make'. [d] Semantic shift to mean 'branch'. [e] Means 'understand, know'. [f] We reconstruction a doublet lexeme here due to the presence of both "halves" across more than one subgroup. [g] This reflex has been semantically bleached and is synchronically a postposition used to encode locations; the initial /g/ is a fossilised prefix. [h] This verb appears to have had a very general meaning such as 'impact', 'make/be contact with' and 'adversely touch'. [i] This verb has semantically specified to mean 'hit'; the form _uta_ is an irregular reflex used for plural patients, and _uca_ is the expected reflex used for singular patient. [j] The loss of *p is irregular; it may be conditioned by fossilised morphology.

	meaning	Bunaq	Makasae	Makalero	Fataluku	Oirata	Pᴛɪᴍ
61	long, tall	esen [a]	asan	asan	––	––	*esan
62	long (time)	––	muʔir	muʔit	muʔit(e)	––	(*mukir)
63	meat	sael [b]	seu	seur	leura [c]	leura [d]	*seor [e]
64	mountain 1	op	afa	(war)afa [f]	apa	––	*apa
65	mountain 2	lolo	lari [g]	larin	––	––	*laru(n)
66	moon	hul	uru	uru	uru	uru	*huru
67	mouth	–agar	aʔa	haʔa ~ aʔa	oʔo	oʔo	*aka(r)
68	name	–ini(l)	naj	nej	ne	ne:(ne)	*–nei
69	narrow	muel [h]	––	––	male(te)	male(te)	*mole
70	new	tip [i]	sufa	hofar	––	––	*(t, s)ifa
71	old	legul [j]	laʔir	laʔit	laʔit(e)	lait(e)	*lakir
72	owl	tupi	tutuguru [k]	toutou [k]	tupukuru	––	*tupi–guru [l]
73	path	hik	hiʔa	hiʔa	iʔa	ia(ra)	*hika
74	pea, bean	ho	wa: [m]	wa [m]	––	wa	*wa
75	pierce, pound	––	––	tafa ~ dafa	tafa	––	(*tafa)
76	pig	––	baj	paj	paj	haj	*baj
77	*P. indicus*	mazoʔ [n]	mater	mater	matar(ia)	––	*ma(t,T)ar
78	price, expense	bol	bura [o]	pura [o]	pura [o]	hura	*bura
79	rare	lesa	––	lesa [p]	––	––	*lesa
80	rat	zul	dura	sura	cura	ʈura	*Dura

[a] This item has semantically shifted to mean 'high, above, upper'. [b] This verb has semantically specified to mean 'shoot'. [c] Semantically specified to mean 'pig'. [d] Irregular reflexes; the initial /l/ is unexpected. [e] This probably originally meant both 'meat' and 'wild game'. [f] This item only occurs in a lexical doublet meaning 'sacred place on top of a mountain'; the first, non–cognate item in the compound, *war*, means 'stone'. [g] Semantic specification to mean 'slope'. [h] This item displays vowel–liquid metathesis. [i] This form is irregular; initial /t/ is unexpected, perhaps it was conditioned by the following high vowel. [j] Semantic shift to mean 'long, tall'. [k] These items show harmonization of the original medial bilabial consonant, *p, with the initial /t/ consonant; see also 83. [l] We reconstruct a doublet lexeme here due to the presence of both "halves" across more than one subgroup. [m] Means 'egg'; semantic flip with *uta* 'bean, pea'; see 33. [n] This form is irregular; a medial *t is expected. [o] Semantic shift to a verb meaning 'sell'. [p] Semantic shift to 'not exist'.

