Lio and the Central Flores languages

Alexander Elias

A thesis submitted in partial fulfilment of the
requirements for the degree of
Research Master of Arts
in
Linguistics
by
Alexander Elias

Student ID: s1915444
Supervisor: Prof. dr. M.A.F. Klamer
Second Reader: Dr. E.I. Crevels
Date: 13th November, 2018
Faculty of Humanities, Leiden University

Universiteit
Leiden

Abstract

This thesis consists of two interconnected parts: a synchronic section dealing with Lio phonology, and a diachronic section dealing with the internal relations of the Central Flores language group, of which Lio is a member.

The first section is a description of the phonetics and phonology of Lio (Austronesian), a language spoken in Flores, an island in the Lesser Sunda island chain of eastern Indonesia. I describe the phonemic inventory, phonotactics, stress system and adaptation of loanwords into Lio. This is based on fieldwork carried out in Central Flores in July-August 2017 which focused mainly on Lio. This is a contribution to the state of linguistic documentation in Central Flores, which remains relatively poorly documented. This will also set the stage for the second part of the thesis, because Lio is an important language for reconstructing aspects of Proto-Central Flores.

The second section is a historical analysis of the relations of the Central Flores languages, and a reconstruction of Proto-Central Flores. I present evidence that the Central Flores languages form a valid innovation-defined subgroup, which underwent a period of splitting and isolation at the level of Proto-Central Flores. Then I address the internal relations of the Central Flores group and the process of differentiation from Proto-Central Flores to the modern Central Flores languages. Lio is one of the more conservative members of the Central Flores group, and is crucial for distinguishing the reflexes of certain Proto-Central Flores phonemes. The Central Flores group forms a linkage, with patterns of intersecting isoglosses which are not easily captured in a tree diagram. Therefore, the findings of this section will be cast in the framework of Historical Glottometry, a wave model-based methodology which is better equipped to represent and model the relations holding between linkages.

Acknowledgements

I would like to thank the following people who made this thesis possible.

First and foremost, I would like to thank Tedyis Sofa for all of her help. This includes all the time spent on transcribing and explaining the meanings of Lio words, as well as the great company and delicious food I shared with her. It was thanks to her that I felt so welcomed to Flores and integrated with the community, and this research would not have happened without her. I hope to one day be able to return the hospitality that she showed me, and I hope we remain in touch for a long time to come.

My next acknowledgement is to Edy Dole and his mother Madalena, as well as to his wife and children. Edy was one of the first people I met in Flores and was happy to welcome me into his family, leading to the greatly productive time I spent recording with Madalena. He showed me life in his village of Kanganara and we spent many great evenings together. I am proud to call him *ka'é* (older brother) and to be his *aji* (younger brother).

I would like to particularly acknowledge the following people I met in Flores for their friendship and support in large and small ways: Mama and Bapak Fanny, Mama and Bapak Gilbert, Pak Frumen Dole, Pak Alfons Mbuu, Pak Stan Jaya, Pak Gregorius Bito, Pak Dominikus, Ibu Vero, Pak Sipri, Pak Ali, Bartolomeus and his family, Eriko Aoki, Satoshi Nakagawa. There are many people whose names are not listed who I remember fondly from my time in Flores and who made it the adventure that it was.

I would like to acknowledge Dr. Marian Klamer, my supervisor, for her support and confidence in me, and for guiding me through the process of planning and carrying out this plan. I would also like to thank Antoinette Schapper for helping me settle on central Flores as a fieldwork destination, and I am grateful to Owen Edwards, Hanna Fricke and George Saad for their linguistic input, friendship and advice at many stages of this research.

I would like to thank Pak Wayan Arka (although I had not met him at the time) for putting me in touch with many useful local contacts in Flores. I would also like to thank Pak Yosef Bataona (who I have still not met in person) for helping me with my visa application and introducing me to the people in Ende who would turn out to be crucial to the success of this thesis.

Finally, I would like to acknowledge the funding bodies which made this research possible. The Leids Universiteits Fonds generously provided a 1100 euro LISF grant (LUF International Study Fund) which covered my international travel costs and living costs in the field. A second source of support was a 300 euro contribution from Marian Klamer's research project 'Reconstructing the past through languages of the present: The Lesser Sunda Islands' which covered the cost of internal flights and transport within Indonesia. Together, these funding sources made this project possible and I thank them for their support.

List of Tables

Abbreviations

The abbreviations used in this thesis are listed below.

1	first person
2	second person
3	third person
1PL.E	1st plural exclusive
1PL.I	1st plural inclusive
ADV	adverbial
C:NAG	Central Nage (Boawae)
C:NGD	Central Ngadha (So'a)
CF	Central Flores
CL	classifier
COM	comitative
COND	conditional
CONJ	conjunction
DEM	demonstrative
DISTR	distributive
E:NAG	Eastern Nage (Nangamboa)
E:NGD	Eastern Ngadha (Takatunga)
E:NGO	Eastern Nga'o (Oja)
EMPH	emphatic
END	Ende
EXCLAM	exclamation
EXI	existential
F	feminine
FUT	future
HG	Historical Glottometry
IMP	imperative
INTENS	intensifier
IPA	International Phonetic Alphabet
ITER	iterative
KEO	Kéo (Udiworowatu)
LIO	Lio
LOC	locative
M	masculine
MAN	Manggarai
NAG	Nage
NEG	negative
NGD	Ngadha
NGO	Nga'o
NOM	nominalizer
PAL	Palu'e
PAN	Proto-Austronesian
PCF	Proto-Central Flores
PCMP	Proto-Central Malayo-Polynesian
PERF	perfective
PL	plural
PMP	Proto-Malayo-Polynesian
PROH	prohibition
RECIP	reciprocal
REL	relativizer
RON	Rongga
SG	singular
SIK	Sika
SVC	serial verb construction
VOT	voice onset time
W:NGD	Western Ngadha (Bajawa)
W:NGO	Western Nga'o (Watumite)

Contents

1 Introduction

1.1 Aim of research

This thesis consists of two sections. In the first, I will describe the phonetics and phonology of Lio as spoken in Detukeli, based on fieldwork carried out in 2017. In the second section, I will compare Lio with the other Central Flores languages, reconstructing aspects of Proto-Central Flores and analyzing the historical relationships which hold between them. The second section is based on both my own fieldwork, as well as data collected from published sources on other CF languages.

This thesis makes a contribution to the linguistic literature because Lio and the other Central Flores languages are still very poorly described. While there is a Lio-German dictionary available (Arndt 1933), and an Indonesian-language bachelor's thesis on Lio grammar (Levi 1978) this is the first work in English which treats aspects of the structure of Lio. I will provide a synchronic analysis of Lio phonology, supplemented by glossed Lio texts for future researchers to consult. Central Flores has lagged behind eastern and western Flores in the availability of descriptive materials, and eastern Indonesia as a whole remains rather poorly known to linguists despite its great linguistic diversity.

This thesis will also contribute to Austronesian historical linguistics because there has not yet been an in-depth study of the relations holding between the Central Flores languages. Fernandez (1996) is an attempted classification of all languages of Flores, but the scope is larger and it takes in a larger sample of languages than that considered here. This thesis seeks only to show that the Central Flores languages form an innovation-defined subgroup and to analyze the internal relations among them.

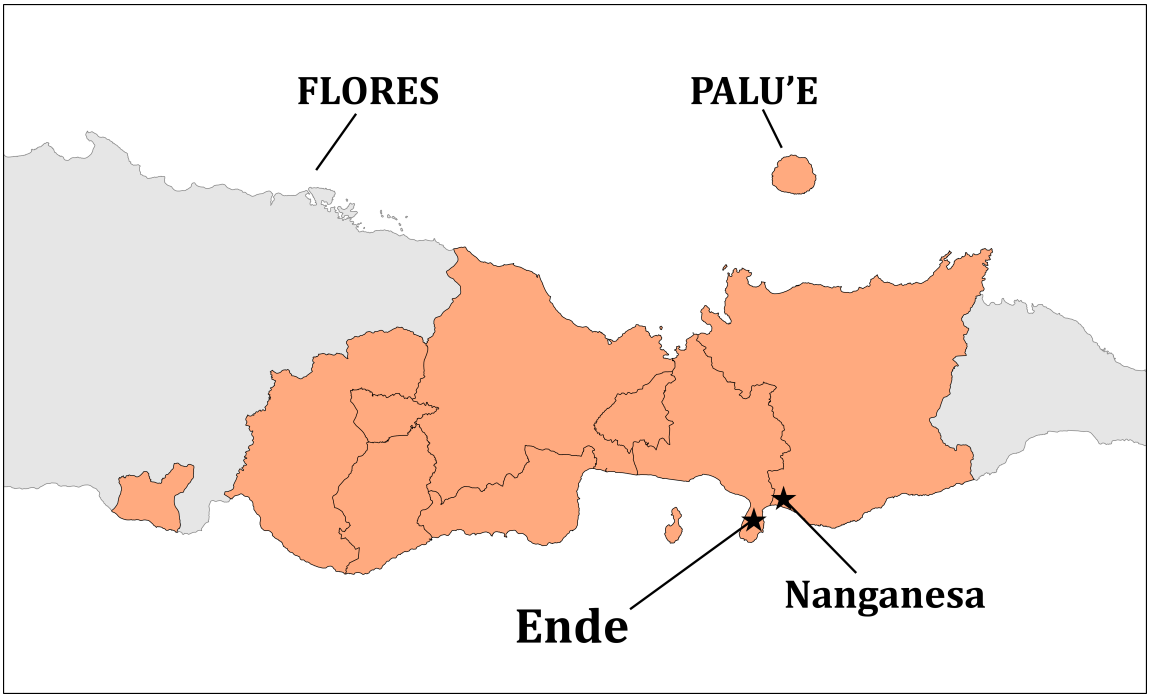
The Central Flores group is notable for being a linkage, meaning that the usual tree diagram which is used to represent language history does not capture the relations between languages well. I have chosen to use the recent framework of Historical Glottometry (a methodology based on the wave model, put forth in François 2014, François 2017, Kalyan and François 2018) in order to characterize the relations between CF languages in this work, rather than rely on the tree model. Historical Glottometry is still a work in progress and has yet to see wide application to the analysis of non-tree-like language groups. Thus, this thesis will also contribute to the development and implementation of Historical Glottometry as a viable alternative to the tree model.

1.2 Fieldwork methodology

1.2.1 Fieldwork situation

The total amount of time that I spent in the field was 61 days (July - August 2017). When I arrived in Indonesia on the 2nd of July, I spent my first week in Yogyakarta, Java, taking Indonesian and Javanese classes and getting accustomed to life in Indonesia. I arrived in Ende, Flores on the 8th of July, and I remained in that area until the 28th of August,

Figure 1: Map indicating the islands of Flores and Palu'e, and the locations of the towns of Ende and Nanganesa on Flores. Areas of Flores which are colored in are those which constitute Central Flores linguistically, while Palu'e is the closest relative of the Central Flores languages



with the exception of a two-day trip to Singapore for visa renewal purposes (30 July - 1 August).

Ende is one of the largest cities in Flores, with around 100,000 inhabitants (2010 Indonesian census). It is home to a small airport, a university and the only soccer stadium in Flores. Ende lies on the border between two language areas: the eponymous Ende language is spoken to the west of the city and in the city proper, while the Lio-speaking area lies just to the east, and stretches right up the outskirts of the city itself.

On the eastern outskirts of Ende, across the Wolowona River, lies the village of Nanganesa. It is about a half hour's walk from Ende town center and has a more rural feel. Nanganesa lies in the Lio speaking area but contains a mix of different language users because of its proximity to the urban area and work opportunities. I rented a room in a boarding house in Nanganesa, and spent most of my days in a communal courtyard nearby transcribing with my main language helper, Tedyis Sofa. She was a 37 year-old Lio speaker from the village of Wolopau who also spoke excellent English, in addition to Mandarin and Cantonese (having worked abroad in Australia and Hong Kong).

While Lio was the primary focus of my fieldwork, I collected some data on other languages in the region as well. Since I was near the city of Ende, I took advantage to record some data in Ende. I also met speakers of Sika, Ngadha, Nga'o and Nage in my time there, so I took the opportunity to elicit data in those languages. I went to the village of Nangamboa to record wordlists and texts in Eastern Nage, and I located a group of five houses in Ende where everyone came from two Nga'o-speaking villages called Watumite and Oja (Tendambepa) in order to collect data on Nga'o.

1.2.2 Participants and recordings

The bulk of my recordings were recorded from Madalena Lero, a 62 year-old Lio speaker from Detukeli. She was the mother of Ferdinandus Dole, a professor of education at the University of Flores. Ferdinandus (or Edy) helped me out greatly with logistics and I was soon introduced to his family, including his mother. Madalena had a thorough knowledge of traditional Lio practices and history and was always willing to make recordings. Therefore, the results presented in this thesis are based substantially on the speech of Madalena Lero.

Another rich source of Lio data is the Lio prayer book *Jala da Ghéta Surga*, a collection of Catholic prayers. Some are translated into Lio but others are clearly original compositions by a native speaker, although there is no authorship attribution. Tedyis was kind enough to read the entire book out loud (which took nearly 6 hours in total), providing good material for phonetic and phonological analysis but with a few intonational problems due to the nature of reading out loud.

Table 1 is a complete list of the recordings produced during my time in Ende, along with the speaker, length, genre, language, and whether or not transcription was completed before I left. The genres represented include: wordlists, counting, historical texts, personal anecdotes, frog stories (elicited with the wordless picture book Frog, Where Are You? by Mercer Mayer), prayers (drawn from the Lio prayer book *Jala da Ghéta Surga*), procedural texts (such recipes and instructions for performing rituals), ritual language (performed by a skilled orator on a ceremonial occasion)

The abbreviations for speakers are as follows: ML (Madalena Lero), 62 years old, female Lio speaker; SJ (Stan Jaya), 42 years old, male Lio speaker; TS (Tedyis Sofa), 37 years old, female Lio speaker; AM (Alfons Mbuu), 42 years old, male Ende speaker; BM (Bapak Mbuu, father of AM) first name and age unknown, male Ende speaker; UM (Urbanus Mbu'u), upwards of 60 years old, male Nage speaker; BB (Bapak Bartolomeus), first name and age unknown, male Nage speaker; VM (Veronika Mbewu), 53 years old, female Nga'o speaker; DD (Dominikus Dei), 59 years old, male Nga'o speaker; BS (Bapak Siprianus), age unknown, male Nga'o speaker; AF (Aurelius Fredimento), 39 years old, male Sika speaker; YF (Yosef Fra'e), 36 years old, male Ngadha speaker; GB (Gregorius Bito), 38 years old, male Ngadha speaker.

Table 1: List of recordings

	Filename	Speaker	Time	Genre	Lang	Transcribed
1	Lio Bridewealth	ML	23:02	Procedural	LIO	Yes
2	Lio Frog Story 1	ML	13:45	Frog Story	LIO	Yes
3	Lio Frog Story 2	SJ	19:20	Frog Story	LIO	Yes
4	Gawi Kelimutu	Multiple	1:43	Ritual	LIO	Yes
5	Lio History	ML	14:48	History	LIO	Yes
6	Lio Numbers	SJ	5:25	Counting	LIO	Yes
7	Lio Positional	SJ	12:20	Positional	LIO	Yes
8	Lio Prayer Book	TS	350:20	Prayers	LIO	Yes
9	Lio Sambal Recipe	ML	2:48	Procedural	LIO	Yes
10	Lio Scorpion Story	ML	16:02	Anecdote	LIO	Yes
11	Ende Frog Story	AM	19:24	Frog Story	END	Yes
12	Ende Positional	AM	12:54	Positional	END	Yes
13	Ende Numbers	AM	5:13	Counting	END	Yes
14	Ende History 1	BM	31:57	History	END	No
15	Ende Building Prayer	BM	0:34	Ritual	END	Partly
16	Ende Blessing	BM	2:05	Ritual	END	Partly
17	Ende History 2	BM	1:58	History	END	No
18	Nage Wordlist	UM	57:28	Wordlist	E:NAG	Yes
19	Nage Frog Story	BB	2:30	Frog Story	E:NAG	No
20	Nage History	UM	30:44	History	E:NAG	No
21	Nage Farm Rituals	UM	7:41	Procedural	E:NAG	No
22	Nage Scary Story	UM	11:04	Anecdote	E:NAG	No
23	Nga'o Wordlists	VM, DD	79:09	Wordlist	W:NGO, E:NGO	Yes
24	Nga'o Farm Rituals	BS	58:22	Procedural	E:NGO	No
25	Nga'o Frog Story	DD	6:31	Frog Story	E:NGO	No
26	Nga'o Scary Story	VM	5:55	Anecdote	W:NGO	No
27	Sika Frog Story	AF	17:16	Frog Story	SIK	No
28	So'a Wordlist	YF	68:42	Wordlist	C:NGD	Yes
29	Bajawa Frog Story	GB	7:50	Frog Story	W:NGD	No

1.2.3 Data collection, analysis and storage

All recordings were made on a Sony HDR-PJ350 camcorder. Once the recording equipment was set up and ready to record, I would give the participant a prompt, such as 'Tell me a story that happened to you as a child' or 'Tell me about the history of this area,' and record the resulting narrative. I did not rely much on structured elicitation, because the language is morphologically very simple and there were no paradigms to exhaustively fill in. However, I did ask people to simply count up from one to one hundred, and I used the Frog Story (presented on my laptop's screen) as an elicitation tool in a number of cases.

The resulting recordings were saved on a hard drive and transcribed with the help of Tedyis. Transcriptions in languages other than Lio were made with the help of the original participant who recorded the text. I made the transcriptions and glossed them in Excel as it was much faster to run on my laptop than ELAN, but the drawback is that they are not synchronized with the videos.

Upon my return to Leiden, the recordings were placed into an archive storing materials related to Marian Klamer's current NWO Vici grant project 'Reconstructing the past through languages of the present: The Lesser Sunda Islands'. In addition, the wordlists which I collected were incorporated into LexiRumah, a lexical database of languages of the Lesser Sunda Islands which is one part of the research output of Marian Klamer's project.

1.2.4 Ethics

Informed consent was obtained verbally from all participants in recording sessions. This was obtained by asking their permission to record them, to store the recordings, to show the recordings to others, and to include the contents of the recording in any research results. Consent was never withheld, and all participants were enthusiastic about participating in such a research project. As such, I felt it would be inappropriate to introduce a written consent form into what was an otherwise relaxed and informal recording situation. This would have been potentially intimidating, especially to participants with a low level of literacy. In some cases, the statement of consent is included at the beginning of the recording, but in other cases I had not yet turned on the recording equipment when consent was obtained.

2 Lio phonetics and phonology

I will begin the first section of this thesis by giving an overview of Lio's typology (section 2.1) and the orthography specific to Lio which will be used in this section (section 2.2). I will then cover the Lio consonant inventory (section 2.3) and vowel inventory (section 2.4) from a phonetic and phonological standpoint. I will describe Lio phonotactics (section 2.5) and briefly touch on Lio stress (section 2.6), although much more remains to be said about stress. Finally, I will cover the phonology of loanwords in Lio (section 2.7).

2.1 Typological overview of Lio

Typologically, Lio stands out from its Austronesian relatives because it is highly isolating. There are no bound affixes whatsoever in Lio. Grammatical relations are marked by constituent order, which is SV/AVO in an unmarked clause. However, as in many isolating languages, ellipsis is extremely common and rearrangement of the constituents is possible for pragmatic reasons. Context is therefore crucial to disambiguate grammatical relations, and a Lio sentence taken out of context is often highly ambiguous.

Lio is an excellent example of a serializing language. Serial verb constructions (SVC's), consisting of a single predicate made up of more than one verb, play a very important role at all levels of Lio grammar. Serial verbs can refine the meaning of a main verb and thus allow the derivation of new meanings from the stock of existing verbs. Serial verbs often serve to mark the function of peripheral constituents, playing a role similar to prepositions or case marking clitics/affixes in other languages. Serial verbs can add arguments in causative or benefactive constructions, or can provide directional specification to a verb involving motion.

Word classes in Lio are rather flexible, in two senses. First, each word class has a broad range of grammatical uses; for instance, nouns can serve as the head of a predicate in Lio, taking almost the full range of markers available to verbs in that position. Second, each lexeme is often assigned to multiple word classes; for instance, the form /ləma/ is a noun meaning 'tongue', a verb meaning 'to say', and a numeral classifier which serves to count units of speech. This flexibility both on the grammatical and lexical level (not to mention its extremely isolating profile) makes Lio a particularly difficult case for the establishment of word classes. However, there is no need to posit 'pre-categorial lexemes' or the absence of word classes entirely; with a bit of careful analysis, it is possible to disentangle the complicating factors.

Lio has a rich system of numeral classifiers: these are morphemes which must appear in numeral phrases and categorize objects according to various properties when counting. For instance, *kolo* is used with humans, *éko* with land animals, and *esa* serves as a general default classifier. There are several hundred such classifiers, but there are fewer than 10 common ones which account for majority of instances. This system resembles the numeral classifiers found in isolating South-East Asian languages such as Vietnamese, and is rather

unusual in an Austronesian context.

These typological features are shared with the other CF languages, which differ mainly in lexicon and phonology.

2.2 Orthography

When writing Lio, speakers use a conventional system based on Indonesian orthography but this has no official status. This orthography is well-suited for representing Lio phonology and will be used in this thesis, but several distinctions (such as that between /e/ and /ə/) are often overlooked in practice. The Lio-German dictionary compiled by the German missionary Paul Arndt (1933) uses its own conventions, which are described in case the reader needs to refer to the dictionary. The few other extant publications in Lio employ varying orthographic conventions which resemble Arndt's system rather than the current conventional orthography.

IPA Value	Conventional	Arndt (1933)
/a/	a	a
/e/	é	é
/ə/	e	e
/i/	i	i
/o/	o	o
/u/	u	u
/p/	p	p
/b/	b	b
/ᵐb/	mb	mb
/ɓ/	bh	b’
/m/	m	m
/f/	f	f
/ʋ/	w	v
/t/	t	t
/d/	d	d
/ⁿd/	nd	nd
/ɗ/	dh	d’
/n/	n	n
/s/	s	s
/r/	r	r
/l/	l	l
/dʒ/	j	dz
/k/	k	k
/g/	g	g
/ᵑg/	ngg	n’g
/ŋ/	ng	n’
/ɯ/	gh	y
/ʔ/	’	(trema)
/h/	h	h

Since the conventional Lio orthography is not standardized or regulated, there are variations from user to user. In particular, the distinction between *é* /e/ and *e* /ə/ is not always respected, and both can be symbolized with the letter *e*. Another area of variation is in the representation of the glottal stop /ʔ/. This can be represented by an apostrophe, a period, a comma, or left unmarked. Thus, the word /hoʔo/ ’yes’ can be spelled *ho’o*, *ho.o*, *ho,o*, or simply *hoo*.

The main differences between the conventional orthography and Arndt’s are as follows. The conventional orthography is more prone to using digraphs, representing the phonemes /ɓ ɗ ŋᵑg/ with the digraphs *bh dh ng ngg* where Arndt chooses a diacritic apostrophe *b’ d’*

n' n'g. In the case of /dʒ/, Arndt uses a digraph *dz* whereas the conventional orthography uses *j*, in line with Indonesian. For the phoneme /ʊ/, the conventional orthography uses *w* (seeing that segment as the closest equivalent to the Indonesian phoneme /w/) while Arndt uses *v*. One striking difference between the systems is the treatment of /ɯ̨/: the conventional orthography uses *gh*, but Arndt uses *y*. The treatment of the glottal stop is much more systematic in the conventional orthography, where it is represented everywhere by an apostrophe. In Arndt's orthography, a glottal stop is symbolized with a dieresis over the following vowel, as in *hoö* /hoʔo/ 'yes', so there is an imperfect mapping between symbols and phonemes.

2.3 Consonant inventory

		LABIO-		APICO-	LAMINO-		DORSO-	GLOTTAL
		LABIAL	DENTAL	ALVEOLAR	DENTAL	PALATAL	VELAR	
STOP	VOICELESS	p ⟨*p*⟩			t ⟨*t*⟩		k ⟨*k*⟩	ʔ ⟨*'*⟩
	VOICED	b ⟨*b*⟩		d ⟨*d*⟩			g ⟨*g*⟩	
	PRENASAL	ᵐb ⟨*mb*⟩		ⁿd ⟨*nd*⟩			ᵑg ⟨*ngg*⟩	
	IMPLOSIVE	ɓ ⟨*bh*⟩		ɗ ⟨*dh*⟩				
AFFRICATE						dʒ ⟨*j*⟩		
NASAL		m ⟨*m*⟩		n ⟨*n*⟩			ŋ ⟨*ng*⟩	
FRICATIVE			f ⟨*f*⟩	s ⟨*s*⟩				[h] ⟨*h*⟩
APPROXIMANT			ʊ ⟨*w*⟩	l ⟨*l*⟩			ɯ̨ ⟨*gh*⟩	
TRILL				r ⟨*r*⟩				

Table 2: Lio consonant phonemes

Lio has 23 consonant phonemes, listed in Table 2. Each phoneme is shown in IPA, and the orthography used in this thesis is provided between the angled brackets. The glottal fricative /h/ is a marginal phoneme in Detukeli Lio, present only in one word: *ho'o* 'yes'. In the south-eastern Lio dialects, namely the area centered on the village of Lisédetu, a semi-regular shift of word-initial /k-/ to /h-/ has introduced /h/ as a robust member of the phonemic system, but only in word-initial position.

2.3.1 Voiceless stops

The voiceless stops /p, t, k/ may appear in both word-initial and word-medial position. They are phonetically aspirated in almost all cases, with the heaviest aspiration in stressed word-initial position. In word-medial position, there is sometimes pre-aspiration as well as post-aspiration.

Voiceless bilabial stop /p/ The voiceless bilabial stop /p/ appears both word-initially and word-medially. It is usually aspirated, with the heaviest aspiration in stressed word-initial position:

paru	[ˈp^{h}aru]	‘to run’
apu	[ˈʔap^{h}u]	‘dew’

In word-medial position, there is sometimes pre-aspiration in addition to the usual post-aspiration:

apa	[ˈʔa^{h}p^{h}a]	‘what’

The following minimal pairs show that /p/ contrasts with all other voiceless stops and labial phonemes, as well as the absence of any phoneme:

/p/ : /∅/	initial	*pau*	‘mango’	medial	*nipa*	‘snake’
		au	‘bamboo’		*nia*	‘face’
/p/ : /t/	initial	*pau*	‘mango’	medial	*napa*	‘to wait’
		tau	‘to make, to do’		*nata*	‘to chew betel’
/p/ : /k/	initial	*pai*	‘to call’	medial	*apu*	‘dew’
		kai	‘he, she, it; 3SG’		*aku*	‘I; 1SG’
/p/ : /ʔ/	initial	N/A	N/A	medial	*sopi*	‘to scoop’
		N/A	N/A		*so’i*	‘to release’
/p/ : /b/	initial	*péré*	‘door’	medial	*lapa*	‘to chop finely’
		béré	‘to flow’		*laba*	‘hive’
/p/ : /^{m}b/	initial	*pana*	‘to shoot’	medial	*kapa*	‘thick’
		mbana	‘to go’		*kamba*	‘buffalo’
/p/ : /ɓ/	initial	*peni*	‘to feed birds’	medial	*sopé*	‘large sailboat’
		bheni	‘wonderful’		*sobhé*	‘to cover’
/p/ : /dʒ/	initial	*jata*	‘eagle’	medial	*napa*	‘to wait’
		pata	‘cloth’		*naja*	‘name’
/p/ : /m/	initial	*pai*	‘to call’	medial	*lepa*	‘small house’
		mai	‘to come’		*lema*	‘tongue’
/p/ : /f/	initial	*pai*	‘to call’	medial	N/A	N/A
		fai	‘wife’		N/A	N/A
/p/ : /ʋ/	initial	*pati*	‘to give’	medial	*apu*	‘dew’
		wati	‘basket’		*awu*	‘ash, dust’

Voiceless lamino-dental stop /t/ The voiceless lamino-dental stop /t/ [t̪] is produced slightly further forward than the other coronal consonants, such as /d, ^{n}d, ɗ, n/. The tip and blade of the tongue make a relatively wide contact with both the back of the teeth

and the alveolar ridge, whereas the other coronal consonants are produced only with the tip of the tongue, and do not involve dental contact. It appears both word-initially and word-medially. It is usually aspirated in both positions, with the heaviest aspiration in stressed word-initial position:

tolo	[ˈt̪^{h}olo]	‘summit’
ata	[ˈʔat̪^{h}a]	‘person’

In word-medial position, there is sometimes pre-aspiration in addition to the usual post-aspiration:

latu	[ˈla^{h}t̪^{h}u]	‘there is; EXI’

The mismatch in place of articulation between the voiceless stop /t/ and the other coronal consonants is not unique to Lio. Blust (2013) has the following to say on the Austronesian languages in general: ‘the *t:d* contrast, which is signaled primarily by different values for the feature [voice], is redundantly signaled in many languages by a difference of place: *t* is postdental, while *d* (like *n* and *l*) is alveolar (Blust 2013; p172)’. This observation certainly applies to Lio, where the voiceless stop /t/ involves contact with the teeth, but the other coronal consonants do not.

The following minimal pairs show that /t/ contrasts with the other voiceless stops and coronal phonemes, as well as the absence of any phoneme:

/t/ : /∅/	initial	*tebé*	'cliff'	medial	*até*	'liver'
		ebé	'they; 3PL'		*aé*	'water'
/t/ : /p/	initial	*tau*	'to make, to do'	medial	*nata*	'to chew betel'
		pau	'mango'		*napa*	'to wait'
/t/ : /k/	initial	*to'o*	'to arise'	medial	*beta*	'to snap'
		ko'o	'slave'		*beka*	'to increase'
/t/ : /ʔ/	initial	N/A	N/A	medial	*pata*	'cloth'
		N/A	N/A		*pa'a*	'to offer'
/t/ : /d/	initial	*téi*	'to find'	medial	*keta*	'cold'
		déi	'to rejoice'		*keda*	'to step on'
/t/ : /ⁿd/	initial	*tu'a*	'hard'	medial	*pata*	'cloth'
		ndu'a	'forest, highlands'		*panda*	'short'
/t/ : /ɗ/	initial	*toka*	'to pierce'	medial	*mota*	'betel vine'
		dhoka	'deaf'		*modha*	'wet'
/t/ : /dʒ/	initial	*toka*	'to pierce'	medial	*até*	'liver'
		joka	'to push'		*ajé*	'string'
/t/ : /n/	initial	*ta'u*	'to fear'	medial	*ata*	'person'
		na'u	'to arrange'		*ana*	'child'
/t/ : /s/	initial	*ta'o*	'to gather up'	medial	*keta*	'cold'
		sa'o	'house'		*kesa*	'to add more'
/t/ : /l/	initial	*ta'i*	'excrement'	medial	*até*	'liver'
		la'i	'to lick'		*alé*	'to ask'
/t/ : /r/	initial	*ta'i*	'excrement'	medial	*jata*	'eagle'
		ra'i	'to approach'		*jara*	'horse'

Voiceless velar stop /k/ The voiceless velar stop /k/ appears both word-initially and word-medially. It is usually aspirated, with the heaviest aspiration in stressed word-initial position:

kami [ˈkʰami] 'we (exclusive); 1PL.E'
laka [ˈlakʰa] 'to help'

In word-medial position, there is sometimes pre-aspiration in addition to the usual post-aspiration:

lako [ˈlaʰkʰo] 'dog'

The following minimal pairs show that /k/ contrasts with the other voiceless stops and dorsal phonemes (including glottal stop /ʔ/), as well as the absence of any phoneme:

/k/ : /∅/	initial	*kapa*	'thick'	medial	*nuka*	'to rise'
		apa	'what'		*nua*	'village'
/k/ : /p/	initial	*kai*	'he, she, it; 3SG'	medial	*aku*	'I; 1SG'
		pai	'to call'		*apu*	'dew'
/k/ : /t/	initial	*ko'o*	'slave'	medial	*beka*	'to increase'
		to'o	'to arise'		*beta*	'to snap'
/k/ : /ʔ/	initial	N/A	N/A	medial	*laka*	'to help'
		N/A	N/A		*la'a*	'to visit'
/k/ : /g/	initial	*kola*	'to chase'	medial	*beka*	'to increase'
		gola	'to roll'		*bega*	'comrade'
/k/ : /ᵑg/	initial	*kéu*	'areca nut'	medial	*laka*	'to help'
		nggéu	'left'		*langga*	'to surpass'
/k/ : /ŋ/	initial	*kaju*	'wood'	medial	*weki*	'body'
		ngaju	'to chew'		*wengi*	'when'
/k/ : /ɯ/	initial	*ka*	'to eat'	medial	*lagha*	'to split off'
		gha	'here'		*laka*	'to help'

2.3.2 Glottal stop

Glottal stop /ʔ/ The glottal stop /ʔ/ (represented orthographically with an apostrophe ') is only contrastive in word-medial position between two vowels. Furthermore, the first vowel may not be schwa /ə/ but must instead be one of the five full vowels /i, u, e, o, a/. The glottal stop /ʔ/ is the only consonant phoneme which may not begin a word, and which has restrictions on the adjacent vowels. The following are examples of well-formed words containing glottal stop /ʔ/:

ma'u	[ˈmaʔu]	'beach'
ké'a	[ˈkʰeʔa]	'coconut shell'
ji'é	[ˈdʒiʔe]	'excellent'
lo'o	[ˈloʔo]	'small'
tu'u	[ˈt̪ʰuʔu]	'to stop'

The glottal stop /ʔ/ is not always produced with complete closure of the glottis in connected speech. Instead, there can be a brief glottal constriction whose primary phonetic cue is creakiness on the surrounding vowels (akin to the creakiness observed before implosive stops) and a momentary dip in intensity. This can be transcribed with the symbol [ʔ̰], a creaky-voiced glottal approximant:

no'o	[ˈnoʔ̰o]	'with, and; COM'
ema du'a	[ˌʔəma ˈduʔ̰a]	'father's elder brother (FeB)'

Although the glottal stop /ʔ/ is not contrastive in word-initial position, a glottal stop or

approximant is inserted before vowel-initial words. This occurs even in connected speech, meaning that phonetic vowel sequences almost never occur across a word boundary:

(1) *iné* *ema* *aku*
['ʔinẽ 'ʔə̰mə̰̃ ˌʔə̰̃aku]
mother father 1SG
'my mother and father (my parents)'

The following minimal pairs show that /ʔ/ contrasts with the voiceless stops, implosive stops, and dorsal phonemes, as well as the absence of any phoneme, but only in word-medial position:

/ʔ/ : /∅/	*pa'u*	'to throw'
	pau	'mango'
/ʔ/ : /p/	*so'i*	'to release'
	sopi	'to scoop'
/ʔ/ : /t/	*pa'a*	'to offer'
	pata	'cloth'
/ʔ/ : /k/	*la'a*	'to visit'
	laka	'to help'
/ʔ/ : /g/	*pa'a*	'to offer'
	paga	'to nurture'
/ʔ/ : /ᵑg/	*la'a*	'to visit'
	langga	'to surpass'
/ʔ/ : /ɓ/	*la'a*	'to visit'
	labha	'dung heap'
/ʔ/ : /ɗ/	*ko'o*	'slave'
	kodho	'to look'
/ʔ/ : /ŋ/	*ra'o*	'next, to arrive'
	rango	'fever'
/ʔ/ : /ɯɻ/	*wo'a*	'to split off'
	wogha	'piece, portion'

2.3.3 Voiced stops

The plain voiced stops /b d g/ appear both word-initially and word-medially. Voicing begins before the moment of release in word-initial position, so these are true voiced stops.

Voiced bilabial stop /b/ The voiced bilabial stop /b/ appears both word-initially and word-medially. Voicing begins before the moment of release in word-initial position, so this

is a true voiced stop:

bara	[ˈbaɾa]	‘white’
tébo	[ˈt̪ʰebo]	‘to strike’

The following minimal pairs show that /b/ contrasts with the other voiced stops (including the voiced affricate /dʒ/) and labial phonemes, as well as the absence of any phoneme:

/b/ : /∅/	initial	*boka*	‘to fall’	medial	*tébo*	‘to strike’
		oka	‘lime’		*téo*	‘to hang’
/b/ : /p/	initial	*béré*	‘to flow’	medial	*laba*	‘hive’
		péré	‘door’		*lapa*	‘to chop finely’
/b/ : /d/	initial	*bowa*	‘to untie’	medial	*eba*	‘father’s sister’
		dowa	‘already; PERF’		*eda*	‘father’s brother’
/b/ : /g/	initial	*boro*	‘gebang palm’	medial	*tebé*	‘cliff’
		goro	‘to pull’		*tegé*	‘to remove’
/b/ : /ᵐb/	initial	*boko*	‘short’	medial	*laba*	‘hive’
		mboko	‘fruit’		*lamba*	‘drum’
/b/ : /ⁿd/	initial	*bara*	‘white’	medial	*laba*	‘hive’
		ndara	‘thin’		*landa*	‘large fishing net’
/b/ : /ᵑg/	initial	*béu*	‘far’	medial	*laba*	‘hive’
		nggéu	‘left’		*langga*	‘to surpass’
/b/ : /ɓ/	initial	*boro*	‘gebang palm’	medial	*dobé*	‘sloped’
		bhoro	‘to bake’		*dobhé*	‘to pour in’
/b/ : /ɗ/	initial	*boka*	‘to fall’	medial	*laba*	‘hive’
		dhoka	‘deaf’		*ladha*	‘to hang down’
/b/ : /dʒ/	initial	*bara*	‘white’	medial	*laba*	‘hive’
		jara	‘horse’		*laja*	‘sail’
/b/ : /m/	initial	*bara*	‘white’	medial	*eba*	‘father’s sister’
		mara	‘to cherish’		*ema*	‘father’
/b/ : /f/	initial	*bai*	‘too, excessively’	medial	N/A	N/A
		fai	‘wife’		N/A	N/A
/b/ : /ʋ/	initial	*bara*	‘white’	medial	*laba*	‘hive’
		wara	‘shoulder’		*lawa*	‘to fight’

Voiced apico-alveolar stop /d/ The voiced apico-alveolar stop /d/ is produced with the tip of the tongue against the alveolar ridge, with no dental contact. This makes it an apico-alveolar stop, in contrast to its voiceless counterpart /t/ which is lamino-dental. It occurs in both word-initial and word-medial position. Voicing begins before the moment of release in word-initial position, so this is a true voiced stop:

dala [ˈdala] ‘star’
gudu [ˈgudu] ‘to startle, to be startled’

The following minimal pairs show that /d/ contrasts with the other voiced stops (including the voiced affricate /dʒ/) and coronal phonemes, as well as the absence of any phoneme:

/d/ : /∅/	initial	*dau*	‘must’	medial	*gadi*	‘wounded’
		au	‘bamboo’		*gai*	‘elephant grass’
/d/ : /t/	initial	*déi*	‘to rejoice’	medial	*keda*	‘to step on’
		téi	‘to find’		*keta*	‘cold’
/d/ : /b/	initial	*dowa*	‘already; PERF’	medial	*eda*	‘father’s brother’
		bowa	‘to untie’		*eba*	‘father’s sister’
/d/ : /g/	initial	*demi*	‘if, when; COND’	medial	*budu*	‘to banish’
		gemi	‘nice, kind’		*bugu*	‘diligent’
/d/ : /^{m}b/	initial	*deru*	‘to gasp’	medial	*eda*	‘father’s brother’
		mberu	‘to eat raw’		*emba*	‘which’
/d/ : /^{n}d/	initial	*dala*	‘star’	medial	*seda*	‘to weave’
		ndala	‘fishing net’		*senda*	‘trap’
/d/ : /^{ŋ}g/	initial	*da*	‘towards’	medial	*pida*	‘to press down’
		ngga	‘cool, fresh’		*pingga*	‘dish’
/d/ : /ɓ/	initial	*da*	‘towards’	medial	*bedu*	‘to stamp (feet)’
		bha	‘plate’		*mbebhu*	‘small chunk’
/d/ : /ɗ/	initial	*du*	‘to reach’	medial	*kodo*	‘hole in tree’
		dhu	‘to pound’		*kodho*	‘to look’
/d/ : /dʒ/	initial	*dala*	‘star’	medial	*pida*	‘to press down’
		jala	‘path’		*pija*	‘how many’
/d/ : /n/	initial	*deké*	‘stick, rod’	medial	*woda*	‘bell’
		neké	‘to hide’		*wona*	‘moringa’
/d/ : /s/	initial	*du'u*	‘to stop’	medial	*eda*	‘father’s brother’
		su'u	‘to carry on head’		*esa*	‘seed, fruit’
/d/ : /l/	initial	*dau*	‘must’	medial	*gudu*	‘to startle’
		lau	‘seaward’		*gulu*	‘thorn’
/d/ : /r/	initial	*da*	‘towards’	medial	*woda*	‘bell’
		ra	‘blood’		*wora*	‘foam’

Voiced velar stop /g/ The voiced velar stop /g/ appears both word-initially and word-medially. Voicing begins before the moment of release in word-initial position, so this is a true voiced stop:

gaé [ˈgae] ‘to search for’
paga [ˈp^{h}aga] ‘to nurture’

The following minimal pairs show that /g/ contrasts with the other voiced stops (including the voiced affricate /dʒ/) and dorsal phonemes (including glottal stop /ʔ/), as well as the absence of any phoneme:

/g/ : /∅/	initial	*gena*	'to affect'	medial	*buga*	'morning'
		ena	'sand'		*bua*	'body hair'
/g/ : /k/	initial	*gola*	'to roll'	medial	*bega*	'comrade'
		kola	'to chase'		*beka*	'to increase'
/g/ : /ʔ/	initial	N/A	NA	medial	*paga*	'to nurture'
		N/A	N/A		*pa'a*	'to offer'
/g/ : /b/	initial	*goro*	'to pull'	medial	*tegé*	'to remove'
		boro	'gebang palm'		*tebé*	'cliff'
/g/ : /d/	initial	*gemi*	'nice, kind'	medial	*bugu*	'diligent'
		demi	'if, when; COND'		*budu*	'to banish'
/g/ : /ᵐb/	initial	*gana*	'manger, stall'	medial	*legu*	'to thunder, to roar'
		mbana	'to go'		*lembu*	'sheep'
/g/ : /ⁿd/	initial	*gawi*	'to dance'	medial	*paga*	'to nurture'
		ndawi	'to hold hands'		*panda*	'short'
/g/ : /ᵑg/	initial	*gana*	'manger, stall'	medial	*logo*	'to roll down'
		nggana	'right (side)'		*longgo*	'back'
/g/ : /ɓ/	initial	*geto*	'to chop up'	medial	*tugé*	'to stick in fire'
		bheto	'large bamboo'		*tubhé*	'to fight'
/g/ : /ɗ/	initial	*gana*	'manger, stall'	medial	*rega*	'market'
		dhana	'to hold up'		*redha*	'to tug'
/g/ : /dʒ/	initial	*gedho*	'good, delicious'	medial	*kago*	'to grab'
		jedho	'to limp'		*kajo*	'ancestor'
/g/ : /ŋ/	initial	*gana*	'manger, stall'	medial	*rega*	'market'
		ngana	'termite'		*renga*	'to withstand'
/g/ : /ɰ/	initial	*galé*	'to choose'	medial	*laga*	'to wriggle'
		ghalé	'downwards, west'		*lagha*	'to split off'

2.3.4 Prenasalized voiced stops

The prenasalized voiced stops /ᵐb ⁿd ᵑg/ appear both word-initially and word-medially. Given the absence of any other consonant clusters in Lio, the prenasalized stops are best analyzed as single phonemes rather than clusters of nasal plus voiced stop. In word-initial position, the prenasalization is realized as a relatively lengthy period of nasal airflow before the release of the oral closure. In word-medial position, the prenasalization is phonetically identical to a sequence of nasal plus homorganic stop, but based on Lio speakers' behavior when breaking speech into individual syllables, the nasal remains as part of the onset of the second syllable (it does not resyllabify to the coda of the preceding syllable).

Prenasalized voiced bilabial stop /^{m}b/ The prenasalized voiced bilabial stop /^{m}b/ (spelled with the digraph *mb*) appears both word-initially and word-medially. Word-initially, the prenasalization is realized as a period of nasal airflow beginning well before the release of the stop, and can be transcribed [^{m}b]. Word-medially, this is phonetically identical to a sequence [mb]:

mbalé	[ˈ^{m}bale]	'to become'
lamba	[ˈlamba]	'drum'

The following minimal pairs show that /^{m}b/ contrasts with the other voiced stops (including the voiced affricate /dʒ/), nasals, and labial phonemes, as well as the absence of any phoneme:

/ᵐb/ : /∅/	initial	*mbana*	'to go'	medial	*lambu*	'shirt'
		ana	'child'		*lau*	'seaward'
/ᵐb/ : /p/	initial	*mbana*	'to go'	medial	*kamba*	'buffalo'
		pana	'to shoot'		*kapa*	'thick'
/ᵐb/ : /b/	initial	*mboko*	'fruit'	medial	*lamba*	'drum'
		boko	'short'		*laba*	'hive'
/ᵐb/ : /d/	initial	*mberu*	'to eat raw'	medial	*emba*	'which'
		deru	'to gasp'		*eda*	'father's brother'
/ᵐb/ : /g/	initial	*mbana*	'to go'	medial	*lembu*	'sheep'
		gana	'manger, stall'		*legu*	'to thunder, to roar'
/ᵐb/ : /ⁿd/	initial	*mbé'o*	'to know'	medial	*tembu*	'to grow'
		ndé'o	'to sing'		*tendu*	'to follow'
/ᵐb/ : /ᵑg/	initial	*mbana*	'to go'	medial	*lamba*	'drum'
		nggana	'right (side)'		*langga*	'to surpass'
/ᵐb/ : /ɓ/	initial	*mbalé*	'to become'	medial	*lamba*	'drum'
		bhalé	'to return'		*labha*	'dung heap'
/ᵐb/ : /ɗ/	initial	*mbana*	'to go'	medial	*kamba*	'buffalo'
		dhana	'to hold up'		*kadha*	'to yell'
/ᵐb/ : /dʒ/	initial	*mbulu*	'ten'	medial	*kambu*	'womb'
		julu	'to walk in a row'		*kaju*	'wood'
/ᵐb/ : /m/	initial	*mbaku*	'wing'	medial	*emba*	'which'
		maku	'strong, hard'		*ema*	'father'
/ᵐb/ : /n/	initial	*mbola*	'large basket'	medial	*emba*	'which'
		nola	'to damage'		*ena*	'sand'
/ᵐb/ : /ŋ/	initial	*mbana*	'to go'	medial	*tembu*	'to grow'
		ngana	'termite'		*tengu*	'neck'
/ᵐb/ : /f/	initial	*mboko*	'fruit'	medial	N/A	N/A
		foko	'throat'		N/A	N/A
/ᵐb/ : /ʋ/	initial	*mbulu*	'ten'	medial	*lamba*	'drum'
		wulu	'thin bamboo'		*lawa*	'to fight'

Prenasalized voiced apico-alveolar stop /ⁿd/ The prenasalized voiced apico-alveolar stop /ⁿd/ (spelled with the digraph *nd*) appears both word-initially and word-medially. Word-initially, the prenasalization is realized as a period of nasal airflow beginning well before the release of the stop, and can be transcribed [ⁿd]. Word-medially, this is phonetically identical to a sequence [nd]:

ndaté	[ˈⁿdatʰe]	'heavy'
panda	[ˈpʰanda]	'short'

The following minimal pairs show that /ⁿd/ contrasts with the other voiced stops (in-

cluding the voiced affricate /dʒ/), nasals, and coronal phonemes, as well as the absence of any phoneme:

/ⁿd/ : /∅/	initial	*ndaté*	'heavy'	medial	*kando*	'jewelry'
		até	'liver'		*kao*	'to scoop'
/ⁿd/ : /t/	initial	*ndu'a*	'forest, highlands'	medial	*panda*	'short'
		tu'a	'hard'		*pata*	'cloth'
/ⁿd/ : /b/	initial	*ndara*	'thin'	medial	*landa*	'large fishing net'
		bara	'white'		*laba*	'hive'
/ⁿd/ : /d/	initial	*ndala*	'fishing net'	medial	*senda*	'trap'
		dala	'star'		*seda*	'to weave'
/ⁿd/ : /g/	initial	*ndawi*	'to hold hands'	medial	*panda*	'short'
		gawi	'to dance'		*paga*	'to nurture'
/ⁿd/ : /ᵐb/	initial	*ndé'o*	'to sing'	medial	*tendu*	'to follow'
		mbé'o	'to know'		*tembu*	'to grow'
/ⁿd/ : /ᵑg/	initial	*ndua*	'to descend'	medial	*wanda*	'to dance (with cloth)'
		nggua	'ritual'		*wangga*	'to carry on shoulders'
/ⁿd/ : /ɓ/	initial	*ndongo*	'kind, loving'	medial	*landa*	'large fishing net'
		bhongo	'stupid'		*labha*	'dung heap'
/ⁿd/ : /ɗ/	initial	*ndu*	'to follow, to obey'	medial	*landa*	'large fishing net'
		dhu	'to pound'		*ladha*	'to hang down'
/ⁿd/ : /dʒ/	initial	*ndara*	'thin'	medial	*kando*	'jewelry'
		jara	'horse'		*kajo*	'ancestor'
/ⁿd/ : /m/	initial	*ndara*	'thin'	medial	*landa*	'large fishing net'
		mara	'to cherish'		*lama*	'fast'
/ⁿd/ : /n/	initial	*ndu*	'to follow, to obey'	medial	*panda*	'short'
		nu	'smoke, cloud'		*pana*	'to shoot'
/ⁿd/ : /ŋ/	initial	*ndala*	'fishing net'	medial	*bhondo*	'much, many'
		ngala	'can, to be able'		*bhongo*	'stupid'
/ⁿd/ : /s/	initial	*ndu'a*	'forest, highlands'	medial	*mondo*	'to gather up'
		su'a	'iron'		*moso*	'to spoil, to rot'
/ⁿd/ : /l/	initial	*ndu*	'to follow, to obey'	medial	*mondo*	'to gather up'
		lu	'teardrop'		*molo*	'correct'
/ⁿd/ : /r/	initial	*ndaté*	'heavy'	medial	*wanda*	'to dance (with cloth)'
		raté	'tomb'		*wara*	'shoulder'

Prenasalized voiced velar stop /ᵑg/ The prenasalized voiced velar stop /ᵑg/ (spelled with a trigraph *ngg*) appears both word-initially and word-medially. Word-initially, the prenasalization is realized as a period of nasal airflow beginning well before the release of the stop, and can be transcribed [ᵑg]. Word-medially, this is phonetically identical to a

sequence [ŋg]:

nggua	[ˈᵑgu^{w}a]	‘ritual, culture’
wangga	[ˈʋaŋga]	‘to carry (something) on the shoulders’

The following minimal pairs show that /ᵑg/ contrasts with the other voiced stops (including the voiced affricate /dʒ/), nasals, and dorsal phonemes (including glottal stop /ʔ/), as well as the absence of any phoneme:

/ᵑg/ : /∅/	initial	*nggana*	‘right (side)’	medial	*pongga*	‘to hit’
		ana	‘child’		*poa*	‘early’
/ᵑg/ : /k/	initial	*nggéu*	‘left’	medial	*langga*	‘to surpass’
		kéu	‘areca nut’		*laka*	‘to help’
/ᵑg/ : /ʔ/	initial	N/A	N/A	medial	*langga*	‘to surpass’
		N/A	N/A		*la’a*	‘to visit’
/ᵑg/ : /b/	initial	*nggéu*	‘left’	medial	*langga*	‘to surpass’
		béu	‘far’		*laba*	hive‘’
/ᵑg/ : /d/	initial	*ngga*	‘cool, fresh’	medial	*pingga*	‘dish’
		da	‘towards’		*pida*	‘to press down’
/ᵑg/ : /g/	initial	*nggana*	‘right (side)’	medial	*longgo*	‘back’
		gana	‘manger, stall’		*logo*	‘to roll down’
/ᵑg/ : /ᵐb/	initial	*nggana*	‘right (side)’	medial	*langga*	‘to surpass’
		mbana	‘to go’		*lamba*	‘drum’
/ᵑg/ : /ⁿd/	initial	*nggua*	‘ritual’	medial	*wangga*	‘to carry on shoulders’
		ndua	‘to descend’		*wanda*	‘to dance (with cloth)’
/ᵑg/ : /ɓ/	initial	*ngga*	‘cool, fresh’	medial	*langga*	‘to surpass’
		bha	‘plate’		*labha*	‘dung heap’
/ᵑg/ : /ɗ/	initial	*nggana*	‘right (side)’	medial	*ngonggo*	‘to feast’
		dhana	‘to hold up’		*ngodho*	‘to receive’
/ᵑg/ : /dʒ/	initial	*ngga*	‘cool, fresh’	medial	*langga*	‘to surpass’
		ja	‘to shine’		*laja*	‘sail’
/ᵑg/ : /m/	initial	*ngga’é*	‘master’	medial	*langga*	‘to surpass’
		ma’é	‘do not; PROH’		*lama*	‘fast’
/ᵑg/ : /n/	initial	*nggua*	‘ritual’	medial	*wongga*	‘to boil’
		nua	‘village’		*wona*	‘moringa’
/ᵑg/ : /ŋ/	initial	*nggana*	‘right (side)’	medial	*langga*	‘to surpass’
		ngana	‘termite’		*langa*	‘to take away’
/ᵑg/ : /ɯ/	initial	*ngga*	‘cool, fresh’	medial	*wongga*	‘to boil’
		gha	‘here’		*wogha*	‘piece, portion’

2.3.5 Voiced implosive stops

There are only two voiced implosive stops /ɓ ɗ/, with no equivalent voiced velar implosive. The implosives appear both word-initially and word-medially.

