



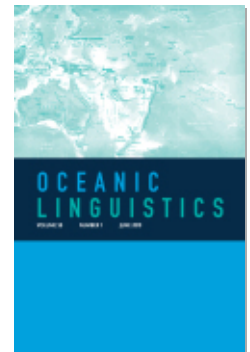
PROJECT MUSE®

Reintroducing Welaun

Owen Edwards

Oceanic Linguistics, Volume 58, Number 1, June 2019, pp. 31-58 (Article)

Published by University of Hawai'i Press



➔ For additional information about this article

<https://muse.jhu.edu/article/730687>

Reintroducing Welaun

Owen Edwards

LEIDEN UNIVERSITY AND LANGUAGE & CULTURE UNIT (UBB), KUPANG

In this article I provide an overview of Welaun, an Austronesian language of central Timor. Despite previous documentation, Welaun is mostly unknown to the scientific community. Based on data gathered during original fieldwork, I show that Welaun is a distinct language according to both linguistic and social criteria. I investigate the historical phonology of Welaun and show that it, Kemak, Tokodede, and Mambae form a Central Timor subgroup apart from other languages of the region. I also provide an initial description of the phonology and morphology of Welaun along with a glossed text and two wordlists.

1. INTRODUCTION.¹ In this paper I introduce Welaun (a.k.a. Bekais/Wekais) to the scientific community. Welaun is an Austronesian language of central Timor, straddling the border between Timor-Leste and the Indonesian regency of Belu.

Welaun has been described by Hull (2003) and a 189 page Welaun–Tetun dictionary written by a native speaker has been published (da Silva 2012). However, Hull (2003) is written in Tetun Dili and da Silva (2012) is not widely available outside of Timor. These facts, combined with a misidentification of Welaun as a Tetun dialect by de Almeida (1976) which was followed by Hull (1998) (but self corrected in Hull [2003]), means that the status and existence of Welaun remains unclear in the scientific community. At the time of writing neither Glottolog (Hammarström et al. 2018) nor Ethnologue (Simons and Fennig 2018) had entries for Welaun in any part of their databases.

Based on original fieldwork in Indonesia, I show that Welaun is a distinct language. This is evident in both its lexicon and morphology. Based on shared sound changes it can be classified as part of the Central Timor subgroup alongside Kemak, Tokodede, and Mambae. However, Welaun has a number of sound changes which are also found in Tetun.²

1. I would like to thank the Language and Culture Unit (UBB) in Kupang with whose support the data in this article were gathered, as well as the financial support of an Endeavour Australia Fellowship. I would also like to thank two anonymous reviewers who provided comments on earlier drafts. All errors or deficiencies remain my own.

Publication of this paper was supported by the VICI research project “Reconstructing the past through languages of the present: the Lesser Sunda Islands”, funded by the Netherlands Organisation for Scientific Research, project number 277-70-012.

2. Throughout this paper all uses of the name *Tetun* should be taken to refer to ISO 693-3 [tet]. Tetun Dili (ISO 693-3: [tdt]) is referred to as *Tetun Dili*.

The structure of this paper is as follows. In section 2, I provide an initial overview of the sociolinguistic context/profile of Welaun. In section 3, I compare the lexicon of Welaun with other regional languages. This shows clearly that Welaun is a separate language. In section 4, I provide a brief overview of the phonology and morphology of Welaun. I conclude in section 5 with an overview of the historical phonology and subgrouping of Welaun. Two appendices are provided which give a glossed Welaun text and wordlists respectively.

2. SOCIOLINGUISTIC PROFILE. Welaun is spoken in an area which straddles northern parts of the border between Indonesia and Timor-Leste. At the beginning of December 2017, while conducting a survey of Kemak in the Indonesian regency of Belu, I collected a 270-item Welaun wordlist and text from the hamlet (*kampung*) of Oele'u in Bauhoo village (*desa*).

Having been alerted to the presence of Welaun in Belu, I returned in January 2018 to collect additional data. I collected a 270-item wordlist and two texts from the hamlet of Mahein in Bauda'ok <Baudaok> village. I also went back to Oele'u and collected a 630-item wordlist and worked through the text I had collected the first time.³

In addition, I went to Haekesak and met with Fillipe Pereira and Dominggus Pereira (both Welaun speakers) and collected a list of places where Welaun is spoken. This is the longest list of locations for Welaun I collected. Other reports I collected are consistent with this list, but it does not seem to be complete. This list is given in table 1 with village (*desa/suco*) locations and

TABLE 1. LOCATIONS OF WELAUN

Country	Village (<i>desa/suco</i>)	Hamlet (<i>dusun/aldeia</i>)	Notes
Timor-Leste	Leohito	Aiasa	
		Mohaak	
		Ferik Katuas	
		Falohai	
		Raiulun	Raiulun dialect
		Balibo Vila	Fatululik
Indonesia	Cowa		mostly Tetun [tet]
			mostly Welaun
			less Welaun
			mostly Tetun [tet]
	Tohe	Kotafoun A	mostly Tetun [tet]
		Kotafoun B	mostly Tetun [tet]
	Asumanu	Balah	
		Niinluli	
	Sarabau	Motamoru	Raiulun dialect
	Bauhoo	Oele'u/Onosorun	Raiulun dialect
Sakalo'on			
Bauda'ok	Mahein	Raiulun dialect	
Tulakadi	Salore		
Silawan	Weehaas [†]	only a small number of speakers	

[†] My consultants reported that Weehaas was in Tulakadi, but it appears to be in neighboring Silawan.

3. This data has been archived with PARADISEC at <http://catalog.paradisec.org.au/repository/OE8>

MAP 1. LOCATIONS OF WELAUN



hamlet (*dusun/aldeia*) locations shown in Map 1. Not all hamlets have been located.

Leohito in Timor-Leste is identified as the Welaun center (*induk*). Speakers report three main migrations of people from Leohito into Indonesia: (i) early 20th century fleeing conflict with the Portuguese, (ii) during the Second World War fleeing conflict with the Japanese, (iii) at the turn of the 21st century during the troubles leading to independence in Timor-Leste. My consultants in Oele'u and Mahein reported that their ancestors came during the second wave of migration.

Fillipe Pereira and Domingus Pereira identified the speech of Raiulun in southern Leohito as a distinct dialect. They reported that this difference is mainly in accent and pronunciation. Speakers in Motamoru, Oele'u, and Mahein are said also to speak the Raiulun dialect.

Comparison of my Oele'u and Mahein data with that of da Silva (2012) does not reveal any significant lexical differences. There is one morphological difference in the possessive paradigm which may be part of this dialect difference (see section 4.2.1).

Welaun speakers think of their speech as a distinct language. While some Kemak speakers claim that Welaun is a dialect of Kemak (which they call Kemak Belau — Kemak has no /w/, hence $w > b$), Welaun speakers deny this. They say that while some words in Kemak and Welaun are similar, Kemak speakers do not understand Welaun. That Welaun is not a dialect of Kemak

is supported by the fact that only about half of all words on a basic wordlist are cognate (see section 4, table 3).

I played the Welaun text in Appendix A to a group of Kemak speakers and this showed that they had only a limited comprehension of parts of it. Their limited comprehension was probably aided by the fact that many Welaun words are the same as Tetun (due to parallel development and/or borrowing). All Kemak (and Welaun) speakers I met in Indonesia were bilingual in Tetun.⁴

2.1 POPULATION AND VITALITY. Williams-van Klinken and Williams (2015), summarizing data from the 2010 census, give 3,887 speakers of Welaun in Timor-Leste. Welaun speakers are the majority in the village (*suco*) of Leohito and are minorities in Cowa and Balibo Vila. The 2015 Timor-Leste census gives the number of speakers of Welaun as 4,075 ([Direcção-Geral de Estatística 2016](#)).

According to the 2010 census villages in Indonesia where Welaun speakers are reported to live have a total population of 13,449 ([Badan Pusat Statistik 2012](#)).⁵ Given that Welaun speakers are a minority in all these villages, at a guess, it is unlikely that they contribute more than 1,500 to the total population of those who speak Welaun. I would thus estimate that there are approximately 5,575 Welaun speakers in both Timor-Leste and Indonesia.

In both Oele'u and Mahein where I collected my Welaun data the language was in vigorous use. In Oele'u I observed that children were acquiring the language.⁶ Whatever the exact reasons for the increase in the reported number of speakers in the Timor-Leste census, this indicates that Welaun remains vigorous in Timor-Leste.

2.2 LANGUAGE NAME(S). All Welaun people I spoke to called their language *Welaun*. When asked, they acknowledged the name *Bekais* but always identified it as a Tetun exonym. Hull (2003) also notes in his abstract that *Welaun* is the name used by the speakers themselves.