	meaning	Bunaq	Makasae	Makalero	Fataluku	Oirata	Pᴛɪᴍ
81	RECP	tV– [a]	ta	ta	(ni)ta	ta	*ta
82	right–hand	––	tane	tanen	tenen(u)	tenen(e)	(*tanen)
83	run	tʃiwal [b]	ditar [c]	titar [c]	tifar(e)	tipar(e)	*tifar
84	scared	––	––	masan	mahan(e)	man(e)	(*masan)
85	sea	mo	––	––	mata	mata [d]	*mata
86	see	haza(l) ~ gaza(l)	––	––	aci	asi	*aDa
87	seek	saga(l) ~ gaga(l)	haga	haka ~ –saka	haʔa [e]	(aʈa)hoko	*saga
88	shoot, release	holi ~ goli [f]	suri ~ duri	huri ~ –suri	suri(te) [g]	––	*suri
89	sick	heser [h]	sisir	hisit ~ –sisit	––	––	*seser
90	sing, howl	––	kaul	kaul	kol(e)	––	(*kaul)
91	sit.PL	––	diar	diar	cuar(e)	rua [i]	(*diar)
92	sit	mit	mit ~ mi [j]	mit	(i)mir(e)	mir(e)	*mit 'sit'
93	small	––	kaʔu(lai)	kaʔu	kaʔu(sila)	––	(*kaʔu)
94	spoon	sulu	sulu	hulu	hula	––	*sula
95	sleep	tʃier	taʔe [k]	tia	taia	taja [l]	*tia(r)
96	stand	net	na [j]	nat	(a)nat(e)	nat(e)	*nat
97	stand.PL	––	nahar	naser	neher(e)	––	(*naser)
98	star	bi [m]	ifibere	ifi	ipi(naka)	ihi	*ipi–bere
99	stone	hol	––	uar	––	uar(aha)	*war
100	sugarcane	up	ufa	ufa	upa	uha	*upa

[a] The vowel of this innovative prefix has become unspecified, harmonising with the first vowel of the root to which the prefix attaches. [b] Semantic shift to mean 'run away, flee'. [c] These items show harmonization of the original medial bilabial consonant, *p, with the initial /t/ consonant; see also 72. [d] Semantic shift specifying to mean "sea which is sheltered and quiet". [e] Semantic shift to mean 'see'. [f] Semantic shift to mean 'hunt'. [g] Semantic shift to mean 'squirt'. [h] Semantic shift to mean 'die, dead'; see also 25. [i] This form is irregular, having undergone consonant metathesis (rua < *ruad < *duar < *diar) in the process of which the reflex of initial *d has been lost. [j] The (optional) deletion of the final consonant on these posture verbs is irregular. [k] The glottal stop here appears to be the result of a phonemicisation of a non–phonemic segment epenthetically inserted between syllables after metathesis. [l] The glide here appears to be the result of fortition of original /i/ as synchronically present in Fataluku. [m] We reconstruction a doublet lexeme here due to the presence of both "halves" across more than one subgroup. Apocope of final CV has occurred in this item in Bunaq.

	meaning	Bunaq	Makasae	Makalero	Fataluku	Oirata	Pᴛɪᴍ
101	sun	hot	watu	watu	vacu	waʈu	*waTu
102	sweat	sil	––	––	her(u)	––	*sir
103	swell	pe ᵃ	fatu	fatu	fatu	patu	*fatu
104	tail	–uloʔ ᵇ	ula	ula	ula(fuka)	ula(pua)	*ula(ʔ)
105	take	––	ma ᶜ	mej	me	me	(*mej)
106	tick, leech	kata	––	––	kocu	ko:ʈo	*koTu
107	tie	ti	sil	sil	(i)sil(e)	i:l(e)	*(t, s)il
108	throw	––	liʔan	liʔan	(li)liʔen(e) ᵈ	len(e)	(*liken)
109	together with	––	––	––	acu	asu	––
110	tooth	–ewe ᵉ	wasi	wasi	vahi(nu)	wai(ni)	*wasi
111	tongue	–up	ifi	ifil	epul(u)	uhul(u)	*ipul
112	top, above	wa	gua	kua– ~ wa–	––	––	*wa
113	tree	hotel	ate	ate	ete	ete	*hote
114	urinate	sele ᶠ	––	irih ᵍ	iris(e) ᵍ	iris(e) ᵍ	*rese
115	vagina	––	aru	aru	aru	aru	(*aru)
116	wake (intrans.)	––	dur	dur	cur(e)	ʈu:r(e)	(*dur)
117	wake (trans.)	otin ʰ	tane	tane	tani ~ cani	––	*Tani
118	walk	lagor	laʔa	laʔa	laʔa	lara ⁱ	*lakor
119	wall	sirin	sidi	hir	––	––	*siri(n)
120	war	asu (en) ʲ	sala	hala	hal(u)	al(a)	*(a)solo
121	water	il	ira	ira	ira	ira	*ira
122	weave, plait	sien ᵏ	sina	hina	hina	––	*sina
123	where	teo	––	tau–	te	te(nai)	*teu
124	wife	––	da	sa	––	––	––
125	wind	bel	––	pare ˡ	pari	hari	*beri
126	3	g–	g–	k–	––	––	*g–
127	3.ALIEN	gie	gi	ki	i	ue	*gie