Implosive or preglottalized stops are an areal feature of Flores and the Lesser Sunda islands (Donohue 2006, Blust 2013; p87-88), and it is very common to have a gap at the velar place of articulation. In certain languages, such as Komodo (Verheijen 1982), Bimanese (Jonker 1896, Ismail, Azis, Yakub, Taufik and Usman 1985), Kambera (Onvlee 1984, Klamer 1998) and Palu'e (Donohue 2003), the unmarked series of voiced oral stops are implosives, and the plain voiced stops /b, d/ may appear only rarely or in loanwords. Rongga, a close relative of Lio in western Flores, has the full series of implosive stops /ɓ ɗ ɠ/ (Arka 2016).

These segments are produced with simultaneous oral and glottal closure, hence the label 'preglottalized' sometimes used by other authors to describe nearby languages with comparable phonemes (see Baird 2002 on Kéo). However, they have a prominent release burst which suggests that intra-oral pressure is lowered during the closure, leading to an influx of air upon release. Hence, I prefer the label 'implosive', because a purely preglottalized sound would not be expected to show any significant difference in its release burst. It is worth noting that there are good articulatory reasons for having a velar gap in the implosive series. It is easier to expand a large cavity of air than a small one, so anterior implosives like /ɓ/ and /ɗ/ are easier to produce than ones further back like /ɠ/ (see Ladefoged and Maddieson 1996; p. 82-83).

Due to the glottal constriction involved in producing them, implosives often cause allophonic creakiness in the preceding vowel.

Voiced bilabial implosive stop /ɓ/ The voiced bilabial implosive stop /ɓ/ (spelled with the digraph *bh*) appears both word-initially and word-medially:

bhondo	[ˈɓondo]	'much, many'
nibho	[ˈniɓo]	'to jump, to be startled'

In word-medial position, it may cause creakiness on the preceding vowel:

nggebhé	[ˈᵑgɐ̯ɓːe]	'to lie flat'

The following minimal pairs show that /ɓ/ contrasts with the other voiced stops (including the voiced affricate /dʒ/), the glottal stop /ʔ/ and the other labial phonemes, as well as the absence of any phoneme:

/ɓ/ : /∅/	initial	*bhalé*	'to return'	medial	*nibho*	'to jump, to startle'
		alé	'to ask'		*nio*	'coconut'
/ɓ/ : /p/	initial	*bheni*	'wonderful'	medial	*sobhé*	'to cover'
		peni	'to feed birds'		*sopé*	'large sailboat'
/ɓ/ : /ʔ/	initial	N/A	N/A	medial	*labha*	'dung heap'
		N/A	N/A		*la'a*	'to visit'
/ɓ/ : /b/	initial	*bhoro*	'to bake'	medial	*dobhé*	'to pour in'
		boro	'gebang palm'		*dobé*	'sloped'
/ɓ/ : /d/	initial	*bha*	'plate'	medial	*mbebhu*	'small chunk'
		da	'towards'		*bedu*	'to stamp (feet)'
/ɓ/ : /g/	initial	*bheto*	'large bamboo'	medial	*tubhé*	'to fight'
		geto	'to chop up'		*tugé*	'to stick in fire'
/ɓ/ : /^{m}b/	initial	*bhalé*	'to return'	medial	*labha*	'dung heap'
		mbalé	'to become'		*lamba*	'drum'
/ɓ/ : /^{n}d/	initial	*bhongo*	'stupid'	medial	*labha*	'dung heap'
		ndongo	'kind, loving'		*landa*	'large fishing net'
/ɓ/ : /^{ŋ}g/	initial	*bha*	'plate'	medial	*labha*	'dung heap'
		ngga	'cool, fresh'		*langga*	'to surpass'
/ɓ/ : /ɗ/	initial	*bhoa*	'valley'	medial	*labha*	'dung heap'
		dhoa	'to sink'		*ladha*	'to hang down'
/ɓ/ : /dʒ/	initial	*bha*	'plate'	medial	*nibho*	'to jump, to startle'
		ja	'to shine'		*nijo*	'to spit'
/ɓ/ : /m/	initial	*bhesu*	'hefty'	medial	*labha*	'dung heap'
		mesu	'to fall'		*lama*	'fast'
/ɓ/ : /f/	initial	*bhu*	'to rise (of sun)'	medial	N/A	N/A
		fu	'hair'		N/A	N/A
/ɓ/ : /ʋ/	initial	*bhanda*	'rich'	medial	*labha*	'dung heap'
		wanda	'to dance (with cloth)'		*lawa*	'to fight'

Voiced apico-alveolar implosive stop /ɗ/ The voiced apico-alveolar implosive stop /ɗ/ (spelled with the digraph *dh*) appears both word-initially and word-medially:

dhoka [ˈɗok^{h}a] 'deaf'
jedho [ˈdʒəɗo] 'to limp'

In word-medial position, it may cause creakiness on the preceding vowel:

kodho [ˈk^{h}o̰ɗo] 'to look'

The following minimal pairs show that /ɗ/ contrasts with the other voiced stops (including the voiced affricate /dʒ/), the glottal stop /ʔ/ and the other coronal phonemes, as

well as the absence of any phoneme:

/ɗ/ : /∅/	initial	*dhana*	'to hold up'	medial	*modha*	'wet'
		ana	'child'		*moa*	'thirsty'
/ɗ/ : /t/	initial	*dhoka*	'deaf'	medial	*modha*	'wet'
		toka	'to pierce'		*mota*	'betel vine'
/ɗ/ : /ʔ/	initial	N/A	N/A	medial	*kodho*	'to look'
		N/A	N/A		*ko'o*	'slave'
/ɗ/ : /b/	initial	*dhoka*	'deaf'	medial	*ladha*	'to hang down'
		boka	'to fall'		*laba*	'hive'
/ɗ/ : /d/	initial	*dhu*	'to pound'	medial	*kodho*	'to look'
		du	'to reach'		*kodo*	'hole in tree'
/ɗ/ : /g/	initial	*dhana*	'to hold up'	medial	*redha*	'to tug'
		gana	'manger, stall'		*rega*	'market'
/ɗ/ : /ᵐb/	initial	*dhana*	'to hold up'	medial	*kadha*	'to yell'
		mbana	'to go'		*kamba*	'buffalo'
/ɗ/ : /ⁿd/	initial	*dhu*	'to pound'	medial	*ladha*	'to hang down'
		ndu	'to follow, to obey'		*landa*	'large fishing net'
/ɗ/ : /ᵑg/	initial	*dhana*	'to hold up'	medial	*ngodho*	'to receive'
		nggana	'right (side)'		*ngonggo*	'to feast'
/ɗ/ : /ɓ/	initial	*dhoa*	'to sink'	medial	*ladha*	'to hang down'
		bhoa	'valley'		*labha*	'dung heap'
/ɗ/ : /dʒ/	initial	*dhoka*	'deaf'	medial	*ladha*	'to hang down'
		joka	'to push'		*laja*	'sail'
/ɗ/ : /n/	initial	*dhawé*	'to work (field)'	medial	*pedhé*	'to cook rice'
		nawé	'lemongrass'		*pené*	'door'
/ɗ/ : /s/	initial	*dhawé*	'to work (field)'	medial	*kadha*	'to yell'
		sawé	'to finish'		*kasa*	'chest'
/ɗ/ : /l/	initial	*dhu*	'to pound'	medial	*ngadha*	'to look around'
		lu	'teardrop'		*ngala*	'can, to be able'
/ɗ/ : /r/	initial	*dhua*	'misfortune'	medial	*kodho*	'to look'
		rua	'two'		*koro*	'chili pepper'

2.3.6 Affricate

Voiced lamino-palatal affricate /dʒ/ There is only one affricate in Lio, the voiced lamino-palatal affricate /dʒ/ (spelled *j*). It has no voiceless counterpart, nor does it align with any other segment in place of articulation. This phoneme has a number of allophones in free variation with each other, with no clear conditioning. The fricative portion of the affricate may be post-alveolar or palatal, as in [dʒ] or [dʑ], or it may be closer to a true alveolar, as in [dz]. Furthermore, the initial stop may be entirely gone, leaving only a

voiced fricative such as [z, ʑ]:

joka	[ˈdʒokʰa]	‘to push’
kaju	[ˈkʰadʒu]	‘wood’
jara	[ˈdʑaɾa]	‘horse’
jata	[ˈzatʰa]	‘eagle’
gajo	[ˈgʰadzo]	‘to scoop’

The following minimal pairs show that /dʒ/ contrasts with the voiced stops and the other coronal phonemes, as well as the absence of any phoneme:

/dʒ/ : /∅/	initial	*jata*	‘eagle’	medial	*ngaji*	‘to pray’
		ata	‘person’		*ngai*	‘breath’
/dʒ/ : /t/	initial	*joka*	‘to push’	medial	*ajé*	‘string’
		toka	‘to pierce’		*até*	‘liver’
/dʒ/ : /b/	initial	*jara*	‘horse’	medial	*laja*	‘sail’
		bara	‘white’		*laba*	‘hive’
/dʒ/ : /d/	initial	*jala*	‘path’	medial	*pija*	‘how many’
		dala	‘star’		*pida*	‘to press down’
/dʒ/ : /g/	initial	*jedho*	‘to limp’	medial	*kajo*	‘ancestor’
		gedho	‘good, delicious’		*kago*	‘to grab’
/dʒ/ : /ᵐb/	initial	*julu*	‘to walk in a row’	medial	*kaju*	‘wood’
		mbulu	‘ten’		*kambu*	‘womb’
/dʒ/ : /ⁿd/	initial	*jara*	‘horse’	medial	*kajo*	‘ancestor’
		ndara	‘thin’		*kando*	‘jewelry’
/dʒ/ : /ᵑg/	initial	*ja*	‘to shine’	medial	*laja*	‘sail’
		ngga	‘cool, fresh’		*langga*	‘to surpass’
/dʒ/ : /ɓ/	initial	*ja*	‘to shine’	medial	*nijo*	‘to spit’
		bha	‘plate’		*nibho*	‘to jump, to startle’
/dʒ/ : /ɗ/	initial	*joka*	‘to push’	medial	*laja*	‘sail’
		dhoka	‘deaf’		*ladha*	‘to hang down’
/dʒ/ : /n/	initial	*joka*	‘to push’	medial	*peju*	‘bladder’
		noka	‘to swear’		*penu*	‘to fill’
/dʒ/ : /s/	initial	*jala*	‘path’	medial	*maja*	‘dry’
		sala	‘wrong’		*masa*	‘clean’
/dʒ/ : /l/	initial	*jengi*	‘to burn (land)’	medial	*ajé*	‘string’
		lengi	‘oil’		*alé*	‘to ask’
/dʒ/ : /r/	initial	*ja*	‘to shine’	medial	*maja*	‘dry’
		ra	‘blood’		*mara*	‘to cherish’

2.3.7 Nasals

The nasals /m n ŋ/ appear both word-initially and word-medially. These nasal phonemes occasionally cause phonetic nasalization on the following vowel or sequence of vowels, but this is not always the case:

mai	[ˈmãĩ, ˈmai]	'to come'
méa	[ˈmẽã, ˈmea]	'shy, ashamed'

Bilabial nasal /m/ The bilabial nasal /m/ appears both word-initially and word-medially:

ma'é	[ˈmaʔe]	'do not, prohibitive marker; PROH'
lama	[ˈlama]	'fast'

The following minimal pairs show that /m/ contrasts with the other nasal phonemes (including prenasalized stops) and labial phonemes, as well as the absence of any phoneme:

/m/ : /∅/	initial	*maé*	'soul'	medial	*lima*	'hand, arm'
		aé	'water'		*lia*	'hole, cave'
/m/ : /p/	initial	*mai*	'to come'	medial	*lema*	'tongue'
		pai	'to call'		*lepa*	'house'
/m/ : /b/	initial	*mara*	'to cherish'	medial	*ema*	'father'
		bara	'white'		*eba*	'father's sister'
/m/ : /ᵐb/	initial	*maku*	''	medial	*ema*	'father'
		mbaku	''		*emba*	'which'
/m/ : /ⁿd/	initial	*mara*	'to cherish'	medial	*lama*	'fast'
		ndara	'thin'		*landa*	'large fishing net'
/m/ : /ᵑg/	initial	*ma'é*	'do not; PROH'	medial	*lama*	'fast'
		ngga'é	''		*langga*	'to surpass'
/m/ : /ɓ/	initial	*mesu*	'to fall'	medial	*lama*	'fast'
		bhesu	'hefty'		*labha*	'dung heap'
/m/ : /n/	initial	*miu*	'you (plural); 2PL'	medial	*lima*	'hand, arm'
		niu	'to invite'		*lina*	'clear, bright'
/m/ : /ŋ/	initial	*mara*	'to cherish'	medial	*lama*	'fast'
		ngara	'owner'		*langa*	'to pick up'
/m/ : /f/	initial	*mai*	'to come'	medial	N/A	N/A
		fai	'wife'		N/A	N/A
/m/ : /ʊ/	initial	*ma'u*	'beach'	medial	*kumu*	'fingernail'
		wa'u	'to go down'		*kuwu*	'hut'

Apico-alveolar nasal /n/ The apico-alveolar nasal /n/ appears both word-initially and word-medially:

niu	[ˈniʲu]	‘coconut’
ani	[ˈʔani]	‘bee’

The following minimal pairs show that /n/ contrasts with the other nasalized phonemes (including prenasalized stops) and coronal phonemes, as well as the absence of any phoneme:

/n/ : /∅/	initial	*nua*	‘village’	medial	*minu*	‘to drink’
		ua	‘rattan’		*miu*	‘you (plural); 2PL’
/n/ : /t/	initial	*na'u*	‘to arrange’	medial	*ana*	‘child’
		ta'u	‘to fear’		*ata*	‘person’
/n/ : /d/	initial	*neké*	‘to hide’	medial	*wona*	‘moringa’
		deké	‘stick, rod’		*woda*	‘bell’
/n/ : /ᵐb/	initial	*nola*	‘to damage’	medial	*ena*	‘sand’
		mbola	‘large basket’		*emba*	‘which’
/n/ : /ⁿd/	initial	*nu*	‘smoke, cloud’	medial	*pana*	‘to shoot’
		ndu	‘to follow, to obey’		*panda*	‘short’
/n/ : /ᵑg/	initial	*nua*	‘village’	medial	*wona*	‘moringa’
		nggua	‘ritual’		*wongga*	‘to boil’
/n/ : /ɗ/	initial	*nawé*	‘lemongrass’	medial	*pené*	‘door’
		dhawé	‘to work (field)’		*pedhé*	‘to cook rice’
/n/ : /dʒ/	initial	*noka*	‘to swear’	medial	*penu*	‘to fill’
		joka	‘to push’		*peju*	‘bladder’
/n/ : /m/	initial	*niu*	‘to invite’	medial	*lina*	‘clear, bright’
		miu	‘you (plural); 2PL’		*lima*	‘hand, arm’
/n/ : /ŋ/	initial	*nai*	‘to climb’	medial	*ani*	‘bee’
		ngai	‘breath’		*angi*	‘wind’
/n/ : /s/	initial	*nai*	‘to climb’	medial	*ena*	‘sand’
		sai	‘who’		*esa*	‘seed, fruit’
/n/ : /l/	initial	*nu*	‘smoke, cloud’	medial	*mina*	‘oil, grease’
		lu	‘teardrop’		*mila*	‘dark’
/n/ : /r/	initial	*nia*	‘face’	medial	*ngana*	‘termite’
		ria	‘large’		*ngara*	‘owner’

Velar nasal /ŋ/ The velar nasal /ŋ/ (spelled with the digraph *ng*) appears both word-initially and word-medially:

ngangé	[ˈŋaŋe]	‘to not want’
ngaju	[ˈŋadʒu]	‘to chew’
angi	[ˈʔaŋi]	‘wind’

The following minimal pairs show that /ŋ/ contrasts with the other nasalized phonemes (including prenasalized stops) and dorsal phonemes (including glottal stop /ʔ/), as well as the absence of any phoneme:

/ŋ/ : /∅/	initial	*ngana*	‘termite’	medial	*runga*	‘sweat’
		ana	‘child’		*rua*	‘two’
/ŋ/ : /k/	initial	*ngaju*	‘to chew’	medial	*wengi*	‘when’
		kaju	‘wood’		*weki*	‘body’
/ŋ/ : /ʔ/	initial	N/A	N/A	medial	*rango*	‘fever’
		N/A	N/A		*ra’o*	‘next, to arrive’
/ŋ/ : /g/	initial	*ngana*	‘termite’	medial	*renga*	‘to withstand’
		gana	‘manger, stall’		*rega*	‘market’
/ŋ/ : /^{m}b/	initial	*ngana*	‘termite’	medial	*tengu*	‘neck’
		mbana	‘to go’		*tembu*	‘to grow’
/ŋ/ : /^{n}d/	initial	*ngala*	‘can, to be able’	medial	*bhongo*	‘stupid’
		ndala	‘fishing net’		*bhondo*	‘much, many’
/ŋ/ : /^{ŋ}g/	initial	*ngana*	‘termite’	medial	*langa*	‘to pick up’
		nggana	‘right (side)’		*langga*	‘to surpass’
/ŋ/ : /m/	initial	*ngara*	‘owner’	medial	*langa*	‘to pick up’
		mara	‘to cherish’		*lama*	‘fast’
/ŋ/ : /n/	initial	*ngai*	‘breath’	medial	*angi*	‘wind’
		nai	‘to climb’		*ani*	‘bee’
/ŋ/ : /ɯ/	initial	*nga*	‘to leak’	medial	*wonga*	‘flower’
		gha	‘here’		*wogha*	‘piece, portion’

2.3.8 Fricatives

There are two fricatives phonemes /f, s/ in Detukeli Lio, plus a third fricative /h/ which is marginal in Detukeli but common in south-eastern Lio dialects. These three sounds all share the property of being voiceless, and both /f/ and /h/ are limited to word-initial position. The voiceless fricatives /f, s/ are occasionally produced with aspiration in stressed initial position, with a period of voicelessness after the end of the fricative ends but before the voicing of the vowel begins.

Voiceless labio-dental fricative /f/ The voiceless labio-dental fricative /f/ only appears word-initially. It is the only consonant which is barred from word-medial position. It appears that some instances of earlier initial /ʊ-/ underwent a shift to /f-/, with the conditioning unclear beyond the fact that only initial /ʊ-/ was affected. This results in the

current situation, where /f/ only appears word-initially:

fai	[ˈfai]	‘wife’
foko	[ˈfok^{h}o]	‘throat’

There are only two instances of word-medial /f/ of which I am aware. The first, listed in Arndt (1933) is an onomatopoeia with exceptional phonotactics: *fufu* ‘to hiss, to whistle’.

The second was provided by ML and FD during elicitation of a wordlist: *béfa* ‘a kind of large building used for traditional gatherings’. Many nearby languages (Verheijen 1982 on Komodo; Jonker 1896, Ismail 1985 on Bimanese; Arka 2016 on Rongga) have a medial /f/ which corresponds to Lio /ʋ/ in many cases. This suggests that *béfa* in Lio is a borrowing from one of these nearby languages, cognate to the native Lio word *béwa* ‘long, tall’. The semantic connection stems from the parallelistic phrase used for the most important traditional house in the village: *sa’o ria ténda béwa* {house + large + platform + tall} (lit. ‘large house, tall platform’).

The voiceless labio-dental fricative /f/ is sometimes aspirated in stressed word-initial position, with a brief period of voicelessness after the fricative ends but before the voicing of the vowel begins:

fua	[ˈf^{h}u^{w}a]	‘wasp’

Although they are share some phonetic overlap, this phoneme does not form a voiced-voiceless pair with the voiced labio-dental approximant /ʋ/. The frication in /f/ is much stronger than in /ʋ/, and they have different restrictions on their distributions.

The following minimal pairs show that /f/ contrasts with the other continuants and labial phonemes, as well as the absence of any phoneme, but only appears in word-initial position:

/f/ : /ɸ/	*fua*	‘wasp’
	ua	‘rattan’
/f/ : /p/	*fai*	‘wife’
	pai	‘to call’
/f/ : /b/	*fai*	‘wife’
	bai	‘too, excessively’
/f/ : /^{m}b/	*foko*	‘throat’
	mboko	‘fruit’
/f/ : /ɓ/	*fu*	‘hair’
	bhu	‘to rise (of sun)’
/f/ : /m/	*fai*	‘wife’
	mai	‘to come’
/f/ : /s/	*sa’o*	‘house’
	fa’o	‘to whip’
/f/ : /ʋ/	*fau*	‘shade’
	wau	‘to smell’
/f/ : /l/	*fau*	‘shade’
	lau	‘seaward’
/f/ : /r/	*fua*	‘wasp’
	rua	‘two’
/f/ : /ɰ/	*fa*	‘to blow gently’
	gha	‘here’

Voiceless apico-alveolar fricative /s/ The voiceless apico-alveolar fricative /s/ appears both word-initially and word-medially:

su’a	[ˈsuʔa]	‘iron’
moso	[ˈmoso]	‘to spoil, to rot’

The apico-alveolar fricative /s/ is occasionally aspirated in stressed word-initial position, with a brief period of voicelessness after the frication ends but before the voicing of the vowel begins:

sa’o	[ˈs^{h}aʔo]	‘house’

The following minimal pairs show that /s/ contrasts with the other continuants and coronal phonemes, as well as the absence of any phoneme:

/s/ : /∅/	initial	*séo*	'to fry'	medial	*pasu*	'cheek'
		éo	'that, which; REL'		*pau*	'mango'
/s/ : /t/	initial	*sa'o*	'house'	medial	*kesa*	'to add more'
		ta'o	'to gather up'		*keta*	'cold'
/s/ : /d/	initial	*su'u*	'to carry on head'	medial	*esa*	'seed, fruit'
		du'u	'to stop'		*eda*	'father's brother'
/s/ : /ⁿd/	initial	*su'a*	'iron'	medial	*moso*	'to spoil, to rot'
		ndu'a	'forest, highlands'		*mondo*	'to gather up'
/s/ : /ɗ/	initial	*sawé*	'to finish'	medial	*kasa*	'chest'
		dhawé	'to work (field)'		*kadha*	'to yell'
/s/ : /dʒ/	initial	*sala*	'wrong'	medial	*masa*	'clean'
		jala	'path'		*maja*	'dry'
/s/ : /n/	initial	*sai*	'who'	medial	*esa*	'seed, fruit'
		nai	'to climb'		*ena*	'sand'
/s/ : /f/	initial	*sa'o*	'house'	medial	N/A	N/A
		fa'o	'to whip'		N/A	N/A
/s/ : /ʋ/	initial	*so'o*	'more'	medial	*ngasu*	'hundred'
		wo'o	'bow'		*ngawu*	'object, treasure'
/s/ : /l/	initial	*sia*	'bright, daylight'	medial	*kesa*	'to add more'
		lia	'hole, cave'		*kela*	'to split'
/s/ : /r/	initial	*sia*	'bright, daylight'	medial	*masa*	'clean'
		ria	'large'		*mara*	'to cherish'
/s/ : /ɰ/	initial	*sawa*	'large snake'	medial	*léghu*	'clumsy'
		ghawa	'downwards'		*lésu*	'cloth headdress'

Glottal fricative /h/ The glottal fricative /h/ is found only in one high frequency word in Detukeli Lio, the word for 'yes':

ho'o [ˈhoʔo] 'yes'

In south-eastern Lio dialects of Lio, centered around the village of Lisédetu, there has been a shift where certain instances of initial /k-/ became /h-/, with no clear conditioning environment. As a result, in the south-eastern Lio dialects where /h/ is a full phoneme, it is still restricted to word-initial position (like /f/). The following words are pairs drawn from my corpus of Detukeli Lio compared with entries from Ardnt (1933), which reflect the south-eastern variant:

DETUKÉLI LIO	SOUTH-EASTERN LIO	GLOSS
ki	*hi*	'thatch'
kiwa	*hiwa*	'year'
keku	*heku*	'tender'
kaki	*haki*	'husband, male'
ko'o	*ho'o*	'slave'

It is not possible to show a full set of minimal pairs since /h/ appears in only one word, but it is possible to show that other words ending in the sequence /-oʔo/ contrast with *ho'o* 'yes':

/h/ : /p/	*ho'o*	'yes'
	po'o	'bamboo joint'
/h/ : /t/	*ho'o*	'yes'
	to'o	'to arise'
/h/ : /k/	*ho'o*	'yes'
	ko'o	'slave'
/h/ : /b/	*ho'o*	'yes'
	bo'o	'full (of food)'
/h/ : /m/	*ho'o*	'yes'
	mo'o	'to want; FUT'
/h/ : /n/	*ho'o*	'yes'
	no'o	'with, and; COM'
/h/ : /s/	*ho'o*	'yes'
	so'o	'more, very'
/h/ : /ʋ/	*ho'o*	'yes'
	wo'o	'bow'
/h/ : /l/	*ho'o*	'yes'
	lo'o	'small'
/h/ : /r/	*ho'o*	'yes'
	ro'o	'fast'

2.3.9 Approximants

Voiced labio-dental approximant /ʋ/ The voiced labio-dental approximant /ʋ/ (spelled *w*) is produced with constriction between top teeth and the bottom lip. This constriction may sometimes be strong enough to cause frication, so the realization ranges from an approximant [ʋ] to a true fricative [v]. It appears both word-initially and word-medially:

watu	[ˈʋatʰu]	'stone'
ghawa	[ˈɰaʋa]	'downwards'

Although they are share some phonetic overlap, this phoneme does not form a voiced-voiceless pair with the voiceless labio-dental fricative /f/. The frication in /f/ is much stronger than in /ʋ/, and /f/ has additional restrictions on its distribution which /ʋ/ does not share.

The following minimal pairs show that /ʋ/ contrasts with the other continuants and labial phonemes, as well as the absence of any phoneme:

/ʋ/ : /∅/	initial	*wau*	'to smell'	medial	*béwa*	'long, tall'
		au	'bamboo'		*béa*	'to repent'
/ʋ/ : /p/	initial	*wati*	'basket'	medial	*awu*	'ash, dust'
		pati	'to give'		*apu*	'dew'
/ʋ/ : /b/	initial	*wara*	'shoulder'	medial	*lawa*	'to fight'
		bara	'white'		*laba*	'hive'
/ʋ/ : /^{m}b/	initial	*wulu*	'thin bamboo'	medial	*lawa*	'to fight'
		mbulu	'ten'		*lamba*	'drum'
/ʋ/ : /ɓ/	initial	*wanda*	'to dance (with cloth)'	medial	*lawa*	'to fight'
		bhanda	'rich'		*labha*	'dung heap'
/ʋ/ : /m/	initial	*wa'u*	'to go down'	medial	*kuwu*	'hut'
		ma'u	'beach'		*kumu*	'fingernail'
/ʋ/ : /f/	initial	*wau*	'to smell'	medial	N/A	N/A
		fau	'shade'		N/A	N/A
/ʋ/ : /s/	initial	*wo'o*	'bow'	medial	*ngawu*	'object, treasure'
		so'o	'more'		*ngasu*	'hundred'
/ʋ/ : /l/	initial	*wau*	'to smell'	medial	*kiwa*	'year'
		lau	'seaward'		*kila*	'to flash'
/ʋ/ : /r/	initial	*wé'é*	'close, near'	medial	*pawé*	'good'
		ré'é	'bad'		*paré*	'uncooked rice'
/ʋ/ : /ɰ/	initial	*wéa*	'gold'	medial	*lawa*	'to fight'
		ghéa	'there'		*lagha*	'to split off'

Voiced apico-alveolar lateral approximant /l/ The voiced apico-alveolar lateral approximant /l/ appears both word-initially and word-medially:

lé'é [ˈleʔe] 'arrow, sharp'
kéli [ˈk^{h}eli] 'mountain'

The following minimal pairs show that /l/ contrasts with the other continuants and coronal phonemes, as well as the absence of any phoneme:

/l/ : /∅/	initial	*lau*	'seaward'	medial	*alé*	'to ask'
		au	'bamboo'		*aé*	'water'
/l/ : /t/	initial	*la'i*	'to lick'	medial	*alé*	'to ask'
		ta'i	'excrement'		*até*	'liver'
/l/ : /d/	initial	*lau*	'seaward'	medial	*gulu*	'thorn'
		dau	'must'		*gudu*	'to startle'
/l/ : /ⁿd/	initial	*lu*	'teardrop'	medial	*molo*	'correct'
		ndu	'to follow, to obey'		*mondo*	'to gather up'
/l/ : /ɗ/	initial	*lu*	'teardrop'	medial	*ngala*	'can, to be able'
		dhu	'to pound'		*ngadha*	'to look around'
/l/ : /dʒ/	initial	*lengi*	'oil'	medial	*alé*	'to ask'
		jengi	'to burn (land)'		*ajé*	'string'
/l/ : /n/	initial	*lu*	'teardrop'	medial	*mila*	'dark'
		nu	'smoke, cloud'		*mina*	'oil, grease'
/l/ : /f/	initial	*lau*	'seaward'	medial	N/A	N/A
		fau	'shade'		N/A	N/A
/l/ : /s/	initial	*lia*	'hole, cave'	medial	*kela*	'to split'
		sia	'bright, daylight'		*kesa*	'to add more'
/l/ : /ʋ/	initial	*lau*	'seaward'	medial	*kila*	'to flash'
		wau	'to smell'		*kiwa*	'year'
/l/ : /r/	initial	*la'i*	'to lick'	medial	*alé*	'to ask'
		ra'i	'to approach'		*aré*	'cooked rice'
/l/ : /ɰ/	initial	*la'i*	'to lick'	medial	*wola*	'again, to return'
		gha'i	'leg, foot'		*wogha*	'piece, portion'

Voiced velar approximant /ɰ/ The voiced velar approximant /ɰ/ (spelled with the digraph *gh*) appears both word-initially and word-medially, but is quite rare word-medially:

ghawa	[ˈɰaʋa]	'downwards'
wogha	[ˈʋoɰa]	'piece, portion'

There is a rather wide range of possible realizations for this phoneme. Most of the time, there is no frication whatsoever and it is a true approximant [ɰ]. However, it may also involve some light frication, in which case [ɣ] or [ɣ̟] would be a more appropriate transcription.

ghéta	[ˈɣetʰa]	'upwards'
ghélé	[ˈɣele]	'towards the mountains'
gha	[ɣa]	'here'

This phoneme has low type frequency but high token frequency, appearing in a handful of the most frequent spatial deictic markers but few other Lio words. Two sets of spatial deictics, the demonstrative adverbs and the directional adverbs, contain initial /ɰ-/ in

most members of the paradigm, but this falls short of being a productive process. For instance, the three demonstrative adverbs all begin with /ɯɻ/:

gha	'here (proximal)'
ghéa	'there (medial)'
gharu	'over there (distal)'

However, of the six directional adverbs, only four begin with /ɯɻ/:

ghéta	'upwards; towards the mountains'
ghalé	'downwards; west along the coast'
ghélé	'upwards'
ghawa	'downwards'

The two members of the paradigm which do not contain an initial /ɯɻ-/ are *mena* 'east along the coast' and *lau* 'towards the sea'. The latter is a very ancient term derived from Proto-Austronesian **lahud* 'downstream, towards the sea' and cognate with Malay *laut* 'sea'.

In other languages of Flores with a comparable phoneme, it is usually analyzed as a voiced velar fricative /ɣ/ (see Arndt 1961 on Ngadha, McDonnell 2009 on Ende, Arka 2016 on Rongga). I prefer to analyze it as an approximant for two reasons. The first is that frication is the exception rather than the norm when producing this phoneme. The second is that under my analysis, all Lio fricatives are voiceless while all approximants are voiced. However, in the second section where I treat the CF languages comparatively, I will use the unified symbol /ɣ/ for the equivalent phoneme in all languages.

The following minimal pairs show that /ɯɻ/ contrasts with the other continuants and dorsal phonemes (including glottal stop /ʔ/), as well as the absence of any phoneme:

/ɯ/ : /∅/	initial	*ghalé*	'downwards, west'	medial	*wogha*	'piece, portion'
		alé	'to ask'		*woa*	'to lie'
/ɯ/ : /ʔ/	initial	N/A	N/A	medial	wogha	'piece, portion'
		N/A	N/A		*wo'a*	'to split off'
/ɯ/ : /g/	initial	*ghalé*	'downwards, west'	medial	*lagha*	'to split off'
		galé	'to choose'		*laga*	'to wriggle'
/ɯ/ : /ᵑg/	initial	*gha*	'here'	medial	*wogha*	'piece, portion'
		ngga	'cool, fresh'		*wongga*	'to boil'
/ɯ/ : /ŋ/	initial	*gha*	'here'	medial	*wogha*	'piece, portion'
		nga	'to leak'		*wonga*	'flower'
/ɯ/ : /f/	initial	*gha*	'here'	medial	N/A	N/A
		fa	'to blow gently'		N/A	N/A
/ɯ/ : /s/	initial	*ghawa*	'downward'	medial	*léghu*	'clumsy'
		sawa	'large snake'		*lésu*	'cloth headdress'
/ɯ/ : /ʋ/	initial	*ghéa*	'there'	medial	*lagha*	'to split off'
		wéa	'gold'		*lawa*	'to fight'
/ɯ/ : /l/	initial	*gha'i*	'leg, foot'	medial	*wogha*	'piece, portion'
		la'i	'to lick'		*wola*	'again, to return'
/ɯ/ : /r/	initial	*gha'i*	'leg, foot'	medial	*wogha*	'piece, portion'
		ra'i	'to approach'		*wora*	'foam'

2.3.10 Trills

Voiced apico-alveolar trill /r/ The voiced apico-alveolar trill /r/ is the only rhotic and it appears both word-initially and word-medially. It is in free variation everywhere between a tap [ɾ] and a trill [r], but the tap realization is more common than the trill:

ra'i	['ɾaʔi, 'raʔi]	'to approach'
koro	['kʰoɾo, 'kʰoro]	'chili pepper'

The following minimal pairs show that /r/ contrasts with the other continuants and coronal phonemes, as well as the absence of any phoneme:

/r/ : /∅/	initial	*rua*	'two'	medial	*paru*	'to run'
		ua	'rattan'		*pau*	'mango'
/r/ : /t/	initial	*ra'i*	'to approach'	medial	*jara*	'horse'
		ta'i	'excrement'		*jata*	'eagle'
/r/ : /d/	initial	*ra*	'blood'	medial	*wora*	'foam'
		da	'towards'		*woda*	'bell'
/r/ : /ⁿd/	initial	*raté*	'tomb'	medial	*wara*	'shoulder'
		ndaté	'heavy'		*wanda*	'to dance (with cloth)'
/r/ : /ɗ/	initial	*rua*	'two'	medial	*koro*	'chili pepper'
		dhua	'misfortune'		*kodho*	'to look'
/r/ : /dʒ/	initial	*ra*	'blood'	medial	*mara*	'to cherish'
		ja	'to shine'		*maja*	'dry'
/r/ : /n/	initial	*ria*	'large'	medial	*ngara*	'owner'
		nia	'face'		*ngana*	'termite'
/r/ : /f/	initial	*rua*	'two'	medial	N/A	N/A
		fua	'wasp'		N/A	N/A
/r/ : /s/	initial	*ria*	'large'	medial	*mara*	'to cherish'
		sia	'bright, daylight'		*masa*	'clean'
/r/ : /ʋ/	initial	*ré'é*	'bad'	medial	*paré*	'uncooked rice'
		wé'é	'close, near'		*pawé*	'good'
/r/ : /l/	initial	*ra'i*	'to approach'	medial	*aré*	'cooked rice'
		la'i	'to lick'		*alé*	'to ask'
/r/ : /ɰ/	initial	*ra'i*	'to approach'	medial	*wora*	'foam'
		gha'i	'leg, foot'		*wogha*	'piece, portion'

2.4 Vowel inventory

Lio has a total inventory of 6 vowels (see Table 3) which can be divided into the 5 'full' vowels /i, u, e, o, a/ plus the schwa /ə/. Their orthographic representations are given in angled brackets as well. Note that the unaccented symbol *e* represents IPA schwa /ə/, while the acute accent *é* represents the IPA mid vowel /e/.

	FRONT	CENTRAL	BACK
CLOSE	i ⟨*i*⟩		u ⟨*u*⟩
MID	e ⟨*é*⟩	ə ⟨*e*⟩	o ⟨*o*⟩
OPEN		a ⟨*a*⟩	

Table 3: Lio vowel phonemes

2.4.1 Full vowels

The five full vowels /i, u, e, o, a/ are different from the schwa /ə/ (which is spelled with *e*) in that there are no restrictions on their distribution. They may appear in either syllable of a disyllable, or as the sole vowel in a monosyllable. They can form vowel sequences with each other, and all possible sequences of full vowels are attested (see section 2.5.3). When two vowels form a sequence, they always remain two separate syllables in careful speech.

The full vowels are predictably lengthened in monosyllabic lexical items, in contrast to unstressed grammatical words which do not undergo lengthening. This lengthening serves to fulfill the bimoraic requirement, which requires that the surface forms of all lexical words consist of exactly two moras (see section 2.5.4).

The close vowels /i, u/ and mid vowels /e, o/ form natural classes and share certain properties. The pair /i, u/ are optionally lowered in word-final position, approaching the cardinal vowels [e, o]. The pair /e, o/ range from mid [e̬, o̬] to open-mid [ɛ, ɔ]. They are especially open before glottal stop /ʔ/.

In sequences of a close vowel followed by a non-high vowel, a phonetic glide is inserted between the two. However, there is still a clear phonetic contrast between the presence and absence of an approximant such as *w* /ʊ/ in that position:

nia	[ˈniʲa]	‘face’
lio	[ˈliʲo]	‘Lio (people, language)’
nua	[ˈnuʷa]	‘village’
nuwa	[ˈnuʊa]	‘to grow, age’

Close front unrounded vowel /i/ The high front unrounded vowel /i/ may appear in either syllable of a disyllable, or as the sole vowel in a monosyllable. This vowel is predictably lengthened in lexical monosyllables and stress-bearing grammatical words:

kita	[ˈkʰitʰa]	‘we (inclusive); 1PL.I’
ra'i	[ˈraʔi]	‘to approach’
ki	[kʰiː]	‘thatch’

It can be further lowered in word-final position, approaching cardinal close-mid [e]:

kawi	[ˈkʰaʊe]	‘to marry’

The following minimal pairs show that /i/ contrasts with the other vowels in penultimate position, final position (second vowel of a disyllable), and as the sole vowel of a monosyllable:

	PENULT		FINAL		MONO	
/i/ : /u/	*mila*	‘dark’	*weli*	‘price’	*di*	‘also’
	mula	‘to plant’	*welu*	‘to put’	*du*	‘to reach’
/i/ : /e/	*kinga*	‘ear’	*aji*	‘younger sibling’	*ki*	‘thatch’
	kénga	‘to lie on back’	*ajé*	‘string’	*ké*	‘to cry’
/i/ : /o/	*dhika*	‘pure’	*toki*	‘to bite’	*mi*	‘sweet’
	dhoka	‘deaf’	*toko*	‘bone’	*mo*	‘tired’
/i/ : /a/	*piré*	‘forbidden’	*laki*	‘chief’	*ki*	‘thatch’
	paré	‘uncooked rice’	*laka*	‘to help’	*ka*	‘to eat’
/i/ : /ə/	*lika*	‘fireplace’	N/A	N/A	N/A	N/A
	leka	‘in, at; LOC’	N/A	N/A	N/A	N/A

Close back rounded vowel /u/ The high front unrounded vowel /u/ may appear in either syllable of a disyllable, or as the sole vowel in a monosyllable. This vowel is predictably lengthened in lexical monosyllables and stress-bearing grammatical words:

ru’é	[ˈruʔe]	‘to chew something hard’
ngawu	[ˈŋaʊu]	‘object, treasure’
tu	[tʰuː]	‘to bring’

It can be further lowered in word-final position, approaching cardinal close-mid [o]:

paru	[ˈpʰaro]	‘to run’

The following minimal pairs show that /u/ contrasts with the other vowels in penultimate position, final position (second vowel of a disyllable), and as the sole vowel of a monosyllable:

	PENULT		FINAL		MONO	
/u/ : /i/	*mula*	‘to plant’	*welu*	‘to put’	*du*	‘to reach’
	mila	‘dark’	*weli*	‘price’	*di*	‘also’
/u/ : /e/	*bua*	‘body hair’	*ma’u*	‘beach’	*bu*	‘drunk’
	béa	‘to repent’	*ma’é*	‘do not; PROH’	*bé*	‘to pour’
/u/ : /o/	*nuka*	‘to rise’	*lu*	‘teardrop’	*talu*	‘to reply’
	noka	‘to curse’	*lo*	‘body’	*talo*	‘cannot’
/u/ : /a/	*pu’u*	‘trunk’	*lambu*	‘shirt’	*du*	‘to reach’
	pa’u	‘to throw’	*lamba*	‘drum’	*da*	‘towards’
/u/ : /ə/	*tulu*	‘to help’	N/A	N/A	N/A	N/A
	telu	‘three’	N/A	N/A	N/A	N/A

Mid front unrounded vowel /e/ The mid front unrounded vowel /e/ may appear in either syllable of a disyllable, or as the sole vowel in a monosyllable. This vowel is predictably lengthened in lexical monosyllables and stress-bearing grammatical words:

méko	[ˈmekʰo]	‘to move’
garé	[ˈgare]	‘to speak’
ké	[kʰe:]	‘to cry’

The usual realization of this vowel ranges from mid [e̬] to open-mid [ɛ], and it is especially lowered in the vicinity of glottal stop /ʔ/:

ré’é	[ˈrɛʔɛ]	‘bad, evil’

The following minimal pairs show that /e/ contrasts with the other vowels in penultimate position, final position (second vowel of a disyllable), and as the sole vowel of a monosyllable:

	PENULT		FINAL		MONO	
/e/ : /i/	*kénga*	‘to lie on back’	*ajé*	‘string’	*ké*	‘to cry’
	kinga	‘ear’	*aji*	‘younger sibling’	*ki*	‘thatch’
/e/ : /u/	*béa*	‘to repent’	*ma’é*	‘do not; PROH’	*bé*	‘to pour’
	bua	‘body hair’	*ma’u*	‘beach’	*bu*	‘drunk’
/e/ : /o/	*néa*	‘just now’	*ka’é*	‘older sibling’	*ré*	‘to invite’
	noa	‘easy’	*ka’o*	‘to cradle’	*ro*	‘painful’
/e/ : /a/	*lé’é*	‘arrow’	*até*	‘liver’	*ké*	‘to cry’
	la’é	‘not yet’	*ata*	‘person’	*ka*	‘to eat’
/e/ : /ə/	*ésa*	‘to pull’	N/A	N/A	N/A	N/A
	esa	‘seed, fruit’	N/A	N/A	N/A	N/A

Mid back rounded vowel /o/ The mid back rounded vowel /o/ may appear in either syllable of a disyllable, or as the sole vowel in a monosyllable. This vowel is predictably lengthened in lexical monosyllables and stress-bearing grammatical words:

ndota	[ˈⁿdotʰa]	‘to chop up’
talo	[ˈtʰalo]	‘cannot’
nggo	[ᵑgo:]	‘gong’

The usual realization of this vowel ranges from mid [o̬] to open-mid [ɔ], and it is reliably lowered in the vicinity of glottal stop /ʔ/:

lo’o	[ˈlɔʔɔ]	‘small’

The following minimal pairs show that /o/ contrasts with the other vowels in penultimate position, final position (second vowel of a disyllable), and as the sole vowel of a monosyllable:

	PENULT		FINAL		MONO	
/o/ : /i/	*dhoka*	'deaf'	*toko*	'bone'	*mo*	'tired'
	dhika	'pure'	*toki*	'to bite'	*mi*	'sweet'
/o/ : /u/	*noka*	'to curse'	*talo*	'cannot'	*lo*	'body'
	nuka	'to rise'	*talu*	'to reply'	*lu*	'teardrop'
/o/ : /e/	*noa*	'easy'	*ka'o*	'to cradle'	*ro*	'painful'
	néa	'just now'	*ka'é*	'older sibling'	*ré*	'to invite'
/o/ : /a/	*tolo*	'summit'	*boko*	'short'	*ro*	'painful'
	talo	'cannot'	*boka*	'to fall'	*ra*	'blood'
/o/ : /ə/	*gomo*	'hole'	N/A	N/A	N/A	N/A
	gemo	'to touch'	N/A	N/A	N/A	N/A

Open central unrounded vowel /a/ The open central unrounded vowel /a/ may appear in either syllable of a disyllable, or as the sole vowel in a monosyllable. This vowel is predictably lengthened in lexical monosyllables and stress-bearing grammatical words:

sawé	['saʋe]	'to finish; completive aspect marker'
buga	['buga]	'morning'
ka	[k^{h}a:]	'to eat'

The usual realization of this vowel in stressed position is slightly further back than the cardinal front [a], closer to to open central [ä]. In unstressed position it often reduced to [ɐ]. In the following example, the unstressed reciprocal marker *rapa* 'each other; RECIP has reduction in both syllables because neither is stressed, and is realized the same as the stressed schwa /ə/ in the following word *bego*:

(2)

rapa	*bego*
[rɐpɐ	'bɐg:o]
RECIP	play

'(They) are playing together'

The following minimal pairs show that /a/ contrasts with the other vowels in penultimate position, final position (second vowel of a disyllable), and as the sole vowel of a monosyllable:

	PENULT		FINAL		MONO	
/a/ : /i/	*paré*	'uncooked rice'	*laka*	'to help'	*ka*	'to eat'
	piré	'forbidden'	*laki*	'chief'	*ki*	'thatch'
/a/ : /u/	*ta'u*	'to fear'	*lamba*	'drum'	*da*	'towards'
	tu'u	'dry'	*lambu*	'shirt'	*du*	'to reach'
/a/ : /e/	*la'é*	'not yet'	*ata*	'person'	*ka*	'to eat'
	lé'é	'arrow'	*até*	'liver'	*ké*	'to cry'
/a/ : /o/	*talo*	'cannot'	*laka*	'to help'	*ra*	'blood'
	tolo	'summit'	*lako*	'dog'	*ro*	'painful'
/a/ : /ə/	*laka*	'to help'	N/A	N/A	N/A	N/A
	leka	'in, at; LOC'	N/A	N/A	N/A	N/A

2.4.2 Schwa

Mid central unrounded vowel /ə/ The mid central unrounded vowel, or schwa /ə/ has three restrictions which the other vowels do not share: it may not be the final syllable of a word, it may not form a vowel sequence, and it may not come before a glottal stop. Since it is barred from the final syllable, this means it cannot be the sole vowel of a monosyllable and it must appear as the first vowel in a $(C_1)VC_2V$ disyllable where C_2 is not glottal stop /ʔ/. Stress is somewhat unpredictable on words containing schwa /ə/, with some clearly stressed on the final syllables, others clearly stressed on the initial syllable, and others stressed on either syllable depending on prosodic context. Stressed schwa is generally realized as to [ɐ], while unstressed schwa ranges from [ɐ] to [ə]:

ebé	[ˈʔɐbːe, ʔəˈbːe]	'they; 3PL'
ema	[ˈʔɐmːa, ʔəˈmːa]	'father'
welu	[ˈʋɐlːu]	'to put, to leave'
kema	[ˈkʰɐmːa]	'to work'
nggebhé	[ˈᵑgɐβːe]	'to lie flat, to cover'
peri	[pəˈɾi]	'kind of bamboo'
seré	[səˈɾe]	'to boil (rice)'

The following minimal pairs show that /ə/ contrasts with the other vowels, but only as the penultimate syllable in a disyllable:

/ə/ : /i/	*leka*	'in, at; LOC'
	lika	'fireplace'
/ə/ : /u/	*bega*	'friend'
	buga	'morning'
/ə/ : /e/	*esa*	'seed, fruit'
	ésa	'to pull'
/ə/ : /o/	*kela*	'to split'
	kola	'to chase'
/ə/ : /a/	*leko*	'frog'
	lako	'dog'

2.4.3 Phonetic properties of schwa

There is some phonetic overlap between /a/ and /ə/, which can both be realized as [ɐ]. However, the difference in quality between the two phonemes is reinforced by secondary phonetic cues. While the full vowels /i, u, e, o, a/ are of approximately equal length under constant conditions, the schwa /ə/ is only half as long as the others. In addition, word-medial consonants are lengthened by roughly 25% after schwa /ə/ (although this does not apply to the consonant /r/).

In order to quantify the phonetic differences between schwa and the other vowels, I took the minimal pair *leko* /ləko/ 'frog' and *lako* /lako/ 'dog' as a case study. It is a fortunate coincidence that two of the main characters in the Frog Story are a dog and a frog, so I had access to a large number of tokens of both words recorded in a single sitting under comparable conditions.

In order to compare these two vowels phonetically, I first identified and isolated 56 tokens of *lako* /lako/ 'dog' and 38 tokens of *leko* /ləko/ 'frog' from the recording. I coded these tokens into four categories: 'utterance-initial', 'utterance-medial', 'utterance-final' and 'isolated' (that is, single-word utterances). I discarded all outlier tokens with length greater than 500ms, as these were usually the result of hesitation. For my final analysis, I considered only utterance-medial tokens, because these were the most numerous. In the end, I was left with 18 utterance-medial tokens of *lako* /lako/ 'dog' and 23 utterance-medial tokens of *leko* /ləko/ 'frog'. Each token has the phonetic structure [$C_1V_1(^h)C_2{}^hV_2$], with regular post-aspiration and optional pre-aspiration of C_2. Any pre-aspiration, if present, was counted as part of the duration of the preceding vowel. The mean duration of each segment, in milliseconds, is shown in Table 4.

As seen in Table 4, the mean duration of the two words is roughly equal: 308ms for *leko* /ləko/ 'frog' and 312ms for *lako* /lako/ 'dog'. Thus, the difference between the two words is not one of overall duration, but rather of the proportion of duration that each segment takes up. This can be better visualized with the 100% stacked bar chart in Figure 2, where the duration of each segment is shown as a proportion of the total duration.

The most salient difference between the two words is the duration of V_1 itself. The

C_1	V_1	C_2	^{h}	V_2	TOTAL
l	ə	k	^{h}	o	
73ms	47ms	73ms	41ms	74ms	308ms
l	a	k	^{h}	o	
67ms	93ms	58ms	30ms	64ms	312ms

Table 4: Mean duration of each segment in milliseconds

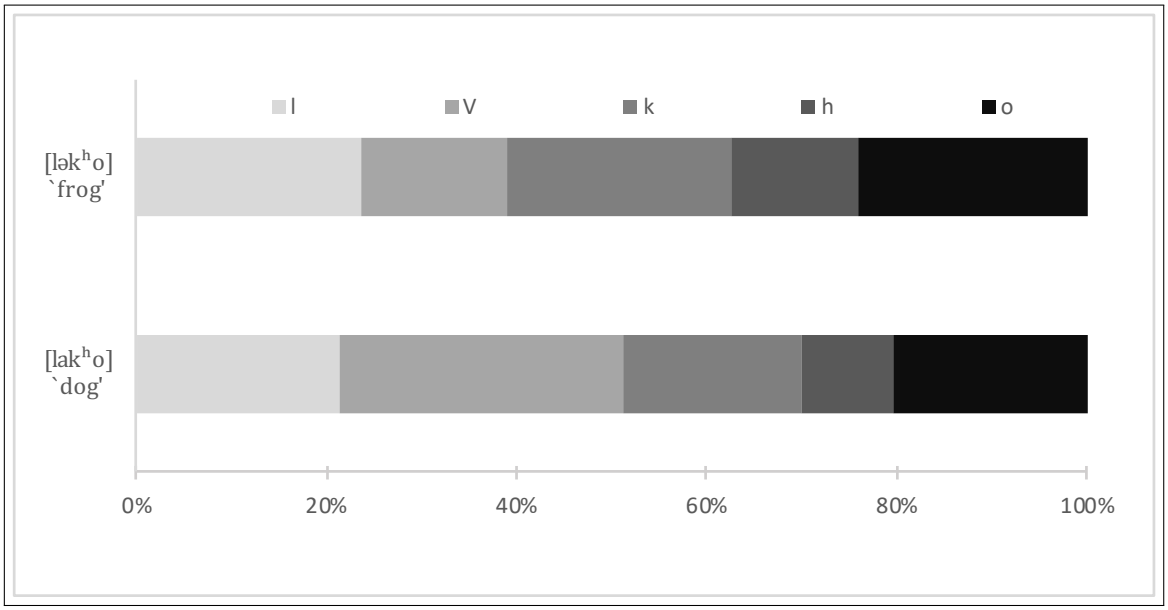


Figure 2: Duration of each segment as a proportion of total duration

full vowel /a/ lasts 106% longer than schwa /ə/ in the same position, a highly salient perceptual difference (97ms for /a/ vs. 47 ms for /ə/). Given that the overall duration stays roughly the same, other segments in *leko* /ləko/ 'frog' are lengthened to compensate for the reduction of the vowel.

This lengthening affects primarily C_2 (/k/) and its accompanying aspiration. The results show that after schwa /ə/, the medial /k/ is held 26% longer (73ms in /ləko/ vs. 58ms in /lako/), and the aspiration lasts 37% times longer (41ms in /ləko/ vs. 30ms in /lako/).

On the other hand, the initial /l/ of *leko* /ləko/ is only 9% times longer than that of *lako* /lako/ (73ms in /ləko/ vs. 67ms in /lako/), and the final /o/ is only 16% times longer (74ms in /ləko/ vs. 64ms in /lako/). These slight increases are not nearly as perceptually salient as the changes in duration of V_1 and C_2.