The name *Wekais* in da Silva (2012) was not used by any of my consultants. It probably results from analysis of *Bekais* as a Welaun word, based on the etymology proposed by da Silva (2012:viii) whereby the first part comes from Welaun *wee* 'water' (Tetun *bee* ~ *wee*) and the second part from Welaun *kais* 'dig with claw (i.e. of birds)', with reference to the ancestors of the Welaun people finding water where birds had been looking for food.

4. One Welaun speaker in Oele'u explained to me that just as Indonesian is the unifying language (*bahasa persekutuan*) of the country, Tetun is the unifying language of Belu regency. Near universal Tetun bilingualism has doubtlessly contributed to the presence of Kemak and Welaun in this region not being fully appreciated.

5. The breakdown of the population of villages with Welaun speakers is as follows: Tohe 5,517, Silawan 3,313, Asumanu 1,770, Tulakadi 1,043, Sarabau 661, Bauhoo 610, Bauda'ok 535.

6. I was not in Mahein long enough to know whether Welaun was being acquired by children there. I strongly suspect that it is based on the fact it was in vigorous use among adults.

3. LEXICAL COMPARISON. The lexicon of Welaun clearly shows that it is a distinct language. An approximation of the level of lexical difference between Welaun and other regional languages can be gauged by a lexicostatistical comparison.

I compared the 270-item wordlists I collected in Oele'u and Mahein (given in [Appendix 2](#)), as well as the equivalent wordlist extracted from da Silva (2012) with equivalent lists in other languages of the region. These other languages were: two varieties of Kemak (Leolima from Umaklaran village and Lemia from Sadi village), Northwest Mambae, South Mambae, Tokodede, Tetun [tet], Tetun Dili [tdt], Kusa-Manea (Meto cluster), and Kotos Amarasi (Meto cluster).⁷

I entered these lists into WordSurv and made three different comparisons: similarity, cognacy, and identical forms. The results of these comparisons are given in table 2, table 3, and table 4, respectively.

TABLE 2. LEXICAL SIMILARITY

Welaun (Oele'u)											
89	Welaun (Mahein)										
93	91	Welaun (Leohito)									
41	40	41	Kemak (Leolima)								
40	39	42	88	Kemak (Lemia)							
33	33	37	46	47	Northwest Mambae						
32	32	35	45	45	67	South Mambae					
28	34	37	50	50	61	49	Tokodede				
57	55	58	39	38	32	33	37	Tetun [tet]			
51	50	55	38	37	37	34	38	83	Tetun Dili [tdt]		
27	27	27	26	24	26	24	25	29	27	Meto (Kusa-Manea)	
25	24	26	24	23	25	24	24	26	25	76	Meto (Kotos Amarasi)

TABLE 3. TRUE COGNATES

Welaun (Oele'u)											
90	Welaun (Mahein)										
93	92	Welaun (Leohito)									
51	51	52	Kemak (Leolima)								
51	49	51	89	Kemak (Lemia)							
45	45	48	53	54	Northwest Mambae						
43	44	46	51	50	68	South Mambae					
37	45	49	55	55	67	51	Tokodede				
63	62	65	51	50	42	44	45	Tetun [tet]			
57	57	62	48	46	46	43	46	84	Tetun Dili [tdt]		
42	43	43	37	37	34	33	34	44	40	Meto (Kusa-Manea)	
38	38	41	34	34	33	31	33	39	37	78	Meto (Kotos Amarasi)

7. Data for these wordlists were compiled from the following sources: Kemak, Kusa-Manea and Kotos Amarasi (own field notes), Northwest Mambae (Fogaça 2017), South Mambae (Fogaça 2017, as well as Grimes et al. 2014), Tokodede (Klamer 2002b), Tetun [tet] (Morris 1984), Tetun Dili [tdt] (Williams-van Klinken 2015).

TABLE 4. IDENTICAL FORMS

Welaun (Oele'u)										
81	Welaun (Mahein)									
87	85	Welaun (Leohito)								
16	15	16	Kemak (Leolima)							
16	16	17	78	Kemak (Lemia)						
12	12	13	24	26	Northwest Mambae					
13	14	15	23	22	40	South Mambae				
8	10	10	24	26	28	21	Tokode			
31	30	32	16	14	13	14	9	Tetun [tet]		
26	26	30	16	15	18	18	10	61	Tetun Dili [tdt]	
8	8	8	6	5	4	5	4	8	8	Meto (Kusa-Manea)
11	7	8	4	3	4	3	5	6	6	51 Meto (Kotos Amarasi)

For calculating similarity, the number of word pairs with equivalent glosses and sufficiently similar forms was summed. This excluded non-similar true cognates, such as Welaun *kiis-aat* Kusa-Manea *nisi-f* ‘teeth’ (both from **ɲisi* ‘grin, show the teeth’). For cognacy, the number of word pairs which were ultimately from the same etymon were summed. Thus, Welaun *kiis-aat* and Kusa-Manea *nisi-f* were not counted as lexically similar, but they were counted as cognate. For identical forms, word pairs which were phonemically identical ignoring productive morphology were summed.

While the threshold for language and dialect needs to be calibrated separately for each study and region (J. Grimes, 1988, 1995), Welaun relates to other languages in the region significantly below the range of 70-80% used in other studies as indicative of different languages (Dyen 1965, Smith 1984, Grimes and Grimes 1987).

Welaun shows the pattern of sporadic convergence with Tetun [tet] (Simons, 1977:122), which is most likely inflated due to contact and bilingualism. It is in the range of 58% lexical similarity, and 65% cognate. Indeed, nearly a third of the vocabulary on my 270-item wordlist is identical in both Welaun and Tetun. This is probably due to many sound changes being shared by each of these languages with borrowing from Tetun also playing a role.

The next closest language to Welaun after Tetun is Kemak, in the range of 41% lexical similarity and 52% cognate. The low amount of vocabulary shared between Welaun and Kemak shows clearly that Welaun is not a dialect of Kemak. This is consistent with the opinions of Welaun speakers but goes against the claims of some Kemak speakers.

4. SKETCH OF WELAUN PHONOLOGY AND MORPHOLOGY. In this section I describe the most salient points of Welaun phonology and morphology which can be gauged from my data.

4.1 PHONOLOGY. Welaun has thirteen consonants: /t k ʔ b d m n f s h r l w/, and five vowels: /i e a o u/. The labio-velar glide /w/ is realized as a bilabial

glide [β] before the front vowels /e/ and /i/. This realization also occasionally occurs before /a/.⁸ As is typical of many languages in this region, /t/ is dental [t̪] while /d/ is apical [d], or occasionally even slightly retroflex [ɖ].

In Mahein, but not in Oele'u, word final /n/ was often realized as velar [ŋ]. Three examples are *iluk-aan* [ʔilʊ'kaŋ] 'nose', *kiis-aan* [ki:'saŋ] 'teeth', and *ikim-aan* [ʔiki'maŋ] 'chin'. In careful speech, such as repeating a word being elicited, this was often changed to final [n].

Welaun has no /p/ phoneme in native words. While Hull (2003:57) claims that /p/ in borrowings is assimilated as /b/, there is no evidence in my data or da Silva (2012) to support this. The only example of assimilation of /p/ is *saʔi* 'cow' from Malay *sapi*. Otherwise, it is unassimilated, as in *dapur* 'kitchen' from Malay *dapur*.

The usual realizations of the front and back vowels is as lax [ɪ ɛ ɔ ʊ]. The back rounded vowel is usually raised to mid-high [o] before another high vowel. Thus *lobu* → ['lobʊ] 'mud', *koʔi* → ['koʔi] 'scratch', and *oi* ['ʔoi] 'sword grass'. The mid front vowel is also sometimes raised before high vowels, though much less so than /o/. Two examples are *enu* → ['ʔɛnʊ] 'drink' and *saresu* → [sə'resʊ] ~ [sə'resʊ] 'work'.

The low vowel /a/ is almost always front [a], [a̠], [æ] and central realizations as [ɐ] are only common antepenultimate (third last) syllables. The fronting of /a/ is most prominent word finally in closed syllables, particularly in the sequence *aaC* which occurs in several suffixes (see section 4.2). Three examples are *fulan* → ['fulaŋ] 'moon', *tikan* → ['tikæŋ] 'year', and *oet-aat* → [ʔɛ'tæ:t̪] 'vein'. Such fronting seemed much more prominent in Oele'u than in Mahein.