ᵃ Apocope of final CV has occurred in this item. ᵇ The origin of the glottal stop is unknown. ᶜ This form is irregular (expected *maj), probably due to phonetic reduction with grammaticalisation into a case marker. ᵈ The initial extra /li/ appears to be the result of CV reduplication. ᵉ Apocope of final CV has occurred in this item. ᶠ This form shows consonant metathesis: sele < *lese < *rese. ᵍ These appear to show vowel metathesis: iris < *risi < *rese. ʰ This form shows metathesis: otin < *toni < *Tani. ⁱ This form is irregular; expected is *laʔa. ʲ This reflex is only found in this compound meaning 'warrior' with the noun *en* 'person'. ᵏ This form shows metathesis: sien < *sine < *sina. ˡ Semantic shift to mean 'dry monsoon'; see also item 17.

Endnotes

[1] The authors would like to thank the participants of the Conference on history, contact and classification of Papuan Languages, especially Mark Donohue, Andy Pawley and Edgar Suter, for their helpful and insightful comments. Schapper's research was supported by grant 08-EuroBABEL-OP-O25 from De Nederlandse Organisatie voor Wetenschappelijk Onderzoek. The following abbreviations are used: ALIEN 'alienable possessor'; AN 'animate gender'; EXCL 'exclusive'; INCL 'inclusive'; P 'patient'; PAP 'Proto-Alor-Pantar'; PETIM 'Proto-Eastern Timor'; PFRATA 'Proto Fataluku-Oirata'; PMAKA 'Proto-Makasae-Makalero'; PTAP 'Proto-Timor-Alor-Pantar'; PTIM 'Proto-Timor'; RECP 'reciprocal'; TAP 'Timor-Alor-Pantar'; TNG 'Trans-New Guinea Phylum'; 1 'first person'; 2 'second person'; 3 'third person'.

[2] Makuva has since been identified as Austronesian (see e.g. Engelenhoven 2010c).

[3] In Capell's and Almeida's materials, Fataluku is referred to as "Dagadá" or "Dàgadá". We use "Fataluku" here as it is the preferred name of community.

[4] Gunn (2007) calls this period an 'ethnographic gap' in which social research was positively discouraged by the Indonesian rulers.

[5] The materials on Makasae Fatumaka are online at: http://www.ling.hawaii.edu/ldtc/languages/makasae_fatum/. The materials on Makasae Osoroa are online at:
http://www.ling.hawaii.edu/ldtc/languages/makasae_ osor/.

[6] These materials on Fataluku are online at: http://www.ling.hawaii.edu/ldtc/languages/fataluku/.

[7] Bahasa Alor is the only Austronesian language spoken on the islands (with the exception of the modern lingua franca, Malay, and a recent immigrant community of Sama Bajau on Pantar). It is spoken exclusively in small pockets on the northwestern coasts of Alor and Pantar and has not had a significant impact on the mountain-dwelling Papuan language speakers.

[8] Note that, whilst for the purposes of ease of comparison in this paper we assume the existence of the PTIM node, it does in fact remain to be seen exactly where PAP fits in the overall family tree. See section 7 for further discussion of this point.

[9] The Ossu dialect differs from all other Makasae dialects in having merged *f and *p to /p/.

[10] Cognates in AP languages also show no sign of this final /l/, e.g., PAP *-ain(i,u) 'name', *mi '(be) in/on' and *tei 'tree'.

[11] Consonant final items in Makasae and Makalero occur with a final echo vowel when the lexeme occurs in phrase-final position. This phenomenon in Makasae and Makalero is very different from the final vowels we observe in Fataluku and Oirata. The final vowel's appearance is entirely predictable in Makasae and Makalero and cannot be said to be part of the underlying representation of the lexeme. Fataluku and Oirata show no such predictability in the form of the vowel or in the position of their syntactic appearance.

[12] Several such doublets are also reconstructed to PTIM; see the Appendix for details.