Thus, the conclusions drawn from this comparison are that, in comparable environments, the schwa /ə/ is about half as long as a full vowel, and causes an increase of around 25% in the length of the following consonant. It is my impression that the schwa /ə/ lengthens different classes of consonants to different degrees, affecting obstruents more than continuants and leaving the alveolar trill /r/ unaffected, but this must be confirmed by further phonetic analysis.

One plausible (but ultimately incorrect) account for the phonetic differences between schwa and the full vowels is that the main vowels can take stress, while the schwa cannot. This would predict that all instances of *leko* take final stress /ləˈo/ 'frog' while all instances of *lako* /ˈlako/ 'dog' have the usual initial stress of Lio, thereby explaining the fact that schwa is so much shorter, and also why the aspiration is heavier on the form *leko* /ləˈo/ with final stress. However, this proposal is empirically false in Lio (see section 2.4.2), and there are cases in my corpus where *leko* 'frog' receives unambiguous initial stress.

2.4.4 Epenthetic schwa analysis

The analysis which I have presented for the Lio vowel inventory is that Lio has 6 phonemic vowels ('the phonemic schwa analysis'). These fall into two classes, the full vowels /i, u, e, o, a/ and the schwa /ə/, which have distinct phonological behavior. This is the analysis adopted by myself, as well as Baird (2002) for Kéo and Arka (2016) for Rongga (closely related languages with similar phonology to Lio). The major drawback of this analysis is that the special properties of schwa /ə/ must be arbitrarily stipulated. It is descriptively valid, but it does not provide a principled reason for why these differences exist. As a reminder, the special phonological and phonetic properties of schwa which call for explanation are the following:

1) Schwa may not appear as the final vowel in a word
2) Schwa may not form vowel sequences with other vowels
3) Schwa is phonetically shorter than the other vowels
4) Schwa may not precede the glottal stop /ʔ/
5) Schwa causes lengthening of the following consonant

One potential analysis is that the schwa is not a phonemic vowel, but is inserted epenthetically to break up underlying consonant clusters ('the epenthetic schwa analysis'). This is the analysis adopted by McDonnell (2009) for Ende and Donohue (2005) for Palu'e. The major benefit of this analysis is that at least the first three special properties of the schwa /ə/ are explained. However, the drawbacks associated with this analysis outweigh the benefits, and I have ultimately chosen not to adopt it for Lio.

Under an epenthetic schwa analysis, Lio has only 5 phonemic vowels /i, u, e, o, a/ and allows /C_1C_2V/ syllables with underlying initial consonant clusters. A ban on consonant clusters in the surface forms leads to schwa epenthesis, yielding surface forms of type [CəCV] from underlying /C_1C_2V/ syllables.

UNDERLYING	SURFACE	GLOSS
/ptu/	[pətu]	'hot'
/kku/	[kəku]	'tender'
/bga/	[bəga]	'friend, comrade'
/ddu/	[dədu]	'to announce'
/ᵑgɓe/	[ᵑgəɓe]	'to lie on stomach'
/ɓni/	[ɓəni]	'wonderful'
/ɗso/	[ɗəso]	'to recede (water)'
/dʒɗo/	[dʒəɗo]	'to pity'
/mra/	[məra]	'to live'
/mŋa/	[məŋa]	'only'
/lko/	[ləko]	'frog'
/ldʒa/	[lədʒa]	'day'
/ʊva/	[ʊəva]	'yard'

The major strength of this analysis is that it reduces the number of vowels from 5 to 6, allowing for a more concise description of the Lio vowel inventory. In doing so, it also accounts for the first three special properties of schwa, but does not readily account for the remaining two properties.

The first special property of schwa, the fact that it cannot appear as the final vowel in a word, follows naturally from the fact that it is inserted only to break up underlying consonant clusters in words of type /C_1C_2V/. All words are still required to end with one of the phonemic vowels, /i, u, e, o, a/, and there is no way to insert the schwa /ə/ such that it will end up in final position.

The second special property of schwa, the fact that it cannot form vowel sequences,

follows for the same reasons. Given that schwa is used to break up underlying /C_1C_2V/ consonant clusters, it will only ever appear between two consonants. A word which contains a vowel sequence must have an empty C_2 slot by definition, so it cannot be targeted by schwa epenthesis.

The third special property of schwa, the fact that it is phonetically shorter than the other vowels, can be explained by the fact that schwa is merely inserted for ease of pronunciation, and is not equal to the phonemic vowels /i, u, e, o, a/ from a prosodic point of view. The reduced length of the schwa would be well-accounted for by the epenthetic schwa analysis.

However, there is no clear explanation for the fourth special property of schwa, the fact that it cannot come before a glottal stop. A surface form of the prohibited type [*CəʔV] (such as [*səʔu]), would follow from an underlying form of type /*CʔV/ (such as /*sʔu/). Why such an underlying form should be prohibited is not explained by this analysis, as there are no other prohibited cluster types.

As for fifth special property of the schwa, the lengthening of the following consonant, the epenthetic schwa analysis holds some promise. The bimoraic requirement (see section 2.5.4) states that all open class words and stressed function words must consist of exactly two moras on their surface forms. Supposing that the epenthetic schwa analysis is correct, that provides a principled explanation for why it does not constitute a full mora. The lengthening of the following consonant then serves to compensate for the moraically deficient schwa.

The epenthetic schwa analysis has further ramifications, both positive and negative. Some Lio words begin with a schwa under my analysis, such as *ebé* /əbe/ 'they; 3PL'. In order to arrive at such a surface form, there must be an underlying /CCV/ consonant cluster which is being broken up. The easiest solution is to propose is that an underlying form of type /ʔCV/ (such as /ʔbe/) yields a surface form of type [ʔəCV] (such as [ʔəbe]), and this is supported by the fact that all vowel-initial words in Lio do indeed begin with a glottal stop, at least phonetically (see section 2.3.2).

UNDERLYING	SURFACE	GLOSS
/ʔbe/	[ʔəbe]	'they; 3PL'
/ʔma/	[ʔəma]	'father'
/ʔmba/	[ʔəmba]	'which'
/ʔmu/	[ʔəmu]	'mosquito'
/ʔna/	[ʔəna]	'sand'
/ʔsa/	[ʔəsa]	'seed, one'

Since an initial phonemic glottal stop must be posited to account for words such as these, the most economical assumption is to assume that all other word-initial glottal stops (which I have treated as a predictable phonetic phenomenon in my analysis) are in fact phonemic (so what I analyze as *ata* /ata/ 'person' would now be analyzed as *'ata*

/ʔata/). If these initial glottal stops are taken to be phonemic in all vowel-initial words, then the entire view of Lio phonotactics changes.

There would no longer be vowel-initial words in Lio, since previously vowel-initial words would now begin with a phonemic glottal stop. Furthermore, it would eliminate the exception that the glottal stop /ʔ/ is the only consonant which is not allowed in word-initial position. All consonants would be allowed in initial position, and all words would have to begin with a consonant (no more vowel-initial words).

The major drawback of the epenthetic schwa analysis is that it greatly complicates the syllable structure of Lio. Under the phonemic schwa analysis, all Lio syllables are simply (C)V open syllables with optional onsets. However, if schwa is epenthetic (and hence phonemic initial glottal stop must be posited), then word-initial syllables would be (C)CV, with at least one onset consonant (and possibly a cluster). All clusters would be allowed in underlying forms except ones of type /*CʔV/, formed by a consonant plus glottal stop. The structure of final syllables would remain (C)V. This syllable structure would be typologically very unusual compared to a (C)V syllable structure, and the diversity of clusters in the underlying forms belies the extreme phonotactic simplicity of the surface forms.

In light of this, I have decided that it is preferable to choose the phonemic schwa analysis. This is a more conservative analysis, in that it does not provide the same level of explanatory power as the epenthetic schwa analysis, but it requires fewer assumptions.

While the epenthetic schwa analysis is a purely synchronic matter, it is clear from historical evidence that the schwa was phonemic in the past. Lio is a phonologically conservative language, and the Proto-Austronesian schwa is usually preserved. Thus, the Proto-Austronesian numeral **təlu* 'three' has passed essentially unchanged into modern Lio as *telu* /təlu/ 'three' (assuming similar phonetic values). Given that the epenthetic schwa analysis is not viable for Proto-Austronesian (see Blust 2013 for proposals on Proto-Austronesian phonology), there must have been a stage at which the schwa was indeed phonemic.

2.5 Phonotactics

2.5.1 Syllable structure

Lio has a simple (C)V syllable structure with optional onsets. Only open syllables are permitted and there are no consonant clusters.

2.5.2 Word structure

In the following section, I will use the word 'word' as equivalent to 'morpheme', defined in a grammatical sense as the smallest meaning-bearing unit. I am not referring to the phonological word, which may be somewhat larger than the grammatical word in Lio because of compounding and clitics. The equivalence between 'word' and 'morpheme' in

Lio is due to the fact that it is completely isolating. All morphemes, both lexical and functional, obey the same phonotactic rules.

Lio words are either $(C_1)V_2$ monosyllables or $(C_1)V_1(C_2)V_2$ disyllables. There are no native Lio words of more than two syllables. The restrictions on word structure can be summarized in the following terms, where C_1 is a word-initial consonant, C_2 is a word-medial consonant, V_1 is the penultimate vowel in a disyllable, and V_2 is a final vowel (whether in a disyllable or monosyllable):

FOR ANY WORD $(C_1)V_2$ OR $(C_1)V_1(C_2)V_2$:

1) C_1 may not be /ʔ/	/*ʔato/
2) C_2 may not be /f/ (nor /h/, in relevant dialects)	/*kafo, *kaho/
3) V_2 may not be /ə/	/*gatə/
4) If V_1 is /ə/, then C_2 may not be /ʔ/ nor null /∅/	/*ɓəʔu, *ɓəu/

Any word form not explicitly banned by these rules is allowed. The range of possible word structures in Lio is shown below, with illustrative examples:

STRUCTURE	EXAMPLE	GLOSS
V	*a*	'crow'
	é	'to think'
CV	*ka*	'to eat'
	ngé	'to increase'
	dhu	'to pound'
VV	*aé*	'water'
	éo	'that, which; REL'
	iu	'shark'
	ua	'rattan'
CVV	*kao*	'to scoop'
	séo	'to fry'
	ria	'large'
	bué	'bean'
VCV	*ajé*	'string'
	éko	'tail'
	iwa	'no, not; NEG'
	oté	'brain'
	uma	'garden'
CVCV	*pati*	'to give'
	méko	'to move'
	nggebhé	'to lie flat'
	polu	'to bark'
	mbulu	'ten'

2.5.3 Vowel sequences

All of the full vowels /i, u, e, o, a/ may form sequences with one another within a word. In careful speech, these always form separate syllables, but may be diphthongized or coalesce completely in rapid speech. There are no restrictions on possible combinations of vowels, and all combinations are attested either in my corpus of Lio speech or in Arndt (1933). However, the sequences /ie/ and /uo/ are extremely rare. Vowel sequences do not occur across word boundaries, because a phonetic glottal stop [ʔ] is predictably inserted before all vowel-initial words.

	/a/	/e/	/i/	/o/	/u/
/a/	N/A	*ghéa*	*lia*	*dhoa*	*nggua*
	N/A	'there'	'hole'	'to pity'	'ritual'
/e/	*aé*	N/A	*bhié*	*koé*	*bué*
	'water'	N/A	'full'	'to dig'	'bean'
/i/	*fai*	*déi*	N/A	*poi*	*dui*
	'wife'	'to rejoice'	N/A	'penalty'	'horn'
/o/	*kao*	*léo*	*nio*	N/A	*guo*
	'to scoop'	'to hover'	'coconut'	N/A	'to call'
/u/	*pau*	*nggéu*	*miu*	*bou*	N/A
	'mango'	'left'	'you (plural)'	'to assemble'	N/A

2.5.4 Bimoraic requirement

The 'bimoraic requirement' is a constraint in Lio that states that all open class words and stress-bearing function words must consist of exactly two moras. A phonological conspiracy exists to ensure that, on the surface, all open class words and stress-bearing function words conform to the bimoraic requirement. The end result of the bimoraic requirement is that Lio words are remarkably consistent in their overall duration, with complementary processes of lengthening and shortening acting to bring words into the preferred bimoraic surface form.

The first rule which conspires to fulfill the bimoraic requirement is the lengthening of vowels in monosyllables. Words which are underlyingly monosyllabic, such as the verb *tu* 'to bring', must be realized with a lengthened vowel in order to conform to the bimoraic requirement (see section 2.4.1):

(3) *ebé* *tu* *nggebhé*
[ʔəˈbːe ˈtʰuː ˈᵑgɐɓːe]
3PL bring lie.flat

'They bring (offerings) on their bellies'

The second rule which acts to bring words into conformity with the bimoraic requirement is the lengthening of consonants after schwa. I propose that the schwa /ə/ is different from the other vowels because it does not constitute a full mora (while still remaining phonemic; see section 2.4.2). This explains why consonants are lengthened after schwa /ə/: the lengthening of the consonant compensates for the much shorter vowel length in the surface form, and conspires to keep the overall duration constant. Indeed, I found that for the minimal pair *lako* 'dog' and *leko* 'frog', the total duration of the surface forms are almost equal (see section 2.4.3).

The bimoraic requirement is related to the strong preference for disyllabic roots in Austronesian languages generally. Austronesian languages have inherited a structural pressure towards canonically disyllabic roots, and Blust (2013; p. 684) gives examples of the restoration of canonical disyllables in the face of sound change, from languages as diverse as Javanese, Malay, Tagalog, Kelabit, Iban, Tuvaluan and Samoan. I interpret the bimoraic requirement in Lio as the synchronic manifestation of this phenomenon: sound change and borrowing has introduced monosyllabic Lio words, but these are repaired in the surface forms to generate bimoraic (not bisyllabic) words.

While all open class words (regardless of prosodic context) and stressed function words are bound by the bimoraic requirement, unstressed function words are exempt from it. This gives a direct window into the process of grammaticalization in Lio, since the majority of function words exist side-by-side with the open class words from which they grammaticalized. These lexical-grammatical word pairs are formally identical but have distinct behavior with respect to the bimoraic requirement, in addition to their differing syntactic and semantic properties. I will give examples of several classes of words where the alternations caused by the bimoraic requirement are clearest.

Lengthening of monosyllables An example of a monosyllable with both lexical and grammatical meanings is *du* 'to reach; until'. As a main verb, it means 'to reach, to arrive at', and must be produced [du:] with a long vowel. In this use, it calls for its own core arguments and may itself take TAM markers. On the other hand, it has grammaticalized into a prepositional marker meaning 'until' which is pronounced [du] with a short vowel, which can be further reduced to [dʊ, də]. In this use, it does not take core arguments and does not take TAM markers of its own.

(4) *mo('o) du ghalé wena tana*
[mo ˈduː ˌɰale ˌʋɐna ˈt̪ʰana]
FUT reach down bottom ground

'[He] is going end up down on the ground'

(5) *manusia mo('o) mesu du ghalé wena*
[mɐnusiˈa mo ˈmɐs:u du ɯaˌlé vəˈna]
human FUT fall reach down bottom
'The human is going to fall to the bottom'

In the following example, *du* 'to reach, to arrive at; until' is used in both senses in two successive clauses. The implicit third plural subject in both clauses is the boy from the Frog Story and his dog, who are already introduced in the discourse and therefore elided. In the first clause, the verbal complex is *nangu nai* {swim + climb}, a sequential serial verb construction meaning 'they swim and then climb up', and the grammaticalized marker *du* introduces the endpoint of their motion. In the second clause there is no other candidate for main verb other than *du*, and this is clearly reflected in the pronunciation:

(6) *nangu nai du ghélé wolo*
[ˈnaŋu ˈnai du ˌɯele ˈvolo]
swim ascend reach up$_2$ hill
'[They] swim [and] climb to the top of a hill'

du ghélé wolo
[ˈdu: ˌɯele ˈvolo]
reach up$_2$ hill

'[They] reach the top of the hill'

Vowel coalescence Another example of the bimoraic requirement at work concerns the coalescence of the vowel sequences /ai, au/ to [e, o]. Vowel sequences in Lio are always underlyingly bisyllabic, and in careful speech the two vowels are produced as separate syllables. In rapid speech, it is normal for these sequences to become monosyllabic [aj, aw], but open class words are blocked from coalescing completely to [e, o] while function words are not. By far the most salient example of this is the word *kai* 'he, she, it; 3SG'. When it appears as the head of an NP, the vowel sequence may not coalesce and it must be produced as either [kai] or [kaj]. However, when it appears after another nominal head as the possessor, the vowel sequence may coalesce to [ke] (although it could equally well be pronounced as [kai], [kaj], or even [kej]). An analogous situation applies to the word

tau which means ‘to make, to do’ as a main verb and is produced [tau, taw]. It has grammaticalized into a conjunction meaning ‘in order to’, in which case it may be optionally reduced to as [to].

(7) *ana po tau apa*
[ˌʔana ˈpʰoː ˈtʰau ˈʔapa]
child owl make what
‘What is the owl doing?’

(8) *leko wulo t(au) paru*
[ˈlɐkːʰo ˈʊulo tʰo ˈpʰaru
frog jump.out make run
‘The frog jumps out in order to run’

Truncation The effects of the bimoraic requirement can be observed in the truncation of bisyllabic words. It is not possible to truncate a stressed function word or an open class word, but unstressed function words normally have a truncated monosyllabic variant with a short vowel. The full bimoraic form produced in careful speech or elicitation is listed first, followed by the truncated version which is more frequently encountered in natural speech.

FULL	TRUNCATED	GLOSS
dowa	*do*	‘perfect marker; PERF’
sai	*si*	‘imperative marker; IMP’
no’o	*no*	‘with, and, comitative marker; COM’
mo’o	*mo*	‘future marker; FUT’
éo	*o*	‘that, which, relativizer; REL’
ina	*na*	‘this, that; DEM’
so’o	*so*	‘to say, quotative marker; QUOT’
roa	*ro*	‘first, before, earlier’

Again, it is possible to observe a difference in phonological behavior between the imperative marker *sai*, which is almost always reduced to *si* in spontaneous speech, versus the interrogative and relative pronoun *sai* ‘who’, which must obey the bimoraic requirement when it forms the head of an NP.

(9) *mai sai kai gui lako*
[ˈmai si ˌkʰaj ˈgui ˈlakʰo]
come IMP 3SG shout dog

		SG	PL
1	INCL	*aku*	*kita*
	EXCL		*kami*
2		*kau*	*miu*
3		*kai*	*ebé*

Table 5: Full forms of Lio pronouns

"Come!' he shouts to the dog'

(10) *sai éo iwa tu sobhé nggebhé*
[ˈsai ʔeo ˌʔiʋa ˈtʰu ˈsoβe ˈᵑgɐβ:e]
who REL NEG bring cover lie.flat

'anyone who does not bring offerings on their belly'

Pronoun reduction The pronouns are a class of functional elements which have two major functions: they can act either as the head of an NP, or they can follow the head of an NP to indicate a pronominal possessor. In each of these two functions, they behave differently with respect to the bimoraic requirements. The changes observed go beyond the predictable phonological processes targeting other classes of function words. This represents the first signs of Lio shifting away from a purely isolating template and developing bound possessive marking. Marking possessors and core verbal arguments with pronominal affixes is one of the areal features of eastern Indonesia identified by Klamer (2002), so it is not surprising that the pronouns are showing the most advanced signs of grammaticalization in Lio.

The full forms of the pronouns are listed in Table 5. When a pronoun must obey the bimoraic requirement because it is the head of an NP, these are the forms which must be used, and reduction is not possible.

When the pronouns are used as possessors, they are simply placed after the possessee in an N + GEN construction. In this context, they do not act as the head of an NP and therefore are exempt from the bimoraic requirement. Some pronouns have reduced variants which are occasionally used in this construction. However, it is important to stress that the use of these reduced pronouns is the exception, rather than the rule, for all pronouns except the third singular *kai* 'he, she, it; 3SG'. I have listed the attested phonological variants of each pronoun in order of decreasing frequency. A single ⟩ symbol means that the first variant is more frequent than the second, and a double ⟩⟩ means that it is much more frequent.

1SG	*aku* ⟩⟩ *ku*
2SG	*kau* ⟩⟩ *ko* ⟩ *gho*
3SG	*ké* ⟩ *kai* ⟩⟩ *ghi*
1PL.I	*kita*
1PL.E	*kami*
2PL	*miu*
3PL	*ebé* ⟩⟩ *bé*

The 1st and 2nd person plural pronouns do not have reduced variants, but are always produced in their full bimoraic form, even when used as postposed possessors.

The 1st singular *aku* and 3rd plural *ebé* can undergo truncation to *ku* and *bé* respectively, but this is very rare. The truncated forms of the 1st singular *ku* and the 3rd plural *bé* each appear twice in my corpus of Lio speech, and the full form is used in all other instances.

The 2nd singular *kau* and 3rd singular *kai* both show lenition of the initial /k/ to /ɰ/, an alternation which is not otherwise well-attested in Lio. The 2nd singular *kau* can undergo vowel coalescence to *ko* and further to *gho*, but this is attested only once in my corpus of Lio speech and once during elicitation, when TY offered three variants for *du'a kau, du'a ko, du'a gho* {self + 2SG} 'yourself'. The full form *kau* was used in all other cases.

The only pronoun whose reduced form is more frequent than the full form is the 3rd singular *kai*, which undergoes vowel coalescence to *ké* in roughly 80% of environments where reduction is possible, and otherwise normally remains *kai*. The form *ghi* which show lenition of initial /k/ to /ɰ/ is very rare. It is attested once in my corpus of Lio speech and once during elicitation, when TY offered three variants for *du'a kai, du'a ké, du'a gho* {self + 3SG} 'himself, herself, itself'.

Ongoing grammaticalization Grammaticalization is a gradual process and the line between lexical and functional elements is crossed gradually. There are several examples of ongoing grammaticalization, where previously open class words are beginning to undergo the same phonological changes which characterize functional words. Conversely, there are also examples of recent grammaticalization, where a clearly functional element is unusually resistant to phonological reduction and still shows some characteristics of open class words.

Lio has rich system of nominal classification. This includes a set of 'class terms', which are a small, closed set of common nouns such as *ata* 'person', *aé* 'water', or *pu'u* 'trunk, tree' which are used semi-productively to form a great variety of semantically related compound nouns.

For instance, compound words with *aé* 'water' denote liquids: *aé susu* {water + breast} 'milk', *aé ani* {water + bee} 'honey', *aé nio* {water + coconut} 'coconut water'. If the latter word *nio* 'coconut' can combined with the class term *pu'u* 'trunk, tree' instead, to predictably yield *pu'u nio* {tree + coconut} 'coconut palm'.

Class terms behave like functional elements with respect to the bimoraic requirement and can undergo similar phonological reduction. They often bear a weak secondary stress or no stress at all, and are cliticized to the following word. It is not hard to see how these class terms could further lose their lexical character over time and grammaticalize into a prefixed noun classification system.

FULL	TRUNCATED	GLOSS	ENGLISH
ata laki	*ta laki*	person + chief	'chief'
ata du'a	*ta du'a*	person + early	'parent'
ata kaki	*ta kaki*	person + male	'man'
ata fai	*ta fai*	person + female	'woman'
pu'u kaju	*pu kaju*	trunk + wood	'tree (in general)'
pu'u lélé	*pu lélé*	trunk + banyan	'banyan tree'

An example of a functional element which is resistant to phonological reduction is *to'o* which is used both as an independent verb meaning 'to arise, to get up' and as a prepositional element meaning 'from'. The word *to'o* is an excellent candidate for truncation to *to*, by analogy to the many other truncated grammatical words of that form: *no'o, no* 'with, and', *mo'o, mo* 'future', *so'o, so* 'to say, quotative'. However, out of fifteen instances in my data of *to'o* used in a prepositional sense, none were truncated to *to*. Furthermore, *to'o* 'from' may co-occur with the general locative marker *leka*, as in example 11. These facts suggest that the verb *to'o* 'to arise' has recently grammaticalized into a functional element marking the source of motion.

The original interpretation as a sequential serial verb construction can still be made out in sentences like example 11, where *to'o* 'to arise; from' co-occurs with *leka* 'in, at; LOC':

(11) *mbana gha'i wé'é to'o leka tebénaka kanganara raka du gha éndé*
go foot just arise LOC pl.name pl.name reach_2 reach here pl.name
'[We] just went on foot from Tebénaka [and] Kanganara until here in Ende (or: 'arising in Tebénaka and reaching Ende')'

2.6 Stress

Stress has a low functional load in Lio and is usually initial in isolated words. This applies also words with schwa in the first syllable, but these words then receive a geminated medial consonant.

watu [ˈʋat$^{\text{h}}$u] 'stone'
lako [ˈlak$^{\text{h}}$o] 'dog'
ebé [ˈʔəbːe] 'they; 3PL'

The main class of exceptions to this are words with medial /r/. Medial /r/ does not become geminated when it follows schwa, but instead remains as it is, and the stress shifts rightward to the final syllable.

peri	[pəˈri]	'bamboo species'
mera	[məˈra]	'to live'

The usual rule of stress assignment to the first syllable of each word, in conjunction with the tendency of Lio words to contain two syllables, leads to a very common phrasal pattern of initial stress, followed by alternating unstressed and stressed syllables.

(12) *mosa laki do'i kago kao*
['mosa 'lak^{h}i 'doʔi 'k^{h}ago 'k^{h}ao]
male$_2$ chief pour grab$_2$ scoop$_3$
'The chief pours, grabs [and] scoops [it]'

Functional morphemes which do not receive stress are attached phonologically to either the preceding word or the following word, acting as clitics. Morphemes which cliticize to the following word include the comitative marker *noʔo* (usually reduced to *no* in this context) and the singular proclitic *sa*, while morphemes which cliticize to the preceding word include the demonstrative *na* and the 3rd singular possessive marker *kai*, often reduced to *ké*.

(13) *kai latu no=mbendi nggo=ké*
['k^{h}ai 'lat^{h}u no='mbəndi 'ŋgo:=ke]
3SG EXI COM=gun gong=3SG
'He had his guns [and] gongs'

(14) *néa-néa sa=éko molo=na*
[ˌnea-'nea sə='ʔek^{h}o 'molo=na]
just.now:ADV SG=tail.CL correct=DEM
'Just now there was only one (animal)'

However, these aforementioned rules do not account for the placement of stress in all Lio phrases, and more work remains to be done to uncover further patterns in Lio stress assignment.

2.7 Loanword phonology

Through intensive contact and bilingualism with Malay (in the form of both Standard Indonesian and local eastern Malay varieties), a great many Malay words have been incorporated in the Lio lexicon. This includes Malay words of Sanskrit, Tamil, Arabic, Persian,

Dutch, English, Latin, Portuguese and Chinese origin (among others), leading to an indirect influence of those languages on the Lio lexicon. A small number of words have been borrowed from directly from Portuguese and are not present in Malay, but for the most part borrowing from major world languages has taken place through the medium of Malay.

GLOSS	LIO	MALAY	ORIGINAL SOURCE
'to reply'	bala	balas	Malay
'gong'	nggo	gong	Malay
'million'	juta	juta	Sanskrit
'king'	raja	raja	Sanskrit
'gun'	mbendi	bedil	Tamil
'barn'	guda	gudang	Tamil
'to think'	piki	pikir	Arabic
'to marry'	nika	nikah	Arabic
'grape, wine'	anggo	anggur	Persian
'book'	buku	buku	Dutch
'money'	doi	duit	Dutch
'mass'	misa	misa	Latin
('priest')	patér	pater	Latin
'ball'	bola	bola	Portuguese
'chair'	kadéra	kadera	Portuguese
''2.5 cent coin'	goba	gobang	Chinese (Min Nan)
'knife'	piso	pisau	Chinese (Cantonese)

In addition to loanwords from major world languages, Lio has also exchanged countless words with neighboring languages. Distinguishing these loanwords is much more difficult, because the other languages of Flores are poorly known, closely related to Lio, and often have similar phonological characteristics. Thus, the vast majority of loans from nearby languages will probably remain undetected.

In some cases, a characteristic sound change may provide a clue that a word was borrowed into Lio from a nearby language. For instance, the plain voiced stops /b d g/ of Lio correspond in most cases to prenasalized stops /^{m}b ^{n}d ^{ŋ}g/ in Ende. In light of this, it is reasonable to assume that the words *mbendi* 'gun' (from Malay *bedil*, originally Tamil), *nggo* 'gong' (from Malay *gong*), and *nggaja* 'elephant' (from Malay *gajah*, originally Sanskrit) were borrowed into Lio through the medium of Ende. Otherwise, the prenasalization of the voiced stops remains unexplained.

Well-established loanwords are nativized to fit Lio phonology, meaning all consonant clusters are broken up and all phonemes belong to the native Lio phoneme inventory. On the other hand, there is no evidence that loanwords are reduced to two syllables in order to conform to the bimoraic requirement. On the other hand, there are many words used in everyday Lio speech which do not conform to Lio phonology. In some cases these can be

considered recent loans (if there is no other native Lio word for the concept), but in others they are simply instances of code-switching with Malay.

2.7.1 Deletion of coda consonants

Since Lio does not allow any consonants in the coda, these are deleted in nativized loanwords. The following examples show deletion of word-final consonants:

ORIGINAL FORM	LIO	GLOSS
adat	ada	'honor, custom'
balas	bala	'to reply'
pikir	piki	'to think'
nikah	nika	'to marry'
binatang	binata	'animal'

Word-internal codas are also deleted, eliminating any consonant clusters. Note that sequences of homorganic nasal plus voiced stop /mb nd ŋg/ are reinterpreted as single Lio phonemes and preserved, as in the case of *gamba* 'picture' (from Malay *gambar*). The word *lampu, lapu* 'lamp' shows variable deletion. Arndt's 1933 dictionary lists only the form *lapu*, and my informant ML gave that form during elicitation of a wordlist. However, the form *lampu* appears twice in my corpus of ML's speech while the form *lapu* does not appear. It appears the older, nativized form of the loanword *lapu* is being supplanted by a more faithful form *lampu*, driven by ever-increasing Malay proficiency among Lio speakers. However, Lio speakers apparently still consider the nativized form *lapu* to be the prescriptively correct one.

ORIGINAL FORM	LIO	GLOSS
lampu	la(m)pu	'lamp'
janji	jaji	'to promise'
untuk	utu	'for, in order to'
contoh	soto	'example'
perintah	peréta	'to command'
periksa	perésa	'to inspect'

2.7.2 Vowel epenthesis

Words which contain a consonant cluster occasionally insert a vowel to break up the cluster. In most cases (like *gela* 'glass' from Malay *gelas*, earlier Dutch), this probably stems from the vowel insertion in the intermediate Malay form. However, there are some cases where Lio has vowel insertion but Malay does not:

ORIGINAL FORM	LIO	GLOSS
serdadu	sorodadu	‘soldier’
preman	peréma	‘thug’

2.7.3 Vowel changes

Sometimes, loanwords undergo vowel changes when they are nativized to Lio phonology. There are many examples of /ai au/ sequences coalescing into /e o/, a feature which is shared with local eastern Malay varieties. The word *pekia* ‘clothing’ from Malay *pakaian* shows coalescence of /ai/ to /i/ exceptionally.

ORIGINAL FORM	LIO	GLOSS
pisau	piso	‘knife’
pulau	pulo	‘island’
hijau	hijo	‘green’
malaikat	maléka	‘angel’
pakai	paké	‘to use, to wear’
panau	pano	‘scabies’
ramai	ramé	‘bustling’
damai	damé	‘peace’

In some cases, a final high vowel /i, u/ is lowered to a mid vowel /e, o/ in the wake of final consonant deletion:

ORIGINAL FORM	LIO	GLOSS
adik	adé	‘younger sibling’
anggur	anggo	‘grape, wine’
atur	ato	‘to arrange’
masuk	maso	‘to enter’
(paling)	(palé)	‘very, the most’
sambut	sambo	‘communion’
tanggung	tanggo	‘to share burden’
kain	kaé	‘cloth’

There are a few examples of a word-medial consonant being deleted and causing a vowel shift as well:

ORIGINAL FORM	LIO	GLOSS
perintah	peréta	‘to command’
periksa	perésa	‘to inspect’

2.7.4 Consonant changes

Phonemes in Malay loanwords which do not occur in Lio are systematically replaced in loanwords. The Malay palatal nasal *ny* /ɲ/ and affricate *c* /tʃ/ become Lio *n* and *s* respectively. The glottal affricate /h/ is usually deleted in Detukeli Lio.

ORIGINAL FORM	LIO	GLOSS
penyakit	penaki	'disease'
contoh	soto	'example'
baca	basa	'to read'
cara	sara	'manner, way'
hukum	uku	'to punish'

2.8 Summary

Lio's phonemic inventory is fairly typical of the languages of Central Flores. It includes the standard 6 vowel inventory /i, e, u, o, a, ə/, where /ə/ has a special set of phonological properties. Lio as spoken in Detukeli has 22 consonant phonemes in total, including four series of stops: plain voiceless stops /p, t, k, ʔ/, plain voiced stops /b, d, g/, prenasalized voiced stops /ᵐb, ⁿd, ᵑg/ and voiced implosive stops /ɓ, ɗ/. Voiceless stops are usually heavily aspirated. There are three nasal phonemes /m, n, ŋ/, three approximants /ʋ, y, ɣ/, an affricate /dʒ/, a rhotic /r/ and a lateral /l/. There are two fricatives /s, f/, with a third fricative /h/ present in some Lio varieties but not in Detukeli.

There are only CV syllables with no consonant clusters. Vowel sequences are allowed unless they involve schwa /ə/, and all possible vowel sequences are attested. Words may consist of one or two syllables, but monosyllabic words are lengthened due to the bimoraic requirement which states that all words must consist of exactly two moras. There are no native Lio words of more than two syllables.

Stress is normally on the first syllable of the word, even when the first syllable contains a schwa /ə/. Functional morphemes may be phonologically reduced, consisting of fewer than two moras and bearing no stress.

Loanwords are adapted to Lio phonology by the deletion of consonant clusters and final consonants, and the replacement of foreign phonemes by phonemes which are part of Lio's phonemic inventory.

3 Part II: Historical relations of the Central Flores languages

The second section of this thesis will place Lio in its wider context as a member of the Central Flores (CF) language group. I will first explain the IPA based orthography used here (see section 3.1), which differs from the Lio-specific orthography used in the first part of the thesis. I will then provide some theoretical background (see section 3.2) which will inform the discussion that follows. I will review the sound systems and basic data on the living CF languages (section 3.3.1), as well as the other languages spoken on Flores and Palu'e (section 3.3.2). I will then show the top-down changes which lead from PMP to PCF (section 3.4), followed by a bottom-up reconstruction of aspects of PCF based on comparison of living CF languages (section 3.5). Finally, I will carry out an analysis of the internal relations between CF languages (section 3.6), using Historical Glottometry as a methodological tool for quantifying their relations rather than the tree model.

3.1 Orthographic conventions

In the following section of the thesis, I have adapted all of the data (drawn from sources with varying transcriptions) into a single unified transcription which is identical to the IPA **except** that 'y' represents a palatal glide (IPA [j]), not a high front rounded vowel. I will use 'dʒ' for the post-alveolar or palatal affricate, and the problematic symbol 'j' will not be used at all in my transcription.

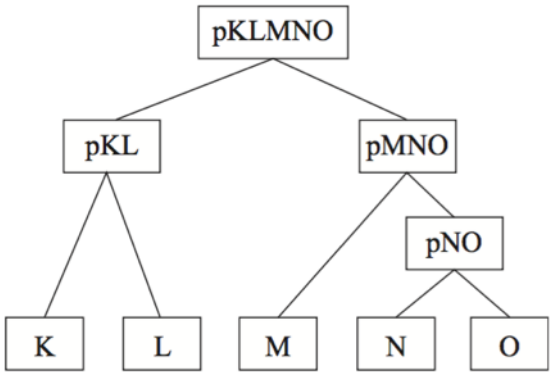
In PMP or PAN forms reconstructed by Blust (Blust and Trussel, ongoing), I have retained his orthography. This entails a few important differences which the reader must be aware of: 1) the symbol 'e' in PMP or PAN forms represents a schwa (presumably [ə]), not the IPA cardinal vowel [e], and 2) the symbol 'j' has unclear phonetic value but is definitely not a palatal glide (which is represented as 'y').

3.2 Theoretical background

There are two main competing approaches to modeling language differentiation: the tree model and the wave model. Each model gives weight to different factors in language differentiation and can be visualized with its own techniques.

The tree model (first proposed by Schleicher 1853) views language differentiation as a primarily a series of splits followed by periods of isolation. An ancestral language community first undergoes a split, which leads to two isolated daughter communities. A split is to be interpreted in the literal sense: it may be a natural disaster, a migration, a period of conflict, or any such event which may lead to a community splitting up and its members losing contact with one another. Once the resulting speech communities have stabilized, they must remain isolated from one another for long enough to undergo a set of distinctive shared innovations which then characterize that speech community. If this period of isola-

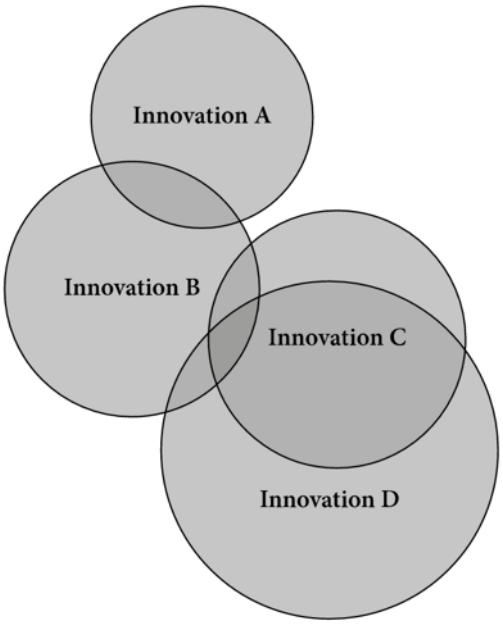
Figure 3: Hypothetical tree diagram; drawn from François (2014)



tion is long enough, and the changes undergone by the daughter communities are extensive enough to compromise mutual intelligibility, the result is the split of one proto-language into multiple daughter languages. After having differentiated, it is possible for the daughter languages to come back into contact, but they are too different for new innovations to travel between languages, since mutual intelligibility is impaired. Furthermore, any such innovations which do manage to spread between previously diverged languages are not considered to be of fundamental significance to their subgrouping. They are considered 'contact effects' or 'horizontal transmission', while 'genetic/genealogical inheritance' or 'vertical transmission' is taken as the only relevant factor for subgrouping. This can be summed up by the tree model maxim 'once a subgroup, always a subgroup'. Thus, English will remain a Germanic language no matter how much lexical or even grammatical material it borrows from French, its long-lost Romance cousin from which it was separated for several thousand years before intensive contact resumed.

Another view of language differentiation is put forth by the wave model, first proposed by Schmidt (1872). The wave model takes the view that all innovations spread horizontally, passing from one group of speakers to another like a ripple spreading in a pond. An innovation must arise in some innovatory center, and then be adopted by those who are in contact with those original innovative speakers. Any given innovation can continue to spread, as long as there is continued contact and some degree of mutual intelligibility between neighboring groups. Even though an earlier innovation has targeted a certain set of languages, there is nothing preventing a later innovation from spreading through a larger or smaller set of languages which intersect the first set. This means that intersecting innovations are not a problem for the wave model, where they would have to be dismissed as 'contact effects' irrelevant for subgrouping under the wave model. A gradual accumulation of innovations over time leads to lower and lower levels of mutual intelligibility, resulting eventually in the fragmentation of the speech area into daughter languages. However, the

Figure 4: Hypothetical set of innovations spreading in a wave-like manner



patterns of innovation which hold between neighboring languages are not necessarily nested, and may intersect freely. This situation cannot be modeled by a tree diagram, and there are other methods of visualization appropriate for the wave model. One wave model-inspired framework, which I will adopt in this thesis, is Historical Glottometry, which provides a way of quantifying the strength of subgroups and visualizing the results as a glottometric diagram or glottometric map (see section 3.6.2 for details). The full details of Historical Glottometry are laid out in a series of recent papers (François 2014, François 2017, Kalyan and François 2018).

The tree and wave models are appropriate in different circumstances; some language groups are more tree-like, and some are more wave-like, although all groups probably show a mix of both characteristics to some degree. Tree-like language groups are those which differentiate primarily by splitting and isolation, and the resulting pattern of non-intersecting innovations can be well-modeled by a tree diagram. Furthermore, the longer periods of isolation which lead to the formation of tree-like language groups usually result in a large number of exclusively shared innovations accumulating. This leads to clear breaks between languages with large bundles of well-supported isoglosses separating neighboring language groups. In such a scenario, distinguishing between vertical and horizontal transmission is normally not problematic. Wave-like language groups tend to form when a language which is spoken over a wide area does not undergo any catastrophic splitting event, but simply differentiates *in situ* into daughter languages. A long and steady decline in mutual intelli-

gibility leads eventually to the formation of separate languages, but these cannot be placed into a tree diagram because they undergo innovations with different subsets of neighbors at different times. Wave-like language groups typically show a pattern whereby they are most similar to their immediate neighbors, and linguistic distance increases steadily with geographic distance. Another pattern which commonly results from wave-like differentiation is that languages in the center of the linguistic area are more innovative, because many different waves of innovation have overlapped there, while the peripheral languages are more conservative because many waves of innovation peter out before they can spread to the very edges of zone. This wave-like scenario of differentiation is what Ross (1988) terms a 'linkage'. In a linkage, the dividing lines between languages are much less clear. There is not generally such an abundance of evidence for each subgrouping, because there are no extended periods of isolation which would allow exclusive innovations to accumulate. In these circumstances, it is much more difficult to distinguish between vertical and horizontal transmission, because the linguistic differences are so much smaller.

It is equally possible to identify tree-like behavior at a higher level of structure, but wave-like behavior at a lower level. For instance the structure of the Indo-European family has tree-like characteristics at a high level: a series of presumed migrations followed by long periods of total isolation resulted in the branching of Indo-European into (among others) the Romance group. This shows a tree-like pattern of non-overlapping innovations and numerous well-attested isoglosses between Romance and the other branches. However, wave-like dynamics dominate the differentiation of Latin (or Proto-Romance) into the modern Romance languages. Romance shows many of the characteristics of a linkage: it was spoken with minor variations over a large, politically unified area with ongoing contact, and the gradual accumulation of innovations led to its eventual breakup into the modern Romance languages. In many regions of the Romance-speaking area, the differences between languages are gradual rather than abrupt, and there is mutual intelligibility even across what are normally thought of as linguistic borders. Thus, Indo-European shows different dynamics at different levels of structure, and similar phenomena can be found across many linguistic groups (including Austronesian).

In relation to the current analysis of the CF languages, the following observations can be made. At a higher level, CF is a tree-like subgroup. There are clear linguistic breaks both to the east and to the west of CF, where a large number of isoglosses line up. Many pieces of strong evidence can be found which bind the CF languages together to the exclusion of the other languages of Flores. This indicates that at some point in history, speakers of Proto-Central Flores did in fact undergo a split from their mother speech community, and were isolated from their sister communities for quite some period of time. On the other hand, once the speakers of PCF were established in central Flores, there were no further episodes of splitting and isolation; wave-like dynamics dominate their differentiation below the level of PCF. The lines between CF languages are not very sharp, mutual intelligibility persists at a low level to the present day, and there are numerous examples of intersecting innovations. There is a pattern of increasing linguistic distance

with increasing geographic distance, and the central CF language are more innovative while the peripheral CF languages are more conservative. In these ways, the CF languages form an archetypal linkage in their low-level structure, but form a well-defined tree-like subgroup at the higher level of PCF.

3.3 Overview of languages sampled

I will review the languages included in this thesis one by one, giving information about the phonology of each language and the sources from which the data used here is drawn.

3.3.1 Central Flores languages

The CF languages share extremely similar morphosyntactic profiles, and a good deal of phonology as well. All CF languages have a strict CV syllable structure, bimoraic word structure, and have implosive or preglottalized stops. All CF languages are highly isolating, with a singular proclitic /sa/ or /ha/ (from PCF *sa) being the only morpheme which has certain characteristics of a bound morpheme. The basic constituent order is SV/AVO, but there is broad freedom to rearrange constituents within a clause (subject to pragmatic factors). Tense, aspect and mood are expressed with independent words (often formally identical to verbs) which come either before or after the main verb, forming a verbal complex. However, the form and placement of these TAM markers does not necessarily correspond between CF languages. NPs are strictly head-initial, and in all languages a genitive relation can be expressed by placing the possessum before the possessor (but some languages have additional strategies; see the next paragraph). Pronominal possession may optionally be expressed with a set of reduced enclitics which are transparently derived from the full pronouns. All CF languages share two innovative features: a highly unusual numeral system with a mixed quinary-decimal base and a set of cognate directionals terms. Both of these can be reconstructed to the level PCF (see sections 3.5.3 and 3.5.4) and provide convincing evidence that the CF languages do indeed form a valid innovation-defined subgroup.

The consonant inventory of each CF language will be dealt with individually, while they share an identical 6 vowel system as shown in table 6.

Table 6: Vowel system of all CF languages

	FRONT	CENTRAL	BACK
CLOSE	i		u
MID	e	ə	o
OPEN		a	

The primary point of difference between the morphosyntax of CF languages lies in alternate strategies for expressing possession. In Lio, the juxtaposed possessum-possessor

construction is the only way of expressing possession within an NP. All other CF languages have a genitive particle /koʔo/, /ko/ or /go/ which can optionally be placed between the possessum and possessor. In Ngadha, it takes the form /go/, and both orders of possessor and possessum are possible when this particle is used (Arndt 1966).

A point of difference in CF languages' phonology is the status of initial glottal stop. In most CF languages, as in PCF, glottal stop is not phonemic in initial position but is predictably inserted before initial vowels. However, in a few CF languages, later developments have created a phonemic distinction between true V-initial, onsetless words (often described as 'breathy') and words which contain an initial glottal stop. Such a phonemic distinction will be noted in languages where it occurs (Ende, some Nage, some Kéo, and Ngadha).

The closest relative of the CF group is Palu'e, which differs in some important respects from the others and is spoken on a small island of the same name off the northern coast of Flores. It has a possessor-possessum order in genitive constructions, and the possessum is marked with a bound morpheme /-n/ which has no counterpart in any other CF language (and also generates closed syllables, which are absent in the other languages). It shares neither the mixed base quinary-decimal system which is so distinctive in CF languages, nor the cognate set of directional terms. For these reasons, it falls outside the scope of the present analysis.

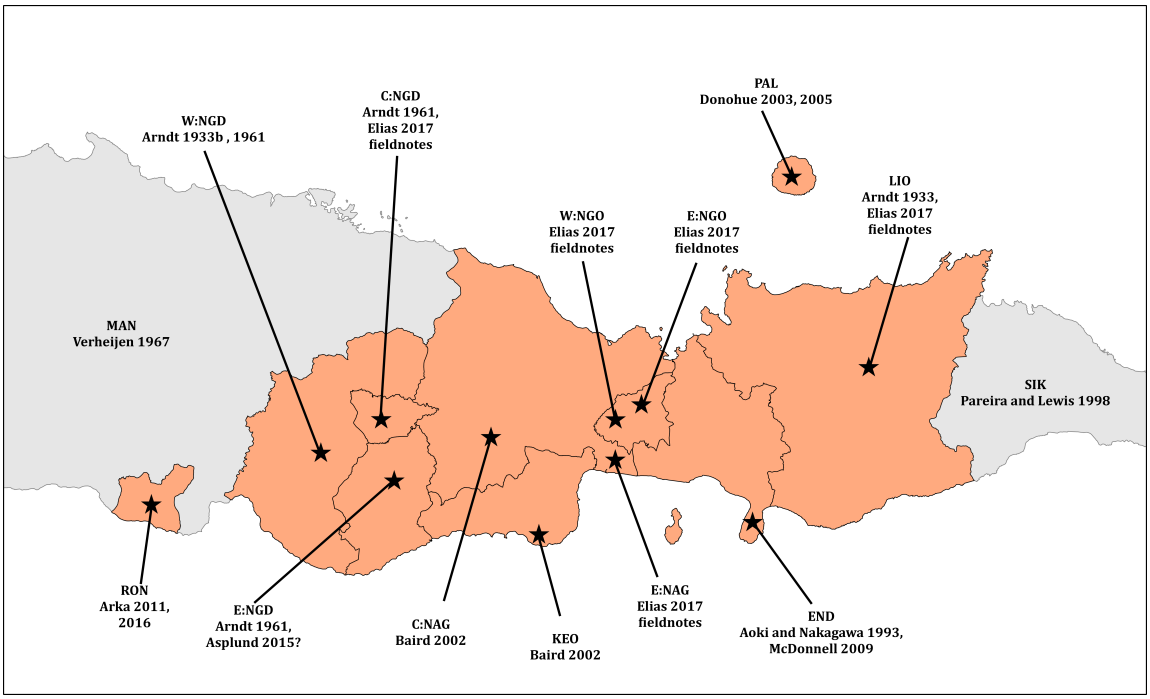
Lio (LIO) Lio is the largest of the CF languages both numerically and geographically, with more than 100,000 speakers (2009 Indonesian census). It the easternmost language of the CF group, spoken throughout the eastern half of Ende Regency. Lio does not have much dialectal variation: the only clear distinction which can be drawn is between Western Lio dialects, which show initial /k-/ in some words where Eastern Lio dialects have an initial /h-/ (a sound correspondence which reflects initial PCF *kl- clusters).

Table 7: Consonant inventory of Lio (Kanganara)

	LABIAL		CORONAL		VELAR		GLOTTAL
STOP	p	b	t	d	k	g	ʔ
AFFRICATE				dʒ			
IMPLOSIVE		ɓ		ɗ			
PRENASAL		ᵐb		ⁿd		ᵑg	
NASAL		m		n		ŋ	
FRICATIVE	f		s			ɣ	(h)
GLIDE		w					
LATERAL				l			
RHOTIC				r			

There are three main sources for my Lio data. The first is the data collected during

Figure 5: Map of CF languages (plus Palu'e, Sika and Manggarai), with names of data sources listed



my fieldwork (Elias 2017 fieldnotes), which also forms the basis of the first section of this thesis. This reflects Lio as spoken in Detukeli, a Western Lio dialect lacking initial /h-/.

The second is a 1933 Lio-German dictionary by the missionary-linguist Paul Arndt. He states in the foreword that the data is based mainly on Lio as spoken in Mbuli and Lisedetu, both located in the modern Subdistrict (or Kecamatan) Wolowaru (Arndt 1933; p5). These are Eastern Lio dialects, which include initial /h-/ as a robust member of the phonemic system. However, in many cases Arndt lists variants from Western Lio dialects without necessarily indicating that this is the case, so this dictionary can be considered a pan-dialectal dictionary.

The third is a collection of Swadesh lists collected from across Ende Regency (Astar, Aritonang, Martis, Kurniawati and Feirizal 2000). Although there are certain inconsistencies in the transcription, this is a valuable source because it provides the only direct evidence regarding dialect variation in Lio. It includes Swadesh lists from the following Lio speaking villages, covering both Eastern and Western Lio dialects: Fataatu, Ngalupolo, Nggela, Ropa, Maubasa, Tou, Watunggere, Wololele A and Wolomage.

Forms cited in this thesis come primarily from my own fieldwork, and therefore reflect Western Lio. Where relevant, Eastern Lio forms from Arndt (1933) will be cited, but this will be explicitly mentioned.

Ende (END) Ende is spoken by somewhat fewer people than Lio, with a figure of 78,000 cited by Wurm and Hattori (1981). It occupies most of the western half of Ende Regency, including the city of Ende itself. Ende and Lio are similar, but not similar enough to be regarded as dialects of one another. For instance, Ende has developed a genitive particle /koʔo/ or /ko/, shared with the other CF languages, while this is missing in Lio. Ende is also notable for the absence of laterals in inherited vocabulary, with the non-lateral /ɹ/ corresponding to /l/ in Lio and most other CF languages. Ende also has a phonemic distinction between an initial glottal stop and true vowel-initial ('breathy') words (such as PCF *iwa ⇒ /ʔiwa/ 'no, not; NEG' versus PCF *kliwa ⇒ /iwa/ 'year').

The data on Ende used here is drawn from an unpublished dictionary created around 1993 by the Japanese anthropologists Eriko Aoki and Satoshi Nakagawa, who have carried out extensive fieldwork in central Flores. The location on which their dictionary is based is not made explicit, but it is nearly identical to the Ende used in the vicinity of the city of Ende.

Another important source of data on Ende is a preliminary phonological description (McDonnell 2009), based on the variety spoken in Onekore, located in the north of the Ende urban area. This variety differs slightly from that of Aoki and Nakagawa's dictionary, and in the one instance where I cite a form from McDonnell which is missing from Aoki and Nakagawa, I have made this explicit.