Vowel initial words usually begin with a glottal stop in isolation, after a pause, and phrase initially. There does not seem to be a contrast between vowel initial and glottal stop initial words. Two examples of glottal stop insertion are: *isa* ['ʔisa] 'one' and *oa-t* ['ʔɔa:t̪] 'child'.

4.1.1 Phonotactics. Welaun has a highly constrained word structure. Words are canonically disyllabic and built off a CV syllable structure. All surface consonant clusters within a single morpheme can be analyzed as resulting from vowel deletion. The voiced plosives /b d/, the glottal consonants /h ʔ/, and the glide /w/ do not occur word finally. All other consonants are attested word finally.

Primary stress is assigned to the penultimate (second last) vowel of a word with secondary stress on every second vowel to the left. Stress is realized primarily by increased pitch and length on that vowel.

Sequences of two identical vowels /ii/, /ee/, /aa/, /oo/ and /uu/ are usually each realized as a single phonetically long vowel. Final /aa/ is often reduced to half long [aː] or occasionally short [a]. However, such reduced double vowels still bear primary stress, as expected for penultimate vowels.

Antepenultimate (third last) vowels and all pre-tonic (before stress) vowels are usually reduced and often entirely deleted in Welaun. This applies to both

8. The glide /w/ does not occur before the back vowels /o/ or /u/ in my data.

single morphemes as well as compounds/phrases and affects all vowels. A selection of examples is given in table 5. When an antepenultimate vowel is deleted the preceding consonant is often lengthened.

TABLE 5. REDUCTION OF ANTEPENULTIMATE VOWELS

Underlying	Phonetic	Gloss
sakulu	[ˈs:kolo]	‘ten’
sakoon	[ˈs:kɔ:n]	‘rainbow’
kabual	[ˈkboal]	‘round’
ulit-aat	[,ʔolˈtæːt]	‘skin’
meʔitaan	[,mɛʔˈtæːn]	‘forest’
	[,mɛʔˈtæːn]	
hinanaan	[hɪˈn:æːn]	‘day’
hare fulin	[,harɔˈfɔlm]	‘rice grain head’
hare wee	[,harˈβeː]	‘wet rice field’
loro mata-n	[,lɔrˈmaʔan]	‘sun’
sarai moli-aan	[səˈrajmɔliˈaːn]	‘livestock’ (lit. living things)
anu kulaʔaan	[ʔanˈkɔlaˈʔaːn]	‘young person’
	[ʔaŋˈkɔlˈʔaːn]	

Antepenultimate vowels which occur in the environment VC_(C)V are particularly vulnerable to such reduction. Vowels which occur before two consonants or vowels which occur after another vowel are not usually reduced. An example of each is *usan milis* [,ʔusaŋˈmɪlɪs] ‘drizzle’ and *oet-aat* → [ʔɛˈtæːt] ‘vein’.

The final vowel of the pronoun *haʔu* ‘1SG’ occasionally metathesizes with the glottal stop when unstressed. An example is *haʔu ta loka* ~ *hauʔ ta loka* ‘I don’t hear’. More commonly the glottal stop is simply deleted. Two examples are *haʔu hanan* ~ *hau hanan* ‘I braid’ and *haʔu nele* ~ *hau nele* ‘I fly’. In all such examples the form *haʔu* is prescriptively correct.

4.2 MORPHOLOGY. Morphologically, Welaun occupies a position intermediate between languages to the west such as the Meto cluster which are morphologically rich and those to the east such as South Mambae which are highly isolating. Welaun has possessive morphology, verbal agreement prefixes, as well as a dependent suffix.

4.2.1 Possession. Welaun has two kinds of possessive constructions. The first construction is typically used with body parts and is formed by the possessed noun taking a suffix which agrees with the possessor.

In both Oele'u and Mahein I elicited partial possessive paradigms for certain body parts. The suffix *-n* is used when the possessor is singular and the suffix *-t* when the possessor is plural. I did not elicit the form with a third person plural

possessor, but both da Silva (2012) and Hull (2003) agree that the suffix *-n* occurs in this slot.

When these suffixes attach to a consonant final stem they have the allomorphs *-aan* and *-aat* respectively. Because stress is assigned to every second vowel to the left, the result of the allomorph with a double vowel is that stress falls on the same syllable as it would on the bare stem without any suffix.

The possessive paradigms I elicited are given in table 6 alongside paradigms from da Silva (2012) and Hull (2003). Both these authors give the suffix *-k* for 1SG possessors while I have *-n*. Use of *-n* rather than *-k* for 1SG possessors may be a feature of the Raiulun dialect.

TABLE 6. POSSESSIVE SUFFIXES

		My field-notes		da Silva (2012)		Hull (2003)
1SG	haʔu	lika-n	iluk-aan	lika-k	iluk-aak	mata-k
2SG	oo	lika-n	iluk-aan	lika-n	iluk-aan	mata-n
3SG	nia	lika-n		lika-n	iluk-aan	mata-n
1PL.EXCL	ami	lika-t		lika-t	iluk-aat	mata-t
1PL.INCL	ita	lika-t	iluk-aat	lika-t	iluk-aat	mata-t
2PL	emi/imi [†]	lika-t				mata-t
3PL	sila			lika-n	iluk-aan	mata-n
		‘ear’	‘nose’	‘ear’	‘nose’	‘eye’

[†] Both emi and imi were recorded for the 2PL pronoun in Oele'u but only imi in Mahein. Da Silva (2012) and Hull (2003) give only imi.

The other possessive construction in my data consists of a pronoun possessor with final *-n* followed by a noun with no suffix. An example from my data is *haʔu-n kaʔa* ‘my sister-in-law’. Whether this construction also occurs with a nominal (rather than pronominal) possessor is currently unknown. Similarly, whether there is a different frame for vocative or referential kin terms cannot be determined based on my limited data.

From a comparative point of view, the genitive suffix *-n* in Welaun has generalized from the slot of 3SG to all other singular persons as well as 3PL, while the suffix *-t* has generalized from 1PL.INCL to plural persons, except 3PL.

These patterns of generalization are different from those seen in most known varieties of Kemak. In Kemak the original 3PL genitive suffix *-r* or *-rV* has generalized to all plural persons while the singular persons each have distinct forms. As an example, Kemak Saneri has the following forms where *mata-* is ‘eye’: 1SG *mata-k*, 2SG *mata-m*, 3SG *mata* (no suffix), 1PL.EXCL *ami mata-r*, 1PL.INCL *ita mata-r*, 3PL *romo mata-r* (data from my own field-notes).

4.2.2 Dependent suffix. The suffix *-aan* occurs on stems which are semantically dependent on another element. This suffix is almost certainly derived ultimately from the *-aan* allomorph of the singular genitive suffix (originally only 3SG), though dependent *-aan* also occurs on vowel final stems.

There are two main uses of this suffix in my data. Firstly, it occurs on words which modify a head. Examples of a number of phrases in which the second part takes the suffix *-aan* are given in table 7.

TABLE 7. PHRASES WITH DEPENDENT *-aan*

Head	Modifier			Phrase
roo	'ship'	+	nele 'fly'	→ roo nele-aan 'aeroplane'
sarai	'thing'	+	moli 'live, life'	→ sarai moli-aan 'livestock'
lai	'earth'	+	botuk 'hole'	→ lai botuk-aan 'hole in ground'
iluk-aat	'nose'	+	botuk 'hole'	→ iluk-aat botuk-aan 'nostril'
lai	'earth'	+	lahuk 'powder'	→ lai lahuk-aan 'dust'
ai fuan	'fruit'	+	dodok 'rotten'	→ ai fuan dodok-aan 'rotten fruit'
uma	'house'	+	luli 'sacred, taboo'	→ uma luli-aan 'sacred house'
anu	'person'	+	badak 'short'	→ anu badak-aan 'short person'

Not all modifiers take this suffix, and of those that do there is sometimes variation in whether it occurs or not. Thus, I have *taku lai* 'peanuts' from *taku* 'bean' and *lai* 'earth, ground' but never **taku lai-aan*. Similarly, both *taku naruk* and *taku naruk-aan* from *taku* 'bean' and *naruk* 'long' were given (by the same consultant) for 'long beans' (*Vigna unguiculata*). The difference between the two could perhaps be translated in English as 'long beans' (*taku naruk*) and 'beans which are long' (*taku naruk-aan*).

Additionally, there are some forms which have only been attested with final *aan*. Such words include *lamaraan* 'inside' and *tasa^aaan* 'ripe'. For both these words da Silva (2012) does not give forms without final *aan*.