Table 8: Consonant inventory of Ende (Ende)

	LABIAL		CORONAL		VELAR		GLOTTAL
STOP	p	b	t	d	k	g	ʔ
AFFRICATE				dʒ			
IMPLOSIVE		ɓ		ɗ			
PRENASAL		ᵐb		ⁿd		ᵑg	
NASAL		m		n		ŋ	
FRICATIVE	f		s			ɣ	h
GLIDE		w					
LATERAL							
RHOTIC			ɹ	r			

Nga'o (E:NGO, W:NGO) Nga'o is a language spoken by roughly 9,000 speakers (Wurm and Hattori 1981) in the westernmost inland portion of Ende Regency. Like Ende, it lacks a plain /l/, having instead a very unusual phoneme which I transcribe /ɮ/ and which may be described as an optionally pre-stopped voiced palatal lateral fricative. Before /i/, this phoneme tends towards a realization as an approximant [l], but in other environments the pre-stopped portion may be dominant and it approaches [d]. This phoneme provides a missing phonetic link between the plain /l/ of Lio, Ngadha and Rongga (inherited from PCF *l) and the corresponding /d/ attested in certain eastern Nage and Kéo dialects, which neighbor the Nga'o speaking area. In the context of Central Flores, Nga'o is unusual because of the presence of a palatal glide /y/ which is otherwise attested only in a few geographically removed Kéo dialects. The status of /h/ and /r/ in inherited vocabulary is marginal.

There is very little published information available on Nga'o, but some sources suggest that it is a western dialect of Ende (Hammarström, Forkel and Haspelmath 2018). My findings suggest that in fact, the differences between Ende and Nga'o in phonology and vocabulary are quite substantial, and Nga'o bears more resemblance to Eastern Nage and Kéo than to Ende.

What I call Western Nga'o (W:NGO) is based on a 600-item wordlist collected from a speaker from the village of Watumite (however, the wordlist was recorded in the city of Ende). It is notable because there has been a denasalization of /mb, ngg/ (but not /nd/), resulting in a stop with a voice onset time of approximately 0, which I have transcribed as /b̥, g̥/. Although phonetically speaking, the symbols /p/ and /k/ would be appropriate, I have chosen not to use them because it would require the reanalysis of erstwhile /p, k/ as aspirated /pʰ, kʰ/ (like in Lio, they are normally produced with heavy aspiration). For the sake of comparison, it is preferable to choose a new symbol and keep the plain stops /p, k/ as they are.

Eastern Nga'o (E:NGO) is based on a 600-item wordlist collected from a speaker from

Oja, Tendambepa (although the wordlist was recorded in the city of Ende). It shows a full series of prenasalized stops /mb, nd, ŋg/. However, it has lost /r/ and /ɹ/ entirely and replaced them with /y/, leaving it with no rhotics in inherited vocabulary.

Table 9: Consonant inventory of Western Nga'o (Watumite)

	LABIAL		CORONAL		VELAR		GLOTTAL
STOP	p	b	t	d	k	g	ʔ
AFFRICATE			tʃ				
IMPLOSIVE		ɓ		ɗ			
PRENASAL	b̥			ⁿd	g̥		
NASAL		m		n		ŋ	
FRICATIVE	f		s			ɣ	(h)
GLIDE		w		y			
LATERAL				ɮ			
RHOTIC				(r)			

Table 10: Consonant inventory of Western Nga'o (Oja)

	LABIAL		CORONAL		VELAR		GLOTTAL
STOP	p	b	t	d	k	g	ʔ
AFFRICATE				dʒ			
IMPLOSIVE		ɓ		ɗ			
PRENASAL		ᵐb		ⁿd		ᵑg	
NASAL		m		n		ŋ	
FRICATIVE	f		s			ɣ	(h)
GLIDE		w		y			
LATERAL				ɮ			
RHOTIC			ɹ	(r)			

Nage (E:NAG, C:NAG) Nage is spoken by around 50,000 people (Forth 2016) in the northern three-quarters of Nagekeo Regency, located just to the west of Ende Regency. Very little linguistic work has been carried out on Nage, but the anthropologist Gregory Forth has carried out extensive fieldwork in Nage-speaking areas, focusing on traditional Nage classification of animals. Nage dialects are extremely diverse and a dialect survey of Nage is crucial to understanding the CF group, but remains to be carried out. Nage is very close to Kéo and they are often grouped into a single Nage-Kéo language cluster. However, the level of internal diversity within both Nage and Kéo is such that the line between Nage and Kéo (which separates two distinct ethnic identities) does not correspond to any

clearly identifiable linguistic criterion. Eastern dialects of Nage and Kéo share more in common with each other than with western or central dialects of what is supposedly the same language.

In this thesis, Eastern Nage (E:NAG) is represented by a 600-item wordlist I collected from a speaker in the hamlet of Nangamboa Dua (located in the administrative village of Ondorea Barat, in the far west of Ende Regency), but it clearly has close affiliations with other varieties identified as eastern dialects of Nage by Baird (2002). It is distinguished by reflecting PCF *l as /d/ (a sound change which it shares with some Kéo dialects). Eastern Nage also reflects PCF *r as /ɹ/, and the phoneme /r/ is marginal. Eastern Nage has a phonemic distinction between initial glottal stop and true vowel-initial ('breathy') words.

Table 11: Consonant inventory of Eastern Nage (Nangamboa)

	LABIAL		CORONAL		VELAR		GLOTTAL
STOP	p	b	t	d	k	g	ʔ
AFFRICATE				dʒ			
IMPLOSIVE		ɓ		ɗ			
PRENASAL		^{m}b		^{n}d		^{ŋ}g	
NASAL		m		n		ŋ	
FRICATIVE	f		s			ɣ	
GLIDE		w					
LATERAL				l			
RHOTIC			ɹ	(r)			

Central Nage (C:NAG), which shows the total loss of PCF *r in all positions, is represented by the 223-item wordlist collected in Boawae which is included at the end of Baird (2002). Central Nage also seems to lack prenasalized phonemes, although there is one instance of /mb/ in the word /mbana/ 'to go, to walk'. Central Nage (in contrast to Eastern Nage) does not have a phonemic distinction between initial glottal stop and true vowel-initial words. Although the wordlist for Central Nage list is significantly shorter than the lists representing other languages, I felt it was important to include a Central Nage dialect because Eastern Nage is quite divergent from the other Nage dialects. In addition to the wordlist in Baird (2002), the wordlist collected by Leif Asplund on the Austronesian Basic Vocabulary Database (Greenhill, Blust and Gray 2008) is also marked as coming from Boawae and shows the same reflexes, so it can provide some additional evidence for Central Nage. In some very rare cases when an item was missing from Baird's list but included in Asplund's, I have cited the form listed in the ABVD. However, the reader should keep in mind that the Central Nage data is considerably shakier than any of the other languages or varieties included in this analysis.

In the north and west of the Nage speaking area, there are serious gaps in knowledge and linguistic fieldwork in these areas is needed to establish a clear picture of the dialect

Table 12: Consonant inventory of Central Nage (Boawae)

	LABIAL		CORONAL		VELAR		GLOTTAL
STOP	p	b	t	d	k	g	ʔ
AFFRICATE				dʒ			
IMPLOSIVE		ɓ		ɗ			
PRENASAL		(^{m}b)					
NASAL		m		n		ŋ	
FRICATIVE	f		s	z		ɣ	h
GLIDE		w					
LATERAL				l			
RHOTIC				(r)			

situation there.

Kéo (KEO) Kéo is spoken by approximately 40,000 people (Baird 2002) in the southern quarter of Nagekeo Regency. As mentioned in the section on Nage (section 3.3.1), Kéo is extremely close to Nage and the line between them reflects the distinction between two ethnic groups more than any single linguistic factor. There is a considerable amount of linguistic diversity within Kéo, but only the dialect of Udiworowatu, which forms the basis of Louise Baird's 2002 PhD thesis, is described well-enough to be included in this thesis. This innovative Central Kéo dialect shows a shift of PCF *l to /d/, but there are other Kéo dialects to the east and west which do not undergo such a change. This dialect of Kéo also has a phonemic distinction between initial glottal stop and true vowel-initial words, but this distinction is lacking in many other Kéo dialects.

Table 13: Consonant inventory of Kéo (Udiworowatu)

	LABIAL		CORONAL		VELAR		GLOTTAL
STOP	p	b	t	d	k	g	ʔ
AFFRICATE				dʒ			
IMPLOSIVE		ɓ		ɗ			
PRENASAL		^{m}b		^{n}d		^{ŋ}g	
NASAL		m		n		ŋ	
FRICATIVE	f		s		x		h
GLIDE		w					
LATERAL				l			
RHOTIC				r			

Ngadha (E:NGD, C:NGD, W:NGD) Ngadha is spoken by roughly 60,000 people (unclear source, cited from Simons and Fennig 2017) located in Ngada Regency (note the slight difference in orthography). There is substantial dialectal variation within Ngadha, with a particularly divergent eastern dialect characterized by the complete loss of PCF *l in all positions. In addition, the towns of Bajawa and So'a speak varieties of Ngadha which are commonly known under the same name as the towns they are spoken in. In this thesis, the variety I call Western Ngadha (W:NGD) corresponds to Ngadha as spoken in Bajawa, what I call Eastern Ngadha (E:NGD) corresponds to that of So'a, and the divergent Eastern Ngadha (E:NGD) is that spoken in Takatunga. Ngadha has a phonemic distinction between initial glottal stops and true vowel-initial words, but initial glottal stops are very rare (in contrast to other CF languages, where the vowel-initial words are rare). Unlike other CF languages, Ngadha lacks prenasalized stops.

My source for Western Ngadha, the variety spoken in Bajawa, is the work of Paul Arndt. He wrote a grammar of Ngadha based on the variety of Bajawa (Arndt 1933) as well as a pan-dialectal Ngadha-German dictionary (Arndt 1961). I have attempted to cite forms from his dictionary which correspond with those found in his grammar, but since he does not explicitly mark which entries belong to which exact region, it is possible that my Western Ngadha data contains certain forms which do not actually occur in Bajawa Ngadha.

Table 14: Consonant inventory of Western Ngadha (Bajawa)

	LABIAL		CORONAL		VELAR		GLOTTAL
STOP	p	b	t	d	k	g	ʔ
AFFRICATE				dʒ			
IMPLOSIVE		ɓ		ɗ			
PRENASAL							
NASAL		m		n		ŋ	
FRICATIVE	f		s	z		ɣ	h
GLIDE		w					
LATERAL				l			
RHOTIC				r			

My source for Central Ngadha, the variety spoken in So'a, is a 600-item wordlist I recorded from a speaker born and raised in So'a, but living and working in Ende at the time of the recording. There is no other published linguistic material on Central (or So'a) Ngadha. While it is listed as a separate language on Glottolog (Hammarström, Forkel and Haspelmath 2018) and Ethnologue (Simons and Fennig 2017), the data I collected shows that it bears so close a resemblance to other Ngadha dialects that it cannot be considered a separate language. It is also clear that forms from So'a are included in Arndt (1961), the Ngadha-German dictionary, and are not considered a separate language by him.

Phonologically, the only difference between Central (So'a) Ngadha and Western (Bajawa) Ngadha is the absence of the affricate /dʒ/ in Central Ngadha, where it corresponds to /z/.

Table 15: Consonant inventory of Central Ngadha (So'a)

	LABIAL		CORONAL		VELAR		GLOTTAL
STOP	p	b	t	d	k	g	ʔ
AFFRICATE							
IMPLOSIVE		ɓ		ɗ			
PRENASAL							
NASAL		m		n		ŋ	
FRICATIVE	f		s	z		ɣ	h
GLIDE		w					
LATERAL				l			
RHOTIC				r			

My source for Eastern Ngadha is limited to a 215-item wordlist collected by Leif Asplund in Takatunga (Golewa Sub-district) and published on the Austronesian Basic Vocabulary Database. This dialect is on the border between Ngadha and Nage and is in some ways transitional between those two languages, and is unique among CF languages in showing the total loss of PCF *l. Another unique feature for Ngadha varieties is the reflex of PCF *dʒ as /r/, a change it shares with many other CF languages but not with other Ngadha varieties. This is a transitional variety important for understanding the CF group as a whole and is included despite the poor level of documentation.

Table 16: Consonant inventory of Eastern Ngadha (Takatunga)

	LABIAL		CORONAL		VELAR		GLOTTAL
STOP	p	b	t	d	k	g	ʔ
AFFRICATE				dʒ			
IMPLOSIVE		ɓ		ɗ			
PRENASAL							
NASAL		m		n		ŋ	
FRICATIVE	f		s			ɣ	h
GLIDE		w					
LATERAL							
RHOTIC				r			

Arndt (1961) records some intriguing dialectal variants which are apparently found in none of the varieties which I include in my thesis, but provide crucial evidence for

the development of certain PCF sounds. For instance, words which I reconstruct with initial PCF *kl- clusters often are listed with three variants in Arndt (1961): PCF *klale 'fly' ⇒ /xale/ (the form recorded for all three varieties of E:NGD, C:NGD and W:NGD) alongside /kale, sale/. These alternative forms will be cited in the discussion on PCF clusters, although it remains unclear which geographical location in the Ngadha-speaking area actually employs them.

Rongga (RON) Rongga is the westernmost CF language, spoken in Manggarai Timur (East Manggarai) Regency and slightly separated from the other CF languages by speakers of Manggarai and languages closely related to it. It is the smallest CF language by number of speakers and geographical extent, spoken by about 5,000 people in the 3 villages of Bamo, Tanah Rata and Komba. The data for Rongga (RON) in this thesis is drawn from work published by I Wayan Arka including a Rongga-English dictionary (Arka, Seda, Geland, Nani and Ture 2011) and a grammatical description (Arka 2016). Rongga has little dialectal variation because of its small geographical extent. This was previously the most poorly known of the CF languages, but thanks to the recent work by Arka it now enjoys a rather good level of documentation.

Table 17: Consonant inventory of Rongga

	LABIAL		CORONAL		VELAR		GLOTTAL
STOP	p	b	t	d	k	g	ʔ
AFFRICATE				dʒ			
IMPLOSIVE		ɓ		ɗ		ɠ	
PRENASAL		^{m}b		^{n}d		^{ŋ}g	
NASAL		m		n		ŋ	
FRICATIVE	f	(v)	s			ɣ	h
GLIDE		w					
LATERAL				l			
RHOTIC			ɹ	r			

It is notable as the only CF language which includes an implosive velar /ɠ/ according to Arka (2016), whereas all other CF languages have a gap in the implosive series at the velar place of articulation. He also describes a voiced labiodental fricative /v/ in addition to the more usual voiceless /f/, but almost all words including /v/ have a doublet in /w/ listed in the dictionary (Arka, Seda, Geland, Nani and Ture 2011), so the status of this sound as a contrastive phoneme is doubtful and it may be an alternate realization of the phoneme /w/.

3.3.2 Languages neighboring the Central Flores group

Palu'e (PAL) Palu'e (PAL), spoken on the small island of Palu'e off the north coast of central Flores, is without a doubt the closest relative of the CF languages but lies outside the scope of some of the key innovations defining CF as a subgroup. The source of the Palu'e data here is a 2003 dictionary by Mark Donohue which is available in electronic format (Donohue 2003). He also wrote a paper on the status of the passive in Palu'e (Donohue 2005), from which some additional facts about the language can be gleaned.

Palu'e lacks a plain voiced /b/, having only an implosive /ɓ/ instead. It also has a voiced-voiceless pair of affricates /tʃ, dʒ/ and lacks /f/, in contrast to the CF languages.

Table 18: Consonant inventory of Palu'e

	LABIAL	CORONAL		VELAR		GLOTTAL
STOP	p	t	d	k	g	ʔ
AFFRICATE		tʃ	dʒ			
IMPLOSIVE	ɓ					
PRENASAL	^{m}b		^{n}d		^{ŋ}g	
NASAL	m		n		ŋ	
FRICATIVE		s				
GLIDE	w					
LATERAL			l			
RHOTIC			r			

The evidence for an innovation-defined subgroup which covers the mainland CF languages but excludes Palu'e is much stronger than the evidence for a subgroup including Palu'e. The key differences in Palu'e are the lack of a quinary numeral system, the lack of cognate directional terms, the retention of some bound morphology and word-final consonants, and the opposite order in genitive phrases.

Western Flores: Manggarai (MAN) Manggarai is a large and internally diverse language cluster spoken over the entire western third of Flores. In the eastern Manggarai-speaking area, where it abuts the CF languages, there are certain poorly known language varieties similar to Manggarai but perhaps divergent enough to be considered separate languages. These include Rembong, Riung, Rajong, Kepo', Kolor, Wae Rana, and/or Manus, but several of these names may refer to the same language or to languages which are extremely close to each other. More descriptive work is needed in these areas before their interrelationships become clear.

Manggarai is well-described, with a comprehensive and etymologically oriented Manggarai-Indonesian-Manggarai dictionary in two volumes (Verheijen 1967), which the source of the data in this thesis. Manggarai is important to reconstruction within PMP because it is

phonologically and lexically conservative and preserves PMP final consonants well. Many entries in the Austronesian Comparative Dictionary (Blust and Trussel, ongoing) are reconstructed to the level of PMP (rather than a lower level like PWMP), based only on evidence from a Manggarai reflex.

The consensus view, originally put forth (without published evidence) by Esser (1938) and conserved in Blust (2008) is that Manggarai and the CF languages belong in a subgroup termed 'Sumba-Hawu' together with the languages of Sumba and Hawu. Since this thesis deals primarily with the internal subgrouping of the CF languages, I will not seek to address the issue of whether or not Manggarai is indeed closer to the CF languages than any other CMP languages. Manggarai forms will be cited when relevant to show that certain innovations are restricted to the CF group and define it as a subgroup.

Manggarai differs from the CF languages in phonology in certain important ways. First of all, it allows final consonants freely, and generally preserves PMP final consonants quite faithfully. Second, it has prenasalized voiceless stop phonemes /mp, nt, ŋk/ which are lacking in the CF languages. Third, it allows prenasalized affricates /ɲtʃ, ɲdʒ/, unlike the CF languages. Finally, it is lacking the implosive stops which are characteristic of the CF languages. Overall, Manggarai has more phonological complexity and is more conservative than the CF languages.

Table 19: Consonant inventory of Manggarai

	LABIAL		CORONAL		VELAR		GLOTTAL
STOP	p	b	t	d	k	g	ʔ
AFFRICATE			tʃ	dʒ			
IMPLOSIVE							
PRENASAL	^{m}p	^{m}b	^{n}t, ^{ɲ}tʃ	^{n}d, ^{ɲ}dʒ	^{ŋ}k	^{ŋ}g	
NASAL		m		n		ŋ	
FRICATIVE			s				h
GLIDE		w	y				
LATERAL				l			
RHOTIC				r			

Eastern Flores: Sika (SIK) Languages belonging to the Flores-Lembata subgroup of CMP are spoken in the eastern third of Flores. The immediate neighbor of the CF languages to the east is Sika, and this is the language which I will use as a point of reference for the languages of Eastern Flores. There are also several varieties of Lamaholot spoken on Eastern Flores and on the islands further east, but these do not come directly into contact with CF languages.

The source of the Sika data here is Pareira and Lewis (1998), a Sika-Indonesian dictionary. In addition, aspects of an eastern dialect of Sika known as Hewa have been described

by Fricke (2014), which provides additional insight into the grammar of Sika as a whole.

Sika has a more complex syllable structure than the CF languages, in that it allows initial clusters of stop + liquid such as /gl, pl, kr/ and final consonants. It also has a set of preglottalized resonants which I represent as /ʔl, ʔr, ʔw/. However, it is lacking the implosives and prenasalized stops which are characteristic of the CF languages, and has no affricates.

The consensus view (as put forth by Esser 1938 but never backed up with published data) is that the Flores-Lembata languages - including Sika - are rather distantly related to the CF languages, forming their own primary subgroup of Central Malayo-Polynesian. Since this thesis deals primarily with the internal subgrouping of the CF languages, I will not seek to address this issue. Forms from eastern Flores will be cited where relevant to show that certain innovations are restricted to CF and define it as a subgroup.

Table 20: Consonant inventory of Sika

	LABIAL		CORONAL		VELAR		GLOTTAL
STOP	p	b	t	d	k	g	ʔ
AFFRICATE				dʒ			
IMPLOSIVE							
PRENASAL							
NASAL		m		n		ŋ	
FRICATIVE			s				h
GLIDE		w, ʔw					
LATERAL				l, ʔl			
RHOTIC				r, ʔr			

3.3.3 Summary

In this section, I have reviewed some basic facts about the Central Flores languages (Lio, Ende, Nga'o, Nage, Kéo, Ngadha, Rongga) and dialects thereof, as well as neighboring languages (Manggarai of western Flores, Sika of eastern Flores, and Palu'e from the nearby island of the same name). The sound systems of each language have been outlined, as well as demographic information such as their location, approximate numbers of speakers and dialectal diversity. For more information on each language, the reader is directed to the published sources which I draw on, cited in the relevant sections.

The sampling of Central Flores languages included in this thesis is quite thorough but there are some notable gaps in the central zone of the Central Flores-speaking area. By the standards of the Central Flores languages, Lio, Endenese and Rongga are well-studied (with at least a proper dictionary for each) and low levels of dialectal diversity simplify the situation for these languages. Nage and Keo are in need of a dialect survey, because there

is known to be a great amount of dialect diversity which cannot be captured in the present thesis due to lack of information. Nga'o, which remains almost completely uninvestigated, would benefit greatly from a period of targeted fieldwork, and the preliminary findings show a great amount of divergence even between neighboring Nga'o dialects. Eastern dialects of Ngadha, which are transitional with western Nage dialects, are also understudied, and appear to be quite divergent from the well-investigated western Ngadha dialects spoken in Bajawa. Central Ngadha (as spoken in So'a) remains almost completely unstudied, but the preliminary data which I have collected suggests that it does not differ in significant ways from western Ngadha. Until such a time as a more detailed dialectal survey of central Flores is carried out, it will be impossible to understand the full picture of the historical relations of CF languages.

3.4 Top-down sound correspondences: Proto-Malayo-Polynesian to Proto-Central Flores

In this section, I will show the regular sound changes which lead from PMP to PCF. These are more straightforward than the changes occurring between PCF and the modern CF languages, and can be dealt with more briefly. The focus in the examples is on the relation between the PMP and the PCF form, but modern cognates from CF languages are cited for illustration. However, note that the cognates do not necessarily correspond exactly with the meaning of the gloss listed. For instance, PMP *habaRat 'west wind', has remained Lio /wara/ 'west wind' alongside PMP *haŋin 'wind' ⇒ Lio /aŋi/), while in Central Ngadha (So'a) /wara/ simply means 'wind'. However, when used as evidence of the development of PMP medial *R ⇒ PCF *r, the cognacy of the forms is what matters, not exact correspondence in meaning.

I will also deal with PMP consonant clusters, the developments of PMP vowels and diphthongs, and the reduction of PMP trisyllables to disyllables.

The regular reflexes of the rare PMP phonemes *g, *r and *c are difficult to determine because of a scarcity of solid cognate sets.

3.4.1 Regular reflexes of PMP phonemes in PCF

Table 21 shows the PMP phonemes which are preserved identically in PCF, while Table 22 shows those PMP phonemes which undergo some change in PCF. These correspondences reflect initial and medial position only; in final position, all PMP consonants are lost, and do not normally produce changes in the preceding vowels.

PMP nasals *m, *n, *ñ, *ŋ The rather rare PMP palatal nasal *ñ is merged to PCF *n, as in PMP *ñaRa ⇒ PCF *nara 'brother (term of address used by women)'. The other PMP nasals are retained *m, *n, *ŋ ⇒ PCF *m, *n, *ŋ everywhere except in final position, where they are deleted. PMP nasals can also undergo special developments

Table 21: Identical reflexes of PMP phonemes in PCF

PMP	*p	*t	*k	*m	*n	*ŋ	*s	*w
PCF	*p	*t	*k	*m	*n	*ŋ	*s	*w
LIO	p	t	k	m	n	ŋ	s	w
END	p	t	k	m	n	ŋ	s	w
E:NGO	p	t	k	m	n	ŋ	s	w
W:NGO	p	t	k	m	n	ŋ	s	w
KEO	p	t	k	m	n	ŋ	s	w
E:NAG	p	t	k	m	n	ŋ	s	w
C:NAG	p	t	k	m	n	ŋ	s	w
E:NGD	p	t	k	m	n	ŋ	s	w
W:NGD	p	t	k	m	n	ŋ	s	w
C:NGD	p	t	k	m	n	ŋ	s	w
RON	p	t	k	m	n	ŋ	s	w

Table 22: Non-identical reflexes of PMP phonemes in PCF

PMP	*q		*b	*z	*j	*ñ	*h	*R	*y
ENVIR	V_V	else							
PCF	*ʔ	*∅	*w	*dʒ	*dʒ	*n	*∅	*r	*y
LIO	ʔ	∅	w	dʒ	dʒ	n	∅	r	dʒ
END	ʔ	∅	w	r	r	n	∅	r	dʒ
E:NGO	ʔ	∅	w	y	y	n	∅	y	dʒ
W:NGO	ʔ	∅	w	ɹ	ɹ	n	∅	y	tʃ
KEO	ʔ	∅	w	r	r	n	∅	l	dʒ
E:NAG	ʔ	∅	w	ɹ	ɹ	n	∅	l	tʃ
C:NAG	ʔ	∅	w	r	r	n	∅	∅	dʒ
E:NGD	ʔ	∅	w	r	r	n	∅	r	dʒ
W:NGD	ʔ	∅	w	dʒ	dʒ	n	∅	r	dʒ
C:NGD	ʔ	∅	w	z	z	n	∅	r	z
RON	ʔ	∅	w	ɹ	ɹ	n	∅	r	dʒ

when they come into contact with stops, sometimes preventing lenition of PMP *b, *d but themselves being deleted in the process.

Table 23: Reflexes of PMP *m, *n, *ñ, *ŋ in PCF

PMP	*m	*n	*ñ	*ŋ
PCF	*m	*n	*ñ	*ŋ
PMP	*li**m**a	*a**n**ak	*mi**ñ**ak	*tali**ŋ**a
PCF	*li**m**a	*a**n**a	*mi**n**a	*kli**ŋ**a
LIO	li**m**a	a**n**a	mi**n**a	ki**ŋ**a
END	ɹi**m**a	ʔa**n**a	mi**n**a	i**ŋ**a
E:NGO	ɻi**m**a	a**n**a	mi**n**a	yi**ŋ**a
W:NGO	ɻi**m**a	a**n**a	mi**n**a	yi**ŋ**a
KEO	di**m**a	ʔa**n**a	mi**n**a	i**ŋ**a
E:NAG	di**m**a	a**n**a	mi**n**a	li**ŋ**a
C:NAG	li**m**a	a**n**a	mi**n**a	hi**ŋ**a
E:NGD	i**m**a	a**n**a	mi**n**a	hi**ŋ**a
W:NGD	li**m**a	a**n**a	mi**n**a	hi**ŋ**a
C:NGD	li**m**a	a**n**a	mi**n**a	hi**ŋ**a
RON	li**m**a	a**n**a	mi**n**a	hi**ŋ**a
GLOSS	five	child	fat	ear

PMP voiceless stops *p, *t, *k, *q The PMP voiceless stops undergo very simple developments. PMP *p, *t, *k ⇒ PCF *p, *t, *k everywhere except in coda position, where they are deleted. The only exception is when they come into contact with a nasal phoneme and become voiced or prenasalized.

PMP *q ⇒ PCF *ʔ between vowels, but only root-internally as in PMP *taqi 'excrement' ⇒ PCF *taʔi. It is deleted in initial and final position, and it is deleted when it comes between vowels across a morpheme boundary, as in PMP *qetaq 'to eat raw' ⇒ PMP *ma-qetaq 'green, raw, unripe' ⇒ PCF *məta 'green, blue'. This is related to the fact that morphologically complex PMP words usually have at least three syllables, and these are systematically reduced to disyllables in PCF (see section 3.4.3).

PMP *b There is a regular change of PMP *b ⇒PCF *w, as in all the other languages of Flores. This can be blocked by the presence of a nasal consonant before *b, yielding either *b or *mb.

Two words with solid PMP etymologies reflect initial PMP *b ⇒ *f. These are PMP *buhek 'hair' ⇒ PCF *fu, and PMP *bahi 'woman' ⇒ PCF *fai. The reasons for this are

Table 24: Reflexes of PMP *p, *t, *k, *q in PCF

PMP	*p	*t	*k	*q	
PCF	*p	*t	*k	*ʔ	*∅
ENVIR				V_V	else
PMP	*ha**p**uy	*ma**t**a	***k**ami	*wa**q**ay	****q**atay**
PCF	*a**p**i	*ma**t**a	***k**ami	*wa**ʔ**i	*ate
LIO	a**p**i	ma**t**a	**k**ami	ɣa**ʔ**i	ate
END	ʔa**p**i	ma**t**a	**k**ami	a**ʔ**i	ʔate
E:NGO	a**p**i	ma**t**a	**k**ami	a**ʔ**i	ate
W:NGO	a**p**i	ma**t**a	**k**ami	a**ʔ**i	ate
KEO	ʔa**p**i	ma**t**a	**k**ami	a**ʔ**i	ʔate
E:NAG	a**p**i	ma**t**a	**k**ami	a**ʔ**i	ate
C:NAG	a**p**i	ma**t**a	**k**ami	a**ʔ**i	ate
E:NGD	a**p**i	ma**t**a	**k**ami	a**ʔ**i	ate
W:NGD	a**p**i	ma**t**a	**k**ami	wa**ʔ**i	ate
C:NGD	a**p**i	ma**t**a	**k**ami	wa**ʔ**i	ate
RON	a**p**i	ma**t**a	**k**ami	wa**ʔ**i	ate
GLOSS	fire	eye	1PL.E	leg	liver

not clear, but initial /f-/ is part of the phonemic system of nearly all modern CF languages, corresponding to /w/ in nearby languages.

Table 25: Reflexes of PMP *b in PCF

PMP	*b-	*-b-	*b-	
PCF	*w-	*-w-	*f-	
PMP	***b**ulan	*qa**b**u	***b**uhek	***b**ahi
PCF	*w**w**ula	*a**w**u	*f**f**u	*f**f**ai
LIO	**w**ula	a**w**u	**f**u	**f**ai
END	**w**uɹa	ʔa**w**u	**h**u	**h**ai
E:NGO	**w**uʒa	a**w**u	**f**u	**f**ai
W:NGO	**w**uʒa	a**w**u	**f**u	**f**ai
KEO	**w**uda	ʔa**w**u	**f**u	**f**ai
E:NAG	**w**uda	a**w**u	**f**u	**f**ai
C:NAG	**w**ula	a**w**u	**f**u	**f**ai
E:NGD	**w**ua	a**w**u	**f**u	**f**ai
W:NGD	**w**ula	a**w**u	**f**u	**f**ai
C:NGD	**w**ula	a**w**u	**f**u	**f**ai
RON	**w**ula	a**w**u	**f**u	**f**ai
GLOSS	moon	ash	hair	woman

PMP *z, *j, *d In all languages, the PMP phonemes *z, *j ⇒ PCF *dʒ, which then undergoes extensive changes in daughter languages (section 3.5.1). PMP *d also merges with these two phonemes in all languages except Lio, which reflects PMP *d ⇒ /r/ but PMP *z, *j ⇒ /dʒ/. Therefore PMP *d ⇒ *dʒ cannot have occurred at the level of PCF, and a retention of PMP *d ⇒ PCF *d is the simplest way to account for the well-supported Lio reflexes. Additionally, the reflexes of PMP *d and *r remain distinct in most CF languages, so these two must remain distinct at the level of PCF.

As with the voiceless stops, PMP *d may show different developments when contact with a nasal.

PMP *l, *R PMP *l is retained as PCF *l, which then undergoes changes in some daughter languages.

PMP *R ⇒ PCF *r in medial position, and probably also in initial position but the evidence is somewhat weak.

There are at least three candidates for initial PMP *R but each has some complications: PMP *Ribu 'thousand' ⇒ PCF *riwu is probably not a direct inheritance but a Malay

Table 26: Reflexes of PMP *z, *j, *d in PCF

PMP	*z		*j	*d	
PCF	*dʒ		*dʒ	*d	
PMP	***z**alan	*qu**z**an	*ŋa**j**an	***d**aki	*ma-qu**d**ip
PCF	***dʒ**ala	*u**dʒ**a	*ŋa**dʒ**a	***d**aki	*mu**d**i
LIO	**dʒ**ala	u**dʒ**a	na**dʒ**a	**r**aki	mu**r**i
END	**r**aɹa	ʔu**r**a	ŋa**r**a	**r**aki	mu**r**i
E:NGO	**y**aɮa	u**y**a	ŋa**y**a	**y**aki	mu**y**i
W:NGO	**ɹ**aɮa	u**ɹ**a	ŋa**ɹ**a	**ɹ**aki	mu**ɹ**i
KEO	**r**ada	ʔu**r**a	ŋa**r**a	**r**aki	mu**r**i
E:NAG	**ɹ**ada	u**ɹ**a	ŋa**ɹ**a	(koto)	mu**ɹ**i
C:NAG	**z**ala	u**z**a	ŋa**z**a	**z**aki	mu**z**i
E:NGD	**r**a	u**r**a	ŋa**r**a	**r**aki	mu**r**i
W:NGD	**z**ala	u**z**a	ŋa**z**a	**z**aki	mu**z**i
C:NGD	**z**ala	ʔu**z**a	ŋa**z**a	**z**aki	mu**z**i
RON	**l**aɹa	nu**ɹ**a	ŋa**ɹ**a	**ɹ**aki	mu**ɹ**i
GLOSS	road	rain	name	dirty	alive

loan. In PMP *Ramut 'root' ⇒ PCF *kramu, there is the development of an initial cluster (cf Sika /ʔramut/), which is then resolved as /k-/ in all daughter languages except some dialects of Lio which show /h-/. In PMP * sa-ŋa-Ratus 'one hundred' ⇒ PCF *sa ŋasu there is the retention of the nasal linker *ŋa and deletion of the following *R, as well as metathesis of the final two consonants (which must have happened before final consonant deletion in PCF): PMP *sa ŋa Ratus ⇒ **sa ŋa Rasut ⇒ PCF *sa ŋasu 'one hundred'. Similar developments occur in Sika /ŋasu ha/ 'one hundred', Savu/Hawu /heŋahu/ and many languages of Sumba (Kambera /haŋahu/, Kodi /saŋasu/, Wewewa /ngau/), so this is also probably the result of a loan distribution.

On the other hand, PMP initial *R- may be reflected after the irregular loss of the first syllable in PMP *daRaq ⇒ CPF *ra 'blood' (instead of expected *daRa, which is the form reflected in Palu'e /ladʒa/ and Manggarai /dara/). This also occurs Savu/Hawu /ra/ and the languages of Sumba (Anakalang, Baliledo, Lamboya, Mamboru /ra/, Wewewa /raʔa/), and suggests that the reflexes of initial PMP *R- may be the same as medial *-R-.

PMP *w, *y The PMP glides are retained as PMP *w, *y ⇒ PCF *w, *y, although PCF *y then undergoes changes in all daughter CF languages.

Rare PMP phonemes *r, *g, *c The rare PMP phonemes *r, *g and *c are poorly represented in my data.

Table 27: Reflexes of PMP *l, *R in PCF

PMP	*-R-	*-R-	*l-	*-l-
PCF	*r-	*-r-	*l-	*-l-
PMP	*da**R**aq	*pa-la**R**iw	*l**l**ima	*bu**l**an
PCF	***r**a	*pa**r**u	***l**ima	*wu**l**a
LIO	**r**a	pa**r**u	**l**ima	wu**l**a
END	**r**a	pa**r**u	**ɹ**ima	wu**ɹ**a
E:NGO	**y**a	pa**y**u	**ɮ**ima	wu**ɮ**a
W:NGO	**y**a	pa**y**u	**ɮ**ima	wu**ɮ**a
KEO	**l**a	pa**l**u	**d**ima	wu**d**a
E:NAG	**l**a	pa**l**u	**d**ima	wu**d**a
C:NAG	ai	pau	**l**ima	wu**l**a
E:NGD	**r**a	-	ima	wu**d**a
W:NGD	**r**aʔa	pa**r**u	**l**ima	wu**l**a
C:NGD	**r**aʔa	pa**r**u	**l**ima	wu**l**a
RON	**r**aʔa	pa**r**u	**l**ima	wu**l**a
GLOSS	blood	to run	five	moon

Table 28: Reflexes of PMP *w, *y in PCF

PMP	*w-	*-w-	*-y-	
PCF	*w-	*-w-	*-y-	
PMP	***w**aRi	*ta**w**a	*ka**y**u	(← *kahiw)
PCF	***w**ari	*ta**w**a	*ka**y**u	
LIO	**w**ari	ta**w**a	ka**dʒ**u	
END	**w**ari	ta**w**a	ka**dʒ**u	
E:NGO	**w**ayi	ta**w**a	ka**dʒ**u	
W:NGO	**w**ayi	ta**w**a	ka**tʃ**u	
KEO	**w**ali	ta**w**a	ka**dʒ**u	
E:NAG	(maɹa)	ta**w**a	ka**tʃ**u	
C:NAG	**w**ai	ta**w**a	ka**dʒ**u	
E:NGD	-	ta**w**a	ka**dʒ**u	
W:NGD	**w**ari	ta**w**a	ka**dʒ**u	
C:NGD	**w**ari	ta**w**a	ka**z**u	
RON	**w**ari	ta**w**a	ka**dʒ**u	
GLOSS	to dry	to laugh	wood	

For PMP *r, I have one cognate set PMP *periŋ‘bamboo species’ ⇒ PCF *pəri which appears to show PMP *r merging with *R ⇒ PCF *r.

For PMP *g, I have one cognate set PMP *gatel ‘itchy’ ⇒ PCF *kate which appears to show PMP *g ⇒ PCF *k.

I have no solid cognate sets giving evidence for the developments of PMP *c, but this phoneme is thought to have merged with PMP *s in all but a few WMP languages.

Table 29: Reflexes of PMP *r, *g in PCF

PMP	*r?	*g?
PCF	*r?	*k?
PMP	*periŋ	***g**atel
PCF	*pə**r**i	***k**ate
LIO	pə**r**i	**k**ate
END	pə**r**i	**k**ate
E:NGO	pə**y**i	**k**ate
W:NGO	pə**ɹ**i	**k**ate
KEO	pə**r**i	**k**ate
E:NAG	pə**ɹ**i	**k**ate
C:NAG	-	-
E:NGD	-	-
W:NGD	pə**r**i	**k**ate
C:NGD	pə**r**i	**k**ate
RON	pə**r**i	(mbisa)
GLOSS	bamboo sp.	itchy

PMP vowels PMP *a, *i, *u are straightforwardly reflected as PCF *a, *i, *u in most cases. PMP *e ⇒ PCF *e in final position, and is retained as PCF *əelsewhere (schwa /ə/ is barred from the final syllable in all modern CF languages).

The PMP ‘final diphthongs’ *-ay, *-aw (which are in fact VC sequences) typically behave like other VC sequences, and the final PMP *-y, *-w are dropped (PMP *matay ‘to die’ ⟩ PCF *mata, PMP *nakaw ‘to steal’ ⟩ PCF *naka). There are quite a few exceptions to this, where PMP *-ay, *-aw coalesce to PCF *-e, *-o instead. This occurs in some obvious loanwords (PMP *pisaw ⇒ PCF *piso ‘knife’, loaned from Malay), and words which are suspect due to other phonological irregularities (PMP *buay ⇒ PCF *bue ‘bean’, but lenition to **wue is expected). The main exception which resists explanation in this way is PMP *qatay ⇒ PCF *ate ‘liver’. There is one exceptional reduction in PMP *waqay ⇒ PCF *waqi ‘leg’.

Table 30: Reflexes of PMP vowels in PCF

PMP	*a	*i	*u	*e		*ay	*aw	*iw	*uy
envir.				_(C)#	else				
PCF	*a	*i	*u	*e	*ə	*a	*a	*u	*i
Lio	a	i	u	e	ə	a	a	u	i
Ende	a	i	u	e	ə	a	a	u	i
Nga'o (OJ)	a	i	u	e	ə	a	a	u	i
Nga'o (WM)	a	i	u	e	ə	a	a	u	i
Kéo (KK)	a	i	u	e	ə	a	a	u	i
Kéo (ND)	a	i	u	e	ə	a	a	u	i
Kéo (LL)	a	i	u	e	ə	a	a	u	i
Kéo (UW)	a	i	u	e	ə	a	a	u	i
Nage (NM)	a	i	u	e	ə	a	a	u	i
Nage (UT)	a	i	u	e	ə	a	a	u	i
Nage (BW)	a	i	u	e	ə	a	a	u	i
E Ngadha	a	i	u	e	ə	a	a	u	i
Ngadha	a	i	u	e	ə	a	a	u	i
So'a	a	i	u	e	ə	a	a	u	i
Rongga	a	i	u	e	ə	a	a	u	i

PMP *uy is attested 3 times, yielding PCF *i twice in PMP *hapuy ⇒ PCF *api 'fire' and PMP *babuy ⇒ PCF *wawi 'pig', and PCF *u once in PMP *naŋuy ⇒ PCF *naŋu 'to swim'. PMP *iw is attested three times, once yielding PCF *o in PMP *baliw ⇒ PCF *walo 'to return' and once yielding PCF *u in PMP *pa-laRiw ⇒ PCF *paru 'to run'. In the case of PMP *kahiw ⇒ PCF *kayu 'wood', there was most likely already a reduction to *kayu at the level of PCEMP.

3.4.2 Deletion of coda consonants

All PMP final consonants are deleted unconditionally in PCF, and do not regularly cause any changes in the preceding vowel. The total loss of final consonants is illustrated in Table 31, with a subset of possible final phonemes.

Table 31: Deletion of PMP final consonants in PCF

PMP	*-k	*-b	*-m	*-s	*-j	*-R	*-q
PCF	*-∅	*-∅	*-∅	*-∅	*-∅	*-∅	*-∅
PMP	*anak	*ma-huab	*ma-qitem	*nipis	*pusej	*tuquR	*um-utaq
PCF	*ana	*moa	*mite	*nipi	*puse	*tuʔu	*muta
LIO	ana	moa	mite	nipi	puse	tuʔu	muta
END	ʔana	moa	mite	nipi	puse	tuʔu	muta
E:NGO	ana	moa	mite	nipi	puse	(maya)	muta
W:NGO	ana	moa	mite	nipi	puse	(maɹa)	muta
KEO	ʔana	(ŋaŋa)	mite	(baɓi)	puse	tuʔu	muta
E:NAG	ana	moa	mite	nipi	puse	tuʔu	muta
C:NAG	ana	moa	mite	-	-	tuʔu	muta
E:NGD	ana	moa	mite	nipi	-	tuʔu	muta
W:NGD	ana	moa	mite	nipi	puse	tuʔu	muta
C:NGD	ana	moa	mite	nipi	(ŋalu ate)	tuʔu	muta
RON	ana	(ŋaŋa)	mite	nipi	puse	(maɹa)	muta
GLOSS	child	to yawn	black	thin	navel	dry	to vomit

PMP had a few types of word-internal consonant clusters: 1) nasal + homorganic stop such as in PMP *empu 'grandparent, grandchild', 2) clusters arising from reduplication, such as PMP *tektek 'gecko', and 3) a few clusters of *R and another consonant, as in PMP *beRŋi 'night'.

Nasal + voiced stop clusters such as *mb are reanalyzed as unitary phonemes, while nasal + voiceless stop clusters such as PMP *mp are merged to PCF *b or *mb (as in PMP *empu ⇒ PCF *əmbu 'grandparent, grandchild'). Clusters formed by reduplication are usually resolved by deleting the first member (as in PMP *pihpih ⇒ PCF *pipi 'cheek')

but also sometimes the second (as in PMP *tektek 'gecko' ⇒ PCF *təke). Clusters with *R in PMP are resolved by deleting the *R in PCF (as in PMP *qaRta 'stranger, outsider' ⇒ PCF *ata 'person' or PMP *beRŋi 'night' ⇒ PCF *wəŋi 'when').

Table 32: Reflexes of PMP consonant clusters in PCF

PMP	*-mp-	*-Rŋ-	*-Rt-	*-kt-	*-hp-
PCF	*-mb-	*-∅ŋ-	*-∅t-	*-k∅-	*-∅p-
PMP	*empu	*beRŋi	*qaRta	*tektek	*pihpih
PCF	*əmbu	*wəŋi	*ata	*təke	*pipi
LIO	əmbu	wəŋi	ata	təke	pipi
END	ʔəmbu	wəŋi	ʔata	təke	pipi
E:NGO	əmbu	-	ata	təke	pipi
W:NGO	əbu	-	ata	təke	pipi
KEO	ʔembu	wəŋi	ʔata	teke	pipi
E:NAG	əbu	-	ata	təke	pipi
C:NAG	əbu	wəŋi	ata	təke	-
E:NGD	-	wəŋi	ata	-	-
W:NGD	buʔu	wəŋi	ata	təke	pipi
C:NGD	buʔu	wəŋi	ata	təke	pipi
RON	əmbu	wəŋi	ata	təke	pasu
GLOSS	grandparent	when?	person	gecko	cheek

While final consonants are normally deleted without changing the final vowel, there are nonetheless a few examples where a final vowel is lowered upon deletion of a final consonant. This occurs particularly often with final PMP *-R, although cases without lowering such as PMP *tuquR ⇒ PCF *tuʔu 'dry' and PMP *bibiR ⇒ PCF *wiwi 'lip' show that the lowering is not fully predictable. There are only two instances of final PMP *-eq in my data, both of which develop into PCF *-a (in PMP *taneq ⇒ PCF *tana and PMP *ma-baseq ⇒ PCF *(m)basa 'wet'), so this may constitute a regular development.

3.4.3 Reduction of PMP trisyllables

PCF and the modern CF languages have a strong constraint which requires that words consist of a maximum of two syllables. Therefore, all PMP forms consisting of three (or more) syllables are reduced to disyllabic forms in PCF. There are several strategies by which trisyllabic PMP forms may be reduced to PCF disyllables.

The most common strategy is deletion of the initial syllable, as in PMP *qiteluR ⇒ PCF *təlo 'egg'. When the initial syllable has the form PMP *qV-, this is always the strategy employed, but other initial syllables may be deleted as well. One PMP trisyllable,

Table 33: Reflexes of final PMP *-R, *-q in PCF

PMP	*-uR	*-uR	*-uR	*-iR	*-ud	*-eq	*-eq
PCF	*-o	*-o	*-o	*-e	*-o	*-a	*-a
PMP	*niuR	*qiteluR	*ikuR	*wahiR	*bulud	*taneq	*ma-baseq
PCF	*nio	*təlo	*eko	*wae	*wolo	*tana	*(m)basa
LIO	nio	təlo	eko	ae	wolo	tana	(ndəmo)
END	nio	təɹo	ʔeko	ʔae	woɹo	tana	mbasa
E:NGO	nio	təɬo	eko	ae	woɬo	tana	mbasa
W:NGO	nio	təɬo	eko	ae	woɬo	tana	b̥asa
KEO	nio	tədo	ʔeko	ʔae	wodo	tana	(bita)
E:NAG	nio	tədo	eko	ae	wodo	tana	(pisa)
C:NAG	nio	təlo	eko	ae	-	tana	basa
E:NGD	-	təo	eko	ae	-	tana	(rasi)
W:NGD	nio	təlo	eko	wae	wolo	tana	basa
C:NGD	nio	təlo	eko	wae	wolo	tana	basa
RON	nio	təlo	eko	wae	wolo	tana	mbasa
GLOSS	coconut	egg	tail	water	hill	land	wet

*taliŋa ‘ear’ shows a unique development: the first vowel is deleted and the initial *tal- sequence becomes a PCF *kl- cluster, yielding PCF *kliŋa.

If the PMP trisyllable is formed by the verbal PMP *ma- prefix, the vowel of the prefix is usually dropped and the initial nasal segment either causes prenasalization of the following segment, or protects it from the normal processes of lenition targeting PMP phonemes (such as PMP *b ⇒ PCF *w). This process applies mainly to labial initial PMP roots (such as PMP *ma-paqit ⇒ PCF *baʔi ‘bitter’), but is also observed in some non-labial roots (PMP *ma-tuqah ⇒ PCF *nduʔa ‘primary forest’)

In PMP trisyllables formed with *ma- where the root begins with PMP *q-, *h-, the initial nasal of the prefix is retained and the *q-, *h- is deleted, often with coalescence of adjacent vowels (as in PMP *ma-quhaw ⇒ PCF *moa ‘thirsty’).

3.4.4 Summary

In this section, I have gone over the regular reflexes of PMP phonemes as they are inherited in PCF, as well as the regular deletion of PMP coda consonants and the reduction of PMP trisyllables.

Most PMP phonemes have very straightforward reflexes in PCF. Some sound changes which occur between PMP and PCF, such as PMP *q ⇒ PCF *ʔ (between vowels), PMP *b ⇒ PCF *w, PMP *ñ ⇒ PCF *n, PMP *R ⇒ PCF *r, and the deletion of PMP *h are

Table 34: Reflexes of PMP trisyllables in PCF

PMP	*qV-	*qV-	*qV-	*bi-	*ba-	*ka-	*tal-
PCF	*∅-	*∅-	*∅-	*∅-	*∅-	*∅-	*kl-
PMP	*qalejaw	*qalima	*qiteluR	*bituka	*banua	*kamu-yu	*taliŋa
PCF	*lədʒa	*lima	*təlo	*tuka	*nua	*miu	*kliŋa
LIO	lədʒa	lima	təlo	tuka	nua	miu	kiŋa
END	ɹəra	ɹima	təɹo	tuka	nua	miu	iŋa
E:NGO	ɻəya	ɻima	təɻo	tuka	(mboʔa)	miu	yiŋa
W:NGO	ɻəɹa	ɻima	təɻo	tuka	(b̥oʔa)	miu	yiŋa
KEO	dəra	dima	tədo	tuka	nua	miu	iŋa
E:NAG	dəɹa	dima	tədo	tuka	(mboʔa)	miu	liŋa
C:NAG	ləza	lima	tələ	tuka	(boʔa)	miu	hiŋa
E:NGD	əra	ima	təo	tuka	-	miu	hiŋa
W:NGD	ləza	lima	tələ	tuka	nua	miu	hiŋa
C:NGD	ləza	lima	tələ	tuka	nua	miu	hiŋa
RON	ləɹa	lima	tələ	tuka	nua	meu	hiŋa
GLOSS	day, sun	arm, hand	egg	belly	village	2PL	ear

Table 35: PCF reflexes of PMP trisyllables beginning in stative *ma-

PMP	*ma-b-	*ma-p-	*ma-qi-	*ma-qu-	*ma-qu-	*ma-he-
PCF	*b-	*b-	*mi	*mu-	*mo-	*me-
PMP	*ma-baRani	*ma-penuq	*ma-qitem	*ma-qudip	*ma-quhaw	*ma-heyaq
PCF	*bani	*bənu	*mite	*mudi	*moa	*mea
LIO	bani	bənu	mite	muri	moa	mea
END	mbani	mbənu	mite	muri	moa	mea
E:NGO	(dʒiɣa)	mbənu	mite	muyi	moa	mea
W:NGO	(tʃiɣa)	b̥ənu	mite	muɹi	moa	mea
KEO	(ŋasi)	mbənu	mite	muri	moa	mea
E:NAG	(ŋasi)	mbənu	mite	muɹi	moa	mea
C:NAG	-	bənu	mite	muzi	-	-
E:NGD	-	-	mite	muri	-	mea
W:NGD	(nadʒi)	bənu	mite	muzi	moa	mea
C:NGD	(nazi)	bənu	mite	muzi	moa	mea
RON	mbani	mbənu	mite	muɹi	(maɹa)	mea
GLOSS	angry	full	black	alive	thirsty	shy

extremely common and do not provide very good evidence towards the unity of PCF as a clade. One of the more distinctive sound changes to occur is the merger of PMP *z, *j ⇒ PCF *dʒ.

The deletion of all PMP coda consonants is a systematic change which usually leaves no trace and results in only open syllables in all modern CF languages. This is a somewhat distinctive feature, because all other languages of Flores retain at least some coda consonants. The reduction of PMP trisyllables to PCF disyllables is not unique to the CF languages, but it is carried out with exceptional regularity in the CF languages (for which not a single word of more than two syllables can be reconstructed).

3.5 Bottom-up reconstruction of Proto-Central Flores

In this section I will compare the CF languages and determine what can be reconstructed about their most recent common ancestor. At a high level, CF is a tree-like subgroup: the evidence shows that speakers of PCF underwent a significant period of isolation from their sister speech communities. This evidence justifies the reconstruction of PCF as a valid node, because there is reason to believe that there was indeed a speech community ancestral to all and only the CF languages in central Flores at some point in the past.

This evidence takes the form of a large number of exclusively shared innovations which link the CF languages to the exclusion of all other Austronesian languages. The strongest pieces of evidence are the innovation of a mixed base quinary-decimal system of numbers (see section 3.5.3) and the development of a shared paradigm of directionals (see section 3.5.4). Other key pieces of evidence are the total loss of final consonants and bound morphology, as well as a large number of lexical replacements, irregular sound changes, and semantic shifts which affect only the CF languages (see section 3.5.5). Taken together, these provide convincing evidence that there was a split between PCF and its nearest relatives, including its closest relative Palu'e (which does not participate in these key innovations), followed by a period of isolation which allowed exclusive innovations to accumulate.

3.5.1 PCF phonology

I will now examine the sound correspondences which hold between CF languages. I will propose a reconstructed PCF proto-phoneme for each and laying out the sequence of regular sound changes which derives the attested CF reflexes from the reconstructed PCF forms.

When there is more than one entry for a given CF language, I choose the one that is cognate. If there is a word attested for that meaning but it is not cognate, it is listed in (parentheses), while if there is simply no word attested, this is marked as a hyphen -.

The consonant inventory which I reconstruct for PCF is shown in Table 36. The reconstructed vowel inventory of PCF is shown in Table 37. The vowel system of PCF needs no further explanation, because all modern CF languages share a 6-vowel system identical to the one shown and all vowels correspond identically across CF languages.

Table 36: Reconstructed consonant inventory of PCF

	LABIAL		CORONAL		VELAR		GLOTTAL
STOP	*p	*b	*t	*d	*k	*g	*ʔ
AFFRICATE				*dʒ			
IMPLOSIVE		*ɓ		*ɗ			
PRENASAL		*mb		*nd		*ŋg	
NASAL		*m		*n		*ŋ	
FRICATIVE	*f		*s				
GLIDE		*w		*y			
LATERAL				*l			
RHOTIC				*r			

Table 37: Reconstructed vowel inventory of PCF

	FRONT	CENTRAL	BACK
CLOSE	*i		*u
MID	*e	*ə	*o
OPEN		*a	

Reconstructed PCF consonants which correspond identically across all modern CF languages are summarized in Table 38. Furthermore, all PCF vowels correspond identically across the modern CF languages. The reflexes of PCF consonants which undergo changes in the daughter CF languages are summarized in Table 39. All of these correspondences will be treated in more detail in the section that follows.

PCF *p, *t, *k, *ʔ The PCF voiceless stops *p, *t, *k are retained across all CF languages, and can be reconstructed unproblematically. The PCF glottal stop *ʔ only occurs in medial position, where it is retained in all CF languages and can be reconstructed unproblematically.

PCF *m, *n, *ŋ The PCF nasals *m, *n, *ŋ are retained across all CF languages, and can be reconstructed unproblematically for PCF.

PCF *b, *mb, *g, *ŋg First, I will describe the sound correspondences which reflect PCF *b, *mb in detail. The changes affecting PCF *g, *ŋg are exactly analogous to these, and will be discussed more briefly at the end.

There are two distinct sound correspondences which hold between /b : mb : b̥/ in the CF languages, which I reconstruct as two PCF phonemes: *b and *mb. These two sound correspondences are identical except for the Lio reflex. This is because Lio is the only

Table 38: Reflexes of PCF consonants which correspond identically across CF languages

PCF	*p	*t	*k	*ʔ	*m	*n	*ŋ	*ɓ	*ɗ	*w	*s
LIO	p	t	k	ʔ	m	n	ŋ	ɓ	ɗ	w	s
END	p	t	k	ʔ	m	n	ŋ	ɓ	ɗ	w	s
E:NGO	p	t	k	ʔ	m	n	ŋ	ɓ	ɗ	w	s
W:NGO	p	t	k	ʔ	m	n	ŋ	ɓ	ɗ	w	s
KEO	p	t	k	ʔ	m	n	ŋ	ɓ	ɗ	w	s
E:NAG	p	t	k	ʔ	m	n	ŋ	ɓ	ɗ	w	s
C:NAG	p	t	k	ʔ	m	n	ŋ	ɓ	ɗ	w	s
E:NGD	p	t	k	ʔ	m	n	ŋ	ɓ	ɗ	w	s
W:NGD	p	t	k	ʔ	m	n	ŋ	ɓ	ɗ	w	s
C:NGD	p	t	k	ʔ	m	n	ŋ	ɓ	ɗ	w	s
RON	p	t	k	ʔ	m	n	ŋ	ɓ	ɗ	w	s

Table 39: Reflexes of PCF consonants which undergo changes in CF languages

PCF	*b	*mb	*g	*ŋg	*nd	*d	*dʒ	*l	*y	*f	*r
LIO	b	mb	g	ŋg	nd	r	dʒ	l	dʒ	f	r
END	mb	mb	ŋg	ŋg	nd	r	r	ɹ	dʒ	h	r
E:NGO	mb	mb	ŋg	ŋg	nd	y	y	ɬ	dʒ	f	y
W:NGO	b̥	b̥	g̥	g̥	nd	ɹ	ɹ	ɬ	tʃ	f	y
KEO	mb	mb	ŋg	ŋg	nd	r	r	d	dʒ	f	l
E:NAG	mb	mb	ŋg	ŋg	nd	ɹ	ɹ	d	tʃ	f	l
C:NAG	b	b	g	g	d	z	z	l	dʒ	f	∅
E:NGD	b	b	g	g	d	r	r	∅	dʒ	f	r
W:NGD	b	b	g	g	d	z	z	l	dʒ	f	r
C:NGD	b	b	g	g	d	z	z	l	z	f	r
RON	mb	mb	ŋg	ŋg	nd	ɹ	ɹ	l	dʒ	f	r

Table 40: Reflexes of PCF voiceless stops *p, *t, *k, *ʔ

PCF	*p-	*-p-	*t-	*-t-	*k-	*-k-	*-ʔ-
PMP	***p**iliq	*a**p**a	*bi**t**uka	*ki**t**a	*k**k**ahu	*hi**k**an	*ta**q**i
PCF	*p**p**ili	*a**p**a	*t**t**uka	*ki**t**a	*k**k**au	*i**k**a	*ta**ʔ**i
LIO	**p**ili	a**p**a	**t**uka	ki**t**a	**k**au	i**k**a	ta**ʔ**i
END	**p**iɹi	ʔa**p**a	**t**uka	ki**t**a	**k**au	ʔi**k**a	ta**ʔ**i
E:NGO	**p**iɬi	a**p**a	**t**uka	ki**t**a	**k**au	i**k**a	ta**ʔ**i
W:NGO	**p**iɬi	a**p**a	**t**uka	ki**t**a	**k**au	i**k**a	ta**ʔ**i
KEO	**p**idi	ʔa**p**a	**t**uka	ki**t**a	**k**au	ʔi**k**a	ta**ʔ**i
E:NAG	**p**idi	a**p**a	**t**uka	ki**t**a	**k**au	i**k**a	ta**ʔ**i
C:NAG	**p**ili	a**p**a	**t**uka	ki**t**a	**k**au	i**k**a	ta**ʔ**i
E:NGD	**p**i	a**p**a	**t**uka	ki**t**a	**k**au	i**k**a	ta**ʔ**i
W:NGD	**p**ili	a**p**a	**t**uka	ki**t**a	**k**au	i**k**a	ta**ʔ**i
C:NGD	**p**ili	a**p**a	**t**uka	ki**t**a	**k**au	i**k**a	ta**ʔ**i
RON	**p**ili	a**p**a	**t**uka	ki**t**a	**k**au	i**k**a	ta**ʔ**i
GLOSS	to pick up	what	belly	1PL.I	2SG	fish	excrement

Table 41: Reflexes of PCF nasals *m, *n, *ŋ

PCF	*m-	*-m-	*n-	*-n-	*ŋ-	*ŋ
PMP	*m**m**a-qitem	*ka**m**i	*n**n**aŋuy	*a**n**ak	-	*tali**ŋ**a
PCF	*m**m**ite	*ka**m**i	*n**n**aŋu	*a**n**a	*ŋ**ŋ**iʔi	*kli**ŋ**a
LIO	**m**ite	ka**m**i	**n**aŋu	a**n**a	**ŋ**iʔi	ki**ŋ**a
END	**m**ite	ka**m**i	**n**aŋu	ʔa**n**a	**ŋ**iʔi	i**ŋ**a
E:NGO	**m**ite	ka**m**i	**n**aŋu	a**n**a	**ŋ**iʔi	yi**ŋ**a
W:NGO	**m**ite	ka**m**i	**n**aŋu	a**n**a	**ŋ**iʔi	yi**ŋ**a
KEO	**m**ite	ka**m**i	**n**aŋu	ʔa**n**a	**ŋ**iʔi	i**ŋ**a
E:NAG	**m**ite	ka**m**i	**n**aŋu	a**n**a	**ŋ**iʔi	li**ŋ**a
C:NAG	**m**ite	ka**m**i	**n**aŋu	a**n**a	**ŋ**iʔi	hi**ŋ**a
E:NGD	**m**ite	ka**m**i	**n**aŋu	a**n**a	**ŋ**iʔi	hi**ŋ**a
W:NGD	**m**ite	ka**m**i	**n**aŋu	a**n**a	**ŋ**iʔi	hi**ŋ**a
C:NGD	**m**ite	ka**m**i	**n**aŋu	a**n**a	**ŋ**iʔi	hi**ŋ**a
RON	**m**ite	ka**m**i	**n**aŋu	a**n**a	**ŋ**iʔi	hi**ŋ**a
GLOSS	black	1PL.E	to swim	child	teeth	ear

language which does not have a merger of PCF *b, *mb, instead reflecting them faithfully as /b, mb/ respectively.

Table 42: Reflexes of PCF *b, *mb

	*b-	*-b-	*mb-	*-mb-	*-mb-
PMP	***m**a-**p**enuq	-	-	(*tu**mb**uq?)	-
PCF	***b**ənu	*ko**b**e	***mb**e?o	*tə**mb**u	*ə**mb**a
LIO	**b**ənu	ko**b**e	**mb**e?o	tə**mb**u	ə**mb**a
END	**mb**ənu	ko**mb**e	**mb**e?o	tə**mb**u	?ə**mb**a
E:NGO	**mb**ənu	ko**mb**e	**mb**e?o	tə**mb**u	ə**mb**a
W:NGO	**b̥**ənu	ko**b̥**e	**b̥**e?o	tə**b̥**u	ə**b̥**a
KEO	**mb̥**ənu	ko**mb̥**e	**mb̥**e?o	tə**mb̥**u	?ə**mb̥**a
E:NAG	**mb**ənu	ko**mb**e	**mb**e?o	-	ə**mb**a
C:NAG	**b**ənu	ko**b**e	**b**e?o	tə**b**u	**b**a
E:NGD	-	ko**b**e	**b**e?o	-	(de)
W:NGD	**b**ənu	ko**b**e	**b**e?o	tə**b**u	(de)
C:NGD	**b**ənu	ko**b**e	**b**e?o	tə**b**u	(de)
RON	**mb**ənu	ko**mb**e	**mb**e?o	tə**mb**u	(nde)
GLOSS	full	night	to know	to grow	which

In Central Nage (C:NAG) and the three Ngadha dialects (E:NGD, C:NGD, W:NGD), prenasalized phonemes are absent, and all PCF *mb ⇒ /b/. In the other languages except Lio (END, E:NGO, W:NGO, E:NAG, KEO, RON), all PCF *b are instead prenasalized to /mb/. While plain voiced /b/ is not absent from most of these varieties, it is very rare in inherited vocabulary.

The prenasalization of PCF *b ⇒ /mb/ targets two separate portions of the CF group: Rongga in the far west, as well as Ende, Nga'o, Kéo and Nage dialects in the central-eastern part. Rongga is separated from these other dialects not only by languages similar to Manggarai, but also by the non-prenasalizing Ngadha and central Nage varieties. It is more likely that this sound change arose in two places, rather than spreading from the central-eastern zone to the far western periphery without affecting the Ngadha and Nage-Kéo dialects in between.

After the merger of PCF *b, *mb ⇒ mb, there is an additional change which targets only Western Nga'o (W:NGO) whereby the phoneme /mb/ shifts its voice onset time (VOT) to be very close to 0 (an analogous change also affects /ŋg/; see next section). For this reason, I transcribe this as a de-voiced stop [b̥]. This phoneme still contrasts with the highly aspirated /p/ (with a large positive VOT) and the true voiced /b/ (with a small negative VOT), so it is very different from the merger of PCF *mb, *b ⇒ /b/ in

Central Nage and the Ngadha dialects. The latter change eliminates a stop series from the phonemic inventory entirely, while Eastern Nga'o preserves three distinct phonemic series. Synchronically, it would make more sense to analyze Western Nga'o as having a three-way opposition between aspirates /p^{h}/ (current /p/), true voiceless stops /p/ with VOT close to 0 (current /b̥/) and voiced stops /b/, but for comparative purposes is is preferable to keep the symbols as they are here.

The situation with PCF *g, *ŋg is exactly analogous to *b, *mb. There are two correspondences involving /g : ŋg : g̊/, and Lio is the only language showing a difference. The same languages which prenasalize *b also do so with *g, and those which eliminate prenasalization on *mb also do so on *ŋg. Again, Western Nga'o has undergone first the merger of PCF *g, *ŋg ⇒ *ŋg, then a subsequent shift of *ŋg ⇒ /g̊/, but this remains phonologically contrastive with /k/ (aspirate) and /g/ (true voiced). Only one solid cognate set which shows the behavior of Eastern Ngadha (E:NGD) was found, PCF *po(ŋ)go 'to cut off' ⇒ /pogo/, but the lack of a Lio reflex means that the medial consonant cannot be disambiguated between *g and *ŋg.

Table 43: Reflexes of PCF *g, *ŋg

	*g-	*g-	*-g-	*ŋg-	*-ŋg-	*-ŋg-	*-(ŋ)g-
PMP	-	(*hetek?)	-	-	-	-	(*puŋgul?)
PCF	***g**ae	***g**əte	*ya**g**a	***ŋg**ake	*lo**ŋg**o	*e**ŋg**e	*po(**ŋ**)**g**o
LIO	**g**ae	**g**əte	sa**g**a	**ŋg**ake	lo**ŋg**o	e**ŋg**e	-
END	**ŋg**ae	**ŋg**əte	dʒa**ŋg**a	**ŋg**ake	ɹo**ŋg**o	ʔe**ŋg**e	po**ŋg**o
E:NGO	**ŋg**ae	**ŋg**əte	dʒa**ŋg**a	**ŋg**ake	ɧo**ŋg**o	e**ŋg**e	po**ŋg**o
W:NGO	**g̊**ae	**g̊**əte	tʃa**g̊**a	**g̊**ake	ɧo**g̊**o	e**g̊**e	po**g̊**o
KEO	**ŋg**ae	**ŋg**əte	-	**ŋg**ake	do**ŋg**o	ʔe**ŋg**e	po**ŋg**o
E:NAG	**ŋg**ae	**ŋg**əte	tʃa**ŋg**a	**ŋg**ake	do**ŋg**o	e**ŋg**e	po**ŋg**o
C:NAG	**g**ai	-	(dəɗi)	-	lo**g**o	e**g**e	-
E:NGD	-	-	-	-	(kawa tona)	-	po**g**o
W:NGD	**g**ae	**g**əte	dʒa**g**a	**g**ake	lo**g**o	e**g**e	po**g**o
C:NGD	**g**ae	**g**əte	za**g**a	(kupu kupu)	lo**g**o	e**g**e	po**g**o
RON	(pita)	**ŋg**əte	dʒa**ŋg**a	**ŋg**ake	lo**ŋg**o	(reʔa)	po**ŋg**o
GLOSS	to search	to chop	height	butterfly	back	waist	to cut off

PCF *nd There is a regular correspondence between /nd : d/ in the CF languages which I reconstruct as the reflex of PCF *nd. Like *mb and *ŋg, PCF *nd is subject to denasalization in Central Nage and all three varieties of Ngadha, while it is preserved in the other languages.

Table 44: Reflexes of PCF *nd

	*nd-	*nd-	*-nd-
PMP	-	-	-
PCF	***nd**ate	***nd**ena	*na**nd**e
LIO	**nd**ate	**nd**ena	na**nd**e
END	**nd**ate	**nd**ena	na**nd**e
E:NGO	**nd**ate	**nd**ena	(eyu)
W:NGO	**nd**ate	**nd**ena	(eɹu)
KEO	**nd**ate	(baɓi)	na**nd**e
E:NAG	**nd**ate	**nd**ena	(eɹu)
C:NAG	-	-	na**d**e
E:NGD	**d**ate	-	na**d**e
W:NGD	**d**ate	**d**ena	na**d**e
C:NGD	**d**ate	**d**ena	na**d**e
RON	**nd**ate	(dʒere)	na**nd**e
GLOSS	heavy	flat	to sleep

PCF *d, *dʒ There are two complex but well-attested sound correspondences involving the phonemes /dʒ: z : ɹ: r : y/ which I reconstruct as the reflexes of two phonemes, PCF *d and *dʒ. Only Lio shows distinct reflexes with PCF *d ⇒ /r/ while PCF *dʒ ⇒ /dʒ/. All other languages show a merger of PCF *d ⇒ *dʒ and then further changes leading to the wide range of attested reflexes.

I propose the following development from PCF *dʒ to the daughter languages.

Lio is the most conservative CF language with respect to this phoneme. It has a reflex /dʒ/, an alveolar affricate realized as [dʒ] or [dz] which is more or less equal to the original phonetic value which I reconstruct for PCF. It is also the only language in which PCF *d does not merge with *dʒ and yields /r/ instead.

Central Nage, Western Ngadha and Central Ngadha (C:NAG, W:NGD, C:NGD) then underwent the loss of the initial stop portion, leaving only an alveolar fricative /z/, which they reflect to this day. This change also affected the other CF languages, but they then went on to rhotacize the intermediate /z/ to /ɹ/ (preserved in W:NGO, E:NAG, RON) and further to /r/ (preserved in END, KEO, E:NGD). Uniquely among CF languages, E:NGO went on to shift all rhotics to /y/ and thus reflects PCF *dʒ as /y/.

In summary, Lio preserves the closest thing to the original phonetic value of PCF *dʒ. The first shift (/dʒ/ ⇒ /z/) affected all other languages (END, E:NGO, W:NGO, KEO, E:NAG, C:NAG, W:NGD, C:NGD, E:NGD, RON). The second shift to a rhotic (/z/ to /ɹ/) then affected most of these languages (END, E:NGO, W:NGO, KEO, E:NAG, E:NGD, RON). This shift continued onwards from /ɹ/ to /r/ in END, KEO and E:NGD, while /ɹ/

Table 45: Reflexes of PCF *d, *dʒ

	*d-	*-d-	*dʒ-	*-dʒ-	*-dʒ-
PMP	***d**uha	*ma-qu**d**ip	***z**alan	*qu**z**an	*qale**j**aw
PCF	***d**ua	*mu**d**i	***dʒ**ala	*u**dʒ**a	*lə**dʒ**a
LIO	**r**ua	mu**r**i	**dʒ**ala	u**dʒ**a	lə**dʒ**a
END	**r**ua	mu**r**i	**r**aɹa	ʔu**r**a	ɹə**r**a
E:NGO	**y**ua	mu**y**i	**y**aɮa	u**y**a	ɮə**y**a
W:NGO	**ɹ**ua	mu**ɹ**i	**ɹ**aɮa	u**ɹ**a	ɮə**ɹ**a
KEO	**r**ua	mu**r**i	**r**ada	ʔu**r**a	də**r**a
E:NAG	**ɹ**ua	mu**ɹ**i	**ɹ**ada	u**ɹ**a	də**ɹ**a
C:NAG	**z**ua	mu**z**i	**z**ala	u**z**a	lə**z**a
E:NGD	**r**ua	mu**r**i	**r**a	u**r**a	ə**r**a
W:NGD	**z**ua	mu**z**i	**z**ala	u**z**a	lə**z**a
C:NGD	**z**ua	mu**z**i	**z**eta	ʔu**z**a	lə**z**a
RON	**ɹ**ua	mu**ɹ**i	laɹa	nuɹa	ləɹa
GLOSS	two	alive	road	rain	sun

shifted to /y/ in E:NGO alone.

PCF *ɓ There is a regular correspondence between /ɓ/ in all CF languages which I reconstruct as PCF *ɓ. However, Lio is often in disagreement with the other languages, reflecting a plain /b/ where the other languages show implosive /ɓ/. I interpret this as sporadic loss of implosion in Lio rather than two separate PCF phonemes with differing reflexes.

PCF *ɗ There is a regular correspondence between /ɗ/ in all CF languages which I reconstruct as PCF *ɗ. In a small number of words, some CF languages show /nd/ corresponding with /ɗ/. The significance of this remains unclear, but it seems likely that at least some instances of /ɗ/ derive from earlier /nd/.

PCF *l There is a regular sound correspondence between /l : d : ɹ : ɮ : ∅/ for which I reconstruct PCF *l.

The PCF phoneme *l was affected by at least two independent sets of innovations.

First, there is an change of PCF *l ⇒ /∅/ in Eastern Ngadha alone. This is one of the defining characteristics of the Eastern Ngadha, and it does not target any other CF languages or dialects. Central Nage dialects, which are spoken in an area near Eastern Ngadha, show the complete loss of /r/, but these cannot be unified into a single sound change because there is no regular correspondence of /r : l/ in these languages and so

Table 46: Reflexes of PCF *ɓ

	*ɓ-	*ɓ-	*ɓ-	*ɓ-	*ɓ-	*-ɓ-
PMP	***b**etuŋ	-	***b**alik	-	-	-
PCF	***ɓ**əto	***ɓ**anda	***ɓ**ale	***ɓ**oko	***ɓ**ara	*lo**ɓ**u
LIO	**ɓ**əto	**ɓ**anda	**ɓ**ale	**b**oko	**b**ara	lo**b**u
END	**ɓ**əto	**ɓ**anda	**ɓ**aɹe	**ɓ**oko	**ɓ**ara	ɹo**ɓ**u
E:NGO	**ɓ**əto	**ɓ**anda	(koso)	**ɓ**oko	**ɓ**aya	do**ɓ**u
W:NGO	**ɓ**əto	**ɓ**anda	(koso)	**ɓ**oko	**ɓ**aya	do**ɓ**u
KEO	**ɓ**əto	**ɓ**anda	**ɓ**ade	**ɓ**oko	**ɓ**ala	-
E:NAG	**ɓ**əto	**ɓ**anda	(roka)	**ɓ**oko	**ɓ**ala	do**ɓ**u
C:NAG	-	-	-	**ɓ**oko	**ɓ**a	-
E:NGD	-	-	-	**ɓ**oko	**ɓ**ara	-
W:NGD	**ɓ**əto	**ɓ**ada	**ɓ**ale	**ɓ**oko	**ɓ**ara	lo**ɓ**u
C:NGD	**ɓ**əto	-	(kusu)	**ɓ**oko	**ɓ**ara	-
RON	**ɓ**əto	(ɓora)	**ɓ**ale	**ɓ**oko	**ɓ**ara	(ŋəmbu)
GLOSS	bamboo sp.	buffalo	to turn	short	white	dolphin

Table 47: Reflexes of PCF *ɗ

	*ɗ-	*-ɗ-	*nd-/*ɗ-	*nd-/*ɗ-
PMP	-	-	-	-
PCF	***ɗ**oka	*pə**ɗ**e	***nd-/ɗ**oʔa	***nd-/ɗ**əmo
LIO	**ɗ**oka	pe**ɗ**e	(dəke)	**nd**əmo
END	**ɗ**oka	pe**ɗ**e	-	**ɗ**əmo
E:NGO	**ɗ**oka	pə**ɗ**e	**ɗ**oʔa	(mbasa)
W:NGO	**ɗ**oka	pə**ɗ**e	**ɗ**oʔa	(b̥asa)
KEO	-	pə**ɗ**e	-	(bita)
E:NAG	**ɗ**oka	pə**ɗ**e	**ɗ**oʔa	(pisa)
C:NAG	(hiŋo)	pə**ɗ**e	-	(basa)
E:NGD	-	pə**ɗ**e	-	(rasi)
W:NGD	(hiŋo)	pə**ɗ**e	-	**ɗ**əmo
C:NGD	(hiŋo)	(zaka)	**ɗ**oʔa	(basa)
RON	(hiŋo)	pə**ɗ**e	**nd**oʔa	**ɗ**əmo
GLOSS	deaf	to cook	stick	wet

Table 48: Reflexes of PCF *l

PCF	*l-	*l-	*-l-	*-l-
PMP	***l**ima	***l**aku	*bu**l**an	*manta**l**aq
PCF	***l**ima	***l**ako	*wu**l**a	*nda**l**a
LIO	**l**ima	**l**ako	wu**l**a	da**l**a
END	**ɹ**ima	**ɹ**ako	wu**ɹ**a	nda**ɹ**a
E:NGO	**ɮ**ima	**ɮ**ako	wu**ɮ**a	nda**ɮ**a
W:NGO	**ɮ**ima	**ɮ**ako	wu**ɮ**a	nda**ɮ**a
KEO	**d**ima	**d**ako	wu**d**a	nda**d**a
E:NAG	**d**ima	**d**ako	wu**d**a	nda**d**a
C:NAG	**l**ima	**l**ako	wu**l**a	da**l**a
E:NGD	ima	ako	wua	da
W:NGD	**l**ima	**l**ako	wu**l**a	da**l**a
C:NGD	**l**ima	**l**ako	wu**l**a	da**l**a
RON	**l**ima	**l**ako	wu**l**a	nda**l**a
GLOSS	five	dog	moon	star

they target different forms. Thus, Eastern Ngadha reflects PCF *wula 'moon', *wara 'west wind' ⇒ /wua, wara/, while Central Nage shows /wula, wa/, with the opposite pattern of deletion.

The second set of changes affecting PCF *l operate in a classic wave-like fashion, emanating from an innovatory center somewhere in the central CF zone (Eastern Nage, Kéo, Nga'o) and crossing between languages. I propose that the reflexes of PCF *l attested in Eastern and Western Nga'o /ɮ/, Kéo and Eastern Nage /d/ (and possibly Ende /ɹ/) are the result of a connected series of changes.

A first, larger wave of innovation saw PCF *l ⇒ *ɮ in Eastern Nage, Kéo and both Nga'o dialects, before a more restricted wave targeted Eastern Nage and Kéo, shifting the intermediate *ɮ ⇒ /d/. These changes must have taken place more recently than the extensive changes affecting PCF *d, because the resulting instances of /d/ do not undergo the same changes.

This proposed chain is quite unusual, but it is strongly supported by the attestation of the phoneme /ɮ/ in the two Nga'o wordlists I recorded. This sound (which was not previously known to occur in Flores) is, phonetically speaking, the missing link between /l/ and /d/. It is a voiced, strongly fricated palatal lateral [ɮ] with optional pre-stopping [dɮ]. In some cases, particularly after /u/, only the pre-stopped [d] portion can be heard. Before /i/, it is produced with less frication and approaches plain [l]. I propose that in the Nage and Kéo dialects which show the shift PCF *l ⇒ /d/ once had a similar *ɮ phoneme, but the pre-stopped portion of the phoneme *ɮ began obligatory and the fricative portion

disappeared entirely, yielding /d/.

A third change affects Ende alone and results in the shift from PCF *l ⇒ /ɹ/, the currently attested reflex. I have taken the conservative stance that the Ende change of PCF *l ⇒ /ɹ/ is an independent change which affects only Ende because phonetically, the modern Nga'o /ɮ/ and Ende /ɹ/ are quite dissimilar. Furthermore, Eastern Nga'o also includes the phoneme /ɹ/ but this corresponds to PCF *dʒ instead of PCF *l. However, it would be possible to argue that the general trend of shifting away from a plain lateral to a more fricative or non-lateral approximant realization is an areal tendency which has spread amongst the affect dialects.

PCF *w There is a regular correspondence between /w/ in all CF languages which I reconstruct unproblematically as PCF *w. However, in two well-attested words, PCF *waʔi 'leg' and *wae 'water', there is irregular loss of the initial /w-/ in Western Ngadha, Central Ngadha and Rongga.

Table 49: Reflexes of PCF *w

	*w-	*-w-	*w-	*w-	*w-
PMP	***b**ulan	*qa**b**u	***w**aRi	***w**ahiR	***w**aqay
PCF	***w**ula	*a**w**u	***w**ari	***w**ae	***w**aʔi
LIO	**w**ula	a**w**u	**w**ari	ae	**ɣ**aʔi
END	**w**uɹa	ʔa**w**u	**w**ari	ʔae	aʔi
E:NGO	**w**uɮa	a**w**u	**w**ayi	ae	aʔi
W:NGO	**w**uɮa	a**w**u	**w**ayi	ae	aʔi
KEO	**w**uda	ʔa**w**u	**w**ali	ʔae	aʔi
E:NAG	**w**uda	a**w**u	(maɹa)	ae	aʔi
C:NAG	**w**ula	a**w**u	**w**ai	ae	aʔi
E:NGD	**w**ua	a**w**u	-	ae	aʔi
W:NGD	**w**ula	a**w**u	**w**ari	**w**ae	**w**aʔi
C:NGD	**w**ula	a**w**u	**w**ari	**w**ae	**w**aʔi
RON	**w**ula	a**w**u	**w**ari	**w**ae	**w**aʔi
GLOSS	moon	ash	to dry	water	leg, foot

PCF *y There is a sound correspondence between /dʒ : tʃ: z/ which I reconstruct as PCF *y, even though no modern languages reflect this phoneme as /y/. .

All CF languages, and many outside of CF as well, undergo a shift of *y ⇒ *dʒ, and then Western Nga'o and Eastern Nage devoice this to /tʃ/. In So'a, there is a shift of *dʒ ⇒ /z/.

The change PCF *y ⇒ dʒ can be ordered with respect to two other sound changes. It must have occurred after the developments from PCF *dʒ ⇒ /z, ɹ, r/ in most CF languages, because /dʒ/ which reflect earlier *y are not affected by these changes. In languages which have changed PCF *dʒ to these other phonemes, the only source of the phoneme /dʒ/ is PCF *y.

On the other hand, this change must have happened before the shift of /l/ ⇒ /y/ in Nga'o (and also /r/ ⇒ /y/ in Eastern Nga'o), because these instances of /y/ are not affected by the shift PCF *y ⇒ /dʒ/.

Table 50: Reflexes of PCF *y

	*y-	*y-	*-y-	*-y-	*-y-
PMP	(***z**awa)	-	*kah**i**w	-	*ba**y**u
PCF	***y**awa	***y**aga	*ka**y**u	*ŋa**y**u	*wa**y**u
LIO	**dʒ**awa	**s**aga	ka**dʒ**u	ŋa**dʒ**u	wa**dʒ**u
END	**dʒ**awa	**dʒ**aŋga	ka**dʒ**u	ŋa**dʒ**u	wa**dʒ**u
E:NGO	**dʒ**awa	**dʒ**aŋga	ka**dʒ**u	ŋa**dʒ**u	wa**dʒ**u
W:NGO	**tʃ**awa	**tʃ**a̰ga	ka**tʃ**u	ŋa**tʃ**u	wa**tʃ**u
KEO	**dʒ**awa	-	ka**dʒ**u	-	wa**dʒ**u
E:NAG	**tʃ**awa	**tʃ**aŋga	ka**tʃ**u	ŋa**tʃ**u	wa**tʃ**u
C:NAG	-	(dədi)	ka**dʒ**u	ŋa**dʒ**u	(ɗo)
E:NGD	-	-	ka**dʒ**u	ŋa**dʒ**u	-
W:NGD	**dʒ**awa	**dʒ**aga	ka**dʒ**u	ŋa**dʒ**u	wa**dʒ**u
C:NGD	**z**awa	**z**aga	ka**z**u	(ŋeu)	(ŋasi)
RON	**dʒ**awa	**dʒ**aŋga	ka**dʒ**u	(ŋeu)	wa**dʒ**u
GLOSS	maize	height	wood	to chew	to pound

PCF *f There is a regular correspondence between /f/ in all CF languages except Ende (which shows /h/) which I reconstruct as PCF *f. A change of PCF *f ⇒ /h/ has affected only Ende.

PCF *s There is an identical correspondance of /s/ across the CF languages, and this can be reconstructed as PCF *s. There is one case of exceptional lenition in the singular proclitic PCF *sa, which becomes /ha/ in Kéo as spoken in Udiworowatu. However, this change does not affect many other Kéo dialects which are excluded from this analysis for lack of data.

PCF *r There is a sound correspondence between /r : l : y : ∅/ which I reconstruct as PCF *r. All languages except Nga'o, Nage and Kéo retain PCF *r ⇒ /r/.

Table 51: Reflexes of PCF *f

	*f-	*f-	*f-	*f-	*f-
PMP	***b**uhek	***b**ahi	-	-	-
PCF	*f**f**u	*f**f**ai	*f**f**eʔa	*f**f**oko	*f**f**oŋa
LIO	**f**u	**f**ai	**f**eʔa	**f**oko	**f**oŋa
END	**h**u	**h**ai	**h**eʔa	(ŋgora)	**h**oŋa
E:NGO	**f**u	**f**ai	**f**eʔa	-	-
W:NGO	**f**u	**f**ai	**f**eʔa	-	-
KEO	**f**u	**f**ai	**f**eʔa	**f**oko	**f**oŋa
E:NAG	**f**u	**f**ai	**f**eʔa	-	**f**oŋa
C:NAG	**f**u	**f**ai	-	**f**oko	-
E:NGD	**f**u	**f**ai	-	**f**oko	-
W:NGD	**f**u	**f**ai	**f**eʔa	**f**oke	**f**oŋa
C:NGD	**f**u	**f**ai	**f**eʔa	**f**oko	-
RON	**f**u	**f**ai	**f**eʔa	**f**oke	**f**oŋa
GLOSS	hair	woman	light (weight)	throat	to want

Table 52: Reflexes of PCF *s

	*s-	*s-	*s-, *-s-	*-s-	*s-
PMP	-	*s**s**ai	*s**s**usu	*ma-qa**s**in	*s**s**a ŋa puluq
PCF	*s**s**aʔo	*s**s**ai	*s**s**usu	*mə**s**i	*s**s**a mbulu
LIO	s**s**aʔo	s**s**ai	s**s**usu	mə**s**i	s**s**a mbulu
END	s**s**aʔo	s**s**ai	s**s**usu	mə**s**i	s**s**a mbuɹu
E:NGO	s**s**aʔo	s**s**ai	s**s**usu	mə**s**i	s**s**a buɮu
W:NGO	s**s**aʔo	s**s**ai	s**s**usu	mə**s**i	s**s**a buɮu
KEO	s**s**aʔo	s**s**ai	s**s**usu	mə**s**i	**h**a mbudu
E:NAG	s**s**aʔo	s**s**ai	s**s**usu	mə**s**i	s**s**a mbudu
C:NAG	s**s**aʔo	-	s**s**usu	mə**s**i	s**s**a bulu
E:NGD	s**s**aʔo	s**s**ai	s**s**usu	mə**s**i	s**s**a bu
W:NGD	s**s**aʔo	s**s**ei	s**s**usu	mə**s**i	s**s**a bulu
C:NGD	s**s**aʔo	s**s**ei	s**s**usu	mə**s**i	s**s**a bulu
RON	s**s**aʔo	s**s**ei	s**s**usu	mə**s**i	s**s**a mbulu
GLOSS	house	who	breast	salty	ten

Eastern Nage and Kéo shift PCF *r ⇒ /l/. This is the same pair of languages which show the complete shift of PCF *l ⇒ /d/, but the shift of PCF *r ⇒ /l/ must have happened after, because the resulting instances of /l/ are not affected by the shift of *l ⇒ /d/. However, this is a good example of a chain shift, coming about because of the gap created after the shift of PCF *l ⇒ /d/.

In Central Nage alone, PCF *r is deleted entirely, but this cannot have happened through an intermediate stage of /l/ because PCF *l is preserved as such in Central Nage. Therefore, this change affects only Central Nage and is unconnected to the other developments of *l and *r.

Table 53: Reflexes of PCF *r

	*r-	*-r-	*-r-	*-r-	*-r-
PMP	*da**R**aq	-	-	*pa-la**R**iw	*pe**r**iŋ
PCF	***r**a*	*to**r**o	*ɓa**r**a	*pa**r**u	*pə**r**i
LIO	**r**a	to**r**o	ba**r**a	pa**r**u	pə**r**i
END	**r**a	to**r**o	ɓa**r**a	pa**r**u	pə**r**i
E:NGO	**y**a	to**y**o	ɓa**y**a	pa**y**u	pə**y**i
W:NGO	**y**a	to**y**o	ɓa**y**a	pa**y**u	pə**ɹ**i
KEO	**l**a	to**l**o	ɓa**l**a	pa**l**u	pə**r**i
E:NAG	**l**a	to**l**o	ɓa**l**a	pa**l**u	pə**ɹ**i
C:NAG	ai	to	ɓa	pau	-
E:NGD	**r**a	to**r**o	ɓa**r**a	-	-
W:NGD	**r**aʔa	to**r**o	ɓa**r**a	pa**r**u	pə**r**i
C:NGD	**r**aʔa	to**r**o	ɓa**r**a	pa**r**u	pə**r**i
RON	**r**aʔa	to**r**o	ɓa**r**a	pa**r**u	pə**r**i
GLOSS	blood	red	white	to run	bamboo sp.

The status of PCF *ɣ It is unclear whether PCF included the phoneme *ɣ. While all modern CF languages do have such a phoneme, there is no single cognate set which provides evidence which would allow this phoneme to be unambiguously reconstructed for PCF. It is relatively rare in all daughter languages, and may have arisen separately in each. The most usual source of /ɣ/ in daughter languages is lenition from /k/, but this cannot be reconstructed to the level of PCF where it occurs: for instance, PMP *paniki 'fruit bat' gives rise to PCF *niki (reflected regularly as /niki/ in most CF languages), but this undergoes sporadic lenition (along with an irregular change in the first consonant) to yield E:NGO, W:NGO, E:NAG /ŋiɣi/.

PCF initial clusters: *kl-, *kr-, *bl- It is necessary to reconstruct initial consonant clusters (at least *kl-, *kr-, and *bl-) in PCF to account for certain sound correspondences, even though there are no consonant clusters present in any modern CF languages. These reconstructed consonant clusters receive support from cognates outside the CF group, and are thus relatively certain.

Table 54: Reflexes of PCF *kl-

PMP	-	*taliŋa	*laki	*lalej
PCF	*kleu	*kliŋa	*klaki	*klale
LIO	keu	kiŋa	kaki	kale
END	eu	iŋa	aki	aɹe
E:NGO	yeu	yiŋa	yaki	yaɮe
W:NGO	yeu	yiŋa	yaki	yaɮe
KEO	eu	iŋa	aki	ade
E:NAG	leu	liŋa	laki	lade
C:NAG	heu	hiŋa	haki	
E:NGD		hiŋa	haki	
W:NGD	heu	hiŋa	haki	hale
C:NGD	heu	hiŋa	haki	hale
RON	heu	hiŋa	haki	hale
GLOSS	betel nut	ear	man	fly (n.)

The best-attested sound correspondence is between /k : l : y: h : ∅/, and I reconstruct this as the regular correspondence set of initial PCF *kl- clusters. I propose that the attested correspondence comes about because different language varieties reduce the cluster in different ways. A good example of an initial *kl- cluster is in PCF *kliŋa 'ear' (from earlier PMP *taliŋa). Only Western Lio (the variety spoken in Detukeli, and included in this analysis) preserves the initial stop, reflecting /kiŋa/. There are certain Eastern Lio dialects, notably those around Lisedetu, which reflect /hiŋa/ instead, but these are absent from my analysis for lack of data.

Eastern Nage reflects the loss of the initial stop instead, and has the form /liŋa/ for 'ear'. This must have occurred after the shift of PCF *l ⇒ /d/, because these instances of /l-/ are not targeted by that shift.

Both Nga'o varieties show the form /yiŋa/. The exact path from PCF *kl- ⇒ /y-/ is unclear, but it probably involved loss of the initial stop first. Since both Nga'o varieties reflect a shift of *r ⇒ /y/, it is possible that the liquid portion of the *kl- cluster merged with *r- first, and then yielded /y-/. Alternatively, it may have shifted directly from *l- ⇒ /y-/, but in a shift that must have happened later than the shift of PCF *l ⇒ /ɮ/ in Nga'o.

In Ende, Kéo, Central Nage, Ngadha and Rongga, the outcome of PCF *kl- was a fricative such as /h-/. This was then lost in Ende and Kéo, yielding true vowel-initial, or 'breathy' words which contrast with glottal-initial words.

A strong argument in favor of the presence of PCF *kl- comes from reflexes of PMP *lalej 'fly', which are continued in PCF *klale. An initial /ka-/ prefix is attested in nearby languages such as Komodo /kalale/. In Sika, which has undergone a regular change of *k ⇒ /ʔ/, the syncope of the first vowel is also attested in /ʔlale/. Sika also has cognates for a large number of other PCF *kl- initial words such as PCF *kliwa 'year' (Sika /ʔliwa/), PCF *kluʔi 'meat' (Sika /ʔluiŋ/), which raises the possibility that initial clusters were introduced mainly in Sika loanwords, but at the level of PCF.

Other initial PCF clusters are less well-attested than PCF *kl-. Initial PCF *kr- clusters appear to have left distinct reflexes only in Eastern Lio varieties around Lisedetu of the type which are not included in this analysis. For instance, PMP *Ramut 'root' seems to be continued in PCF *kramu, with accretion of an initial *k- of unclear origin. This yields /hamu/ in Eastern Lio but /kamu/ in Western Lio (and most other CF languages). However, the presence of a cluster is again corroborated by the Sika cognate /ʔramut/. Another instance is in PCF *krasa 'chest', which yields /hasa/ in Eastern Lio but /kasa/ in all other CF languages. This is corroborated by Komodo /kərasa/ and Sika /ʔrahaŋ/.

The final PCF cluster for which there is evidence is *bl-. This has yielded an initial stop /b-/ in Lio, but reflects an initial lateral *l- in all other CF languages. The most solid example of this is PCF *blewa 'long', possibly backed up by Sika /blo/ 'classifier for long objects'.

Table 55: Reflexes of PCF *kr-, *bl-

PMP	*Ramut	-	-
PCF	*kramu	*krasa	*blewa
W:LIO	kamu	kasa	bewa
E:LIO	hamu	hasa	bewa
END	kamu	kasa	$_e$wa
E:NGO	kamu	kasa	ɮewa
W:NGO	kamu	kasa	ɮ̵ewa
KEO	kamu	kasa	dewa
E:NAG	kamu	kasa	dewa
C:NAG	kamu	-	lewa
E:NGD	kabu	-	ewa
W:NGD	-	kasa	lewa
C:NGD	kamu	kasa	lewa
RON	kamu	kasa	lewa
GLOSS	root	chest	long

3.5.2 Reconstructed PCF pronouns

I will now compare the pronoun systems of the modern CF languages, reconstruct the PCF pronominal system, and trace its development into the daughter languages. The forms of the pronouns in the CF languages are given in Table 56.

Table 56: CF pronoun paradigms with reconstructed PCF pronouns

PMP	*aku	*kahu	’-	*kami	*kita	*kamu-yu	*si ida
PCF	*aku	*kau	*ka(d,dʒ)i	*kami	*kita	*miu	*sida
LIO	aku	kau	kai	kami	kita	miu	əbe
END	dʒaʔo	kau	kai	kami	kita	miu	ʔəbe
E:NGO	ŋaʔo	kau	imu	kami	kita	miu	siya, imu koʔo
W:NGO	ŋaʔo	kau	imu	kami	kita	miu	siɪa, imu koʔo
KEO	ŋaʔo, dʒaʔo	kau	ʔimu	kami	kita	miu	sira, ʔimu koʔo
E:NAG	ŋaʔo	kau	imu	kami	kita	miu	siɪa
C:NAG	ŋaʔo	kau	imu	kami	kita	miu	demu
E:NGA	ŋaʔo	kau	imu	kami	kita	miu	hoga
W:NGA	dʒaʔo, ŋaʔo	kau	gazi	kami	kita	miu	demu
C:NGA	zaʔo, ŋaʔo	kau	gazi	kami	kita	miu	demu
RON	dʒaʔo	kau	kaɹi	kami	kita	meu	siɪa
GLOSS	1SG	2SG	3SG	1PL.E	1PL.I	2PL	3PL

PCF *aku ‘I, me; 1SG’ I reconstruct the PCF 1SG pronoun as *aku, a straightforward retention of the PMP 1SG *aku. This is only inherited in Lio /aku/, while all other CF languages have replaced the 1SG pronoun with an innovative form.

All CF languages other than Lio languages have reflexes of an earlier form *yaʔo, *ŋaʔo, or both. In the peripheral languages Ende (in the east) and Rongga (in the west), only /dʒaʔo/ (← *yaʔo) is found. In Ngadha, Nage and Kéo, both forms co-exist, but with differences in usage reflecting local identities. In Nga’o only the form /ŋaʔo/ is found, hence the name of the language.

Based on the distribution observed, I propose the following scenario. The PCF 1SG pronoun was *aku, but in all languages except Lio this was replaced by an innovative form *yaʔo. Then, an initial velar nasal was added on to form *ŋaʔo, a form which arose in the central CF zone and spread to Ngadha, Nage, Kéo and Nga’o but did not reach Ende and Rongga on the periphery. The innovatory center from which the form *ŋaʔo spread was most likely in the central CF area (Nga’o, Nage and Kéo), affecting the Nga’o-speaking area especially strongly. This change did not completely displace *ŋaʔo as the 1SG pronoun in Ngadha, Nage and Kéo, but introduced it as an alternative.

A few remarks can be made about these developments. First, the innovative forms

*yaʔo and *ŋaʔo probably also reflect PMP *aku, but inherited indirectly. Since PCF does not normally undergo the change *k ⇒ *ʔ, these forms are probably borrowed into the CF languages from another Austronesian source in which this sound changes is regular. A likely candidate for the initial *y- in *yaʔo is the prefixed PMP form *i-aku, but the source of the initial *ŋ- in *ŋaʔo is less clear.

The source language which contributed the form *yaʔo may well have been one of the languages of South or South-East Sulawesi, which are known to have historical links to Flores. For instance, in Buginese of South Sulawesi (whose writing system greatly resembles the 'lota', or 'lontar' system used in Ende) have the 1SG pronoun /iaʔ/, while 1SG pronoun forms in languages of Buton such as Cia-Cia, Lasalimu /iyaʔu/ and Kumbewaha /iaʔu/ (Greenhill, Blust and Gray 2008) have an even greater resemblance. Both Buginese and the languages of Buton are known to have been used as languages of administration and trade, and could have had an influence on the development of the languages of Central Flores.

PCF *kau 'you; 2SG' The PMP 2SG *kahu is retained as PCF *kau and reflected regularly as /kau/ in all CF languages.

PCF *ka(d,dʒ)i 'he; 3SG' The PCF 3SG pronoun *ka(dʒ)i is the least secure among the reconstructed pronouns because the reflexes in daughter languages are not regular enough to establish the original form with certainty.

The evidence from Rongga and Ngadha suggests an original form *kadi or *kadʒi, but the identity of the medial consonant is not possible to determine exactly because only Lio preserves separate reflexes for PCF *d and *dʒ. However, Lio and Ende do not reflect any medial consonant at all, instead showing reflexes of a form *kai. In the remaining CF languages, there has been a semantic shift whereby *imu, originally meaning 'friend, other person' has taken on the basic 3SG pronoun function.

PCF *kami 'we (exclusive); 1PL.E' The PMP 1PL.E *kami is retained as PCF *kami and reflected regularly as /kami/ in all CF languages.

PCF *kita 'we (inclusive); 1PL.I' The PMP 1PL.I *kita is retained as PCF *kita and reflected regularly as /kita/ in all CF languages.

PCF *miu 'you (plural); 2PL' The PMP 2PL *kamu-yu (reconstructed at a lower level as PCEMP *kamiu) is retained as PCF *miu with loss of the first syllable, and reflected regularly as /miu/ in all CF languages except Rongga. Rongga shows an irregular lowering of the medial vowel to /meu/, a feature it shares in common with Manggarai and may be attributable to Manggarai influence.

PCF *sida 'they; 3PL' The bimorphemic PMP form *si ida, composed of a personal marker *si and a 3PL marker *ida, has yielded the PCF *sida. However, reflexes of PCF *sida are no longer the unmarked way of expressing the 3PL in most CF languages. In the central CF zone (Nga'o, Kéo, some Nage) a bimorphemic form *imu koʔo, meaning something like 'the other ones', has become the unmarked 3PL pronoun, with reflexes of *sida reserved for formal or higher register usage. Ngadha has replaced reflexes of *sida with forms such as /demu/ in Central and Western Ngadha, and /hoga/ in Eastern Ngadha, while Ende and Lio have replaced *sida with forms reflecting *əbe.

3.5.3 Reconstructed PCF numeral system

A very strong piece of evidence in favor of CF is the shared innovation of mixed base quinary-decimal numeral system (termed the 'Flores pattern' of numerals by Schapper and Klamer 2014). The numerals 5 and below are monomorphemic, while 6-9 are constructed from the lower numerals using addition, subtraction and multiplication according to the formula shown in Table 57.

Table 57: Reconstructed PCF numerals and their underlying formulas

[1]	[2]	[3]	[4]	[5]	[5+1]	[5+2]	[2x4]	[10-1]
*sa əsa	*dua	*təlu	*wutu	*lima	*lima əsa	*lima dua	*dua mbutu	*təra əsa

The forms of the morphemes are shared by all CF languages, and they undergo the same regular sound changes which affect other PCF words. This is very strong evidence of a single change at the level of PCF, and is also evidence of a non-Austronesian substrate. If only the underlying logic (but not the actual forms) of the numeral system were shared among the CF languages, the evidence would be much weaker, because it could be attributed to a common substrate rather than shared descent. Indeed, Austronesian languages in the vicinity of New Guinea have been attested to switch to non-decimal systems many times as a result of contact with non-Austronesian speakers (see Schapper and Klamer 2014 on numeral systems in Papuan languages of the region). However, the fact the forms of the numerals themselves are cognate, and are old enough to have undergone regular sound change from PCF to the modern languages, is very strong evidence that this transition from decimal to quinary occurred once at the level of PCF. This is highly unlikely to be an instance of the same change recurring in daughter languages after the break-up of PCF. Palu'e and the languages of Western Flores (Manggarai, Rembong, Komodo) have no trace of this system, and retain the Austronesian decimal system with inherited PMP forms. Therefore, there is no doubt that the mixed base 'Flores pattern' of numerals is an innovative feature.

The numerals 5 and below are reflexes of PMP forms, with the exception of 4: PMP *esa ⇒ PCF *sa ʔesa 'one' (with singular proclitic), PMP *duha ⇒ PCF *dua 'two', PMP

*telu ⇒ PCF *təlu ‘three’, PMP *lima ⇒ PCF *lima ‘five’, but PMP *epat is not cognate with PCF *wutu ‘four’. The PCF form *wutu ‘four’ is apparently cognate with Lamaholot /buto/, Kedang /butu rai/ ‘eight’ (with a semantic shift).

Within CF, all languages straightforwardly reflect PCF *wutu ‘four’ except Lio /sutu/, which can plausibly be derived from *sa wutu ‘one (unit of) four’, with the singular proclitic *sa attached to a word *wutu which means ‘a group of four’. However, the Lio form /sutu/ is also a plausible cognate with Sika /hutu/ ‘four’, because Lio initial /s-/ corresponds regularly to Sika /h-/. The relation between the Sika form and the Kedang and Lamaholot forms remains unclear for the moment, and there may have been some borrowing of words for ‘four’ between CF languages and their eastern Flores-Lembata neighbors.

Above 5, numerals are no longer monomorphemic. The numerals 6 and 7 are formed by addition, with a quinary base: PCF *lima əsa {five one} ‘six’ and PCF *lima dua {five two} ‘seven’. The numeral 8 is formed by multiplication, taking 4 as its base: PCF *dua mbutu {two four} ‘eight’. The numeral 9 is formed by subtraction as [10-1], but the [10] is implicit, and the first element in the numeral means something like ‘less, lack’, hence ‘10 lacking 1’: PCF *təra əsa {less one} ‘nine’.

Table 58: CF numeral systems and reconstructed PCF numerals (1 to 5)

PMP	*esa	*duha	*telu	*epat	*lima
PCF	*sa əsa	*dua	*təlu	*wutu	*lima
LIO	sa əsa	rua	təlu	sutu	lima
END	sa ʔəsa	rua	təɹu	wutu	ɹima
E:NGO	sa əsa	yua	təɬu	wutu	ɬima
W:NGO	sa əsa	ɹua	təɬu	wutu	ɬima
KEO	ha ʔəsa	rua	tədu	wutu	dima
E:NAG	sa əsa	ɹua	tədu	wutu	dima
C:NAG	sa əsa	zua	təlu	wutu	lima
E:NGD	sa əsa	rua	təu	wutu	ima
W:NGD	sa əsa	zua	təlu	wutu	lima
C:NGD	sa əsa	zua	təlu	wutu	lima
RON	sa əsa	ɹua	təlu	wutu	lima
PAL	a	rua	təlu	ɓa	lima
MAN	tʃa	sua	təlu	pat	lima
SIK	ha	rua	təlu	hutu	lima
GLOSS	one	two	three	four	five

Table 59: CF numeral systems and reconstructed PCF numerals (6 to 10)

PMP	*enem	*pitu	*walu	*siwa	*sa ŋa puluq
PCF	*lima əsa	*lima dua	*dua mbutu	*təra əsa	*sa mbulu
LIO	lima əsa	lima rua	rua mbutu	təra əsa	sa mbulu
END	ɹima ʔəsa	ɹima rua	rua mbutu	təra ʔəsa	sa mbuɹu
E:NGO	ɻima əsa	ɻima ɹua	ɹua butu	təra əsa	sa buɻu
W:NGO	ɻima əsa	ɻima rua	rua butu	təra əsa	sa buɻu
KEO	dima ʔəsa	dima rua	rua mbutu	təra ʔəsa	ha mbudu
E:NAG	dima əsa	dima ɹua	ɹua mbutu	təra əsa	sa mbudu
C:NAG	lima əsa	lima zua	zua butu	ta əsa	sa bulu
E:NGD	ima əsa	ima rua	rua butu	təra əsa	sa bu
W:NGD	lima əsa	lima zua	rua butu	təra əsa	sa bulu
C:NGD	lima əsa	lima zua	rua butu	təra əsa	sa bulu
RON	lima əsa	lima ɹua	ɹua mbutu	tara əsa	sa mbulu
PAL	ʔəne	ɓitu	walu	iwa	a pulu
MAN	ənəm	pitu	alo	tʃiok	tʃa mpulu
SIK	əna	pitu	walu	hiwa	pulu
GLOSS	six	seven	eight	nine	ten

3.5.4 Reconstructed PCF directionals system

The CF languages share a set of directional terms which correspond in form and in function across the CF languages and provide strong evidence of the validity of PCF as a clade. Directional systems are frequent across Austronesian, and at least the pair *lahud 'towards the sea' and *daya 'inland' can be reconstructed all the way to PAN. However, in the case of the CF languages, the similarities go beyond the expected resemblances between Austronesian languages and must be reconstructed as shared innovations at the level of PCF. An shared grammatical paradigm (including irregularities) is the gold standard for language classification, and this alone would be enough to convincingly argue for the unity of CF as a clade and the validity of PCF as a target of reconstruction.