The other main use of this suffix is as a nominalizer. Examples include *ahenu* 'hang around the neck' → *ahenu-aan* 'necklace', *aan* 'eat' → *aan-aan* 'cooked rice', and *bukar* 'wounded' → *bukar-aan* 'wound'. Indeed, nominalization may be the primary function of this suffix with the attributive use illustrated above marking various types of verbs as functioning within a nominal phrase.

Da Silva (2012) gives many other examples of words with the suffix *-aan* (transcribed *-án*) which he usually translates with "*ne'ebé ...*" 'that which is ...'. Examples include *abusin* 'leave behind' → *abusin-aan* 'that which is left behind', *badak* 'short' → *badak-aan* 'that which is short', and *dale* 'speak' → *dale-aan* 'that which was spoken'.

4.2.3 Verb prefixes. I have found two different kinds of verbal prefixes in my Welaun data: subject agreement prefixes and the causative prefix *a-*. I discuss each in turn. Welaun appears to be unique among the languages of the Central Timor subgroup (also including Kemak, Tokodede and Mambae) in having productive verb agreement.

There are two kinds of verbs which take agreement prefixes in Welaun: all vowel initial verbs and some *h* initial verbs. The following prefixes occur: 1SG *k-*, 2SG *m-*, 3SG/3PL *n-*. When an *h* initial verb takes an agreement prefix,

this prefix replaces the initial *h*. Examples of verbs with agreement prefixes are given in table 8. The Welaun agreement prefixes are the same as those found in Tetun Suai (van Klinken 1999:173) and Tetun Foho, the variety with which Welaun has the most contact.

TABLE 8. VERB AGREEMENT

		#V 'from'	#aC 'know'	#h 'stay'	#h 'sniff'	∅-C 'die'
1SG	haʔu	k-osi	k-alaa	k-ein	k-oran	mate
2SG	oo	m-osi	m-alaa	m-ein	m-oran	mate
3SG	nia	n-osi	n-alaa	n-ein	n-oran	mate
1PL.EXCL	ami	osi	alaa	hein	horan	mate
1PL.INCL	ita	osi	alaa	hein	horan	mate
2PL	emi/imi	osi	alaa	hein	horan	mate
3PL	silá	n-osi	n-alaa	n-ein	n-oran	mate

Two different vowel initial verb stem shapes are attested. Firstly, there are disyllables for which /e a o/ are all attested as the initial vowel. Examples include *enu* 'drink', *ali* 'dig', and *osi* 'from'.

Secondly, there are trisyllables which begin with /a/ followed by a consonant. In many cases the initial vowel of such roots may be a historic prefix, though is no longer analyzable as such because the remainder of the stem does not occur independently. Examples include *awai* 'dry in the sun', *alaa* 'know', and *anilu* 'spit'.

I have so far identified ten *h* initial verbs which take agreement prefixes. Most are extracted from da Silva (2012): *hasu* 'scoop', *hesa* 'kill', *hesaʔat* 'commit suicide' (first element from *hesa* 'kill'), *heu* 'eat (esp. vegetables)', *hili* 'choose', *hoar* 'eat (esp. meat)', *horan* 'smell, sniff', *hosa* 'hit', *hoʔi* 'want', and *huʔu* 'wear'.⁹

Such verbs contrast with *h* initial verbs which do not take agreement, such as *hanan* 'braid, plait' (among many others), for which I collected the phrase *haʔu hanan* 'I plait/braid'.

Welaun has a causative prefix *a-* which attaches to both nouns and verbs to derive causative verbs. This prefix is highly productive. A number of examples are given in table 9. This causative prefix functions very similarly to South Mambae *a-* or Tetun *ha-*.

Some causative derivations also have an additional final *n*, though it is unclear whether this should be analyzed as a productive suffix. Because the causative prefix *a-* is vowel initial, derived causative verbs obligatorily take agreement prefixes.

9. Additionally, the verb *-osi* 'from' may belong to this second class in Mahein. In Mahein the citation form without any pronoun was given as *hosi*. I also collected the Mahein sentences *haʔu k-osi likal mai* 'I've come from (my) field' and *oo m-osi bee?* 'Where are you coming from?'. However, I did not collect any examples of the form *hosi* in a sentence, and the use of this form as the citation form may be due to influence from Tetun *hosi* 'come'.

TABLE 9. CAUSATIVE VERBS

	Base		Causative	
'short'	badak	→	a-badak	'shorten'
'big'	bein	→	a-bein	'enlarge'
'small'	disi	→	a-disi	'shrink, make small'
'hot'	fanas	→	a-fanas	'heat up'
'good'	loil	→	a-loil	'make good, improve'
'go'	laʔo	→	a-laʔo	'send, make go'
'empty'	maman	→	a-maman	'empty out'
'die'	mate	→	a-maten	'kill, make die'
'dry'	mara	→	a-maran	'dry out'
'fall'	mout	→	a-mout	'drop'
'cold'	suman	→	a-suman	'cool down'

5. HISTORICAL PHONOLOGY. The historical phonology of Welaun is somewhat obscured by the fact that it has undergone a number of sound changes which are also found in Tetun, as well as the fact that Welaun, like other languages of the region, has borrowed from Tetun. However, a careful examination of the historical phonology focusing on Welaun words which are not identical to Tetun, shows that Welaun subgroups with Kemak, Mambae, and Tokodede.¹⁰

The most important sound change shared by these languages is the change of Proto-Malayo-Polynesian (PMP) *ŋ to an oral stop, either *g* or *k*. In addition, a number of other changes, while found more widely in the region, can be assigned to Proto-Central Timor and thus support the validity of this subgroup. These changes are: *b > *f, *map > *mp, *d > *r, and *j > *l.

The sound correspondences in Welaun and other Central Timor languages are shown in table 10. Sound correspondences for which I have so far identified only one clear unambiguous example are given in parentheses. No clear reflexes of words with *aya have so far been attested in Welaun.

All reconstructions in this section come from Blust and Trussel (ongoing) and are at the level of PMP unless otherwise indicated. Glosses are for the Welaun reflexes, with semantic changes indicated where necessary.

Throughout this section Kemak data are from Leolima Kemak of Umaklaran village and come from my own field notes. Tokodede data are from Klamer (2002b). Northwest Mambae data are from Railaco sub-district and South Mambae data are from Letefoho village in the sub-district of Same. Mambae data are from Fogaça (2017) and Grimes et al. (2014). Note that many Mambae forms show final CV > VC metathesis when compared with PMP etyma.

10. While Hull (2003) provides an overview of Welaun historical phonology his conclusions are suspect as he uses many reconstructions that have never been previously proposed. He also does not systematically distinguish between native words and words which may be Tetun loans. The data in Hull (2003) also differs from my own data in several respects. The reasons for these discrepancies are not clear.

TABLE 10. CORRESPONDENCES IN CENTRAL TIMOR

PMP		Welaun	Kemak	Tokodede	NW. Mambae	C. Mambae	S. Mambae
*p	#_	h	p	p	p	f	p>f
	V_V	(b)	p	p	p	f	p>f
*ma-p		f	b	b	b	b	b
*t		t	t	t	t	t	t
*k	#_	∅	∅	k	∅	∅	∅
	V_V	?	?/∅	k	k/∅	k/∅	k/∅
*q		∅	∅	∅	∅	∅	∅
*b	#_	f	h	h	h	h	h
	V_V	h	h	h	h	h	h
*d		l	r	r	r	r	r
*g		(k)	(∅)				
*j [g]		l/r	l	l	l	l	l
*z [dʒ]		s	s	s/r	s	s	s
*n		n	n	n	n	n	n
*ñ		(n)	(n)				(n)
*ŋ		k	g	g	g	k	-g-, -k
*s		s	s	s	s	s	s
*h		∅	∅	∅	∅	∅	∅
*R [r]		∅	∅	∅	∅/r	∅/r	∅/r
*r [r]		(r)	(r)				
*l		l	l	l	l	l	l
*w		w	b	∅	∅	∅	∅
*i		i	i	i	i	i	i
*ə		o	e	e	e	e	e
*a		a	a	a	a	a	a
*u		u	u	u	u	u	u
*aw#		o	o	o	o	o	o
*ay#		e	e	e	e	e	e
*uy#		i	i	i	i	i	i
*yu#		i	i	i	i	i	i

The PMP consonants *t, *m, *n, *s, *l, *w, are retained unchanged in Welaun and *q, *h and *R [r] have been lost in all environments. An example of each is given in table 11. PMP final consonants are usually retained in Welaun.

The remainder of this section proceeds as follows. In section 5.1, I discuss the sound changes which provide evidence that Welaun belongs to the Central Timor subgroup. This is followed by a discussion of the other consonants in section 5.2 and vowels in section 5.3.