There are 6 directional terms which form a total of 4 oppositional axes, meaning that certain terms appear in more than one opposition. For example, the Lio term *ghalé* means 'west' in the axis *ghalé :: mena* 'west :: east', but it means 'vertically down' in the axis *ghalé :: ghéta* 'vertically down :: vertically up'. The semantic connection between 'west' and 'vertically down' (both derived from PMP *dalem 'inside, deep') is due to the sun's position below the horizon when it sets in the west. The other term which appears in two different oppositions is *ghéta*, meaning 'vertically up' in the aforementioned axis *ghalé :: ghéta* 'vertically down :: vertically up', but also means 'inland' in the axis *ghéta :: lau* 'inland :: seaward'. The exact same patterns of colexification hold across all CF languages

Table 60: CF directional systems and reconstructed PCF directionals

PMP	-	-	*lahud	*i atas	*i dalem	-
PCF	*d(ʒ)ili	*d(ʒ)ele	*lau	*d(ʒ)eta	*d(ʒ)ale	*məna
LIO	ɣawa	ɣele	lau	ɣeta	ɣale	məna
END	riɹi	reɹe	ɹau	reta	raɹe	məna
E:NGO	-	-	-	yeta	yaɮe	-
W:NGO	-	-	-	ɹeta	ɹaɮe	-
KEO	ridi	rede	dau	reta	rade	məna
E:NAG	ɹidi	-	-	ɹeta	ɹade	-
C:NAG	zili	zele	lau	zeta	zale	məna
E:NGD	-	-	-	reta	rae	-
W:NGD	zili	zele	lau	zeta	zale	məna
C:NGD	-	-	-	zeta	zale	-
RON	ɹili	ɹele	lau	ɹeta	ɹale	məna
GLOSS	down	up	to sea	to mt.; above	below; west	east

for which there is data.

1)	east	↔	west	*məna	↔	*dale
2)	above, upwards	↔	below, downwards	*deta	↔	*dale
3)	to the mountains	↔	towards the sea	*deta	↔	*lau
4)	uphill, upwards	↔	downhill, downwards	*dele	↔	*dili

Based on the available grammatical descriptions, the directional system works identically in all CF languages. All terms are cognate in all languages, with the exception of one Lio term *ghawa* 'downwards', which in other languages is a reflex of PCF *d(ʒ)ili. It is not possible to determine whether the PCF forms contained *d or *dʒ because only the Lio reflex would allow disambiguation of PCF *d and *dʒ, but Lio has undergone an irregular development whereby the 4 of the 6 directional terms begin with /ɣ-/. The origin of the initial /ɣ-/ in the Lio forms is unclear, but it probably spread by analogy to most of the paradigm.

In Table 60, forms from Palu'e, Manggarai and Sika are excluded because these languages have different directional systems, and there are not necessarily exact equivalents for the terms listed here.

3.5.5 Other evidence in favor of CF as a clade

In addition to the very strong evidence provided by the numerals (section 3.5.3) and the directional system (section 3.5.4), there is a substantial body of evidence linking CF as

an innovation defined subgroup. This evidence includes one regular sound change, one morphological innovation, 7 irregular sound changes, 11 lexical replacements and 2 semantic shifts.

The regular sound change which affects all CF languages exclusively is the total loss of all final consonants. There are no final consonants in any modern CF language. Palu'e, which is otherwise very close to the CF languages, does not undergo a complete loss of all final consonants: a final /-n/ marks the possessum in genitive constructions in Palu'e, and creates closed syllables. There are no other regular sound changes which can be identified as exclusively targeting the CF languages, probably because PCF was quite phonologically conservative. The few regular sound changes which could tentatively be assigned to the level of PCF are so common (for example, PMP *b ⇒ PCF *w) that they can be found in most or all nearby languages, and may well have occurred at a higher level.

The morphological innovation which binds all CF languages is the total loss of bound morphology (with the possible exception of the singular proclitic /sa-/). No CF languages has anything that can be unambiguously identified as a bound morpheme. This contrasts with Palu'e, which still maintains the final /-n/ mentioned in the previous paragraph (although Palu'e is otherwise very morphologically simple). The languages of eastern Flores (such as Sika), and western Flores (such as Manggarai) are much more morphologically complex than those of central Flores, and contain bound morphology with both derivational and inflectional functions.

The 7 irregular sound changes which affect only the CF languages are as follows. 1) The initial consonant in PMP *buhek 'hair' yields PCF *fu, instead of expected **wu. 2) PMP *bahi 'female' yields PCF *fai, instead of expected **wai. 3) PMP *buŋa 'flower' undergoes irregular lowering to PCF *woŋa instead of expected **wuŋa. 4) PMP *budaq 'foam' undergoes irregular lowering to PCF *woda instead of expected **wuda. 5) PMP *laku 'civet' undergoes irregular lowering to PCF *lako 'dog', instead of expected **laku. 6) Both vowels in PMP *ikuR 'tail' are irregularly lowered to yield PCF *eko instead of expected **iku. 7) Both vowels in PMP *zaqat 'bad, evil' are irregularly raised to yield PCF *dʒeʔe instead of expected **dʒaʔa.

The 11 lexical replacements which affect only the CF languages are as follows. These occur mainly in the domain of basic vocabulary, which makes them quite convincing evidence of the unity of PCF as an innovation-defined subgroup. 1) PMP *beRŋi 'night' is replaced by a form reflecting *kobe instead. 2) PMP *taqu 'to know' is replaced by a form reflecting *mbeʔo instead. 3) PMP *ma-iRaq 'red' is replaced by a form reflecting *toro instead. 4) PMP *ma-beReqat 'heavy' is replaced by a form reflecting *ndate instead. 5) PMP *buaq 'areca nut' is replaced by a form reflecting *kleu instead. 6) PMP *laŋit 'sky' is replaced by a form reflecting *lidu instead. 7) PMP *kali 'to dig' is replaced by a form reflecting *koe instead. 8) PMP *likud 'back' is replaced by a form reflecting *loŋgo instead. 9) PMP *diqaq 'good' is replaced or supplemented with a form reflecting *pawe instead. 10) PMP *tepiR 'mat' is replaced by a form reflecting *teʔe instead. 11) PMP *adani 'near' is replaced by a form reflecting *weʔe instead.

Table 61: Lexical replacements affecting whole CF group

PMP	*beRŋi	*taqu	*ma-iRaq	*ma-beReqat	*buaq
PCF	*kobe	*mbeʔo	*toro	*ndate	*kleu
LIO	kobe	mbeʔo	toro	ndate	keu
END	kombe	mbeʔo	toro	ndate	eu
E:NGO	kombe	mbeʔo	toyo	ndate	yeu
W:NGO	kob̥e	b̥eʔo	toyo	ndate	yeu
KEO	kombe	mbeʔo	tolo	ndate	eu
E:NAG	kombe	mbeʔo	tolo	ndate	leu
C:NAG	kobe	beʔo	to	-	heu
E:NGD	kobe	beʔo	toro	date	-
W:NGD	kobe	beʔo	toro	date	heu
C:NGD	kobe	beʔo	toro	date	heu
RON	kombe	mbeʔo	toro	ndate	heu
PAL	məre	tʃuʔu	rede	pədʒa	wua
MAN	wie	pətʃiŋ	malo	məndo	ratʃi
SIK	gumaŋ	raʔintaŋ	merak	bərat	wua
GLOSS	night	to know	red	heavy	betel nut

Table 62: Lexical replacements affecting whole CF group (cont.)

PMP	*laŋit	*kali	*likud	*diqaq	*tepiR	*adani
PCF	*lidu	*koe	*loŋgo	*pawe	*teʔe	*weʔe
LIO	liru	koe	loŋgo	pawe	teʔe	weʔe
END	ɹiru	koe	ɹoŋgo	pawe	teʔe	weʔe
E:NGO	ɮiyu	koe	ɮoŋgo	pawe	teʔe	weʔe
W:NGO	ɮiɹu	koe	ɮog̥o	pawe	teʔe	weʔe
KEO	diru	koe	doŋgo	pawe	teʔe	weʔe
E:NAG	diɹu	koe	doŋgo	pawe	teʔe	weʔe
C:NAG	lizu	koe	logo	modɗe	teʔe	weʔe
E:NGD	liru	koe	(kawa tona)	(modɗe)	-	weʔe
W:NGD	lizu	koe	logo	pawe	teʔe	weʔe
C:NGD	lizu	koe	logo	(modɗe)	teʔe	weʔe
RON	liɹu	koe	loŋgo	pawe	teʔe	weʔe
PAL	kəle	kali	tola	mbola	təɓe	təni
MAN	awaŋ	tʃake	toni	diʔa	lotʃe	baliŋ
SIK	wula wutu	gali	tuʔe	ʔəpaŋ	ʔoha	roʔo
GLOSS	sky	to dig	back	good	mat	near

The 2 semantic shifts which affect only the CF languages are as follows. 1) PMP *beRŋi 'night' has shifted semantically to become the question word 'when?', and an innovative word reflecting *kobe has taken on the meaning 'night' instead. The development of 'night' ⇒ 'when' through a construction such as 'how many nights' is attested elsewhere on Flores, as in Sika /rəma pira/ 'when, how many nights?'. 2) PMP *laku 'civet' has yielded PCF *lako 'dog' through semantic shift.

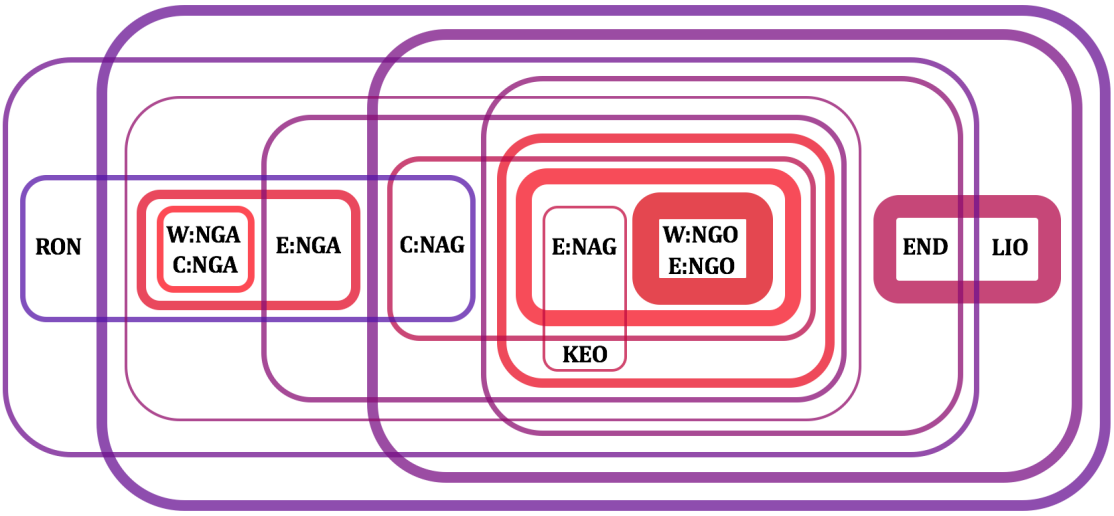
The latter semantic shift from 'civet' to 'dog' in PCF is illuminated by the discussion of Austronesian fauna terms in Blust (2002). Flores has very few native mammals (just bats and rats) because it lies east of the Wallace Line, the eastern border past which most Asian mammals were unable to cross, but too far west to be within the range of marsupials and other Australian mammals. The economic importance of the dog to the early Austronesians is clear from the stability of PMP *asu (PAN *aSu) in most daughter languages, because highly salient animals known by all members of the community are less likely to be replaced. However, the dog lost much of its economic importance in Flores (where there were no large mammals to hunt) and therefore lost much of its salience in the eyes of PCF speakers. Due to the absence of civets in Flores, PMP *laku 'civet' was free to take on the meaning of PMP *asu 'dog', which was lost entirely.

3.5.6 Summary

In this section, I have gone over the regular sound correspondences between CF languages which allow the reconstruction of the PCF phoneme inventory, as well as reconstructed certain aspects of PCF such as its system of pronouns, numerals and directional morphemes. The latter three systems are strong evidence of the unity of CF as a clade and the reality of PCF, as they are unlikely to have arisen more than once in the history of the CF languages. Finally, I have shown some evidence that CF forms a clade in the form of irregular sound changes, lexical replacements, semantic shifts and the loss of all bound morphology.

The more remarkable sound correspondences which hold between CF languages are the following. PCF *dʒ undergoes a merger with PCF *d in all languages except Lio, followed by a series of extensive changes resulting in a range of reflexes such as /dʒ/, /z/, /ɹ/, /r/, /y/ in modern CF languages. PCF *l is subject to various changes, yielding either /r/, /ɹ/, /ʁ/, /d/ or being deleted entirely in certain modern CF languages. In a connected series of changes, PCF *r is shifted to /l/, /y/, or deleted entirely in those CF languages which also show changes in PCF *l. Finally, there is evidence for PCF initial clusters, with the best-supported clusters being *kl-, *kr- and *bl-. These clusters are well-supported by outside evidence, and some instances may represent early loans from eastern Flores languages such as Sika, which allow stop + liquid clusters. Interestingly, even though a velar fricative /ɣ/ is found in all modern CF languages, these do not form any regular sound correspondences and appear to have arisen independently in all daughter languages, and so cannot be reconstructed for PCF.

Figure 6: Glottometric diagram of the CF language group



3.6 Internal subgrouping below the level of PCF

In this section, I will analyse the internal subgrouping of the CF languages by identifying shared innovations among subsets of them. The large number of solid isoglosses which define CF as a clade is typical for a subgroup in the tree-like sense (one which is formed by a split and subsequent period of isolation and shared innovation in a single ancestral speech community). I propose that the later differentiation of the CF languages below the level of PCF is better captured by the wave model, with ongoing contact and mutual intelligibility between neighboring languages leading to a pattern of overlapping innovation which cannot be accurately represented by a tree diagram. These innovations do not tend to fall into nested bundles which could support a clear tree-like internal structure for the CF languages. Instead, neighboring pairs of languages tend to innovate together, while more distant languages do so less often in inverse proportion to their distance. Such a pattern of innovation arises when there are no abrupt splits and periods of isolation, but rather an ongoing situation of contact and mutual intelligibility.

The methodology which I will use to represent my findings is Historical Glottometry (HG), a framework designed specifically to handle linkages and dialect chains and overcome some of the shortcomings of the tree model. The details of the calculations which go into the glottometric analysis are laid out in section 3.6.2. The results can be represented in the form of a glottometric diagram, as shown in Figure 6.

3.6.1 Evidence for different groupings

I will review different groupings which receive support from exclusively shared innovations. None of these receive the overwhelming support that PCF receives, and each proposed innovation-defined subgroup is contradicted by other innovations, which target different subsets of languages. This reflects the fact that the internal structure of CF is non-tree-like, because it did not differentiate through the process of splitting and isolation. Instead, PCF was (and continues to be) spoken in a nearly unbroken stretch of central Flores, with vigorous contact between neighboring groups and numerous cultural similarities.

The basic units which are being compared in this analysis and their abbreviations are as follows: Lio (LIO), Ende (END), Eastern Nga'o - Tendambepa (E:NGO), Western Nga'o - Watumite (W:NGO), Kéo - Udiworowatu (KEO), Eastern Nage - Nangamboa (E:NAG), Central Nage - Boawae (C:NAG), Eastern Ngadha - Takatunga (E:NGD), Central Ngadha - So'a (C:NGD), Western Ngadha - Bajawa (W:NGD), and Rongga (RON).

When an abbreviation such as NGD is used without further geographical specification, it includes all three dialects (covering E:NGD, C:NGD, W:NGD). The same applies to NAG (covering E:NAG and C:NAG) and NGO (covering E:NGO and W:NGO).

I provide a brief summary of the quantity and type of evidence in favor of each of the 15 subgroupings which follow. The abbreviations used as the following: LR (lexical replacement), ISC (irregular sound change), RSC (regular sound change), SI (syntactic innovation), and SS (semantic shift).

Summary of evidence for internal subgroups:
1) LIO, END, NGO, KEO, NAG, NGD: 5 innovations (4 LR, 1 ISC)
2) LIO, END, NGO, KEO, NAG: 4 innovations (2 LR, 2 ISC)
3) LIO, END: 8 innovations (7 LR, 1 ISC)
4) END, NGO, KEO, NAG, NGD, RON: 5 innovations (1 LR, 1 ISC, 2 RSC, 1 SI)
5) END, NGO, KEO, E:NAG: 2 innovations (2 RSC)
6) NGO, KEO, NAG, NGD: 1 innovation (1 ISC)
7) NGO, KEO, NAG, E:NGD: 2 innovations (1 LR, 1 SS)
8) NGO, KEO, E:NAG: 3 innovations (1 LR, 2 RSC)
9) NGO, NAG: 2 innovations (2 LR)
10) NGO, E:NAG: 5 innovations (3 LR, 1 ISC, 1 RSC)
11) NGO: 7 innovations (7 LR)
12) KEO, E:NAG: 1 innovation (1 RSC)
13) C:NAG, NGD, RON: 2 innovations (1 ISC, 1 SS)
14) NGD: 3 innovations (2 LR, 1 ISC)
15) W:NGD, C:NGD: 2 innovations (1 LR, 1 ISC)

LIO, END, NGO, NAG, KEO, NGD There are 5 innovations which affect all CF languages except Rongga (i.e., Lio, Ende, Nga'o, Nage, Kéo and Ngadha). These include 4 lexical replacements and one irregular sound change.

The 4 lexical replacements are as follows. LR1) PMP *hawak 'waist' is replaced by a form reflecting *eŋge instead. LR2) PMP *qilih or *bulud 'mountain' is replaced by a form reflecting *keli instead. LR3) PMP *dataR 'flat, level' is replaced by a form reflecting *ndena instead. LR4) PMP *dukut, *zukut 'grass' is replaced by a form reflecting *kuru instead.

The irregular sound change (ISC1) is an instance of metathesis in the form for 'wing'. The original PMP forms *qelad and *panid are replaced at some higher level by a form reflecting *ləbe, which occurs not only in the CF languages but also in Palu'e, Manggarai, Rembong and Palu'e. This form *ləbe (which is also innovative, but is not the piece of evidence under discussion here) then undergoes irregular metathesis to reflect a form *bəle in all CF languages except Rongga.

Table 63: Innovations affecting LIO, END, NGO, KEO, NAG, NGD

	LR1	LR2	LR3	LR4	ISC1
PMP	*hawak	*qilih, *bulud	*dataR	*dukut	*panid, *qelad
PCF	-	*wolo	-	-	*ləbe
LIO	eŋge	keli	ndena	kuru	bəle
END	ʔeŋge	keɹi	ndena	kuru	mbəɹe
E:NGO	eŋge	keɮi	ndena	kuyu	mbəɮe
W:NGO	e̥ge	keɮi	ndena	kuyu	b̥əɮe
KEO	ʔeŋge	kedi	(baβi)	kulu	-
E:NAG	eŋge	kedi	ndena	kulu	mbəde
C:NAG	ege	-	-	-	bəle
E:NGD	-	-	-	kuru	(beɣa)
W:NGD	ege	keli	dena	kuru	bəle
C:NGD	ege	keli	dena	(witu)	bəle
RON	reʔa	wolo	dʒere	kəri	ləmbe
GLOSS	waist	mountain	flat	grass	wing

LIO, END, NGO, KEO, NAG There are 4 innovations which affect all CF languages except Rongga and the Ngadha dialects (i.e., Lio, Ende, Nga'o, Nage and Kéo). These include 2 lexical replacements and 2 irregular sound changes.

The 2 lexical replacements are as follows. LR5) PMP *deŋeR 'to hear, to listen' is replaced by a form reflecting *lele instead. LR6) PMP *qaluR 'flow, current, river' is replaced by a form reflecting *lowo instead.

The 2 irregular sound changes occur in the words for 'left' and 'right' and are potentially connected, as they both involve accretion of an initial *ŋg- to the forms in question. ISC2) PMP *wanan 'right' yields forms reflecting *ŋgana in the affected languages. ISC3) PMP *ka-wiRi 'left' is replaced by a form reflecting *leu at a higher level, including all CF languages, Manggarai and some languages of Sumba. This then receives an unexpected initial *ŋg to yield forms reflecting *ŋgeu in the affected languages.

Table 64: Innovations affecting LIO, END, NGO, KEO, NAG

	LR5	LR6	ISC2	ISC3
PMP	*deŋeR	*qaluR	*wanan	*ka-wiRi
PCF	*dəŋe	*alo	*wana	*leu
LIO	lele	lowo	ŋgana	ŋgeu
END	ɹeɹe	ɹowo	ŋgana	ŋgeu
E:NGO	ʁeʁe	ʁowo	ŋgeto	ŋgeu
W:NGO	ʁeʁe	ʁowo	(b̥eʔo)	g̥eu
KEO	dede	dowo	-	(mbiri)
E:NAG	dede	dowo	ŋgana	ŋgeu
C:NAG	rele	lowo	gəna	geu
E:NGD	heti	-	wana	eu
W:NGD	zəŋe	alo	wana	leu
C:NGD	zəŋe	ləŋo	wana	leu
RON	ɹəŋe	alo	wana	leu
GLOSS	to hear	river	right	left

LIO, END There are 8 innovations which target the neighboring pair of languages Lio and Ende. These include 7 lexical replacements and one irregular sound change.

The 7 lexical replacements are as follows. LR7) The PMP form *bayawak 'monitor lizard' is replaced by a form reflecting *dəgi in Lio and Ende. This appears to have replaced another form *(w)əti which is attested in other CF languages, Palu'e, Manggarai, and Sika. LR8) The PMP pronoun *si ida 'they; 3PL' is replaced by forms reflecting *əbe instead. LR9) The form for 'to bite', which reflect *kiki in other CF languages and Sika, are replaced by innovative forms reflecting *toki instead. LR10) PMP *diRi 'to stand' is replaced by various forms throughout CF (reflecting *deri in Nga'o, Nage and Kéo, *duge in Ngadha) but shows the innovative form *dari in Ende and Lio. This may also be considered an irregular sound change, depending if this form is cognate with PMP *diRi and Nga'o, Nage and Kéo forms reflecting *deri. LR11) The basic locative preposition 'in, at' is /ləka/ in Lio and /ndəka/ in Ende (Aoki and Nakagawa 1993), but McDonnell (2009) also reports an Ende form /ɹəka/ which corresponds regularly to the Lio form. Thus, an

innovative locative preposition (ending in /-əka/, perhaps /ləka/) is a common innovation of Lio and Ende. LR12) PMP forms *ila, *kita 'to see' (both of which have cognates in other CF languages) are replaced by *kodɗo in Lio and Ende. LR13) PMP *qeneb, *qaneb 'door' is replaced with forms reflecting *pere instead.

The irregular sound change (ISC4) occurs in the 3SG pronoun. Both Lio and Ende reflect a form *kai, which is missing the medial consonant which occurs in Western Ngadha, Eastern Ngadha /gazi/ and Rongga /kaɹi/ (these reflect *kadi instead).

Table 65: Innovations affecting LIO, END

	LR7	LR8	LR9	LR10
PMP	-	*si ida	*a(ŋ)kit	*diRi
PCF	*wəti	*sida	*kiki	*deri?
LIO	dəgi	əbe	toki	dari
END	dəgi	ʔəbe	toki	dari
E:NGO	əti	siya, imu koʔo	kiki	ndeyi
W:NGO	əti	siɹa, imu koʔo	kiki	ndeyi
KEO	-	sira, ʔimu koʔo	kiki	ndeli
E:NAG	əti	siɹa	kiki	ndeli
C:NAG	ɣoʔa	demu	kiki	dei
E:NGD	-	hoga	kiki	duge
W:NGD	ɣora	demu	kiki	duge
C:NGD	ɣora	demu	kiki	duge
RON	ɠora	siɹa	kiki	ndawi
GLOSS	monitor lizard	3pl	to bite	to stand

Table 66: Innovations affecting LIO, END (cont.)

	LR11	LR12	LR13	ISC4
PMP	*i	*ila	*qeneb, *qaneb	*ia
PCF	*one	*ila	*wəwa	*ka(d,dʒ)i
LIO	ləka	koɗo	pere	kai
END	ndəka, ɹəka	koɗo	pere	kai
E:NGO	one	meno	wəsa	imu
W:NGO	one	meɹo	wəsa	imu
KEO	ena	ʔila	wəsa	ʔimu
E:NAG	əna	meno	wəsa	imu
C:NAG	-	gula	wəsa	imu
E:NGD	one	ŋəɗo	-	imu
W:NGD	ləwa	ɣila	wəsa	gazi
C:NGD	ləwa	ɣila	pintu	gazi
RON	one	tei	wəwa	kaɹi
GLOSS	in, at	to see	door	3sg

END, NGO, KEO, NAG, NGD, RON There are 5 innovations which affect all CF languages except Lio (i.e., Ende, Nga'o, Nage, Kéo, Ngadha and Rongga). These include 2 regular sound changes, one lexical replacement, one irregular sound change and one syntactic innovation.

The first regular sound change (RSC1) is a merger of PCF *d, *dʒ ⇒ dʒ in all CF languages except Lio (see section 3.5.1). Lio retains distinct reflexes of these two phonemes (PCF *d ⇒ /r/ while PCF *dʒ ⇒ /dʒ/).

The second regular sound change (RSC2) is a different outcome of PCF *bl- clusters in Lio with regards to the other CF languages (see section 3.5.1. In all CF languages except Lio, PCF *bl- ⇒ *l- (with deletion of the initial element), while in Lio the initial /b-/ is retained instead.

The lexical replacement (LR14) occurs in the 1SG pronoun, which reflects inherited PMP *aku in Lio but is replaced by an innovative form reflecting either *yaʔo or *ŋaʔo in all other CF languages.

The syntactic innovation (SI1) which has taken place in all CF languages apart from Lio is the development of a genitive marker which reflects a form *koʔo. The only strategy for marking nominal possession in Lio is juxtaposition of the possessum and possessor (in that order), whereas all other CF languages have an additional option to place the genitive marker between the two elements of the possessive construction.

The irregular sound change (ISC5) affects the comitative and instrumental marker, which can be reconstructed as PCF *noʔo. In all CF languages except Lio, this has been

replaced or supplemented by an innovative form reflecting *neʔe, with an irregular shift in both vowels. The fact that *neʔe is an innovation rather than a retention is supported by the attestation of the form /noʔo/ in Palu'e.

Table 67: Innovations affecting END, NGO, KEO, NAG, NGD, RON

	RSC1		RSC2	LR14	ISC5
PMP	*quzan	*ma-qudip	*lawaʔ	*aku	*agu
PCF	*udʒa	*mudi	*blewa	*aku	*noʔo
LIO	udʒa	muri	bewa	aku	noʔo
END	ʔura	muri	ɹewa	dʒaʔo	neʔe
E:NGO	uya	muyi	ʝewa	ŋaʔo	neʔe
W:NGO	uɹa	muɹi	ʝewa	ŋaʔo	neʔe
KEO	ʔura	muri	dewa	ŋaʔo, dʒaʔo	neʔe
E:NAG	uɹa	muɹi	dewa	ŋaʔo	neʔe
C:NAG	uza	muzi	lewa	ŋaʔo	-
E:NGD	ura	muri	ewa	ŋaʔo	neʔe
W:NGD	uza	muzi	lewa	ŋaʔo, dʒaʔo	neʔe
C:NGD	ʔuza	muzi	lewa	zaʔo, ŋaʔo	neʔe
RON	nuɹa	muɹi	lewa	dʒaʔo	neʔe
GLOSS	rain	to live	long	1sg	and

END, NGO, KEO, E:NAG There are 2 regular sound changes which target Ende, both varieties of Nga'o, Eastern Nage and Kéo (as spoken in Udiworowatu, but not necessarily all Kéo dialects).

The first regular sound change (RSC3) is the general prenasalization of all PCF *b ⇒ /mb/ and all PCF *g ⇒ /ŋg/ in the affected dialects (see section 3.5.1. There is a later change in Western Nga'o alone which partially undoes this change, yielding a final result of /b̥, g̥/ instead, but this follows the earlier prenasalization.

The second regular sound change (RSC4) is the shift of *z ⇒ /ɹ/ which took place at some point after PCF *d, dʒ had already shifted to an intermediate *z (see section 3.5.1). In Eastern Nage and Western Nga'o, the final result of this remains /ɹ/, but in Eastern Nga'o it shifts further to /y/ and in Kéo it shifts further to /r/. A similar change affects some CF languages further west, such as Rongga (showing /ɹ/) and Eastern Ngadha (showing /r/), but these are geographically separated from the other cluster of languages and probably represent a separate development. Although this development appears unusual at first, it recurs over and over in the Austronesian languages of Flores, targeting many languages outside CF such as Palu'e, Sika, most varieties of Lamaholot, and Alorese. In view of this, it is reasonable to believe it has arisen twice within CF in two geographically distinct areas.

NGO, KEO, NAG, NGD One irregular sound change (ISC6) affects Nga'o, Nage, Kéo and Ngadha. This is the accretion of an initial /ŋ-/ on the 1SG pronoun. The inherited PMP form *aku (preserved only in Lio) is replaced first with *yaʔo, then the group of languages in question innovated a new form, reflecting *ŋaʔo. In the affected dialects, the forms reflecting *yaʔo and *ŋaʔo may exist side-by-side, but languages which reflect *yaʔo never have the two forms alongside each other, which is evidence that *ŋaʔo is a later innovation.

Table 68: Innovations affecting NGO, KEO, NAG, NGD

	ISC6
PMP	*aku
PCF	*aku
LIO	aku
END	dʒaʔo
E:NGO	ŋaʔo
W:NGO	ŋaʔo
KEO	ŋaʔo, dʒaʔo
E:NAG	ŋaʔo
C:NAG	ŋaʔo
E:NGD	ŋaʔo
W:NGD	ŋaʔo, dʒaʔo
C:NGD	zaʔo, ŋaʔo
RON	dʒaʔo
GLOSS	1sg

NGO, KEO, NAG, E:NGD There are 2 innovations which affect Nga'o, Nage, Kéo and Eastern Ngadha, and make Eastern Ngadha appear to be transitional between the central CF languages and the other Ngadha dialects. One is a semantic shift and one is a lexical replacement.

The semantic shift (SS1) involves the development of an innovative 3SG pronoun from an earlier form *imu meaning 'friend, companion'. It retains its original meaning in Ende, Lio and other Ngadha dialects, but has grammaticalized into the regular 3SG pronoun in the languages in question.

The lexical replacement (LR15) involves the replacement of the original PCF negator *iwa with an innovative form *mona.

Table 69: Innovations affecting NGO, KEO, NAG, E:NGD

	SS1	LR15
PMP	*ia	*bak
PCF	*ka(d,dʒ)i	*iwa
LIO	kai	iwa
END	kai	ʔiwa
E:NGO	imu	mona
W:NGO	imu	mona
KEO	ʔimu	mona
E:NAG	imu	mona
C:NAG	imu	mona
E:NGD	imu	mona
W:NGD	gazi	ɓaʔi
C:NGD	gazi	ɓaʔi
RON	kaɹi	mbiwa
GLOSS	3sg	no; not

NGO, KEO, E:NAG There are 3 innovations which target Nga'o, Kéo and Eastern Nage. These include 2 regular sound changes and one lexical replacement.

The first regular sound change (RSC5) is a shift from PCF *l ⇒ *ɮ, a phoneme which is retained only in Nga'o (see section 3.5.1. In Kéo and Eastern Nage, this is then eliminated by a later change of *ɮ ⇒ /d/.

The second regular sound change (RSC6) is a shift of PCF *r ⇒ *l, forming a change chain with the previous sound change (see section 3.5.1). However, in the two Nga'o dialects, all instances of /l/ are then eliminated by a recent change of *l ⇒ /y/ (which must have occurred after PCF *l ⇒ /ɮ/ in Nga'o).

The lexical replacement (LR16) affects the word for 'to come'. While all CF languages retain a reflex of PMP *um-ai, reflected as PCF *mai, the languages in question have also innovated a form reflecting *lemba as well.

Table 70: Innovations affecting NGO, KEO, E:NAG

	RSC5	RSC6	LR16
PMP	*bulan	*ma-iRaq	*um-a(R)i
PCF	*wula	*toro	*mai
LIO	wula	toro	mai
END	wuɹa	toro	mai
E:NGO	wuɮa	toyo	mai, ɮəmba
W:NGO	wuɮa	toyo	mai, ɮə̰ba
KEO	wuda	tolo	mai, demba
E:NAG	wuda	tolo	demba
C:NAG	wula	to	taso
E:NGD	wua	toro	mai
W:NGD	wula	toro	mai
C:NGD	wula	toro	gaɗo
RON	wula	toro	mai
GLOSS	moon	red	to come

NGO, NAG There are 2 changes which target Nga'o and Nage alone, both of which are lexical replacements.

The first lexical replacement (LR17) is in the word for 'a bit'. PMP *dikit, which is retained in Ende /sa ɗiki/, is replaced by a form reflecting *sa yoʔo instead.

The second lexical replacement (LR18) is PMP *hiket 'to tie', which reflects a form *pəte instead in these languages.

Table 71: Innovations affecting NGO, NAG

	LR17	LR18
PMP	*dikit	*hiket
PCF	-	*ike
LIO	sa loʔo	rike
END	sa ɗiki	rike
E:NGO	sa dʒoʔo	pəte
W:NGO	sa tʃoʔo	pəte
KEO	ha goʔo	ʔike
E:NAG	sa tʃoʔo	pəte
C:NAG	sa tʃoʔo	pəte
E:NGD	-	pau
W:NGD	sa kəɗi	-
C:NGD	sa kəɗi	rati
RON	sa ito	ike
GLOSS	a few	to tie

NGO, E:NAG There are 5 innovations which affect Nga'o plus Eastern Nage. These include one regular sound change, 3 lexical replacements, and one irregular sound change.

The regular sound change (RSC7) concerns the outcome of initial PCF *kl- clusters (see section 3.5.1). In the affected languages, PCF *kl- ⇒ *l-, which is the outcome reflected in Eastern Nage. In the two Nga'o dialects, a later change of *l ⇒ /y/ eliminates the resulting instances of /l/.

The 3 lexical replacements are as follows. LR19) PMP *ma-panas 'hot' is replaced by a form reflecting *buli instead. LR20) PMP *kutaña 'to ask' is replaced by a form reflecting *wande instead. LR21) PMP *ma-tuqah 'old growth, primary forest' is replaced by a form reflecting *kləwo instead.

The irregular sound change (ISC7) occurs in PMP *paniki 'fruit bat', which shows an irregular development to a form reflecting *ŋiɣi in the affected languages.

Table 72: Innovations affecting NGO, E:NAG

	RSC7	LR19	LR20	LR21	ISC7
PMP	*taliŋa	*ma-panas	*kutaña	*ma-tuqah	*paniki
PCF	*kliŋa	*(m)bana	*tana	*nduʔa	*niki
LIO	kiŋa	pətu	ale	nduʔa	niki
END	iŋa	pətu	ʔaɹe	nduʔa	-
E:NGO	yiŋa	buɬi	wande	yəwo	ŋiɣi
W:NGO	yiŋa	buɬi	wande	yəwo	ŋiɣi
KEO	iŋa	pətu	ʔade	komba	nixi
E:NAG	liŋa	budi	wande	ləwo	ŋiɣi
C:NAG	hiŋa	bana	-	witu	mete
E:NGD	hiŋa	sume	-	kisa kadʒu	-
W:NGD	hiŋa	bana	tana	kala	mete
C:NGD	hiŋa	bana	tana	witu	mete
RON	hiŋa	mbana	tana	kala	niɠi
GLOSS	ear	hot	to ask	forest	fruit bat

NGO There are 7 lexical replacements which affect Eastern and Western Nga'o exclusively, and provide evidence for the unity of Nga'o as a subgroup.

The 7 lexical replacements are as follows. LR22) PMP *haŋin 'wind' is replaced by a form reflecting *widʒo instead. LR23) PMP *ma-qasin 'salty' is replaced by a form reflecting *kande instead. LR24) PMP *hajek 'to smell, to sniff' - which is already replaced by a form reflecting *ŋudu in the CF languages generally - is replaced by a further innovative form reflecting *pidʒu. LR25) PMP *baliw 'to return' is replaced by a form reflecting *koso instead. LR26) PMP *buka 'to open' is replaced by a form reflecting *ɓaŋge instead. LR27) PMP *apuŋ 'to float' - which is already replaced by a form reflecting *mbawa in the CF languages generally - is replaced by a further innovative form reflecting *leya. LR28) PMP *luheq 'teardrops' is replaced by a form reflecting *ae ŋa instead.

Table 73: Innovations affecting NGO

	LR22	LR23	LR24	LR25
PMP	*haŋin	*ma-qasin	*hajek	*baliw
PCF	*aŋi	*məsi	*ŋuru	*walo
LIO	aŋi	məsi	ŋuru	walo
END	ʔaŋi	məsi	ŋuru	waɹo
E:NGO	wiyo	kande	piyu	koso
W:NGO	wiɹo	kande	piɹu	koso
KEO	ʔaŋi	məsi	səŋu	wado
E:NAG	aŋi	məsi	ŋuɹu	roka
C:NAG	wa	-	səŋu	-
E:NGD	wara	-	ŋido	-
W:NGD	wara	məsi	səŋu	walo
C:NGD	wara	məsi	səŋu	kusu
RON	wara	məsi	ŋuɹu	walo
GLOSS	wind	salty	to smell	to return

Table 74: Innovations affecting NGO (cont.)

	LR26	LR27	LR28
PMP	*buka	*apuŋ	*luheq
PCF	*kai	*mbawa	*wae lu
LIO	kai	mbawa	ae lu
END	kai	mbawa	ʔae ɹu
E:NGO	ɓange	ɠeya	ae ŋa
W:NGO	ɓage̊	ɠeya	ae ŋa
KEO	kai	-	du
E:NAG	buka	mbawa	ae du
C:NAG	-	-	-
E:NGD	kaʔi	-	-
W:NGD	-	ɓawa	wae mata
C:NGD	buka	ɓawa	wae mata
RON	kai	mbawa	luʔu
GLOSS	to open	to float	tear

KEO, E:NAG There is one regular sound change (RSC8), whereby PCF *l ⇒ /d/ in Eastern Nage and Kéo as spoken in Udiworowatu (see section 3.5.1). All instances of PCF *l are reflected as /d/ in these two language varieties, passing through an intermediate *ɮ stage (which is still retained in Nga'o). Thus, the regular sound change PCF *l ⇒ *ɮ ⇒ /d/ is a shared innovation between Eastern Nage and Kéo. There are Kéo dialects which do not undergo this change, but there was too little data available to include them in this analysis.

Table 75: Innovations affecting KEO, E:NAG

	RSC8
PMP	*bulan
PCF	*wula
LIO	wula
END	wuɹa
E:NGO	wuɮa
W:NGO	wuɮa
KEO	wuda
E:NAG	wuda
C:NAG	wula
E:NGD	wua
W:NGD	wula
C:NGD	wula
RON	wula
GLOSS	moon

C:NAG, NGD, RON There are 2 innovations which affect Central Nage, Ngadha and Rongga exclusively. These include one irregular sound change and one semantic shift.

The irregular sound change (ISC8) targets the word for 'to walk', which reflects PMP *ma-panaw in most CF languages. In the Central Nage, Ngadha and Rongga, the most basic word for 'to walk' reflects a form *laʔa instead, which is most likely derived from PMP *lakaw but with an unexpected shift of *k ⇒ ʔ.

The semantic shift (SS2) occurs in the word for 'wind'. Most CF languages retain a reflex of PMP *haŋin 'wind', while the PMP *habaRat 'west wind' has taken on the basic meaning of 'wind' in the affected languages (although it may co-exist with reflexes of *haŋin as well).

Table 76: Innovations affecting C:NAG, NGD, RON

	LR29	SS2
PMP	*panaw	*haŋin
PCF	*mbana	*aŋi
LIO	mbana	aŋi
END	mbana	ʔaŋi
E:NGO	mbana	wiyo
W:NGO	ɓ̥ana	wiɪo
KEO	mbana	ʔaŋi
E:NAG	mbana	aŋi
C:NAG	laʔa	wa
E:NGD	aʔa	wara
W:NGD	laʔa	wara
C:NGD	laʔa	wara
RON	laʔa	wara
GLOSS	to walk	wind

NGD There are 3 innovations which target the three Ngadha varieties exclusively and provide evidence for the unity of Ngadha as a subgroup. These include 2 lexical replacements and one irregular sound change.

The 2 lexical replacements are as follows. LR29) PMP *diRi 'to stand' is replaced by a form reflecting *duge instead. LR30) PMP *balik 'to reverse, to turn' is replaced by a form reflecting *ɣeo instead.

The irregular sound change (ISC9) occurs in the word for 'dull'. The PCF *boŋo (with no clear PMP antecedent) irregularly shows an initial implosive in the Ngadha varieties, reflecting a form *ɓoŋo (although Eastern Ngadha has /ɓuŋu/, with no clear explanation).

Table 77: Innovations affecting NGD

	LR32	LR31	ISC8
PMP	*diRi	*balik	*dumpul
PCF	*deri?	*bale?	*boŋo
LIO	dari	beʔo	boŋo
END	dari	nənde	boŋo
E:NGO	ndeyi	riʔu	boŋo
W:NGO	ndeyi	riʔu	boŋo
KEO	ndeli	bale	-
E:NAG	ndeli	belo	boŋo
C:NAG	dei	liʔu	boŋo
E:NGD	duge	ɣeo	ɓuŋu
W:NGD	duge	ɣeo	ɓoŋo
C:NGD	duge	ɣeo	ɓoŋo
RON	ndawi	ɓale	ɗumbu
GLOSS	to stand	to turn	blunt

W:NGD, C:NGD There are 2 innovations which target only Western and Central Ngadha. One of these is a lexical replacement and one is an irregular sound change.

The lexical replacement (LR31) occurs in the negator morpheme. The PCF negator *iwa is replaced with a form reflecting *ɓaʔi in these two dialects of Ngadha, whereas the usual negator in Eastern Ngadha is /mona/ (as in Nage, Kéo and Nga'o).

The irregular sound change (ISC10) is the accretion of an initial /ɣ-/ on the word for 'to see'. PMP *ila has been replaced by a form reflecting *ɣila in the affected dialects.

Table 78: Innovations affecting W:NGD, C:NGD

	LR32	ISC9
PMP	*bak	*ila
PCF	*iwa	*tei
LIO	iwa	koɗo
END	ʔiwa	koɗo
E:NGO	mona	meno
W:NGO	mona	meɹo
KEO	mona	ʔila
E:NAG	mona	meno
C:NAG	mona	gula
E:NGD	mona	ŋəɗo
W:NGD	ɓaʔi	ɣila
C:NGD	ɓaʔi	ɣila
RON	mbiwa	tei
GLOSS	no; not	to see

3.6.2 Methodology: Historical Glottometry

Historical Glottometry allows one to quantify the strength of evidence in support of various subgroups and then create a visualization of the data in the form of a glottometric diagram or glottometric map. The full details on how to carry out the necessary calculations can be found in Kalyan and François (2018), and will be summarized here.

There are three basic quantities which enter into the final calculation for creating the visualizations. For any subset of languages, the following three quantities can be counted: 1) e = the number of exclusively shared innovations, 2) p = the number of supporting innovations, and 3) q = the number of conflicting innovations.

The number of exclusively shared innovations (e) is simply equal to the number of innovations which target all and only the subset of languages under consideration. The number of supporting innovations (p) is equal to the number of innovations which target all the languages under consideration, but not necessarily in an exclusive way: supporting innovations but may also target languages which are outside the subset under consideration. By definition, the number of supporting innovations (p) is always equal to or greater than the number of exclusively shared innovations (e). The number of conflicting innovations (q) is equal to the number of innovations which target some but not all languages in the subset under consideration, as well as at least one language outside the set. A conflicting innovation must target at least one language outside the subset under consideration, and therefore define an intersecting set of isoglosses; if the two innovations target nested sets of languages, then they are not conflicting innovations.

The innovative part of Historical Glottometry is that it allows conflicting innovations to play a role in quantifying the strength of each subgroup. Under the tree model, intersecting isoglosses present an insurmountable problem. Only nested sets of innovations can be relied upon to construct a tree diagram. One of the two conflicting innovations will have to be thrown out for the purposes of constructing a tree diagram (perhaps by being attributed to contact effects and horizontal transmission). In a linkage scenario, where all transmission can be construed as horizontal transmission, too much information ends up being thrown out to ensure that the resulting diagram is a realistic and comprehensive depiction of the linguistic history under question.

From these three basic quantities (e, p, q) the following two values can be calculated: cohesiveness (κ) and subgroupiness (σ).

Cohesiveness (κ) is defined as p / (p + q): the number of supporting innovations divided by the number of total relevant innovations (i.e., all supporting and all conflicting innovations). Cohesiveness is effectively a measure of how often the subset of languages behave similarly with respect to any given innovation. Cohesiveness will always be between 0 and 1, as it measures the proportion of instances in which the languages behave the same. Since conflicting innovations are not recognized as valid in the tree model, all subgroups in a tree diagram would receive the maximum cohesiveness score of 1.

Subgroupiness (σ) is defined as e * κ: the number of exclusively shared innovations weighted by the cohesiveness. This can be thought of as roughly 'quantity weighted by quality'. A subgroup with very high cohesiveness (whose languages almost always behave the same with respect to any given innovation) will receive a high subgroupiness score even with a modest number of exclusively shared innovations, because these are quite convincing. On the other hand, a less cohesive subgroup (whose members engage in shared innovations with other languages regularly) would require a larger collection of exclusively shared innovations to achieve the same level of subgroupiness.

In order to construct a glottometric diagram from a set of calculated cohesiveness and subgroupiness scores, the following conventions apply. The languages are laid out as a series of points, and each subgroup (defined by a bundle of isoglosses) is enclosed by a line of variable thickness and color. Thickness correlates directly with subgroupiness, allowing the most strongly supported subgroupings to become clear from the sets of languages enclosed by the largest lines. This was implemented in Microsoft Word by simply setting the thickness of the line equal to subrgoupiness; so a group with subgroupiness 8.75 will simply have a line thickness of 8.75.

For representing cohesiveness, previous iterations of Historical Glottometry have proposed a 'redness' or 'red-black' scale, whereby subgroups with cohesiveness of 1 receive pure red borders, and less cohesive subgroups have a darker line around them, approaching black as cohesiveness approaches 0. However, there have been problems with the visual salience of various shades of red, so I have here adopted a 'red-blue' convention. Subgroups with a cohesiveness of 1 still receive a pure red border, but the other end of the continuum is now blue instead of black. A cohesiveness score of 0.5 will result in a purple border

with equal parts red and blue, while the border approaches pure blue as the cohesiveness approaches 0. This was implemented by manipulating the RBG color values of the lines directly in Microsoft Word: a subgroup with perfect cohesiveness would have 255 for red, and 0 for green and blue. When cohesiveness is 0.5, the result would be a value of 128 for red, 128 for blue, and 0 for green. As cohesiveness approaches 0, the blue value approaches 255, and red approaches 0 (with green still at 0).

Finally, the transparency on all lines has been set to 25% in order to allow the reader to follow the lines more easily where they intersect.

3.6.3 Results of Glottometric analysis

The values of cohesiveness (κ) and subgroupiness (σ), as well as the quantities required to calculate them (e = number of exclusively shared innovations, p = number of supporting innovations, q = number of conflicting innovations), are reported in Table 79.

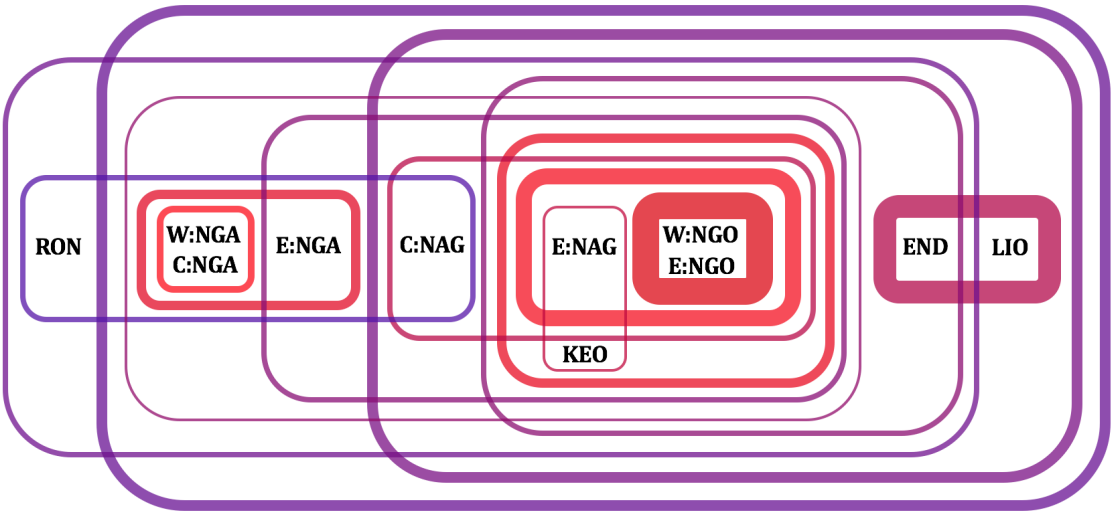
Table 79: Results of glottometric analysis of subgroups listed in order of decreasing subgroupiness (σ)

	e	p	q	κ	σ
NGO	7	36	0	1	7
LIO, END	8	17	7	0.71	5.68
NGO, E:NAG	5	29	1	0.97	4.85
NGO, E:NAG, KEO	3	22	2	0.92	2.76
NGD	3	16	2	0.89	2.67
LIO, END, NGO, NAG, KEO, NGD	5	5	7	0.42	2.1
W:NGD, C:NGD	2	18	0	1	2
LIO, END, NGO, NAG, KEO	4	9	10	0.47	1.88
NGO, NAG	2	19	7	0.73	1.46
END, NGO, NAG, KEO, NGD, RON	5	5	17	0.23	1.15
END, NGO, E:NAG, KEO	2	16	13	0.55	1.1
NGO, NAG, KEO, E:NGD	2	13	11	0.54	1.08
E:NAG, KEO	1	23	7	0.77	0.77
C:NAG, NGD, RON	2	7	14	0.33	0.66
NGO, NAG, KEO, NGD	1	11	8	0.58	0.58

The results of the glottometric analysis, plotted as a glottometric diagram, are shown in Figure 7. A number of observations can be made about this diagram.

1) This glottometric diagram is characterized by a large number of intersecting lines, reflecting a pattern of innovation which is highly incompatible with the tree model of language differentiation. A tree-like language group would yield a glottometric diagram in which all groupings are nested within one another. Instead, the CF languages behave like

Figure 7: Glottometric diagram of the CF language group



a linkage: each language undergoes innovation with its neighbors on either side in a way which cannot be accurately represented in a tree diagram. This results in the chain-like pattern observed in the glottometric diagram, which is typical of linkages generally. Any tree diagram which would attempt to represent the relations between CF languages would lose most of the information contained here in the process of creating a branching structure.

2) The overall number of lines which surround any given language variety can be used a rough proxy for how innovative each variety is. Innovative centers, from which changes spread to neighboring languages, can be identified easily on this diagram by language varieties which are encircled by many lines and share connections with many other varieties. In the case of CF, it is immediately apparent that the central dialects (Nga'o, Eastern Nage and Kéo), are the most innovative of the CF languages. Indeed, this is where the most extreme changes have taken place, and where changes spread most easily across recognized language boundaries. Conversely, the most conservative languages are those with the fewest lines surrounding them: in this case, Rongga and Lio are noticeably more conservative than the other languages. This pattern fits well with the wave model of language differentiation, and further support the idea that CF has differentiated in a non-tree-like manner. Many innovations tend to overlap in the central area, but often do not propagate all the way to the edges of the CF zone.

3) The pairs of languages which are not separated by any lines (Western and Central Ngadha, Western and Eastern Nga'o) do not differ in behavior with respect to any innovation included in my set of innovations. It may still be possible to identify isoglosses which

separate them with a bit more searching, but these pairs of language are indeed extremely similar. In the case of Western and Central Ngadha, these correspond to the Ngadha spoken in Bajawa and in So'a respectively. So'a has been considered a separate language by some sources (Ethnologue, Glottolog), but this provides quantitative evidence to back up my impression that So'a is far too similar to other Ngadha dialects to be considered its own language.

4) Dialects which are transitional between two language zones appear quite clearly in the diagram. For instance, while Western and Central Ngadha are very similar, and are both among the more conservative CF languages, Eastern Ngadha undergoes a number of innovations in common with its eastern neighbors Nage, Kéo and Nga'o. In this sense it is a transitional dialect between the two areas, although the unity of Ngadha as a single language is also well supported.

5) On the other hand, the linguistic entity known as Nage appear quite fragmented in this diagram: Eastern Nage clearly innovates with Kéo and Nga'o much more often than with Central Nage. This also fits well with my impression that the lines between Nage, Kéo and Nga'o are drawn based on identity rather than purely linguistic factors. It would be worth revisiting the idea that Nage forms a unitary language, and it may be more accurate to refer to an Eastern Nage-Kéo-Nga'o cluster on one hand, in contrast with a Central Nage cluster. Much more fieldwork is needed in the Nage-speaking area before this picture will become clearer. It will also be necessary to improve coverage of Kéo, as it is known to be very dialectally diverse but little fieldwork has been carried out so far.

6) Lio and Ende form a cluster which is supported by many individual innovations (hence the thickness of the line surrounding them) but they are not a particularly cohesive subgroup (hence the purplish color of the line, rather than deep red). Ende, the more innovative of the pair, has also participated in innovations with all the other CF languages separately from Lio, bringing the cohesiveness of the Lio-Ende subgroup down. I interpret this to mean that Ende society is more tightly integrated, and has more cultural exchange, with the other CF-speaking groups. In contrast, Lio occupies a more peripheral position and has correspondingly less exchange with the innovative languages of the central CF area.

3.7 Conclusions

The historical scenario which is supported for the CF languages is the following. An ancestral language, PCF, spread rather quickly over an area of Central Flores corresponding roughly to the current spread of CF languages. This language was originally spoken with minor dialectal variations over the whole area, but mutual intelligibility was not impeded and there was ongoing cultural exchange throughout the area. Thus, a newly arisen innovation (for example, the replacement of the PCF 1SG *aku with the forms *yaʔo and *ŋaʔo) was able to spread relatively easily over a large area, intersecting the isoglosses formed by earlier minor innovations. The wave-like spread of innovation also explains why

peripheral languages are more conservative with respect to many innovations because they are on the edge of the CF world (hence Lio being the only CF language to preserve the original form *aku), while central dialects of Nage and Kéo are the most innovative because many waves of innovation have overlapped in their area. Over time, the accumulation of innovations and the development of more localized ethnic identities led to the formation of less porous linguistic boundaries and the gradual loss of mutual intelligibility. However, the earlier situation of mutual intelligibility and contact is preserved in the pattern of overlapping innovations, which are 'impossible' under a strictly tree-like view of language differentiation. This situation is exactly what Ross (1988) defines as a 'linkage', and the pattern of overlapping innovation observed in the modern CF languages is typical of linkages generally.

4 Appendices

4.1 Appendix A: Texts

4.1.1 Frog Story

The following is a transcript of the Frog Story as told by ML.

(15) *manusia lako no'o leko*
human dog COM frog
'A human, a dog and a frog'

(16) *ebé rapa uli*
3PL RECIP company
'They are friends with each other'

(17) *na ebé bego*
DEM 3PL play
'They are playing'

(18) *manusia lako no'o leko rapa bego*
human dog COM frog RECIP play
'The human, the dog and the frog are playing together'

(19) *sawé ola kobé lako no'o manusia nai roké leka ténda roké*
finish NOM night dog COM human ascend sleep LOC platform sleep
'Then, at night, the dog and the human climb up to sleep in a bed'

(20) *leko leka oné stoflés*
frog LOC inside jar
'The frog is inside a jar'

(21) *leko wulo tau paru*
frog jump.out make run
'The frog jumps out in order to run'

(22) *geju leka stoflés*
come.out LOC jar
'[He] comes out of the jar'

(23) *kai paru*
3SG run
'He runs'

(24) *ebé lako no'o manusia roké ndaté*
3PL dog COM human sleep heavy
'They, the dog and the human, are sleeping soundly'

(25) *ebé bébo*
3PL not.know
'They do not know'

(26) *manusia no lako so gudu*
human COM dog so startle
'The human and the dog are startled'

(27) *ebé ngara kodho da ghalé wena ngéré na*
3PL look.around look towards down bottom like DEM
'They lift their heads [and] look down, like this'

(28) *ai ni stoflés penga sawé gha*
EXCLAM EXCLAM jar empty finish here
"Oh no! The jar is completely empty"

(29) *leko iwa do téi*
frog NEG PERF find
'The frog is no longer there

(30) *manusia kai ui*
human 3SG EXCLAM

'The human, he [yells:] 'Oy!"

(31) *ué lako leko iwa do téi ga*
EXCLAM dog frog NEG PERF find friend$_2$
"Hey, dog! The frog is no longer here, my friend!"

(32) *lako mulai gaé*
dog begin search
'The dog begins to search'

(33) *manusia paru*
human run
'The human runs'

(34) *lako paru gaé*
dog run search
'The dog runs [and] searches'

(35) *paru gaé leko*
run search frog
'[He] runs [and] searches for the frog'

(36) *leka emba leko na*
LOC which frog DEM
"Where is that frog?"

(37) *leko iwa téi*
frog NEG find
'[He] doesn't find the frog (or: The frog isn't there)'

(38) *lako ... manusia gui-gui lako dhomé*
dog ... human shout:ITER dog go.in
'The dog ... The human shouts [and] the dog goes in'

(39) *kolo lako maso leka stoflés*
head dog enter LOC jar
'The dog's head enters into the jar'

(40) *stoflés ... kolo lako dhomé daki lako kong-kong-kong ai*
jar ... head dog go.in touch$_2$ dog woof:ITER EXCLAM
'The jar ... The dog's head goes in [and] gets stuck, [and] the dog goes 'Woof woof woof', oh my!'

(41) *mo geju wola taré kolo talo*
FUT come.out return$_2$ pull$_2$ head cannot
'[He] wants to come back out, [but he] cannot pull his head'

(42) *mo ngendo kai ngendo talo do*
FUT move.back 3SG move.back cannot PERF
'[He] wants to move back, [but] he can no longer move back'

(43) *su'u daki no'o stoflés éo pusi leko*
carry.head touch$_2$ COM jar REL put.in frog
'The jar in which [they] put the frog is stuck on his head'

(44) *kai rasa wau*
3SG feel smell
'He smells [its] scent'

(45) *na kai pemilik leko ... lako na kai mbana*
DEM 3SG owner frog ... dog DEM 3SG go
'He, the frog's ... the dog's owner, he goes'

(46) *ai ni delu kau ngéré emba néa kau na*
EXCLAM EXCLAM friend$_4$ 2SG like which JUST.NOW 2SG DEM
"Oh no! My friend, what to do with you?"

(47) *stoflés na kau dhomé daki*
jar DEM 2SG go.in touch$_2$
'You went into this jar [and] got stuck'

(48) *kai goro*
3SG pull
'He pulls'

(49) *kai pati langa stoflés éo wau leko*
3SG give take.away jar REL smell frog
'He takes off the jar which smells of frog'

(50) *kai goro lako kai*
3SG pull dog 3SG
'He pulls his dog'

(51) *kai ka'o lako*
3SG cradle dog
'He cradles the dog'

(52) *delu kau na na*
friend$_4$ 2SG DEM DEM
"You are my friend"

(53) *aku até dhoa dema*
1SG liver sink truly
"I truly have compassion for you"

(54) *na ebé paru gaé paru gaé*
DEM 3SG run search run search
'They run around searching, run around searching'

(55) *leka pu'u kaju sa pu'u latu éro ani*
LOC trunk wood SG trunk.CL EXI bee bee$_2$

'In one tree there are bees'

(56) *éro ani no ...*
bee bee$_2$ COM ...
'Bees and ...'

(57) *ebé mera ghalé wena fau kai*
3PL sit down bottom shade 3SG
'They sit under its shade'

(58) *gaé lako ... leko na leka emba*
search dog ... frog DEM LOC which
'Searching for this dog ... frog, where is it?'

(59) *ngara da ghéta wawo*
lift.head towards up top
'They look upwards'

(60) *éro ani na mulai léla ebé léla*
bee bee$_2$ DEM begin fly 3PL fly
'The bees begin to fly, they fly'

(61) *mulai toki lako no manusia*
begin bite dog COM human
'[They] begin to sting the dog and the human'

(62) *lako no manusia ani toki*
dog COM human bee$_2$ bite
'The bees sting the dog and the human'

(63) *ebé hui paru*
3PL EXCLAM run
'They [yell:] 'Oy!' [and] run'

(64) *paru lako*
run dog
'The dog runs'

(65) *manusia boka*
human fall.over
'The human falls'

(66) *kai boka*
3SG fall.over
'He falls'

(67) *lako polu ou-ou-ou da ghéta tolo leka pu'u kaju*
dog bark woof$_2$:ITER towards up summit LOC trunk wood
'The dog barks 'Bow wow wow!' up towards the top of a tree'

(68) *fua ... pu'u ... apa ... nua fua na*
wasp ... tree ... what ... village wasp DEM
'Wasps ... the tree ... what? ... a wasp's nest'

(69) *fua paru ngai lélé lako polu*
wasp run because hear dog bark
'The wasps run because [they] hear the dog barking'

(70) *lako péko*
dog bark$_2$
'The dog barks'

(71) *kai paru*
3SG run
'He runs'

(72) *léla paru ngé béu-béu da ghéta wawo iwa toki*
fly run increase far:INTENS towards up top NEG bite

'[They] fly further and further downwards [but] do not bite [him]'

(73) *sawé na manusia ghéa*
finish DEM human there
'After that, the person there [says]'

(74) *e: lama sai aku nia bowo sawé*
EXCLAM quick IMP 1SG face swell finish
"Hey! Be quick! My face is completely swollen"

(75) *leko apa tau*
frog what make
'The frog [asks]: 'Why?"

(76) *lako nai da ghéta pu'u kaju*
dog ascend towards up trunk wood
'The dog climbs up the tree'

(77) *kai ngada da ghéta beru polu nua fua ghéta wawo ou-ou-ou*
3SG stand.across towards up then bark village wasp up top bark$_2$:ITER
da ghéa
towards there
'He stands across then barks at the wasp's nest up above 'Bow wow wow"

(78) *akirnya apa*
in.the.end what
'In the end what [happens]?'

(79) *nua fua o lako méko bai raka nua fua mesu du ghalé wena*
village wasp REL dog move too very village wasp fall reach down bottom
'The wasp's nest which the dog moved too much, the wasp's nest falls down'

(80) *fua nua kai mesu*
wasp village 3SG fall

'The wasp's nest falls'

(81) *ana éo ... lako polu ghéa tolo ké'é padahal nua fua dowa ghalé tana*
child cat ... dog bark there summit still besides village wasp PERF down ground
'The kitten ... dog is still barking up there even though the wasp's nest is already on the ground'

(82) *fua-fua kai léla*
wasp:PL 3SG fly
'The wasps fly'

(83) *mulai ebé mo toki lako*
begin 3PL FUT bite dog
'[They] begin to sting the dog'

(84) *tapi manusia no'o leko paru sawé*
but human COM frog run finish
'But the man and the frog have run away'

(85) *menga ... leko so béu do*
only ... frog so far PERF
'Only ... the frog is already far away'

(86) *kai no'o pu'u kaju na béu do*
3SG COM trunk wood DEM far PERF
'He is already far away from the tree'

(87) *kai mulai delu*
3SG begin friend
'He begins [to say]: 'Friend!"

(88) *mai sai kai gui lako*
come IMP 3SG shout dog
"Come!' he shouts to the dog'

(89) *tapi lako polu terus péko terus da ghéta tolo kaju*
but dog bark continue $bark_2$ continue towards up summit wood
'But the dog keeps on barking, keeps on barking towards the top of the tree'

(90) *fua toki kai*
wasp bite 3SG
'The wasps sting him'

(91) *na uli kai o manusia nai ghéta tolo pu'u kaju*
DEM company 3SG REL human ascend up summit trunk wood
'His friend, who is a human, climbs up to the top of the tree'

(92) *pu'u kaju lélé na*
trunk wood banyan DEM
'It is a banyan tree'

(93) *kodho pu'u ria na pu'u lélé*
look trunk large DEM trunk banyan
'[He] sees that big tree, the banyan tree'

(94) *kai nai neké ghéta pu'u kaju*
3SG ascend hide up trunk wood
'He climbs [and] hides up in the tree'

(95) *éi aku na neké leka na wé'é*
EXCLAM 1SG DEM hide LOC DEM just
"Hey! I am just hiding here"

(96) *nebu kai neké ghéta tolo kaju leka sénga kaju lélé fua ra'i*
time 3SG hide up summit wood LOC branch wood banyan wasp approach
'While he hides up in the top of the tree, in the branches of the banyan tree, the wasps come'

(97) *kola*
chase
'[They] chase [him]'

(98) *ebé na demi lélé wau ebé kola terus*
3PL DEM COND hear smell 3PL chase continue
'If they smell [his] scent, they will keep on chasing [him]'

(99) *kola terus*
chase continue
'[They] keep on chasing [him]'

(100) *leka emba manusia kai na*
LOC which human 3SG DEM
'Where is the human?'

(101) *kai mo paru*
3SG FUT run
'He is going to run'

(102) *ta'u ola toki*
afraid lest bite
'[He] is afraid lest [they] bite [him]'

(103) *kai mesu tolo kaju*
3SG fall summit wood
'He falls from the top of the tree'

(104) *tuka kénga*
belly lie.on.back
'[He] lies on his back'

(105) *mesu da ghalé wena*
fall towards down bottom

'[He] falls down'

(106) *oi latu fua toki aku*
EXCLAM EXI wasp bite 1SG
"Oy! There are wasps that stung me"

(107) *ngéré emba néa*
like which just.now
"What to do now?"

(108) *lako si iwa do latu*
dog also NEG PERF EXI
"The dog is no longer here"

(109) *ana éo si iwa latu*
child cat also NEG EXI
"The kitten is no longer here"

(110) *aku menga mésa na*
1SG only alone DEM
"I am all alone"

(111) *ngéré emba aku na*
like which 1SG DEM
"What am I to do?"

(112) *aku mo buat apa la'é na*
1SG FUT do what place DEM
"What will I do in this place?"

(113) *aku mo tau apa la'é na*
1SG FUT make what place DEM
"What will I do in this place?"

(114) *aku bébo do*
1SG not.know PERF
"I no longer know"

(115) *kai menga kénga*
3SG only lie.on.back
'He just lies on his back'

(116) *sawé na lako fua kola*
finish DEM dog wasp chase
'After that the wasps chase the dog'

(117) *fua kola terus lako paru*
wasp chase continue dog run
'The wasps keep chasing [and] the dog runs'

(118) *fua toki terus lako na*
wasp bite continue dog DEM
'The wasps keep stinging that dog'

(119) *sa weki tebo lako do latu fua toki mésa*
SG body $body_2$ dog PERF EXI wasp bite alone
'The dog's entire body is stung by wasps'

(120) *nebu lako paru fua kola manusia neké*
time dog run wasp chase human hide
'While the dog runs [and] the wasps chase [him], the human hides'

(121) *neké lé pu'u kaju ... leka watu na*
hide LOC_2 trunk wood ... LOC stone DEM
'[He] hides in a tree ... at the rock'

(122) *kai neké leka watu*
3SG hide LOC stone

'He hides at the rock'

(123) *latu jata ria sa éko mai ghéta m(a)i wawo*
EXI eagle large SG tail.CL come up come top
'There is an eagle [which] comes from above'

(124) *léla mo sao kai*
fly FUT grab 3SG
'[It] flies [and] wants to grab him'

(125) *téi éro ani*
find bee bee_2
'[It] finds the bees'

(126) *jata mai mo ta'o toki ka pesa éro ani*
eagle come FUT $gather_2$ bite eat eat.meat bee bee_2
'The eagle comes to round up and devour the bees '

(127) *ebé muso aé éro ani*
3PL sip water bee bee_2
'They sip the honey'

(128) *fua néa-néa mo tau ka ebé*
wasp just.now:ADV FUT make eat 3PL
'The wasps just now, [it] wants to eat them'

(129) *pesa ebé pesa*
eat.meat 3PL eat.meat
'Eat them, eat'

(130) *jata pesa fua*
eagle eat.meat wasp
'The eagle eats the wasps'

(131) *kai neké lé pu'u kaju*
3SG hide LOC_2 trunk wood
'He hides in a tree'

(132) *sawé jata tau ... ra'i po*
finish eagle make ... approach owl
'Then the eagle does ... An owl comes'

(133) *leka pu'u kaju latu po sa éko*
LOC trunk wood EXI owl SG tail.CL
'In the tree there is an owl'

(134) *po kodho téi ana ghéa*
owl look find child there
'The owl looks [and] sees the child there'

(135) *kai o tau apa kai na ghéta tolo kaju do na ghéta tolo watu*
3SG REL make what 3SG DEM up summit wood PERF DEM up summit stone
"What is he doing up in the top of that tree... on top of that rock?"

(136) *néa-néa kai ghalé wena na ghéta tolo watu*
just.now:ADV 3SG down bottom DEM up summit stone
'Just now he was down below, [now] he is up on top of the rock'

(137) *padahal da ghalé wena latu lako*
besides towards down bottom EXI dog
'Besides, there is the dog down below'

(138) *lako kodho é*
dog look EXCLAM
'The dog watches [and says:] 'Hey!"

(139) *delu aku ata ... manusia na ghéta tolo na*
$friend_4$ 1SG person ... human DEM up summit DEM

"My friend the human is up on top there"

(140) *aku neké gha*
1SG hide here
"I [will] hide here"

(141) *kai neké ghalé wena watu*
3SG hide down bottom stone
'He hides under the rock'

(142) *sawé manusia na ghéta tolo*
finish human DEM up summit
'Then, that human is up on top'

(143) *uli kai ata manusia na ghéta tolo*
company 3SG person human DEM up summit
'His friend the human is up on top'

(144) *sawé ana po dui na kai ghéta tolo kaju*
finish child owl horn DEM 3SG up summit wood
'Then, that horned owl (?), he is up in the top of the tree'

(145) *kai neké*
3SG hide
'He hides'

(146) *iwa nala latu rusa*
NEG long.time EXI deer
'Not long after, there is a deer'

(147) *rusa ghéta wolo watu*
deer up hill stone
'The deer is on top of the rock'

(148) *rusa ndu'a*
deer forest
'A wild deer'

(149) *apa na na*
what DEM DEM
"What is this?"

(150) *latu apa na*
EXI what DEM
"What is there?"

(151) *latu po*
EXI owl
"There is an owl"

(152) *po na ghéa na kodho apa ghéa*
owl DEM there DEM look what there
"What is that owl there looking at?"

(153) *éo kai kodho na kodho apa*
REL 3SG look DEM look what
"What is it that he is watching?"

(154) *rusa néni*
deer peek
'The deer peeks out'

(155) *apa éo kai kodho*
what REL 3SG look
"What is he looking at?"

(156) *ana po tau apa*
child owl make what

"What is the owl doing?"

(157) *rusa kodho leka po*
deer look LOC owl
'The deer looks at the owl'

(158) *ai latu apa na*
EXCLAM EXI what DEM
"Oy, what is there?"

(159) *kai alé ... alé po*
3SG ask ... ask owl
'He asks ... asks the owl'

(160) *apa éo kau gaé*
what REL 2SG search
"What is it that you are looking for?"

(161) *kau gaé apa*
2SG search what
"What are you looking for?"

(162) *rusa téi lako*
deer find dog
'The deer sees the dog'

(163) *rusa no'o lako mbana*
deer COM dog go
'The deer and the dog walk'

(164) *lako nosi no rusa*
dog say COM deer
'The dog says to the deer'

(165) *delu aku manusia ghéa na o fua toki*
friend$_4$ 1SG human there DEM REL wasp bite
"It is my dear friend, the human there, who the wasps bit"

(166) *lako paru jara mulu*
dog run horse first
'The dog runs ahead'

(167) *lako paru*
dog run
'The dog runs'

(168) *rusa ndu no ... rusa ndu'a na ndu no lako*
deer follow COM ... deer forest DEM follow COM dog
'The deer follows ... That wild deer follows the dog'

(169) *mo mbana laka manusia ghéa*
FUT go help human there
'In order to help the human there'

(170) *so du ghéa leka muru fila dowa*
so reach there LOC cliff abyss PERF
'When they have arrived there at a chasm'

(171) *téi do leka muru fila ghéta wawo watu*
find PERF LOC cliff abyss up top stone
'[They] have found a chasm above the rocks'

(172) *du ghalé wena kai béwa*
reach down bottom 3SG tall
'It is far to the bottom of it'

(173) *ghéta tolo du ghalé wena ngéré na*
up summit reach down bottom like DEM

'From the top to the bottom, like this'

(174) *manusia mo mesu du ghalé wena*
human FUT fall reach down bottom
'The human is going to fall to the bottom'

(175) *mo du ghalé wena tana*
FUT reach down bottom ground
'[He] is going end up down on the ground'

(176) *lako kai tiro jara mulu da ghalé*
dog 3SG jump horse first towards down
'The dog jumps down ahead of him'

(177) *lako ghéta wawo kai tiro da ghalé wena*
dog up top 3SG jump towards down bottom
'The dog, from the top he jumps down to the bottom'

(178) *rusa kola liba leka manusia*
deer chase right.away LOC human
'The deer follows right after the human'

(179) *manusia ... kai ta'u manusia ola mesu*
human ... 3SG fear human lest fall
'The human ... He is afraid lest the human fall'

(180) *kai kola liba*
3SG chase right.away
'He follows right after'

(181) *tapi lako kai jara mulu sawé du ghéa*
but dog 3SG horse first finish reach there
'But the dog goes there first of all'

(182) *mo ngéré emba ngéré emba kita na ebé piki*
FUT like which like which 1PL.I DEM 3PL think
"What are we to do?' they think'

(183) *tahu-tahu ebé na kolo telu-telu ghéa ... lako kai rusa ...*
suddenly:ADV 3PL DEM head.CL three:DISTR there ... dog 3SG deer
'Suddenly, the three of them ... the dog, the deer ...'

(184) *rusa ké'é tolo*
deer still summit
'The deer is still on top'

(185) *lako no'o manusia dowa ghalé oné aé*
dog COM human PERF down inside water
'The dog and the human are already down in the water'

(186) *mesu leka aé*
fall LOC water
'[They] fell in the water'

(187) *ghalé tiwu aé ebé mesu*
down lake water 3PL fall
'They fell into a lake'

(188) *rusa wé'é ko i:*
deer just ?? EXCLAM
'The deer just [says]: 'Eeeek!"

(189) *ebé dowa ghalé oné aé*
3PL PERF down inside water
"They are already down in the water"

(190) *aku ngéré emba na*
1SG like which DEM

"What am I to do?"

(191) *aku tau apa na na*
1SG make what DEM DEM
"What do I do?"

(192) *mo laka ebé na*
FUT help 3PL DEM
"[I] will help them."

(193) *demi aku mo mesu ... tiro du ghalé na aku si ghalé oné aé*
COND 1SG FUT fall ... jump reach down DEM 1SG also down inside water
"If I fall ... jump down there, I too [will be] down in the water"

(194) *tapi manusia téi*
but human find
'But the humans sees ...'

(195) *rusa ... manusia so kodho na*
deer ... human so look DEM
'The deer ... the human looks [and says]'

(196) *ei lako aku si ghalé oné aé*
EXCLAM dog 1SG also down inside water
"Hey dog! I too am down in the water"

(197) *kai nangu*
3SG swim
'He swims'

(198) *manusia nangu*
human swim
'The human swims'

(199) *mbana dha'o walo lako kai*
go help.up return$_3$ dog 3SG
'[He] goes to help his dog back up'

(200) *kai so déo lako*
3SG so hold dog
'So he holds the dog'

(201) *kai pémba leka toko wara kai*
3SG cradle$_2$ LOC bone shoulder 3SG
'He holds [him] on his shoulders'

(202) *ai delu kau na kola aku ngéré emba na*
EXCLAM friend$_4$ 2SG DEM chase 1SG like which DEM
"Oy, my friend! Why are [they] chasing me?"

(203) *kita imu rua-rua gha oné aé mésa*
1PL.I friend.*cl* two:DISTR here inside water alone
"The two of us are here in the water alone"

(204) *akirnya ebé nangu*
in.the.end 3PL swim
'In the end, they swim'

(205) *nangu nai du ghélé wolo*
swim ascend reach up$_2$ hill
'[They] swim [and] climb up to the top of a hill'

(206) *du ghélé wolo*
reach up$_2$ hill
'They reach the top of the hill'

(207) *manusia mera piki do*
human sit think PERF

'The human sits [and] thinks'

(208) *ai* *ngéré emba bega* *na*
EXCLAM like which friend$_3$ DEM
"Oy, what to do, my friend?"

(209) *menga pu* *kita* *na*
only EMPH 1PL.I DEM
"[There is] only us"

(210) *leko kai* *iwa* *téi*
frog 3SG NEG find
"There isn't the frog"

(211) *kau nai* *sai*
2SG ascend IMP
"You go up!"

(212) *kau nai* *du* *gha*
2SG ascend reach here
"You climb up here"

(213) *bega* *leko na* *leka emba*
friend$_3$ frog DEM LOC which
"Friend, where is that frog?"

(214) *kita* *mo* *gaé* *wola* *leka emba* *leko na*
1PL.I FUT search return$_2$ LOC which frog DEM
"Where will we look for the frog again?"

(215) *ga* *téi* *menga kita* *imu* *rua*
friend$_2$ find only 1PL.I friend.CL two
"My friend, there are only the two of us"

(216) *susa na*
unfortunate DEM
"It is unfortunate"

(217) *kita mo bhalé wola du ghéa nua ngéré emba na*
1PL.I FUT return return$_2$ reach there village like which DEM
"How will we go back to the village?"

(218) *akir ... gaé terus*
finally ... search continue
'In the end ... [they] keep on searching'

(219) *ebé mbana gaé leka kodo kaju*
3PL go search LOC hollow wood
'They go look inside the hollow of a log '

(220) *téi ... kaju ... kodo kaju ... kaju o mata sawé tu'u sawé*
find ... wood ... hollow wood ... wood REL die finish dry finish
'There is ... a tree ... a hollow log ... a tree which is dead, dried out'

(221) *[lé kodo] ghéa (léko do ghéa)*
[LOC$_2$ hollow] there
'In the hollow there'

(222) *lako néni*
dog peek
'The dog takes a peek'

(223) *kai tatu wau*
3SG sniff$_2$ smell
'He sniffs the scent'

(224) *wau leko*
smell frog

"It smells like frog"

(225) *dau leko na*
must frog DEM
"It must be that frog"

(226) *kai neké leka na kai na*
3SG hide LOC DEM 3SG DEM
"He's hiding here, he is!"

(227) *neké leka na leka kaju tu'u sa mbé'o*
hide LOC DEM LOC wood dry SG know
"[He] is hiding here, in the dried tree maybe"

(228) *manusia kai néni so sepu so néni ngéré na*
human 3SG peek so end so peek like DEM
"The man takes a peek, in the end [he] peeks like this"

(229) *tapi lako kai nugu da ghalé wena*
but dog 3SG bow.head towards down bottom
'But the dog is leaning his head downwards'

(230) *so kodho téi leko nebu nggénggé ghalé wena fata kaju*
so look find frog time protect down bottom fallen.wood wood
'Then [they] look [and] see the frogs taking shelter underneath the fallen tree'

(231) *ngelo delu kau pu nara kami gha gaé-gaé kau na iwa téi*
comrade friend$_4$ 2SG EMPH brother 1PL.E here search:ITER 2SG DEM NEG find
na
DEM
"You, dear friend, brother, we here have been looking and looking for you [but we] did not find [you]"

(232) *ai molo do kita rapa téi wola*
EXCLAM correct PERF 1PL.I RECIP find return$_2$
"Oh, it's OK, we have found each other again"

(233) *kita rapa téi wola na*
1PL.I RECIP find return$_2$ DEM
"We have found each other again"

(234) *mai gha*
come here
"Come here"

(235) *ngéré emba o apa tau kau gharu*
like which REL what make 2SG over.there
"What are you doing over there?"

(236) *leko ghéa éko rua-rua*
frog there tail.CL two:DISTR
'[There are] two frogs there'

(237) *néa-néa sa éko molo na*
just.now:ADV SG tail.CL correct DEM
'Just now there was only one'

(238) *na éko rua no'o uli kai*
DEM tail.CL two COM company 3SG
'Now [there are] two, [him] and his friend'

(239) *ebé ra'i*
3PL approach
'They come closer'

(240) *ebé rapa téi ghéa kai na*
3PL RECIP find there 3SG DEM

'They meet each other there'

(241) *sawé na manusia [na]*
finish DEM human DEM
'After that, the human [says:]'

(242) *ngéré emba ebé ngéré ngangé ... ngangé wa'u du gha*
like which 3PL like not.want ... not.want go.down reach here
"Why do they not want ... not want to come down here?"

(243) *ngangé mai du gha kita*
not.want come reach here 1PL.I
"Do not want to come here to us?"

(244) *mungkin iwa so leko éo kita piara [na] iwa na*
maybe NEG so frog REL 1PL.I raise DEM NEG DEM
"Maybe it's not the frog which we raised, no"

(245) *na leko éo réwo na*
DEM frog REL other DEM
"This is a different frog"

(246) *kita sala na*
1PL.I wrong DEM
"We are mistaken"

(247) *leko na dau leko (é)o réwo*
frog DEM must frog REL other
"This frog must be a different frog"

(248) *iwa so leko (é)o kita piara na iwa*
NEG so frog REL 1PL.I raise DEM NEG
'[It] is not the frog which we raised, no'

(249) *(é)o kita ... éo uli imu kita iwa so o na si iwa sa mbé'o*
REL 1PL.I ... REL company friend 1PL.I NEG so REL DEM also NEG SG know
"Ours ... the one which is our friend is not this one either, maybe not'

(250) *na kita pai juga ebé iwa geju na*
DEM 1PL.I call also.MALAY 3PL NEG come.out DEM
"We even called [and] they did not come out"

(251) *ebé gharu menga kodho kita*
3PL over.there only look 1PL.I
"They over there are just looking at us"

(252) *sawé gharu latu no uli imu kai*
finish over.there EXI COM company friend 3SG
"Then, over there [he] is with his friend"

(253) *ebé gharu latu [do] éko rua*
3PL over.there EXI [PERF] tail.CL two
"There are two of them over there"

(254) *lako kau lélé ro aku nosi*
dog 2SG listen before 1SG say
"Dog, you listen first, I [will] speak"

(255) *sawé na manusia nosi iwa kita nai du ghélé wolo*
finish DEM human say NEG 1PL.I ascend reach up$_2$ hill
'After that, the human says: 'No! We will climb up to the top of the hill"

(256) *kita kodho ghélé wolo*
1PL.I look up$_2$ hill
"We [will] look on top of the hill"

(257) *gharu kai na*
over.there 3SG DEM

"Over there"

(258) *kau ndéki aku nosi*
2SG listen$_2$ 1SG say
"You listen, I speak"

(259) *aku nai da ghélé wawo*
1SG ascend towards up$_2$ top
"I [will] climb to the top"

(260) *aku nai ghélé wawo*
1SG ascend up$_2$ top
"I [will] climb to the top"

(261) *kau jaga ghalé wena*
2SG guard down bottom
"You keep watch down here"

(262) *kau jaga aku nai*
2SG guard 1SG ascend
"You keep watch, I climb"

(263) *ho'o*
yes
"OK"

(264) *ebé kodho ghawa leka oné aé latu ...*
3PL look down$_2$ LOC inside water EXI
"They see that down in the water there is ...

(265) *ai ... lako ghawa oné aé na*
EXCLAM ... dog down$_2$ inside water DEM
"Oy! ... The dog is down in the water!'

(266) *ghéa tolo gha wawo leka fata kaju néa-néa*
there summit here top LOC fallen.wood wood just.now:ADV
'Up there on top, up in the fallen tree '

(267) *o réwo kai sa mbé'o*
REL other 3SG SG know
'Maybe it is a different one'

(268) *ghawa leka tiwu ghawa na latu o ... apa na*
down$_2$ LOC lake down$_2$ DEM EXI REL ... what DEM
'Down in the lake, down there, there is one which ... what? ...'

(269) *pasti latu leko ghawa na*
certainly EXI frog down$_2$ DEM
'There must be a frog down there'

(270) *o nangu ghawa na leko o kita piara sa mbé'o*
REL swim down$_2$ DEM frog REL 1PL.I raise SG know
'The one which is swimming down there is the frog which we raised, maybe'

(271) *uli imu kita sa mbé'o*
company friend 1PL.I SG know
'Our friend, maybe'

(272) *ai sawé leko iwa mbana no ebé*
EXCLAM finish frog NEG go COM 3PL
'Oy! Then, the frog does not go with them'

(273) *tapi leko ghéa na nai mera ghéta wawo fata kaju*
but frog there DEM come sit up top fallen.wood wood
'But the frog there climbs [and] sits on top of the fallen tree'

(274) *kai mera*
3SG sit

'He sits'

(275) *ebé mera*
3PL sit
'They sit'

(276) *ai néa-néa kolo rua .. néa-néa manusia no'o lako ghéa*
EXCLAM just.now:ADV head.CL two ... just.now:ADV human COM dog there
néa-néa
just.now:ADV
"Oy, those two just now ... the human and the dog there just now"

(277) *o gaé apa ebé ghéa na*
REL search what 3PL there DEM
"What is it that they are searching for there?"

(278) *ebé keliru ga delu néa-néa na*
3PL mistaken friend$_2$ friend$_4$ just.now:ADV DEM
"They were mistaken just now, my friend"

(279) *ebé pai kita mo mbana du ghéa ebé*
3PL call 1PL.I FUT go reach there 3PL
"They are calling us to come there to them"

(280) *kita bébo ebé ghéa na o apa ebé olo molo kai*
1PL.I not.know 3PL there DEM REL what 3PL before$_2$ correct 3SG
"We don't know them, what they are [doing] really"

(281) *ebé pai réwo-réwo ghéa na*
3PL call other:ADV there DEM
"They are calling in vain there"

(282) *iwa so uli kita si iwa ebé na*
NEG so company 1PL.I also NEG 3PL DEM

"[They] are not our friends, no, them"

(283) *sala*
wrong
"[They] are mistaken"

(284) *ebé piki so kita na uli*
3PL think so 1PL.I DEM company
"They think we are [their] friends"

(285) *bega delu [mai sai] kita mbana na ta ebé iwa ebé iwa do*
friend$_3$ friend$_4$ [come IMP] 1PL.I go DEM CONJ 3PL NEG 3PL NEG PERF
"Friend, come, let's go"

(286) *ebé mbana sawé do molo sai*
3PL go finish PERF correct IMP
"They have left, OK"

(287) *kita ngéré emba menga kita kodho ebé du ghéa kai*
1PL.I like which only 1PL.I look 3PL reach there 3SG
"What are we to do, do we just watch them [go] there?"

(288) *ngéré emba*
like which
"What to do?"

(289) *ebé gaé ké'é ta iwa*
3PL search still CONJ NEG
"Are they still searching or not?"

(290) *leko no'o ... lako no'o manusia ghéa iwa do téi*
frog COM ... dog COM human there NEG PERF find
'The frog ... The dog and the human are no longer there'

(291) *mbana sawé do*
go finish PERF
'[They] have gone'

(292) *menga leko mésa*
only frog alone
'[There are] only the frogs alone'

(293) *leko éko rua na ...*
frog tail.CL two DEM
'Those two frogs ...'

(294) *leko éo réwo iwa so leko uli imu ebé si iwa*
frog REL other NEG so frog company friend 3PL also NEG
'The other frog is not their friend the frog, no'

(295) *mesi leko uli ebé ...*
COND_2 frog company 3PL ...
'If the frog were their friend ...'

(296) *hai sawé do*
EXCLAM finish PERF
'Oh, it's finished'

4.1.2 Sambal Recipe

The following is a recipe for fish sambal as told by ML.

(297) *tau koro ipu*
make chili small.fish
'Making fish sambal'

(298) *éo mo kema namo kai du'a ra'i kai ... [leka] ipu*
REL FUT work cook 3SG early approach 3SG ... [LOC] small.fish
'[He] who will prepare it, first of all ... [in the] little fish ...'

(299) *ipu leka ma'u leka aé mesi atau leka lowo*
small.fish LOC beach LOC water salt or LOC river
'The little fish are at the beach, in the sea or in the river'

(300) *ata biasa demi [du do wula dala kai] ipu kai nai dowa*
person usually COND [reach PERF moon star 3SG] small.fish 3SG ascend PERF
ata mbana sopi ipu
person go scoop small.fish
'People usually, when the season where the fish come up has arrived, people go scoop the little fish'

(301) *istila sopi ipu wiki no émbér no'o karu*
term scoop small.fish take COM bucket COM sack
'The term 'scooping little fish' is taking [them with?] a bucket [or with] a sack'

(302) *du ghéa mo gajo*
reach there FUT scoop_2
'[They] arrive there in order to scoop'

(303) *ipu [l(au) ghéa mai] leka lowo atau leka aé mesi na no'o [ka] ...*
small.fish seaward there come LOC river or LOC water salt DEM COM [sack] ...
no'o émbér
COM bucket
'Those little fish from down there in the river or in the sea with a [sack] ... with a bucket'

(304) *sawé beru do'i welu leka karu*
finish then pour put LOC sack
'After that they pour them into a sack'

(305) *demi benu do karu atau émbér kai na bhalé wola da ghéa sa'o*
COND full PERF sack or bucket 3SG DEM return return_2 towards there house
'When the sack or bucket is full, [they] go back home'

(306) *du ghéa sa'o ipu latu ata ... po'o ...*
reach there house small.fish EXI person ... bamboo.joint ...
'[When] they get home [with the] fish there are some people ... cook [it] in a joint of bamboo ...'

(307) *latu o po'o liba*
EXI REL bamboo.joint right.away
'There are some who cook it right away in a joint of bamboo'

(308) *po'o pesa paké po'o peri*
bamboo.joint eat.meat use bamboo.joint bamboo
'Cook [and] eat [it] using a joint of bamboo'

(309) *poa pé'i leka api raka du mami*
thrust lean LOC fire reach$_2$ reach cooked
'[They] stick [it] in the fire until [it] is cooked'

(310) *ata pesa liba*
people eat.meat right.away
'People eat it right away'

(311) *tapi latu ipu éo welu wau wé'é*
but EXI small.fish REL put smell just
'But there are some little fish which [they] just leave to rot'

(312) *pusi leka bhoku*
put.in LOC large.bamboo.joint
'[They] put them into a large joint of bamboo'

(313) *ra'i ipu géwu no'o mesi*
approach small.fish mix COM salt
'First the fish is mixed with salt'

(314) *géwu no'o mesi bhondo méma supaya kai ma'é moso*
mix COM salt much really so.that$_2$ 3SG PROH spoil
'[They] mix [it] with a lot of salt so that it does not spoil'

(315) *supaya ipu éko kai na ngéré la'é kai*
so.that$_2$ small.fish tail 3SG DEM like not.yet 3SG
'So that the fish's tail is like before (?)'

(316) *kai mo tu'a bheni deki mesi so('o) bhondo*
3SG FUT hard wonderful touch$_2$ salt more much
'It becomes hard [and] delicious because there is so much salt'

(317) *welu raka du wau sa minggu*
put reach$_2$ reach smell SG week.CL
'[They] leave [it] until [it] stinks, one week'

(318) *leja lima lima esa welu*
day.CL five five one put
'[They] leave [it] around five [or] six days'

(319) *demi kai wau dowa kita ngala tau koro ipu*
COND 3SG smell PERF 1PL.I can make chili small.fish
'Once it stinks, we can make fish sambal'

(320) *naja kai koro ipu*
name 3SG chili small.fish
'It's name is fish sambal'

(321) *bumbu-bumbu kai mo tau géwu goé kai ghéa na*
ingredient:PL 3SG FUT make mix mix$_2$ 3SG there DEM
'The ingredients in order to make the mixture [are:]'

(322) *koro mesi somu sunga faté nawé wunu ora mo tamba*
chili salt onion garlic basil lemongrass leaf sweet.basil FUT add

'Chillies, salt, onion, garlic, basil, lemongrass, sweet basil leaves, [we] will add'

(323) *beru ruja*
then grind
'Then grind [it]'

(324) *pu'i nilu geru pu'i nilu membu*
pluck sour small.lizard pluck sour ??
'A piece of *nilu geru* (a sour spice), a piece of *nilu membu* (another sour spice)'

(325) *ngala pusi ghéa kai na [na]*
can put.in there 3SG DEM [dem]
'[We] can put in there'

(326) *mo tau koro wau*
FUT make chili smell
'In order to make stinking sambal'

(327) *naja koro wau koro ipu*
name chili smell chili small.fish
'The name is stinking sambal, fish sambal'

(328) *kita ruja bhondo méma*
1PL.I grind much really
'We grind [it] up really a lot'

(329) *kita ngala pusi welu leka bhoku atau leka ...*
1PL.I can put.in put LOC large.bamboo.joint or LOC ...
'We can pour [it] into a large joint of bamboo or in ...'

(330) *nebu na latu no stoflés*
time DEM EXI COM jar
'Nowadays it's with a jar'

(331) *pusi welu*
put.in put
'Pour [it] in [and] leave [it]'

(332) *welu ngéré sa wula si ngala kai*
put like SG month.CL also can 3SG
'[We] can leave it around a month even'

(333) *kai jadi wau tapi kai ngé ...*
3SG become$_2$ smell but 3SG increase ...
'It becomes stinky but it gets more and more ...'

(334) *rasa kai ngé gedho-gedho demi kita pesa*
feel 3SG increase good$_2$:INTENS COND 1PL.I eat.meat
'It's taste gets better and better when we eat [it]'

(335) *kita la'i na*
1PL.I lick DEM
'We lick [it]'

(336) *rasa bheni ai*
feel wonderful EXCLAM
'The taste is delicious, ay! (Or: '[It] tastes delicious, ay!')'

(337) *wau ...*
smell ...
'[It] stinks ...'

(338) *demi kita nguru ngéré na wau*
COND 1PL.I sniff like DEM smell
'When we smell it, like this, [it] stinks'

(339) *tapi demi kita la'i rasa kai so('o) éna na*
but COND 1PL.I lick feel 3SG more delicious DEM

'But when we lick [it], it's taste is very delicious'

(340) *gedho na gedho dema na*
good$_2$ DEM good$_2$ truly DEM
'Excellent, truly excellent'

(341) *na tau koro ipu*
DEM make chili small.fish
'That's how to make fish sambal'

(342) *sawé do*
finish PERF
'[It] is finished'

4.1.3 Kanganara History and Farming Rituals

The following is an excerpt of the history and farming rituals of Kanganara, Detukeli as told by ML. Place names are glossed in parentheses when they are mentioned for the first time.

(343) *kami dema (a)ta kanganara (kanga + nara)*
1PL.E truly person pl.name (obelisk + brother)
'We are truly Kanganara people'

(344) *tapi kami latu nua éo sa'o esa telu*
but 1PL.E EXI village REL house seed.CL three
'But we have a village which is three houses'

(345) *nua sa'o esa telu wé'é*
village house seed.CL three just
'A village of only three houses'

(346) *sa'o mamo iné ema aku*
house grandparent mother father 1SG
'The house of my parents' ancestors'

(347) *sa'o ema du'a aku no'o ...*
house father early 1SG COM ...
'The house of my 'big father' (father's elder brother) and ...'

(348) *sa'o esa rua sa'o esa rua*
house seed.CL two house seed.CL two
'Two houses, two houses'

(349) *sa'o menga esa rua*
house only seed.CL two
'Only two houses'

(350) *sa'o ema du'a no'o sa'o mamo iné ema aku*
house father early COM house grandparent mother father 1SG
'My 'big father's' house and my parents' ancestors' house'

(351) *tapi nua ghéa na sa'o esa rua tapi latu nggua*
but village there DEM house seed.CL two but EXI ritual
'The village there is only two houses, but it has a culture (a ritual)'

(352) *nggua mésa*
ritual alone
'A unique culture (ritual)'

(353) *kami ... nggua tana watu ghéa nggua mésa*
1PL.E ... ritual ground stone there ritual alone
'We ... The culture (ritual) of the territory there is a unique culture (ritual) '

(354) *ata éo kema leka wolo bérébu'u (béré + bu'u)*
person REL work LOC hill pl.name (flow + hearth)
'People who work on Mount Bérébu'u'

(355) *ebé tu nggebhé mo tau tedo mula*
3PL bring lie.flat FUT make plant $plant_2$

'They bring offerings on their bellies in order to plant'

(356) *ebé tu ghéa leka sa'o kami*
3PL bring there LOC house 1PL.E
'They bring offerings there to our house'

(357) *jadi kami ... mamo kami mosa laki*
become$_2$ 1PL.E ... grandparent 1PL.E male$_2$ chief
'So we ... our ancestor was a chief'

(358) *mosa laki*
male$_2$ chief
'A chief (Note: There is a term *metu laki* [female$_2$ chief] for female chiefs)'

(359) *kai na ata fai*
3SG DEM person female
'She was a woman'

(360) *tapi ema kai nebu sa leja mosa laki ria*
but father 3SG time SG day.CL male$_2$ chief large
'But her father, long ago, was a great chief'

(361) *kai latu no'o mbendi nggo kai*
3SG EXI COM gun gong 3SG
'He had his guns [and] gongs'

(362) *kai latu tana tanalima (tana + lima)*
3SG EXI ground pl.name (ground + five)
'He had the land of Tanalima'

(363) *tana éo kai*
ground REL 3SG
'The land which was his'

(364) *tana éo embu kami na latu tanalima*
ground REL ancestor 1PL.E DEM EXI pl.name
'The land which was our forefather's is Tanalima'

(365) *sawé welu leka ana kai ata fai mosa laki mamo kami na*
finish put LOC child 3SG person female male$_2$ chief grandparent 1PL.E DEM
'Then [he] left [it] to his daughter, the chief, our ancestor'

(366) *no'o tana k(a)i tanalima*
COM ground 3SG pl.name
'With his land, Tanalima'

(367) *kami mera leka tana bérébu'u tebénaka (tebé + naka)*
1PL.E sit LOC ground pl.name pl.name (cliff + jackfruit)
'We dwell in the land of Bérébu'u [and] Tebénaka'

(368) *tapi [wérobajo] (wéro + bajo) niobéwa (nio + béwa) ghéa lowoléké (lowo +*
but pl.name (?? + ??) pl.name (coconut + tall) there pl.name (river +
léké) pemombaka (pemo + mbaka) o réwo
vine.sp) pl.name (wallowing.hole + ??) REL other
'But Wérobajo, Niobéwa, there in Lowoléké, Pemombaka, [they] are different'

(369) *na leka ... nosi na'u leka aji ana kai*
DEM LOC ... say arrange LOC younger.sibling child 3SG
'In ... [He] arranged it for his descendants'

(370) *kami mera ghéa*
1PL.E sit there
'We dwell there'

(371) *demi uja wa'u dowa*
COND rain go.down PERF
'When the rains begins to fall'

(372) *apu wa'u*
dew go.down
'The dew falls'

(373) *uja na mai*
rain DEM come
'The rain comes'

(374) *kami mulai tau tedo mula*
1PL.E begin make plant plant$_2$
'We begin planting '

(375) *jadi demi mo tedo mula latu no acara kai*
become$_2$ COND FUT plant plant$_2$ EXI COM ceremony 3SG
'So when [we] are going to plant, there is a ceremony'

(376) *tau nggua*
make ritual
'[We] perform a ritual'

(377) *nggua kami menga po'o aré po'o*
ritual 1PL.E only bamboo.joint cooked.rice bamboo.joint
'Our ritual is just to cook rice in a joint of bamboo'

(378) *kami po'o aré*
1PL.E bamboo.joint cooked.rice
'We cook rice in a joint of bamboo'

(379) *sai-sai kai éo kema ghéa leka bérébu'u na wajib ... dau lo*
who:DISTR 3SG REL work there LOC pl.name DEM must$_2$ ... must bring.down
'Anyone who works there on Bérébu'u must ... has to bring down (offerings)'

(380) *paré sa wati manu sa éko moké sa boti*
uncooked.rice SG basket.CL chicken SG tail.CL palm.wine SG bottle.CL

'One basket of rice, one chicken, one bottle of palm wine'

(381) *demi éo kiwa na kai kai kema kiwa ra'o latu no aré*
COND REL year DEM 3SG 3SG work year come$_2$ EXI COM cooked.rice
ké'a
coconut.shell
'He who works this year, will next year have rice in a coconut shell'

(382) *aré ké'a sa ké'a*
cooked.rice coconut.shell SG coconut.shell.CL
'One coconut shell of rice'

(383) *no kiwa no kiwa ngéré na sai éo iwa tu sobhé nggebhé*
COM year COM year like DEM who REL NEG bring cover lie.flat
'Every year, like this, anyone who does not bring offerings on their belly'

(384) *kai mosa laki poi*
3SG male$_2$ chief fine
'The chief fines him'

(385) *poi*
fine
'[He] fines [him]'

(386) *éo poi apa éo mosa laki mo pati leka ebé*
REL fine what REL male$_2$ chief FUT give LOC 3PL
'The fine is whatever the chief wants to give them'

(387) *mosa laki ... kau kiwa na iwa ngala tedo*
male$_2$ chief ... 2SG year DEM NEG can plant
'The chief [says]: 'You may not plant this year"

(388) *mosa laki ndelé*
male$_2$ chief forbid

'The chief forbids'

(389) *iwa ngala tedo*
NEG can plant
"[You] may not plant"

(390) *kau iwa tu paré ké'a moké boti manu éko*
2SG NEG bring uncooked.rice coconut.shell.CL palm.wine bottle.CL chicken tail.CL
paré wati gha kami
uncooked.rice basket.CL here 1PL.E
"You did not bring a coconut shell of rice, a bottle of palm wine [and] a basket of rice here to us"

(391) *ebé iwa*
3PL NEG
'They did not'

(392) *d(au) ebé pati wola beru ebé ngala*
must 3PL give return$_2$ then 3PL can
'They must give it back, then they may'

(393) *jadi demi po'o latu mosa laki kai mosa laki pu'u*
become$_2$ COND bamboo.joint EXI male$_2$ chief 3SG male$_2$ chief trunk
'So when [they] cook the rice, there is the chief, the main chief'

(394) *mosa laki pu'u mamo kami*
male$_2$ chief trunk grandparent 1PL.E
'The main chief is our ancestor'

(395) *mosa laki ria béwa*
male$_2$ chief large tall
'[There is] the 'great tall' chief'

(396) *ria béwa sa kolo ata kaki*
large tall SG head.CL person male
'The 'great tall' [chief] is a man'

(397) *mosa laki kago kao*
male$_2$ chief grab$_2$ scoop$_3$
'[There is] the 'grabbing [and] scooping' chief'

(398) *kago kao kai na ...*
grab$_2$ scoop$_3$ 3SG DEM ...
'The 'grabbing [and] scooping' [chief], he ...'

(399) *demi ata do'i paré leka wati kai na mo kago ...*
COND person pour uncooked.rice LOC basket 3SG DEM FUT grab$_2$ ...
'When people pour rice into the basket he will scoop ...'

(400) *kago na welu da ghéa*
grab$_2$ DEM put towards there
'[He] scoops it [and] puts it there'

(401) *do'i utu mosa laki*
pour for male$_2$ chief
'[He] pours some for the chief'

(402) *leka wati kai éo sisa welu mo t(au) po'o*
LOC basket 3SG REL remain put FUT make bamboo.joint
'The remainder in the basket is left in order to cook it'

(403) *na po'o na*
DEM bamboo.joint DEM
'[They] cook it'

(404) *kai du'a kai wiki*
3SG self 3SG take

'He takes it himself'

(405) *manu kai riké leka saga*
chicken 3SG tie LOC pole
'[He] ties a chicken to a pole'

(406) *pai ... pai naja*
call ... call name
'[He] calls ... calls a name'

(407) *misalnya pai f. d.*
for.example call first.name last.name
'For example [he] calls: 'F. D.!"

(408) *nai nai ghéta sa'o*
ascend ascend up house
'Climb up, climb up into the house'

(409) *nai no paré kai na sa wati no paré*
ascend COM uncooked.rice 3SG DEM SG basket.CL COM uncooked.rice
ké'a k(a)i
coconut.shell 3SG
'[He] climbs up with his basket of rice, with his coconut shell of rice '

(410) *sawé manu kai moké*
finish chicken 3SG palm.wine
'Then his chicken [and] palm wine'

(411) *ebé alé*
3PL ask
'They ask'

(412) *paré*
uncooked.rice

"[Is there] rice?"

(413) *latu*
EXI
"There is"

(414) *do'i ... do'i paré*
pour ... pour uncooked.rice
'[He] pours ... pours the rice'

(415) *do'i kago kao*
pour $grab_2$ $scoop_3$
'[He] pours, grabs [and] scoops [it]'

(416) *kago welu sawé*
$grab_2$ put finish
'[He] grabs [and] leaves [it], that's it'

(417) *sisa leka wati do'i ghéa mbola kai du'a kai*
remain LOC basket pout there large.basket 3SG self 3SG
'What is left in the basket, [he] pours into his own large basket'

(418) *manu latu ta iwa*
chicken EXI CONJ NEG
"Is there a chicken or not?"

(419) *latu*
EXI
"There is"

(420) *moké moké latu iwa*
palm.wine palm.wine EXI NEG
"Palm wine, is there palm wine [or] not?"

(421) *latu*
EXI
"There is"

(422) *paré* *ké'a*
uncooked.rice coconut.shell
"[Is there] rice in a coconut shell?"