5.1 CHANGES DEFINING CENTRAL TIMOR. There are five changes which define the Central Timor subgroup: *ŋ > *g, *b > *f, *map > *mp, *d > *r, and *j > *l. Of these, the change of the velar nasal to a plosive is most important as it is not commonly found in other languages of the region.¹¹

11. The Meto cluster in western Timor has sporadic examples of Proto-Meto *ŋ > k. However, PMP *ŋ is almost always reflected as n in the Meto cluster.

TABLE 11. CONSONANTS UNCHANGED OR LOST IN WELAUN

PMP	Welaun	gloss	PMP	Welaun	gloss	PMP	Welaun	gloss
*taqi	tee	'feces'	*maqudip	moli	'live'	*nakaw	anaʔo	'steal'
*batu	fatu	'stone'	*qatiməla	tamola	'flea'	*pənuq	honu	'full'
*kulit	ulit-aat	'skin'	*ənəm	inam [†]	'six'	*baŋun	fakun	'wake'
*siku	siʔu-t	'elbow'	*luhəq	luu-t	'tears'	*wakaR	ai waʔan	'roots'
*ma-qasu	masun	'smoke'	*malip [‡]	mali	'laugh'	*qasawa	fetosawa	'cousin' [#]
*mapanas	fanas	'hot'	*gatəl	katal	'itchy'			
*quzan	usan	'rain'	*huaji	walin	'ySi'	*Rumaq	uma	'house'
*bəRəqat	boot	'heavy'	*dahun	ai loon	'leaf'	*daRaQ	laa-t	'blood'
*daRəq	lai	'earth'	*talih	tali	'rope'	*ikuR	hiʔu-n	'tail'

[†] inam 'six' was the only form I recorded in Oele'u. In Mahein I recorded both *inam* and *inan* 'six'. Da Silva (2012) and Hull (1998) only give *inan*.

[‡] *malip is reconstructed to PCEMP

[#] Welaun *fetosawa* refers to a woman's male cross cousin; the ideal marriage partner.

5.1.1 *ŋ > *g. In languages of the Central Timor subgroup PMP *ŋ has become a plosive. In Kemak, Tokodede, and Northwest Mambae the result is voiced *g* while in Central Mambae and Welaun the result is voiceless *k*. Evidence that *ŋ in Welaun and Central Mambae went through an intermediate voiced stage comes from South Mambae where there is variation in the reflexes of *ŋ.

In all varieties of South Mambae PMP *ŋ is usually reflected as *g* word medially and as *k* when final due to CV# → VC# metathesis. It is almost certain such final *k*'s are a result of devoicing (whether historic or synchronic) of an earlier final *g. Devoicing of word final voiced stops is a typologically extremely common sound change.

In South Mambae from Hatu-Udo and the village of Betano PMP *ŋ > *g* word initially, while in South Mambae from the village of Letefoho PMP *ŋ > *k* word initially.

Examples of PMP *ŋ in a number of languages of Central Timor are given in table 12. Forms which are cognate but which do not have a reflex of *ŋ are included in parentheses. Additional examples of Welaun *ŋ > *k* I have identified include: *ŋisi > *kiis-aat* 'tooth', *ŋjuda > *kulaʔaan* 'young', *baŋun > *fakun* 'wake s.o. up', *saŋa > *sakan* 'small branch', *diŋdiŋ > *kalilik* ~ *kanilik*, *udaŋ > *ulak* 'prawn, shrimp', and *bubuŋ > *fuhuk* 'ridgepole'.

TABLE 12. CENTRAL TIMOR *ŋ > g > k

PMP	*təliŋa	*saŋapuluq	*sumaqəd	*iŋuŋ	*ŋajan	*dəŋəR
Welaun	lika-t	sakulu	makar-aat	iluk-aat	kalan-aat	loka
Kemak	liga-r	(sapulu)		iliguu-r	gala-r	rega
Tokodede	təligi roa	sagulu		(ilu)	gala	
N.W. Mambae	tliga	saguul	smaga-n	(ilu-n)		
C. Mambae	tika	sakuul	sumaka-n	(ilu-n)		
S. Mambae (Hatu Udo)	teliga	saguul	samaga	(ilu)		
S. Mambae (Letefoho)	teliga	saguul	samaak	(ilu)	kala	
gloss	'ear'	'ten'	'soul'	'nose'	'name'	'hear'

The change of PMP $*\eta > g > k$ is a mostly unique change within the greater Timor region. Neither of the two other Austronesian subgroups on Timor have this change: the Timor-Babar subgroup has $*\eta > n$, while Helong (which appears to belong to its own subgroup) has $*\eta = \eta$ (Edwards 2018:63f).

5.1.2 *b > *f. Most Central Timor languages have $*b > h$. The exception is Welaun in which $*b- > h$ only occurs word medially with $*b- > f$ initially. The data from Welaun indicate that the change of initial $*b- > h$ in the other Central Timor languages went through an intermediate $*f$ - stage.

Initial $*b- > *f$ and medial $*b- > *h$ can be ascribed to Proto-Central Timor, with languages apart from Welaun subsequently undergoing $*f > h$. Examples of PMP $*b > f > h$ in Central Timor languages are given in table 13. This includes data from Kailaku Kemak in which initial $h > \emptyset$ in some forms.

TABLE 13. CENTRAL TIMOR $*b > f > h$

PMP	*bahi	*baqəRu	*batu	*babuy	*balabaw	*təbuh
Welaun	foi-t	fouʔaan	fatu	fahi	laho	tohu
Kemak (Leolima)	hee-r	heuŋ	hatu		laho	tehu
Kemak (Kailaku)	hee-r	heun	atu	ahi	laho	tehu
Tokodede	hee	heu	hatu			
N.W. Mambae		heu	hato		laho	
S. Mambae		heu	haut	hach	laho	
gloss	‘wife’	‘new’	‘stone’	‘pig’	‘mouse’	‘sugarcane’

The examples of Welaun medial $*b- > h$ in table 13 are identical to their Tetun equivalents and thus could be borrowings. Examples of Welaun $*b- > h$ which are not identical to Tetun include: $*bubuŋ > fuhuk$ ‘ridgepole’, $*bəqəbək > foha-t$ ‘voice’, and $*dabuk$ ‘pulverized, ashes’ (Wolff 2010:778) $> lahuk$ ‘dust’ (Tetun *rai rahun*).

The changes $*b > f$ and $*b > h$ are extremely common regionally. Nonetheless, that these changes have occurred in all the Central Timor languages can be taken as supporting evidence for this subgroup.

5.1.3 *ma-p > *mp. The sequence $*map-$ word initially is reflected as f in Welaun. I have so far collected three clear examples: $*ma-panas > fanas$ ‘hot’, $*qapəju > **ma-pəju > foru$ ‘bitter’, and $*hapəjəs > **ma-pəjəs > foras$ ‘sick, painful’.¹² As discussed in section 5.2.1 below, the normal reflex of initial $*p-$ in Welaun is h .

In the other Central Timor languages the normal reflex of $*map-$ is b . Examples include $*ma-panas > Kemak bənsaj$,¹³ Tokodede *bana*, N.W. Mambae *bana*, all ‘hot’ and $*ma-putiq > Kemak butik$, Tokodede *buti*, S. Mambae *buti*, all ‘white’.

12. Mambae has *moras* ‘sick, painful’ also ultimately from $*ma-pəjəs$. This form is probably a borrowing from Tetun, as attested by irregular $*ə > o$ where we expect $*ə > e$ in Mambae.

13. Leolima Kemak *bənsaj* $< *ma-panas$ has final VC $>$ CV metathesis, regular $n/ \rightarrow [ŋ]$ syllable finally, and addition of a suffix $-ŋ$. Leosibe Kemak has *banas* ‘hot’.

It is almost certain that these reflexes arose via reduction of *map- to a nasal-stop cluster *mp and/or *mb with further reduction of this cluster. There is evidence that indicates that *map- had not become *mb- by the stage of Proto-Central Timor.

This evidence comes from a number of forms in which *b-* in Central Timor languages corresponds to *mb-* in the western Rote languages. Examples include Dela-Oenale, Dengka *mbunut*, Welaun *kabunut* ‘coconut husk’ (ultimately from PMP **bunut*), Dela-Oenale, Dengka *mbua*, Welaun *bu*, Kemak *boo*, Tokodede *buo*, S. Mambae *buu* ‘betel nut’ (ultimately from PMP **buaq*), and Dela-Oenale *mbasa* ‘slap’, Welaun and Kemak *basa*, S. Mambae *baas* ‘slap’.