(423) *paré* *ké'a* *latu*
uncooked.rice coconut.shell EXI
"There is rice in a coconut shell"

(424) *kula* *paké ké'a*
measure use coconut.shell
'[He] measures using a coconut shell'

(425) *na* *na* *o* *ké'a* *welu mésa*
DEM DEM REL coconut.shell put alone
'That which is in the coconut shell is set aside'

(426) *na* *o* *ata* *laki* *... iwa*
DEM REL person chief ... NEG
'That is the chief's ... no

(427) *sawé* *néa-néa* *leka éo* *kao* *ngéré na* *welu mésa*
finish just.now:ADV LOC REL $scoop_3$ like DEM put alone
'Then what [he] just scooped like this is set aside'

(428) *jadi* *wa'u* *ghéa* *mosa* *laki*
$become_2$ go.down there $male_2$ chief
'So the chief goes down there'

(429) *roré* *manu* *leka saga éo* *welu manu*
slaughter chicken LOC pole REL put chicken

'[They] slaughter the chicken at the pole where [they] put the chicken'

(430) *roré manu wa'u no paré bara*
slaughter chicken go.down COM uncooked.rice white
'[He] slaughters the chicken [and] goes down with white rice'

(431) *mbana réwu leka tana watu*
go sprinkle LOC ground stone
'[He] goes [and] sprinkles [it] on the territory'

(432) *na mo po'o bhoro*
DEM FUT bamboo.joint bake
'[They] will bake it underground in a joint of bamboo'

(433) *roré manu no'o paré éo ata laki seré leka bu'u*
slaughter chicken COM uncooked.rice REL person chief boil LOC hearth
'[They] slaughter the chicken with the rice that the chief boiled on the hearth'

(434) *leka sa'o ria sa'o laki na*
LOC house large house chief DEM
'In the great house, in the chief's house'

(435) *leka bu'u papa mbé'o*
LOC hearth side know
'On the right-hand hearth'

(436) *kai na du'a kai ata laki na mo seré aré [o] kao*
3SG DEM self 3SG person chief DEM FUT boil cooked.rice REL scoop_3
'He himself, the chief, will boil the rice which [he] scooped'

(437) *ata kaki kao seré*
person male scoop_3 boil
'The man scoops [and] boils [it]'

(438) *ata kaki seré ... seré aré*
person male boil ... boil cooked.rice
'The man boils ... boils the rice'

(439) *seré aré nasu manu*
boil cooked.rice cook.meat chicken
'[He] boils rice, cooks chicken'

4.1.4 Early Life

The following is an anecdote about ML's early life.

(440) *aku nebu lo'o muri ré'é dema*
1SG time small live bad truly
'When I was young, life was really hard'

(441) *aku sekola mbana béu*
1SG school go far
'I went to school far away'

(442) *to'o tebénaka mbana kanganara*
arise pl.name go pl.name
'From Tebénaka to Kanganara'

(443) *uja ra'i ghéa loworipi (lowo + ripi) aé dagé talo*
rain approach there pl.name (river + sprout) water cross cannot
'The rain came there in Loworipi, the water could not be crossed'

(444) *aku kelas tiga*
1SG class three.MALAY
'I was in third grade'

(445) *aku no uli (a)ku kelas tiga*
1SG COM company 1SG class three.MALAY
'Me and my friends were in third grade'

(446) *sawé nara aku kelas satu SD*
finish brother 1SG class one.MALAY primary.school
'Then my brother was in first grade of primary school'

(447) *wula uja angi aé ghalé loworipi ria*
moon rain wind water down pl.name large
'During the season of rain and wind, the water down at Loworipi was high'

(448) *kami dagé talo*
1PL.E cross cannot
'We could not cross'

(449) *kami mbana buga la'é*
1PL.E go morning not.yet
'We went in the early morning'

(450) *kami dagé talo kami bhalé wola*
1PL.E cross cannot 1PL.E return return$_2$
'We could not cross, [so] we went back'

(451) *kami iwa du ghéa sekola*
1PL.E NEG reach there school
'We did not reach school'

(452) *menga ata nara na kai mbana jam sembila*
only person brother DEM 3SG go hour nine.MALAY
'Only my brother, he went at nine o'clock'

(453) *du ghéa aé dheso sawé kai mbana liba da ghéa sekola*
reach there water recede finish 3SG go right.away towards there school
'[He] arrived there [and] the water had receded [so] he went directly to school'

(454) *du ghéa sekola guru alé emba ebé éo réwo*
reach there school teacher ask which 3PL REL other

'[He] arrived there at school [and] the teacher asked: 'Where are the others?"

(455) *ebé mai sawé néa-néa*
3PL come finish just.now:ADV
"They have left just now"

(456) *pak goka iwa téi*
MISTER ?? NEG find
"Mister, there is no (?)"

(457) *guru piki do kami aé béré sawé*
teacher think PERF 1PL.E water flow finish
'The teacher thought the water had washed us away'

(458) *uja ria na*
rain large DEM
'The rainwater was high'

(459) *ebé na aé béré sawé na na*
3PL DEM water flow finish DEM DEM
"The water has washed them away"

(460) *ra'i menga kau mésa*
approach only 2SG alone
"Only you are coming alone"

(461) *iwa téi do*
NEG find PERF
"[They] aren't there anymore"

(462) *ebé bhalé wola da mena jala*
3PL return return$_2$ towards counterclockwise.along.coast path
"They went back 'east' along the path (lit. counterclockwise around the coast on the path)"

(463) *iwa*
NEG
"Nothing"

(464) *padahal kami imu rua aku no uli imu ku na*
besides 1PL.E friend.CL two 1SG COM company friend 1SG DEM
'(Meanwhile) the two of us, me and my friend'

(465) *mbana du'u leka lepa moké ata ema du'a kami*
go stop LOC house$_2$ palm.wine person father early 1PL.E
'Went [and] stopped in our 'big father's' palm-tapping hut'

(466) *kami mera ghéa kai na minu moké ba'i*
1PL.E sit there 3SG DEM drink palm.wine bitter
'We sat there [and] drank bitter palm wine'

(467) *tunu pesa no'o té'u*
grill eat.meat COM rat
'[We] grilled [and] ate rats'

(468) *té'u ata ema du'a welu senda na*
rat person father early put trap DEM
'[For] rats, [our] 'big father' would put out traps'

(469) *so'i ra'i*
release approach
'[We] took [them] out and brought [them] over'

(470) *kami tunu té'u*
1PL.E grill rat
'We grilled the rats'

(471) *ndota la'i té'u minu no moké ba'i*
chop.up lick rat drink COM palm.wine bitter

'[We] chopped up [and] ate the rats [and] drank bitter palm wine'

(472) *tunu ru'é no ké'o*
grill chew COM corn
'[We] roasted [and] chewed corn'

(473) *kami so kodho ata nara wola do sekola*
1PL.E so look person brother return$_2$ PERF school
'We saw that our brother had come back from school'

(474) *kai nosi miu néa-néa guru alé gaé miu*
3SG say 2PL just.now:ADV teacher ask search 2PL
'He said: 'You, just now the teacher was asking [and] searching for you"

(475) *aku nosi miu mai sawé*
1SG say 2PL come finish
"I said you had set off already"

(476) *sawé na ghalé iwa téi*
finish DEM down NEG find
"After that [he] didn't find you down there"

(477) *guru nosi miu aé béré sawé*
teacher say 2PL water flow finish
"The teacher said the water had washed you away"

(478) *kami piki w(a)i sia dapa ro do*
1PL.E think follow$_2$ bright get painful PERF
'We thought: 'Tomorrow [we] are going to get a beating"

(479) *to'o léi sia kami so mbana du ghéa sekola*
arise ?? bright 1PL.E so go reach there school
'The following morning, when we went to school'

(480) *guru pati duké kami leka nia ana kelas satu*
teacher give kneel 1PL.E LOC face child class one.MALAY
'The teacher made us kneel in front of the first grade children'

(481) *kami kelas tiga*
1PL.E class three.MALAY
'We were in third grade'

(482) *guru pati duké*
teacher give kneel
'The teacher made us kneel'

(483) *pai kami du ghéa kelas satu mai*
call 1PL.E reach there class one.MALAY come
'[She] called us to the first grade class: 'Come!"

(484) *na kai mera mai o mai sekola*
DEM 3sg sit come REL come school
"He is the one who came to school yesterday"

(485) *miu mera mai iwa sekola*
2PL sit come NEG school
"You did not come to school yesterday"

(486) *duké duké leka depan kelas ghá wewa kelas satu na*
kneel kneel LOC front class here yard class one.MALAY DEM
"Kneel, kneel before this class here, this first grade schoolyard"

(487) *miu duké*
2PL kneel
"You kneel"

(488) *kami duké leka wewa ghéa na*
1PL.E kneel LOC yard there DEM

'We knelt in the schoolyard there'

(489) *ghawa wena pené na*
down$_2$ bottom door DEM
'Under the door'

(490) *duké nia da ghéa ebé*
kneel face towards there 3PL
'[We] knelt facing them'

(491) *beru nosi ma'é ndu ngéré kami bopa sekola*
then say PROH follow like 1PL.E disappear school
'Then [we] said: 'Do not imitate us [and] skip school"

(492) *to'o ghéa na raka du kelas enam*
arise there DEM reach$_2$ reach class six.MALAY
'From then until sixth grade'

(493) *ghéa nua kami latu rapa pana*
there village 1PL.E EXI RECIP shoot
'There was fighting in our village'

(494) *aku kelas enam SD*
1SG class six.MALAY primary.school
'I was in sixth grade of primary school'

(495) *rapa pana kami ghawa tebénaka no ebé ghéle nangalé (nanga + lé)*
RECIP shoot 1PL.E down$_2$ pl.name COM 3PL up$_2$ pl.name (estuary + sway)
'We shot each other, us down in Tebénaka and them up in Nangalé'

(496) *nangalé no ata kurupoké (kuru + poké)*
pl.name COM person pl.name (grass + throw)
'Nangalé and the Kurupoké people'

(497) *kami ghawa tebénaka ata kanganara pisa pemowawi (pemo + wawi) léwagaré (léwa + garé)*
1PL.E down$_2$ pl.name person pl.name pl.name pl.name (wallowing.hole + pig) pl.name (away + speak)
'Us down in Tebénaka, the people of Kanganara, Pisa, Pemowawi, Léwagaré'

(498) *sa kunu du ghawa*
SG group. reach down$_2$
'[We] are one unit (clan, family) until today'

(499) *ghélé m(a)i nangalé pana ghélé m(a)i du ghawa*
up$_2$ come pl.name shoot up$_2$ come reach down$_2$
'From up in Nangalé [they] would shoot from above to below'

(500) *gola no watu*
roll COM stone
'[They] would roll stones'

(501) *aku ta'u peté ro'o kelas tiga ghéa*
1SG fear fasten fast class three.MALAY there
'I was terribly afraid there in third grade'

(502) *aku [d(au)] duké leka nia kelas satu*
1SG [must] kneel LOC face class one.MALAY
'When I [had to] keel before the first grade class'

(503) *aku mbana ata ji'é mai ghélé m(a)i nangalé*
1SG go person excellent come up$_2$ come pl.name
'I went [and] people came from up in Nangalé'

(504) *dari dubé pu wé'é no jala*
stand block EMPH just COM path
'[They] just stood blocking the path'

(505) *watu ghélé m(a)i logo*
stone up_2 come $roll_2$
'[They] rolled stones from above'

(506) *gai ghélé m(a)i pana pu no lé'é*
$arrow_2$ up_2 come shoot EMPH COM arrow
'[They] shot arrows from above'

(507) *[ta] buga la'é poa aku mbana sekola*
[CONJ] morning not.yet early.morning 1SG go school
'(But) it was the early morning [and] I was going to school'

(508) *raka du ghélé sekola*
$reach_2$ reach up_2 school
'[I] arrived up at school'

(509) *guru si'i léna kau mai sekola*
teacher say_2 first.name 2SG come school
'The teacher [said]: 'Léna, you came to school"

(510) *kau iwa ta'u wo'o lé'é*
2SG NEG fear bow arrow
'Aren't you afraid of bows and arrows?'

(511) *lé'é watu leka ebé tau*
arrow stone LOC 3PL make
'The arrows and stones when they fight'

(512) *aku so aku ta'u sawé do sa leja éo guru pati duké aku*
1SG so 1SG fear finish PERF SG day REL teacher give kneel 1SG
'I [said]: 'I was already scared on the day that you (Teacher) made me kneel"

(513) *nebu kelas tiga SD leka nia kelas satu*
time class three.MALAY primary.school LOC face class one.MALAY

"When I was in third grade, in front of the first grade class"

(514) *nebu na aku watu kaju si aku iwa do ta'u*
time DEM 1SG stone wood also 1SG NEG PERF fear
'Now I am stone and wood, I am no longer afraid'

(515) *aku menga ta'u guru ola pati duké wola aku kelas enam*
1SG only fear teacher lest give kneel return$_2$ 1SG class six.MALAY
SD
primary.school
"I am only afraid lest [you,] Teacher make me kneel again in sixth grade of primary school"

(516) *sawé mbana sekola*
finish go school
'Then, [I] went to school'

(517) *kami kelas enam latu kolo lima rua*
1PL.E class six.MALAY EXI head.CL five two
'There were seven of us in sixth grade'

(518) *éo lulus menga pu aku mésa sa kolo*
REL graduate only EMPH 1SG alone SG head.CL
'The only one who graduated was me alone'

(519) *ebé kolo lima esa iwa lulus*
3PL head.CL five one NEG graduate
'The six of them did not graduate'

(520) *mbana sekola ghéa nita*
go school there pl.name
'[I] went to school in Nita there'

(521) *maso SKP nita*
enter SKP pl.name

'[I] entered the SKP (Sasaran Kerja Pegawai) Nita'

(522) *menga minggu telu molo paru wola*
only week.CL three correct run return$_2$
'In only three weeks, [I] ran back'

(523) *aku paru wola su'u no té'é*
1SG run return$_2$ carry.head COM mat
'I ran back carrying a mat on my head'

(524) *to'o ghéa nita paru du ghéa lékéba'i (léké + ba'i)*
arise there pl.name run reach there pl.name (vine.sp + bitter)
'From Nita [I] ran until Lékéba'i'

(525) *du lékéba'i mera ghéa no'o tua*
reach pl.name sit there COM priest
'[I] reached Lékéba'i [and] lived with a priest'

(526) *sa tenga tahun wula lima kami mera ghéa tua*
SG half year month.CL five 1PL.E sit there priest
'For half a year [or] five months we lived at the priest's'

(527) *sawé bhalé wola*
finish return return$_2$
'Then [we] came back'

(528) *du ghéa nua aku dhawé bené*
reach there village 1SG work$_2$ grass$_2$
'[I] reached the village [and] I worked the land'

(529) *kema uma ... kema uma ghéa*
work garden ... work garden there
'[I] worked the garden ... worked the garden there'

(530) *sawé beru maso wola SMP kelas satu ghélé detukéli (detu +*
finish then enter return$_2$ junior.high class one.MALAY up$_2$ pl.name (flat +
kéli) tahun tuju pulu
mountain) year.*malay* seven.MALAY ten.MALAY
'Then [I] entered first grade of junior high school again up in Detukeli in '70'

(531) *tahun tuju pulu satu tamat ... tuju*
year.MALAY seven.MALAY ten.MALAY one.MALAY graduate$_2$ ... seven.MALAY
pulu dua tamat
ten.MALAY two.MALAY graduate$_2$
'In '71 [I] graduated ... '72 [I] graduated'

(532) *tahun tuju pulu tiga aku sekola SPG*
year.MALAY seven.MALAY ten.MALAY three.MALAY 1SG school teacher.school
negeri éndé
national pl.name
'In '73 I went to school at the National SPG (Sekolah Pendidikan Guru) Ende'

(533) *tapi kami mbana (gh)a'i hari pertama*
But 1PL.E go foot day.MALAY first.MALAY
'But we went on foot the first day'

(534) *aku mai du ghéa éndé menga mai no ata eda*
1SG come reach there pl.name only come COM person uncle
'I came to Ende, [I] just came to my uncle's'

(535) *éo mera gha kilométér empat belas ghélé wolo*
REL sit here kilometer four.MALAY teen.MALAY up$_2$ hill
'Who lives here on kilometer 14, up the hill'

(536) *mbana (gh)a'i wé('é) to'o leka tebénaka kanganara raka du gha éndé*
go foot just arise LOC pl.name pl.name reach$_2$ reach here pl.name
'[We] just went on foot from Tebénaka [and] Kanganara until here in Ende'

(537) *renga uja to'o kajundara (kaju + ndara) raka du detusoko (detu + soko)*
withstand rain arise pl.name (wood + thin) reach$_2$ reach pl.name (flat + elephant.grass)
'[We] endured rain from Kajundara until Detusoko'

(538) *ndemo mésa*
wet$_2$ alone
'[We] were completely soaked'

(539) *pekia [o] kami menga leka weki*
clothing [REL] 1PL.E only LOC body
'The clothes we had were only on our body'

(540) *aku no'o uli neku sa kolo*
1SG COM company 1SG.POSS SG head.CL
'Me and my one friend'

(541) *ndemo lawo ... kléi ndemo bita kia mésa*
wet$_2$ female.sarong ... dress wet$_2$ mud XX alone
'[Our] sarongs were soaked ... [our] dresses were all soaked and muddy'

(542) *kami maso ghéa kilométér empat belas nuaréko jam dua malam*
1PL.E enter there kilometer four.MALAY teen.MALAY pl.name hour two.MALAY night.MALAY
'We entered there on kilometer 14, in Nuaréko, at two in the morning'

(543) *mbana gha'i to'o detukéli kanganara tebénaka*
go foot arise pl.name pl.name pl.name
'[We] went by foot from Detukeli, Kanganara, Tebénaka'

(544) *jam dua malam roké ghéa*
hour two.MALAY night.MALAY sleep there
'At two in the morning [we] slept there'

(545) *jam tuju pagi kami mbana gha'i wola to'o nuaréko*
hour seven.MALAY morning.MALAY 1PL.E go foot return$_2$ arise pl.name
raka du ghéa kota éndé
reach$_2$ reach there city pl.name
'At seven in the morning we went again by foot from Nuaréko until Ende town'

(546) *maso gha éndé jam dua belas siang*
enter here pl.name hour two.MALAY teen.MALAY afternoon.MALAY
'[We] entered Ende at twelve noon '

(547) *(gh)a'i lima na mami mojo roké si*
foot hand DEM cooked sore sleep also
'Our arms and legs were completely sore and we slept'

(548) *to'o talo sawé*
arise cannot finish
'[We] could no longer get up'

(549) *menga leka sa'o ata ji'é*
only LOC house person excellent
'Just in a person's house'

(550) *tapi to'o léi sia kami dau bhalé wola*
but arise ?? bright 1PL.E must return return$_2$
'But waking up the next morning, we had to go back'

(551) *mbana alé mo daftar mo sekola*
go ask FUT register FUT school
'[We] went to ask to register to go to school'

(552) *ebé so la'é buka*
3PL so not.yet open
'It was not yet open'

(553) *bhalé wola*
return return$_2$
'[We] went back'

(554) *bhalé wola nai no oto*
return return$_2$ ascend COM car
'We went back in a car'

(555) *du ghéa nua mai sekola da gha*
reach there village come school towards here
'[We] reached our village, [we] came from school to here'

(556) *tapi selama aku sekola*
but while 1SG school
'But as long as I was in school'

(557) *baku regu su'u no'o kai kolo mésa*
snack carry carry.head COM 3SG head alone
'[I] carried [my] food on my head only'

(558) *to'o ghéa nua tebénaka raka du detukéli*
arise there village pl.name reach$_2$ reach pl.name
'From the village of Tebénaka until Detukeli'

(559) *nuka wolo kéliwatu (kéli + watu) bhetokalo (bheto + kalo)*
go.up hill pl.name (mountain + stone) pl.name (large.bamboo + winding)
'[I] climbed up the hill to Kéliwatu [and] Bhetokalo'

(560) *golé da ghalé ghawa léléndalé (lélé + ndalé) raka du*
come.down towards down down$_2$ pl.name (banyan + name) reach$_2$ reach
lowodenu (lowo + denu)
pl.name (river + tree.sp)
'[I] came down to Léléndalé until Lowodenu'

(561) *du ékoléta (éko + léta) nebu sa leja oto menga esa rua molo*
reach pl.name (tail + bridge) time SG day car only seed.CL two correct
'Long ago, there were only two cars to Ékoléta'

(562) *oto rokaténda (roka + ténda) no'o oto agung (agung)*
car car.name (push.ENDE + platform) COM car car.name (glorious.MALAY)
'The car 'Rokaténda' and the car 'Agung"

(563) *demi oto iwa kami roké wola ghéta ékoléta sa kobé*
COND car NEG 1PL.E sleep return$_2$ up pl.name SG night.CL
'If there was no car, we slept a night in Ékoléta again'

(564) *ata ... leka sa'o ata ji'é*
person ... LOC house person excellent
'A person ... In a person's house '

(565) *demi iwa kami pémé oto raka du baku mami kami sawé*
COND NEG 1PL.E await car reach$_2$ reach snack cooked 1PL.E finish
'If not, we waited for the car until our packed food was finished'

(566) *leja kobé beru nai oto demi oto ra'i*
day night then ascend car COND car approach
'Day or night, [we] just got in the car when the car came'

(567) *jadi ola muri selama sekola iwa ngéré ana lo'o nebu na*
become$_2$ NOM live while school NEG like child small time DEM
'So life when [I] was in school is not like children nowadays'

(568) *éo oto nai leka sa'o*
REL car ascend LOC house
'For whom the car goes up to the house'

(569) *ta iwa téki kami nebu sa leja*
CONJ NEG bring$_2$ 1PL.E time SG day

'But [it] did not take us there, long ago'

(570) *su'u lepo paré ria ... ria*
carry.head container uncooked.rice large ... large
'[We] carried big ... big containers of rice on our heads'

(571) *paré muku uwi kaju ké'o pusi no lepo*
uncooked.rice banana tuber wood corn put.in COM container
'Rice, bananas, cassava, corn was put in a container'

(572) *su'u to'o leka wa'u sa'o raka du ékoléta*
carry.head arise LOC go.down house $reach_2$ reach pl.name
'[We] carried it on our heads from leaving the house until Ékoléta'

(573) *jadi ola muri susa dema na*
$become_2$ NOM live unfortunate truly DEM
'So life was really hard'

(574) *raka du sekola tamat SPG kami muri menga ngéré na mésa*
$reach_2$ reach school $graduate_2$ teacher.school 1PL.E live only like DEM alone
'Until [we] graduated from SPG (teacher's college), this was the only way we lived'

4.1.5 Our Father

The following is the Our Father prayer as recorded in the prayer book *Jala da Ghéta Surga*.

(575) *ema kami éo ghéta wawo ji'é*
father 1PL.E REL up top excellent
'Our Father, who is up in Heaven'

(576) *ata mbeja tau ada sai naja kau*
person $finish_2$ make honor IMP name 2SG
'May all people honor your name'

(577) *ola uku mumu kau mai sai da gha kami*
NOM rule sprout$_2$ 2SG come IMP towards here 1PL.E
'May your kingdom come here to us'

(578) *ola fonga kau mbalé sai leka tana sawé ngéré ghéta wawo ji'é*
NOM want 2SG become IMP LOC ground finish like up top excellent
'May your will be done on the whole Earth as up in Heaven'

(579) *kaju ka kami no'o leja-leja pati kami leja ina*
wood eat 1PL.E COM day:DISTR give 1PL.E day DEM
'Give us today our daily food'

(580) *soli pati ampu sala kami*
CONJ$_2$ give pardon wrong 1PL.E
'And forgive our wrongdoings'

(581) *ngéré kami ngeni pati ampu leka ata éo sala dowa no'o kami*
like 1PL.E also$_2$ give pardon LOC person REL wrong PERF COM 1PL.E
'As we too forgive people who have wronged us'

(582) *ma'é tu kami leka ola gagi dosa ta tau so'i kami leka éo ré'é*
PROH bring 1PL.E LOC NOM invite sin CONJ make release 1PL.E LOC REL bad
'Do not lead us into temptation but deliver us from that which is evil'

(583) *amin*
AMEN
'Amen'

4.1.6 Hail Mary

The following is the Hail Mary prayer as recorded in the prayer book *Jala da Ghéta Surga*.

(584) *avé maria*
hail Mary
'Hail Mary'

(585) *kau benu no'o ngawu du'a ngga'é*
2SG full COM thing early master
'You are full with God's treasures'

(586) *du'a ngga'é no'o kau*
early master COM 2SG
'God is with you'

(587) *kai berka dowa kau so'o langga ata fai mbeja*
3SG bless PERF 2SG more exceed person woman finish$_2$
'He has blessed you more than all women'

(588) *soli berka dowa ana tuka kau yésus*
CONJ$_2$ bless PERF child belly 2SG Jesus
'And has blessed the child of your belly, Jesus'

(589) *santa maria iné du'a ngga'é*
holy.F Mary mother early master
'Saint Mary, mother of God'

(590) *rina pia sai laka kami ata no'o dosa*
beg XX IMP help 1PL.E person COM sin
'Pray to help us, sinners'

(591) *rina nebu ina soli nebu kami wé'é mata*
beg time DEM CONJ$_2$ time 1PL.E near die
'Pray now and when we are near death'

4.1.7 Credo

The following is the Credo prayer as recorded in the prayer book *Jala da Ghéta Surga*.

(592) *aku tonda no'o du'a ngga'é*
1SG believe COM early master
'I believe in God'

(593) *ema mulé ngala mbeja*
father have.power can finish$_2$
'The almighty Father'

(594) *éo nena dowa wawo ji'é no'o tana watu ina iwa no'o apa-apa*
REL create PERF top excellent COM ground stone DEM NEG COM what:ADV
'Who has created this Heaven and Earth from nothing'

(595) *aku tonda no'o yésus kristus ana kai sa kaki wé'é tuha kita*
1SG believe COM Jesus Christ child 3SG SG male just lord 1PL.I
'I believe in Jesus Christ His only Son, our Lord'

(596) *no'o ola mulé ngai santo kau mbalé dowa manusia*
COM NOM have.power spirit holy.M 2SG become PERF human
'By the power of the Holy Spirit he became a man'

(597) *santa maria ka'o dowa kai*
holy.F Mary cradle PERF 3SG
'Saint Mary has cradled him'

(598) *kai nesi susa nebu pontius pilatus*
3SG endure unfortunate time first.name last.name
'He endured suffering during the time of Pontius Pilate'

(599) *ata toka kai leka salib*
person stab 3SG LOC cross
'People nailed him to the cross'

(600) *so mata sawé ata tané kai*
so die finish person bury 3SG
'When he was dead, people buried him'

(601) *kai wa'u da ghawa ola mera napa*
3SG go.down towards down$_2$ NOM sit wait

'He went down to the grave'

(602) *so dagé leja telu kai to'o muri walo*
so cross day.CL three 3SG arise live return$_3$
'When three days had passed he came back to life'

(603) *kai nai da ghéta wawo ji'é*
3SG ascend towards up top excellent
'He went up to Heaven'

(604) *mera papa nggana du'a ngga'é*
sit side right early master
'[He] sits at the right hand of God'

(605) *ema mulé ngala mbeja*
father have.power can finish$_2$
'The almighty Father'

(606) *néa apa kai mo'o mai walo tau timba ata muri no'o ata mata*
just.now what 3SG FUT come return$_3$ make weigh person live COM person die

'One day he will come back to judge the living and the dead'

(607) *aku tonda no'o ngai santo*
1SG believe COM spirit holy.M
'I believe in the Holy Spirit'

(608) *no'o kunu serani katoli*
COM group christian catholic
'In the Catholic Christian Church'

(609) *no'o kunu santo*
COM group holy.M

'In the communion of saints'

(610) *aku tonda no'o ola bopa dosa*
1SG believe COM NOM disappear sin
'I believe in the forgiveness of sins'

(611) *no'o ata mata to'o*
COM person die arise
'That the dead will rise'

(612) *no'o ola muri iwa du'u-du'u*
COM NOM live NEG stop:ADV
'In eternal life'

(613) *amin*
AMEN
'Amen'

4.1.8 Wedding Vows

The following are Lio wedding vows as recorded in the prayer book *Jala da Ghéta Surga*.

(614) *kita tonda sama-sama*
1PL.I believe together:ADV
'We believe together'

(615) *pu'u ra'i-ra'i tana ina mbalé pu'u leka du'a ngga'é nena*
trunk approach:ADV ground DEM become trunk LOC early master create
'In the beginning, this Earth came to be, because of what God created'

(616) *néa kema namo sawé nena na'u mbeja ji'é-ji'é*
just.now work cook finish create arrange finish_2 excellent:ADV
'He created everything, arranged everything perfectly'

(617) *deki kai nara kema mbalé manusia*
touch$_2$ 3SG desire work become human
'Until he felt like creating mankind'

(618) *ndu dema wangu ango du'a kai du'a ngga'é kai nena manusia*
follow truly appearance shape self 3SG early master 3SG create human
'According to God's own image, he created mankind'

(619) *ata kaki no'o ata fai kai pati ebé*
person male COM person female give 3PL
'Man and woman, he made them'

(620) *du'a ngga'é berka ebé dapi nosi*
early master bless 3PL CONJ$_3$ say
'God blessed them and said'

(621) *ngé bhondo beka kapa sai miu*
increase much multiply thick IMP 2PL
"May you increase greatly [and] multiply thickly!"

(622) *ngé riwu-riwu ngéré dala ghéta liru*
increase thousand:ADV like star up sky
"Increase by the thousands like stars up in the sky"

(623) *tamba ngasu-ngasu ngéré ena lau ma'u*
add hundred:ADV like sand seaward beach
"Add more by the hundreds like sand by the shore"

(624) *ngé wai ngéré ana ngana leka tana watu*
increase follow$_2$ like child termite LOC ground stone
"Increase like termites on the land"

(625) *wé'é miu wiwi lema no'o tana watu ina*
so.that 2PL lip tongue COM ground stone DEM

"So that you may rule this land"

(626) *no'o ika lau aé mesi*
COM fish seaward water salt
"And the fish down in the sea"

(627) *dapi ulé agé ghéta liru*
CONJ_3 worm bird up sky
"And the creatures up in the sky"

(628) *dapi mbeja sa kamba jara wawi manu*
CONJ_3 finish_2 SG buffalo horse pig chicken
"And all buffalo [and] horses, pigs [and] chickens"

(629) *soli apa-apa mbeja éo muri méko leka nia tana ina*
CONJ_2 what:ADV finish_2 REL live move LOC face ground DEM
"And everything which lives [and] moves on the face of this Earth"

(630) *so sawé du'a ngga'é na'u nena adam no'o éva*
so finish early master arrange create Adam COM Eve
'After that, God created Adam and Eve'

(631) *ata kaki dau ji'é wé'é welu léwa iné ema kai*
person male must excellent so.that put away mother father 3SG
"It is good that the man leaves his parents"

(632) *péké longgo kunu woé kai*
turn.away back group wind.around 3SG
"Turns his back on his relatives"

(633) *mbana taka tati no'o fai kai*
go stick XX COM female 3SG
"[And] goes [and] cleaves to his wife"

(634) *soli ebé imu rua dau mbalé sa tebo sa lo wé'é*
CONJ$_2$ 3PL friend.CL two must become SG body$_2$ SG body$_3$ just
"And the two of them must become just one body [and] flesh"

(635) *miu imu rua gha ina éo rupa ngéré adam no'o éva éo muri*
2PL friend.CL two here DEM REL shape$_2$ like Adam COM Eve REL live
'You two here, who are shaped like a living Adam and Eve'

(636) *di lélé sawé éo du'a ngga'é papé*
also$_3$ hear finish REL early master advise
'Have also heard what God has warned'

(637) *du'a miu rapa déi dowa*
self 2PL RECIP rejoice PERF
'You yourselves already love each other'

(638) *ta ola nara até miu no'o ola déi ngai miu*
CONJ NOM desire liver 2PL COM NOM rejoice spirit 2PL
'And the desire of your hearts and the joy of your spirits'

(639) *to'o leka du'a ngga'é pa'a nena leka oné tuka até miu*
arise LOC early master offer create LOC inside belly liver 2PL
'Arises from that which God has laid down inside your hearts'

(640) *du'a kai pai miu tau riké peté negi péré muri*
self 3SG call 2PL make tie fasten strong reach$_3$ live
'He Himself calls you to bind together tightly [and] strongly as long as [you] live'

(641) *lai papa lala du mata boka rapa modha du bopa*
spleen side soft reach die lung RECIP wet reach disappear
'Until death do you part (lit. 'Your spleens grow soft together until you die, your lungs rot together until you disappear')'

(642) *miu mai da gha ina leka nia yésus*
2PL come towards here DEM LOC face Jesus
'You come here before Jesus'

(643) *soli yésus mai dhato da gha ina*
CONJ_2 Jesus come self_2 towards here DEM
'And Jesus himself comes here'

(644) *ngéré kai mbana tau nggua nika leka nua kana*
like 3SG go make ritual marry LOC village pl.name
'Like he goes to give a wedding feast in the village of Cana'

(645) *no'o até kai éo dhasi dhoa kai nara berka miu*
COM liver 3SG REL XX sink 3SG desire bless 2PL
'With his heart which is all-merciful, he wants to bless you'

(646) *soli rina miu dau até papa pawé tuka papa keku*
CONJ_2 beg 2PL must liver side good belly side tender
'And asks that you must love each other tenderly (lit. 'your hearts are good to each other, your bellies are tender to each other')'

(647) *kai péra kai ina no'o soto du'a kai*
3SG teach 3SG DEM COM example self 3SG
'He taught this with His own example'

(648) *ngai ngéré yésus wua mesu geréja kai*
because like Jesus weary fall church 3SG
'Because Jesus loves (lit. 'is weary and falls') his Church'

(649) *péré fonga poké tebo loka ra du'a kai*
reach_3 want throw body_2 spill blood self 3SG
'So much that [he] is willing to throw away His own body [and] spill His own blood'

(650) *rio geréja no'o ra kai*
bathe church COM blood 3SG
'Bathe the Church with His blood'

(651) *wé'é tebo kai léi dhika dowa nia iju léi lina méma*
so.that body$_2$ 3SG completely pure PERF face nose completely shine really
'So that His body is now totally pure [and] His face is completely clean'

(652) *wé'é kai tau déi no'o geréja ina éo saré dhaé*
so.that 3SG make rejoice COM church DEM REL glorious XX
'So that he may rejoice in this Church which is glorious'

(653) *ngai dosa kai dhika dowa raki kai masa dowa*
because sin 3SG pure PERF dirt 3SG clean PERF
'Because His sins have been purified, His filth has been cleaned'

(654) *sala léko mbeja iwa rowa mbé'o réwo*
wrong deviate finish$_2$ NEG PERF$_2$ know other
'All wrongdoings are no longer remembered'

(655) *ngéré ina ata kaki dau wua mesu fai kai*
like DEM person male must weary fall female 3SG
'In this way, a man must love his wife'

(656) *pama tebo kai léi masa*
keep body$_2$ 3SG completely clean
'Keep her body pure'

(657) *no'o ada owi dema dhasi dhoa du bupu muwa*
COM honor respect truly XX sink reach very.old white.hair
'And truly respect [her], have compassion until old age'

(658) *demi miu imu rua di nara muri sara ata kaki / ata fai*
COND 2PL friend.CL two also$_3$ desire live manner person male / person female

'If you two want to live as man / wife'

(659) *ndu ngéré du'a ngga'é papé lawé ina*
follow like early master instruct XX DEM
'According to what God instructed'

(660) *na miu garé sia sai ola nara ngai miu gha ina*
DEM 2PL speak bright IMP NOM desire spirit 2PL here DEM
'So explain the desire of your spirits here'

(661) *leka nia aku dapi ebé ina éo dari tau saksi*
LOC face 1SG CONJ_3 3PL DEM REL stand make witness
'Before me and them, who stand to bear witness'

(662) *deki talu lédhé sai éo aku alé tana*
reach_3 reply clear_2 IMP REL 1SG ask ask_2
'So answer clearly what I ask'

(663) *kau simo lélé ngai du'a kau (naja) éo latu gha ina*
2SG receive hear spirit self 2SG (name) REL EXI here DEM
'Do you, by your own will, take (name) who is here'

(664) *tau fai kau ndu no'o ola ada agama serani*
make female 2SG follow COM NOM honor religion christian
'To be your wife according to the customs of the Christian religion'

(665) *dapi kau tanggo tau paga kai laka kai ada owi no'o wua*
CONJ_3 2SG share.burden make nurture 3SG help 3SG honor respect COM weary
mesu kai
fall 3SG
'And do you accept the responsibility of nurturing her, respecting and loving her'

(666) *soli muri lai papa lala du mata boka rapa modha du bopa*
CONJ_2 live spleen side soft reach die lung RECIP wet reach disappear

'And live until death do you part? (lit. 'your spleens grow soft together until you die, your lungs rot together until you disappear')'

(667) *ho'o aku simo*
yes 1SG receive
"Yes, I do."

(668) *kau simo lélé ngai du'a kau (naja) éo latu gha ina*
2SG receive hear spirit self 2SG (name) REL EXI here DEM
'Do you, by your own will, take (name) who is here'

(669) *tau kaki kau ndu no'o ola ada agama serani*
make male 2SG follow COM NOM honor religion christian
'To be your husband according to the customs of the Christian religion'

(670) *dapi kau jaji tau ada owi wua mesu kai*
CONJ_3 2SG promise make honor respect weary fall 3SG
'And do you promise to respect [and] love him'

(671) *ndu wai kai tulu laka kai*
follow follow_2 3SG help_2 help 3SG
'Obey him [and] help him'

(672) *soli muri lai papa lala du mata boka rapa modha du bopa*
CONJ_2 live spleen side soft reach die lung RECIP wet reach disappear
'And live until death do you part?'

(673) *ho'o aku simo*
yes 1SG receive
"Yes, I do."

(674) *(naja) soli (naja) miu jaji tau paga mara*
(name) CONJ_2 (name) 2PL promise make nurture cherish
'(Name) and (name), you promise to nurture [and] cherish'

(675) *ana lo'o mbeja éo du'a ngga'é mo'o pati miu ngéré iné ema éo*
child small finish$_2$ REL early master FUT give 2PL like father mother REL
ji'é
excellent
'All the children which God will give you as excellent parents'

(676) *uku aja péra pati ebé no'o até dhoa*
rule teach$_2$ teach give 3PL COM liver sink
'Instruct [and] teach them with compassion'

(677) *péré ebé muri mbalé manusia éo pawé*
reach$_3$ 3PL live become human REL good
'So that they live [and] become good people'

(678) *ola déi du'a ngga'é dapi kita ata mbeja*
NOM rejoice early master CONJ$_3$ 1PL.I person finish$_2$
'A joy to God and all us people'

(679) *salib tuha yésus kristus*
cross lord Jesus Christ
'The cross of Lord Jesus Christ'

(680) *kai ina ola negi miu nebu susa dhua*
3SG DEM NOM strong 2PL time unfortunate XX
'This is your strength in times of hardship'

(681) *ola ja miu nebu mila hawo*
NOM shine$_2$ 2PL time dark XX
'Your light in times of darkness'

(682) *ola mété miu nebu muri soli nebu mata*
NOM hope 2PL time live CONJ$_2$ time die
'Your hope during life and during death'

(683) *teké sai salib ina*
hold$_2$ IMP cross DEM
'Hold this cross'

(684) *tau ola nika miu tu'a ngéré su'a maku ngéré watu*
make NOM marry 2PL hard like iron hard$_2$ like stone
'To make your marriage strong as iron, hard as stone'

(685) *iwa bowa iwa beta du mata sai*
NEG untie NEG snap reach die IMP
'Never coming untied, never breaking until death'

(686) *o du'a ngga'é ebé ina riké dowa sa kélé uju sa umu*
EXCLAM early master 3PL DEM tie PERF SG gather.CL bundle.up SG bundle.CL
'O God, they have been tied together into one group, wrapped into one bundle'

(687) *nggénggé sai ana-ana kau gha ina mo'o tebo keta lo ngga*
protect IMP child:PL 2SG here DEM FUT body$_2$ cold body$_3$ cool
'Protect Your children here, that they may be healthy (lit. 'have a cold body and cool flesh (body)')'

(688) *kolo ma'é ro oté ma'é nodé*
head PROH painful brain PROH ache
'May [their] heads never hurt, may [their] brains never ache'

(689) *tau sai ebé tu'a ngéré su'a maku ngéré watu*
make IMP 3PL hard like iron hard$_2$ like stone
'Make them strong as iron, hard as stone'

(690) *sa riwu bhalé talo sa ngasu kébé mona*
SG thousand.CL return cannot SG hundred.CL deflect NEG$_2$
'A thousand cannot turn [them] back, a hundred will not turn [them] away'

(691) *wé'é ebé mera nuwa ké'a nuwa kéna*
so.that 3PL sit age coconut.shell age bowl
'So that they may live happily forever after (lit. 'age of coconut shells, age of bowls')'

(692) *ngé kobé ngé leja*
increase night increase day
'Increase by night, increase by day (Add more nights, add more days)'

(693) *so sai gepa gena*
stretch.out attain feel.around touch
'May they be successful (lit. 'stretch out and reach, feel around for and touch')'

(694) *ngai pu'u kau du'a ngga'é kami éo muri pu'u iwa sepu iwa*
because trunk 2SG early master 1PL.E REL live trunk NEG end NEG
'Because You are our God who lives with no beginning [and] no end'

(695) *amin*
amen
'Amen'

(696) *miu simo dowa sakramén nika*
2PL receive PERF sacrament marry
'You have received the sacrament of marriage'

(697) *miu riké dowa sa kélé uju dowa sa umu*
2PL tie PERF SG gather bundle.up PERF SG bundle
'You have been tied together into one group, rolled into one bunch'

(698) *to'o nebu ina miu dau wora sa wiwi - nunu sa lema*
arise time DEM 2PL must rebuke SG lip - speak$_2$ SG tongue
'From now on you must instruct with one lip, speak with one tongue'

(699) *muri sama-sama*
live together:ADV
'Live together'

(700) *riké ma'é towa bowa dhati ma'é beta data*
tie PROH loosen untie link PROH snap XX
'Be tied [and] do not come loose, link together [and] do not break apart'

(701) *miu jaji sawé*
2PL promise finish
'You have promised'

(702) *jadi pati wai ti'i bili*
become$_2$ give flat give$_2$ round
'So give a plate, serve to the brim'

(703) *pati iwa do lai ti'i iwa do wiki*
give NEG PERF pick.up give$_2$ NEG PERF take
'Give and never take back, grant and never take away'

(704) *pai ma'é talu niu ma'é oé*
call PROH reply invite$_2$ PROH answer
'Do not reply to the call, do not answer the invitation (of others)'

(705) *kasa ma'é langga demi langga kasa osé*
fence PROH exceed COND exceed fence narrow
'Do not cross the fence; if you cross it, the fence will squeeze you'

(706) *kota ma'é dhoa demi dhoa watu pida*
stone.wall PROH sink COND sink stone crush
'Do not take down the stone wall; if you take it down, the stones will crush you'

(707) *péré ma'é ndoré demi ndoré péré piro*
door$_2$ PROH pass COND pass door$_2$ pinch

'Do not pass the door; if you pass it, the door will pinch you'

(708) *nebu ina sawé rowa*
time DEM finish PERF_2
'Now it is finished.'

(709) *sawé iwa ro garé mbeja iwa ro weta*
finish NEG PERF_2 speak finish_2 NEG PERF_2 utter
'Finished, never to be spoken about again - Done, never to be mentioned again.'

(710) *hé du'a ngga'é paga sai soli jaga sai ana-ana kau gha ina leka éo ré'é*
EXCLAM early master nurture IMP CONJ_2 guard IMP child:PL 2SG here DEM LOC REL bad
'Oh God, nurture and protect your children here from evil'

(711) *ju angi kaju aé*
push wind wood water
'Push the wind, wood and water'

(712) *rago béu kola béwa*
expel far chase tall
'Expel it far away, chase it [a] long [distance]'

(713) *ju ghélé ulu joka da ghawa éko*
push up_2 head_2 push_2 towards down_2 tail
'Push up to the top, press down to the bottom'

(714) *laka sai wé'é tedo tembu wésa wonga*
help IMP so.that plant grow sow flower
'Help [them] so that [they] plant [and] grow, sow [and] flourish'

(715) *peni ngé wesi nuwa*
feed.birds increase feed.animals grow_2

'Feed the birds more, fatten the animals'

(716) *tuka ngé kambu wonga*
belly increase womb flower
'Make the belly multiply, their womb flourish'

(717) *tebo keta lo ngga*
$body_2$ cold $body_3$ cool
'Cold body, cool flesh'

(718) *pati sai rahmat kau tau ngé bhondo beka kapa*
give IMP grace 2SG make increase much multiply thick
'Give them Your grace to that [they] increase greatly [and] multiply thickly'

(719) *so ngé dowa ma'é bébo ana*
so increase PERF PROH not.know child
'When [they] have increased, may [they] not forget their children'

(720) *so beka walo temu poké deka*
so multiply $return_3$ $PROH_2$ throw take.care
'When they have multiplied again, may they not cast [it] away'

(721) *paga mara ana péra aja mamo*
nurture cherish child teach $teach_2$ grandchild
'Nurture the children, teach the grandchildren'

(722) *wai papa ndu no'o ada lawa*
$follow_2$ side follow COM honor put.in.order
'To treat each other with honor'

(723) *jaga sai wé'é ebé ndu wai agama kau no'o até ebé léi sawé*
guard IMP so.that 3PL follow $follow_2$ religion 2SG COM liver 3PL completely finish
'Watch that they follow Your religion with their whole hearts'

(724) *ta so'o ngé leka sakramén nika éo ina*
CONJ more increase LOC sacrament marry REL DEM
'And that more and more in this marriage'

(725) *ebé nai ma'é sala tangi tika ma'é sala ténda*
3PL ascend PROH wrong stair leap PROH wrong platform
'May they not climb the steps wrongly, may they not leap onto the platform wrongly'

(726) *ngai pu'u kristus tuha kami*
because trunk Christ lord 1PL.E
'Because Christ is our Lord'

(727) *amin*
amen
'Amen'

5 Bibliography

References

[1] Aoki, Eriko & Satoshi Nakagawa. 1993. *Ende-English dictionary*. Unpublished.

[2] Arka, I Wayan, Fransiscus Seda, Antonius Geland, Yohanes Nani and Ivan Ture. 2011. *A Rongga-English Dictionary with English-Rongga Finderlist*. Jakarta: Atma Jaya University Press.

[3] Arka, I Wayan. 2016. *Bahasa Rongga: Deskripsi, Tipologi dan Teori*. Jakarta: Atma Jaya University Press.

[4] Arndt, Paul. 1933a. *Li'onesisch-Deutsches Wörterbuch*. Ende: Arnoldus-Druckerei.

[5] Arndt, Paul. 1933b. *Grammatik der Ngad'a-Sprache. Verhandelingen van het Koninklijk Bataviaasch Genootschap van Kunsten en Wetenschappen*, LXXII. Bandung: A. C. Nix & Co.

[6] Arndt, Paul. 1961. "Wörterbuch der Ngadhasprache". *Studia Instituti Anthropos*, 15. Ende: Arnoldus-Druckerei.

[7] Astar, Hidayatul, Buha Aritonang, Non Martis, Wati Kurniawati and Ferry Feirizal. 2000. *Monografi Kosakata Dasar Swadesh di Kabupaten Ende*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa Departemen Pendidikan Nasional.

[8] Baird, Louise. 2002. *A Grammar of Kéo: An Austronesian language of East Nusantara*. Canberra: Australian National University (Doctoral dissertation).

[9] Blust, Robert. 2001. "Historical morphology and the spirit world: The *qali/kali- prefixes in Austronesian languages" in Joel Bradshaw and Kenneth L. Rehg (eds): *Issues in Austronesian morphology*. Canberra: Pacific Linguistics.

[10] Blust, Robert. 2002. "The History of Faunal Terms in Austronesian Languages". *Oceanic Linguistics*, Vol. 41, No. 1.

[11] Blust, Robert. 2008. "Is there a Bima-Sumba subgroup?". *Oceanic Linguistics* Vol. 47, No. 1.

[12] Blust, Robert. 2013. *The Austronesian languages, revised edition*. Canberra: Pacific Linguistics.

[13] Blust, Robert and Stephen Trussel. Ongoing. *Austronesian Comparative Dictionary*. http://www.trussel2.com/acd. (Accessed October 19, 2018).

[14] Donohue, Mark. 2003. *Bata Lu'a ca: Palu'e dictionary, electronic edition*. Unpublished.

[15] Donohue, Mark. 2005. "The Palu'e passive: from pragmatic construction to grammatical device." in I Wayan Arka and Malcolm Ross (eds.): *The many faces of Austronesian voice systems: some new empirical studies*. Canberra: Pacific Linguistics.

[16] Donohue, Mark. 2006. "Flores Languages" in Keith Brown and Sarah Ogilvie (eds.): *Concise Encyclopedia of Languages of the World*. Amsterdam: Elsevier.

[17] Esser, Samuel J. 1938. "Talen". in *Atlas van Tropisch Nederland*. Jakarta: Koninklijk Nederlandsch Aardrijkskundig Genootschap.

[18] Fernandez, Inyo Yos. 1990. "Posisi Bahasa Komodo, Rembong, dan Paluqe dalam Kelompok Bahasa Flores". *Linguistik Indonesia* 8.

[19] Fernandez, Inyo Yos. 1996. *Relasi historis kekerabatan bahasa Flores: kajian linguistik historis komparatif terhadap sembilan bahasa di Flores*. Ende: Nusa Indah.

[20] Forth, Gregory. 2016. *Why the Porcupine is Not a Bird: Explorations in the Folk Zoology of an Eastern Indonesian People*. Toronto: University of Toronto Press.

[21] François, Alexandre. 2014. "Trees, waves and linkages: models of language diversification." in Claire Bowern & Bethwyn Evans (eds.): *The Routledge Handbook of Historical Linguistics*: 161-189. New York: Routledge.

[22] François, Alexandre. 2017. "Méthode comparative et chaînages linguistiques: Pour un modèle diffusionniste en généalogie des langues." in Jean-Léo Léonard (ed.): *Diffusion: implantation, affinités, convergence. Mémoires de la Société de Linguistique de Paris*, XXIV: 43-82. Louvain: Peeters.

[23] Fricke, Hanna. 2014. *Topics in the Grammar of Hewa: A variety of Sika in Eastern Indonesia*. Munich: LINCOM.

[24] Greenhill, Simon, Robert Blust and Russel Gray. 2008. "The Austronesian Basic Vocabulary Database: From Bioinformatics to Lexomics". *Evolutionary Bioinformatics*, 4.

[25] Hammarström, Harald, Robert Forkel and Martin Haspelmath. 2018. Glottolog 3.3. Jena: Max Planck Institute for the Science of Human History. (Available online at http://glottolog.org, accessed on October 19, 2018).

[26] Ismail, Mansyur, Muhidin Azis, M. Saleh Yakub, M. Taufik H. and M. Kasim Usman. 1985. *Kamus Bima-Indonesia*. Jakarta: Departemen Pendidikan dan Kebudayaan.

[27] Jonker, Johann Christoph Gerhard. 1896. *Bimaneesch-Hollandsch Woordenboek. Verhandelingen van het Koninklijk Bataviaasch Genootschap van Kunsten en Wetenschappen*, XLVIII. Jakarta: Landsdrukkerij.

[28] Kalyan, Siva and Alexandre François. 2018. "Freeing the Comparative Method from the tree model: A framework for Historical Glottometry" in R. Kikusawa and L. Reid (eds.): *Let's talk about trees: Tackling problems in representing phylogenic relationships among languages*. Osaka: National Museum of Ethnology. 59-89.

[29] Klamer, Marian. 1998. *A Grammar of Kambera*. Berlin/New York: Mouton de Gruyter.

[30] Klamer, Marian. 2002. "Typical Features of Austronesian Languages in Central/Eastern Indonesia". *Oceanic Linguistics*, 41.

[31] Ladefoged, Peter and Ian Maddieson. 1996. *The Sounds of the World's Languages*. Oxford: Blackwell.

[32] Mayer, Mercer. 1969. *Frog, Where Are You?*. New York: Penguin Putnam Inc.

[33] McDonnell, Bradley. 2009. "A Preliminary Description of Ende Phonology". *Journal of the Southeast Asian Linguistics Society*, II.

[34] Onvlee, Louis. 1984. *Kamberaas-Nederlands woordenboek*. Leiden: Koninklijk Instituut voor Taal-, Land- en Volkenkunde.

[35] Pareira, M. Mangalangi and E. Douglas Lewis. 1998. *Kamus Sara Sikka-Bahasa Indonesia*. Ende: Nusa Indah.

[36] Ross, Malcolm. 1988. *Proto-Oceanic and the Austronesian languages of western Melanesia*. Canberra: Pacific Linguistics.

[37] Schmidt, Johannes. 1872. *Die Verwandtschaftsverhältnisse der indogermanischen Sprachen*. Weimar: Hermann Böhlau.

[38] Simons, Gary F. and Charles D. Fennig (eds.). 2017. *Ethnologue: Languages of the World, Twentieth Edition*. Dallas, Texas: SIL International. Online version: http://www.ethnologue.com. Accessed on October 19, 2018.

[39] Schleicher, August. 1853. "Die ersten Spaltungen des indogermanischen Urvolkes." *Allgemeine Monatsschrift für Wissenschaft und Literatur*, pp. 786-787.

[40] Verheijen, Jilis. 1967-1970. *Kamus Manggarai: Manggarai-Indonesia & Indonesia-Manggarai*. The Hague: Martinus Nijhoff.

[41] Verheijen, Jilis. 1982. "Komodo: Het Eiland, Het Volk, en De Taal". *Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde*, 96. The Hague: Martinus Nijhoff.

[42] Wurm, Stephen and Shiro Hattori. 1981. "Language atlas of the Pacific area". *Pacific Linguistics, Series C*, 66 & 67. Canberra: Australian National University.