Given that these correspondences remain distinct from those of *map- in Welaun, the best solution appears to be to propose that PMP *map- > *mp in Proto-Central Timor with subsequent Welaun initial *mp- > *f* but *mb > *b*. The other Central Timor languages have a merger of Proto-Central Timor *mp/*mb > *b*.

Finally, it is worth noting that medial clusters of PMP *mp and *mb have become *b* in all the Central Timor languages, including Welaun. Examples include *umpu ‘grandparent/grandchild’ > Welaun *ubun* ‘owner, master’, Kemak *ubu-r* ‘grandchild’, S. Mambae *ubu* ‘master, lord, owner, person’, as well as PCMP **lemba* > Welaun *ai kaleba*, Kemak *ai kaleban*, both ‘carrying stick’.

The change of initial *map- > *mp can be assigned to Proto-Central Timor and provides supporting evidence for this subgroup. While this is not a particularly unusual change, it is different from languages such as Tetun and Meto which have *map- > *m*, as seen in *mapanas > Tetun *manas* ‘hot’, most Meto *manas* ‘sun’ and *maputiq > Tetun *mutin*, Meto *muti?* ‘white’.

Regionally, Galolen and Lakalei both have *map- > *b* which also may have gone through an intermediate *mp stage. Examples include *maputiq > Galolen *buti* (Cristo Rei and Grimes 2011), Lakalei *butin* (Klamer 2002a) ‘white’ and *mapanas > Galolen, Lakalei *banas* ‘hot’.

5.1.4 *d > *r. The change *d > *r can be assigned to Proto-Central Timor. Welaun has subsequently undergone *r > *l* when initial and intervocalic. Mambae also has sporadic instances of *d > *l*. Examples of PMP *d in the Central Timor languages are given in table 14.

TABLE 14. CENTRAL TIMOR *d > *r

PMP	*duRi [†]	*daRaŋ	*daRəŋ	*dahun	*dəŋəR	*maqudip
Welaun	lui-t	laa-t	lai	ai loon	loka	moli
Kemak	rui-r	raa-r	rae		rega	
Tokodede	rui	raa	rae	kai roa		
N.W. Mambae	rui-n	lara	rai-a			mori
S. Mambae	rui	lara	rae			mori
gloss	‘bone’	‘blood’	‘earth’	‘leaf’	‘hear’	‘live’

[†] PMP *duRi is reconstructed with the meaning ‘thorn’.

Word finally PMP *d > r in Welaun. Two examples are *sumaŋəd > *makar-aat* ‘soul’ and *tuhud > *tuur-aat* ‘knee’. While the change *d > r is an extremely common change cross linguistically and has occurred in many other languages of the region, it provides supporting evidence for the Central Timor subgroup.

5.1.5 *j > *l. The final change which provides evidence for the Central Timor subgroup is *j > l. Examples of *j > l include: *huaji > Welaun *walin*, Kemak *ali-r*, Tokodede *ali*, all ‘younger sibling’, and S. Mambae (Hatu-Udo) *ali* ‘younger brother’, *ijjuŋ > Welaun *iluk-aat*, Kemak *iliguu-r*, Tokodede *ilu*, S. Mambae *ilu*, all ‘nose’, and *ŋajan > Welaun *kalan-aat*, Kemak *gala-r*, Tokodede *gala*, S. Mambae *kala*, all ‘name’. Welaun also has PMP *suja > *sula* ‘pitfall’.

Before and after *ə the reflex of *j in Welaun is r. There are three examples: *hapəjəs > **ma-pəjəs > *foras* ‘sick, pain’, *qapəju > **ma-pəju > *foru* ‘bitter’, and *bujəq > *foren* ‘foam’. Another two possible examples are *qaləjaw > *loro* ‘sun’, and *qapəju > *horu-t* ‘gall bladder’ though these two forms are identical to Tetun and thus are possible borrowings.

In all other instances in which *j > r appears to have taken place in Welaun, the Welaun form is identical to Tetun and is thus probably a Tetun borrowing. Examples include: *maja > *mara* ‘dry’, *pija > *hira* ‘how many, several, plural’, *pajay > *hare* ‘rice plant’, and *pusəj > *husar-aat* ‘navel’.

Examples of PMP *j > l in other Central Timor languages where Welaun has *j > r include *qaləjaw > Kemak, Tokodede, N.W. Mambae *lelo*, S. Mambae *leol mata*, both ‘sun’, PMP *pija > Kemak *pila* ‘how many’, Tokodede *piil* ‘how many’, S. Mambae *arfila* ‘when’, *fiil* ‘how many’, and *maja > Kemak *malaj* ‘dry’.

Most other languages of Timor have *j > r. Exceptions include Helong in which *j > l is regular and Galolen, Lakalei and Idate in which *j > l ~ r. The change of *j > l in most environments in Central Timor provides supporting evidence for the Central Timor subgroup, with a conditioned change of *j > r in Welaun before and after *ə.

5.2 OTHER CONSONANTS. Having discussed the changes which provide evidence for the Central Timor subgroup, I now discuss the other consonant changes which have occurred in Welaun. These are the changes affecting *p, *k, *g, *z, and *ñ. Each of these is discussed in turn.

5.2.1 *p. Word initially PMP *p > h in Welaun, though there are only two examples which are not identical to Tetun: *pənuq > *honu* ‘full’ and *əpat > *hoat* ‘four’. There are many other forms which could be native Welaun forms or Tetun loans, such as *puqun > *huun* ‘tree (trunk)’ and *pitu > *hitu* ‘seven’.

I have so far identified only one putative example of medial *p in a form which is not identical to Tetun: *tapəh-i (Wolff 2010:997) > *tabi* ‘winnow’. All other putative Welaun examples are identical to Tetun and thus possible loans. Examples include: *hapuy > *haʔi* ‘fire’, *mahəpi > *meʔi* ‘dream’ and *niʔis* ‘thin’.¹⁴

14. Medial *p becomes ʔ in Tetun Fehan and Tetun Foho — the varieties of Tetun spoken in central Timor and the varieties from which Welaun has borrowed. In East Tetun spoken on the south coast of Timor-Leste the normal reflex of medial PMP *p is h, as seen in *ahi* ‘fire’ and *mehi* ‘dream’.

5.2.2 *k and *g. Word medially the normal reflex of *k in Welaun is *ʔ*. Examples include: *lakaw > *laʔa* ~ *laʔo* ‘go’, *paniki > *niʔi* ‘bat’, *siku > *siʔu-t* ‘elbow’, *wakaR > *ai waʔan* ‘roots’, *likud > *kaliʔu-t* ‘back’, *laki > *laʔi-t* ‘husband’, *kuhkuh > *huʔu-t* ‘fingernail’, and *ikuR > *hiʔun* ‘tail’.

The last two examples also show insertion of initial *h* after medial *C > *ʔ*. This changes also occurs in Kemak, Tetun, as well as Korbafo, Termanu, Bokai, Ba'a and Lole in central Rote. Kemak examples include: *ikuR > *hiʔu-ŋ* ‘tail’, *wakaR > *haʔa-ŋ* ‘roots’, and *kuhkuh > *lima-r huʔu-ŋ* ‘fingernails’. Tetun examples include: *hapuy > *haʔi* ‘fire’ and *aku > *haʔu* ‘I, 1SG’. Central Rote examples include: *hapuy > *haʔi* ‘fire’ and *kapuR > *haʔo* ‘mineral lime’.

Word initial *k is usually lost in Welaun. Examples include: *kulit > *ulit-aat* ‘skin’, *kali > *ali* ‘dig’, *kaən > *aan* ‘eat’, and *kaRat > *aat* ‘bite’.

Word initially PMP *g > *k*, though there are only two examples in my current data: *gatəl > *katal* ‘itchy’, and *gəmgəm > *kumu* ‘squeeze’.

5.2.3 *z > *s. The normal reflex of PMP *z in Welaun is *s*. Examples include *quzan > *usan* ‘rain’, *haRəzan > *osan* ‘ladder, stairs’, *zalan > *salan* ‘path’, *zauq > *laisoo* ‘far’, and *zəlay ‘Job’s tears’ > *sole* ‘corn, maize’.¹⁵

There are also three examples of PMP *z > *d*. Firstly, there is *tuzuq > *atudu* ‘indicate, show’ which may be a borrowing from Tetun *hatudu*. Secondly there is *zaRum > *daum* ‘needle’, and *zaqat > *daat* ‘bad’ which may ultimately be loans but which are not straightforwardly from Tetun, which has *daun* ‘needle’ and *aat* ‘bad’.

Other Central Timor languages have *z > *s* and/or *z > *r*. Examples of *z > *s* include *zəlay ‘Job’s tears’ > Kemak and Tokodede *sele* ‘maize’, *quzan ‘rain’ > Kemak and N.W. Mambae *usa*, S. Mambae *uus* all ‘rain’, and *zalan > Kemak *sala*, S. Mambae *saal* ‘path’.

Examples of *z > *r* include *quzan > Tokodede *ura*, *zauq > S. Mambae *roo* ‘far’ and possibly the second syllable of Tokodede *koro* ‘far’.

5.2.4 *ñ. PMP *ñ is reflected as *n* in Welaun, though there is only one clear example: *bañən > *bani* ‘sneeze’ (with irregular initial *b = *b*). Another possible example is *ñaRa > *naʔi* ‘older brother’, though this etymology requires irregular *a > *i* and *R/Ø > *ʔ*. Another possible etymology for *naʔi* ‘older brother’ is *laki with irregular initial *l > *n*.

5.3 VOWELS. The vowels *i, *u and *a are mostly retained unchanged in Welaun. Examples include *diŋdiŋ > *kalilik* ~ *kanilik* ‘wall’, *bubuŋ > *fuhuk* ‘ridgepole’, and *kaRat > *aat* ‘bite’.

15. For the semantic shift from ‘Job’s tears’ to ‘maize’, see the discussion in Fox (1991:250) in which new crops in Rote were assimilated as already known crops, with eventual semantic shift. Meto varieties have *sonə* ‘Job’s tears’ which is probably ultimately from Welaun before the change *l > *n* occurred in Meto. Welaun is the only extant Austronesian language in Timor which has both regular *z > *s* and *ə > *o*. This is one of several intriguing lexical connections between Welaun and Meto.

There are a small number of exceptions to these regular changes. Firstly, there are two examples in which *a > o: *bahi > *foi-t* ‘wife’ and *əpat > **paat > *hoat* ‘four’.¹⁶ Secondly, *u normally lowers to o before e. Examples are: *bujəq > *foren* ‘foam’, *buay > *foe* ‘beans’, *quay > *oe* ‘rattan’ and *uRat > *oet-aat* ‘veins’, the latter also with irregular final *a > e. Finally, *qatiməla > *tamola* ‘flea’ has irregular *i > a, probably due to its antepenultimate position (see section 4.1.1).

The normal reflex of *ə in non-final syllables is o. Examples include: *qatiməla > *tamola* ‘flea’, *zəlay ‘Job’s tears’ > *sole* ‘corn, maize’, and *dəŋəR > *loka* ‘hear’. Tetun and Habun (ISO 693-3 hbu) are the only other languages of Timor known to have *ə > o. Other languages of Timor have *ə > e.

The normal reflex of *ə in final syllables is a. Examples include *laləj > *lalan* ‘fly (n.)’, *suməŋəd > *makar-aat* ‘soul’, *daləm > **laram > *lamaraan* ‘inside’, *həpəjəs > **mə-pəjəs > *foras* ‘sick’, *kəən > *aan* ‘eat’, *gatəl > *katal* ‘itchy’, and PCEMP *mansər > *lamara* ‘cuscus’.

There are also two examples of *ə > i / _q# and one of *ə > e / _q#. These are: *basəq > **sabəq > *sahi* ‘wash’, *daRəq > *lai* ‘earth, land, soil’, and *bujəq > *foren* ‘foam’. There is also one example of *ə > i word initially and one of *ə > e; *ənəm > *inam* ‘six’ and *bəRay > *fee* ‘give’.

The normal reflexes of the vowel-glide combinations *ay and *aw are e and o respectively. Examples of each include: *sakay > *saʔe* ‘go up’ and *babaw > *faho* ‘mountain’. The vowel-glide combinations *uy and *iw are reflected as i, though all examples could be from Tetun. Thus, *laRiw > *alai* ‘run’ and *babuy > *fahi* ‘pig’. All the changes affecting the vowel-glide combinations follow the most common patterns among languages in the Timor region.

6. CONCLUSION. In this paper I have provided an overview of Welaun, a language not well known in the scientific community. Based on shared sound changes, Welaun can be classified as a member of the Central Timor subgroup which also includes Kemak, Tokodede and Mambae.

However, Welaun has several unique features when compared to these languages. These include productive verbal agreement, and different possessive suffixes than have been described for other Central Timor languages.

More work is clearly required to more fully understand the linguistic structures of Welaun and it is hoped that this paper will provide a useful starting point for any future work. If no such future work is ever undertaken, then I echo the sentiment of da Silva (2012:vi) that at least this paper will testify to the existence of the Welaun language.

16. An alternate etymology of *əpat > *hoat* ‘four’ is that it went through intermediate *pəat, in which case *ə > o would be regular. However, all other languages of Timor attest intermediate **paat. A clear example is N.W. Mambae *paat* ‘four’ rather than *peat as would be expected from *pəat.

APPENDIX 1. WELAUN TEXT

In this appendix I present a short text about life in Oele'u (Bauhoo village) recorded on the 4th of December 2017. The narrator is Antonius Yoset Mali, a 65 year old man who was born and raised in Oele'u. When I returned to Oele'u in January 2018 I checked most unclear sections with the narrator. This text can be downloaded from PARADISEC at http://catalog.paradisec.org.au/collections/OE8/items/Welaun20171204_OeleuStory.

Clauses are transcribed according to intonation units with the time (in recording) of the beginning of each clause given as the sentence number. Indonesian and Tetun insertions are given in standard orthography in italics.

- 0.15 ya ami laʔa uma ain lamaraan saresu
yes IPL.EXCL LOC village¹⁷ inside work
'We in this village work.'
- 0.22 ahh anu *tani*
person farmer
'We are farmers.'
- 0.25 akiʔak saʔi, akiʔak fahi, tau hare wee, tau likal
nurture cow nurture pig do field.rice water do field
'We raise cows, raise pigs, make wet rice fields, make fields.'
- 0.30 bodik urus uma ain *dan*
in.order.to organize village and
'(We do it) to manage the village and'
- 0.33 bele a-laʔo oan hira sakola
can CAUS-go child PL school
'we can send the children to school'
- 0.38 *dan*,
and
- 0.41 *dalam* iha, laʔa uma ain lamaraan ami saresu
inside LOC LOC village inside IPL.EXCL work
'In the village we work,'
- 0.46 fetu mane saresu uma isa
woman man work house one
'women and men work (to make) households'
- 0.51 *selama* tikan isa ami kai seresu
during year one IPL.EXCL go¹⁸ work
'During a year we go and work'
- 0.57 *hanya* urus likal, urus hare wee
only organize field organize field.rice water
'managing fields, managing wet rice fields'
- 0.59 akiʔak sarai moli-aan, saʔi, fahi, bibi, noo toman
nurture thing live-DEP cow pig goat with normal
'We raise livestock, cows, pigs, goats, as normal.'
- 1.06 *mungkin* duma nia dei ma ami dale
maybe like this just and IPL.EXCL speak
'Maybe that's enough for us to talk about.'
- 1.08 *terima kasi*
'Thank you.'

17. The phrase *uma ain* 'village' is composed of *uma* 'house' and *ain* for which the only identification I can currently make would be *ai-n* 'foot/leg'.

18. There is an audible sound after *ami* 'IPL.EXCL' in line 0.51. Based on the phonetics and context I have guessed that this might be *kai* 'go'. Though it could also be a kind of pause.

APPENDIX 2. WELAUN WORDLISTS

In this [appendix I](#) present the wordlists I collected in Oele'u and Mahein. Words are grouped by semantic fields. The etymology column on right gives the likely source as well as comparisons to likely cognates. All reconstructed forms marked with a single asterisk are P(CE)MP. Forms marked with a double asterisk are my own tentative reconstructions to a lower level.

When a Welaun word can be identified as a Tetun [tet] borrowing due to it displaying irregular sound changes this is indicated as with '<T.'. When a word is identical to Tetun, but there is no comparative basis for identifying it as a borrowing this is marked as '=T.'. Words marked as '=T.' are not necessarily Tetun loans (though this is possible). They could also be cases of parallel development.

Forms which do not show regular sound changes in Welaun and are suspected of being borrowings, but for which a source language has not been identified are marked as 'loan?' with the ultimate PMP etymology then given.

Comparison with likely cognates in other languages of the region are preceded by cf. where such forms are not glossed, they can be assumed to have the same glosses as the Welaun forms. Potential etymologies with problems are preceded by ?.

When I have evidence that a string in a particular word is synchronically an affix, it is separated from the stem with a hyphen. Where such synchronic evidence is lacking I do not separate possible affixes from their stems.

English	Oele'u	Mahein	Etymology
nose	iluk-aat	iluk-aan	*ijuj
eye	mata-t	mata-n	=T. <i>matan</i>
ear	lika-t	lika-n	*təliŋa
mouth	kuku-t	kuku-n	cf. N.W. Mambae <i>gugu-n</i>
tongue	bero-t	bero-t	cf. Helong <i>belo?</i> 'lick'
tooth	kiis-aat	kiis-aan	*ŋisi (CV > VC metathesis)
lips	kuku-t lai-n	kuku-n fihi-n	<i>fihi-n</i> < *biRbiR
cheek	hasan-aat	hasam-aan	
chin	ikam-aat	ikim-aan	
face	luam-aat	loom-aan	
neck	kole-t	kole-n	
head	ulu-t fatu-n	ulu-n (fatu-n)	*quluh, =T. <i>ulun</i>
hair (head)	ulu-t loo-n	ulu-t loo-n	*quluh + *dahun 'leaf'
hand	lima-t	lima-n	*lima, =T. <i>liman</i>
finger nail	hu?u-t	hu?u-n	*kuhkuh
hair (body)	fulu-t	fulu-n	*bulu, =T. <i>fulun</i>
elbow	si?u-t	si?u-n	*siku
shoulder	kabaas-aat	kabaas-aan	cf. Tetun <i>kbaas</i>
belly	ba?i-t	ba?i-n	cf. Kemak <i>ba?iŋ</i>
foot/leg	ai-t	ai-n	?*qaqay, =T. <i>ain</i>
knee	tuur-aat	tuur-aat	*tuhud
thigh	hii-t	hii-n	
skin	ulit-aat	ulit-aan	*kulit
meat/flesh	na?an	na?an	=T.
bone	lui-n	lui-t	*duRi (PMP 'thorn')
fat	bo?ur-aan	bo?ur	cf. Tetun <i>kbokur</i>
blood	laa-t	laa-t	*daRaq
heart	fua-t	fua-t	*buaq, =T. <i>fuan</i>
liver	ate-t	ate-t	*qatay, =T.
saliva	abat bee-n	aba-t	cf. Tetun <i>kaban</i> , Tii <i>ambe</i>
urine	mii	mii	*kəmiq, =T.

English	Oele'u	Mahein	Etymology
feces	tee	tee	*taqi
person	anu	anu	
man/male	mane	mane	*maRuqanay, =T.
female	feto	feto	< Meto/Tetun < *bətaw ¹⁹
husband	laʔi-t	laʔi-t	*laki
wife	foi-t	foi-t	*bahi
father	amakai	amakai	*ama, cf. Raklungu <i>amakai</i> ²⁰
mother	inakai	inakai	*ina, cf. Raklungu <i>inakai</i>
son/daughter	oa-t	oa-t	< T.
grandchild	bei-n oa-n	bei-t oa-n	< T. < *baqi
grandfather	amakai tuan	bei	< T. < *baqi
grandmother	inakai feli	bei	< T. < *baqi
older brother	naʔi	naʔi	?*laki
older sisiter	bii	bii	< T.
younger sibling	wali-n	wali-n	*huaji
brother (of woman)	mane-n		*maRuqanay
sister (of man)	feto-n		< Meto/Tetun < *bətaw
uncle (MB)	baba	baba	cf. Meto <i>baba</i> ²
aunt (FZ)	kiʔi	kiʔi	< T.
friend	maluk-aat	maluk-aan	=T. <i>maluk</i>
guest	bainaka		=T.
slave	ata-n	ata	*qaRta
child	wali oa-n	wali oa-n	
bridewealth	ana foli-n	foli-n	*bəli, =T.
name	kalan-aat	kalan-aan	*ŋajan
1SG	haʔu	haʔu	*aku, =T.
2SG	oo	oo	*kahu, =T.
3SG	nia	nia	*ia, =T.
1PL.EXCL	ami	ami	*kami, =T.
1PL.INCL	ita	ita	*kita, =T.
2PL	emi, imi	imi	*kamuyu, =T.
3PL	silā	silā	*sida
bird	manu	manu	*manuk, =T.
egg	manu tolun	man tolun	*qataluR, =T.
bat	niʔi	niʔi	*paniki
wing	lilaf-aat	lilaf-aan	cf. T. <i>liras</i> among others
snake	saba	saba inan	*sawa (irr. *w > b)
crocodile	naʔibein	bei liurai	cf. Meto <i>naiʔbesi</i>
louse	utu	utu	*kutu, =T.
mosquito	kohu	kohu	
scorpion	ulakai	ulakai	?*qudaŋ 'shrimp' + ai 'wood'
spider	saba dadai	dabadai	cf. T. <i>labadain</i>
ant	kokos	hokos	
rat/mouse	laho	laho	*balabaw, =T.
cuscus	lamara	lamara	*mansər (PCEMP)
dog	asu	asu	*asu, =T.
cow	saʔi	saʔi	<Malay <i>sapi</i>
horse	kuda	kuda	<Malay <i>kuda</i>
tail	hiʔu-n	hiʔu-n	*ikuR

19. While Tetun *feto* 'female' is ultimately from *bətaw, it must have come through an intermediate language with *ə > e, as the regular reflex of *ə in Tetun is o.

20. Raklungu (from Atauro island) data supplied by Charles Grimes (p.c. May 2018).

English	Oele'u	Mahein	Etymology
horn	huuk-aan	huuk-aan	
fish	naʔan tasi	naʔan tasi	lit. 'sea meat'
tree	ai huun		*puqun, =T.
wood	ai	ai	*kahiwi, =T.
leaf	ai loon	ai loon	*dahun
root	ai waʔan	ai waʔan	*wakaR
bark	ai ulit-aan	ai ulit-aan	*kulit
fruit	ai fuan	ai fuan	*buaq, =T.
flower	ai fotun	ai fotun	cf. Kemak <i>hetun</i> < **bətun
thorn	ai koon	ai koon	cf. Dengka <i>ngou-ʔ</i>
bamboo	au	au	*qauR
rattan	oe		*quay, =T.
grass	suʔut	suʔut	cf. Tetun <i>duʔut</i>
pandanus	boorlaku		
banana	hudi	hudi	*punti, =T.
coconut	nuu	nuu	*niuR, =T.
coconut shell	kabunut	kabunut	*bunut (irr. *b = b)
sugarcane	tohu	tohu	*təbuh, =T.
tamarind	anilu	anilu	loan? < *ɲilu
casuarina tree	ai hou	ai hou	*qaRuhu
cassava	uhi loke	uhi loke	*qubi
sweet potato	umalae	umalae	cf. Kemak <i>malae</i> ²¹
beans/legumes	foe	foe	*buay
rice (plant)	hare	hare	< T. < *pajay
rice (uncooked)	foos	foos	< T. < *bəRas
rice (cooked)	etu	etu	=T.
betel-nut	bua	bua	*buaq (irr. *b = b)
betel-vine fruit	buru	buru	*burun (irr. *b = b)
betel-vine leaf	malus	malus	cf. Meto <i>manus</i> among others, =T. <i>malus</i> 'betel leaf/fruit'
mineral lime	haul	haul	loan? < *kapuR
sun	loro	loro matan	*qaləjaw, =T.
moon	fulan	fulan	*bulan, =T.
star	fitun	fitun	*bituqən, =T.
cloud	kaloʔan	kaloʔan	<T. (Tetun Fehan <i>kaloʔan</i>)
sky		laleʔan	<T. (Tetun Fehan <i>laleʔan</i>)
rain	usan	usan	*quzan
wind	kolu	kolu	cf. Tokodede <i>gelu</i> < **ɲəlu, cf. Dhao <i>ɲəlu</i> , Kambera <i>ɲilu</i>
night	fanili	fanili	
day	hilan-aan	hinan-aan	
sea	tasi	tasi	*tasik, =T.
sand	henek	lai henek	cf. Kemak <i>rae henek</i> ?*qənay
ground	lai	lai	*daRəq
dust	lai lahuk-aan	ahu	*dabuk (Wolff 2010:778), *qabu
mud	lobu	lobu	
salt	sia	sia	*qasiRa
sugar	sia milak	sia milak	
water	wee	wee	*wahiR

21. In Tetun *malae* is 'foreign'. Welaun *umalae* is probably historically from *uhi* 'tuber' + *malae*.

English	Oele'u	Mahein	Etymology
mountain	lai lolon-aan, lai naruk-aan	faho	*babaw
forest	meʔitaan	ai matua	
river	mota	mota	=T.
lake	wee lihun	debu	=T. <i>lihun</i> 'body of still water'
fire	haʔi	haʔi	<T.
ashes	lahuk-aan	lai lahuk-aan	*dabuk (Wolff 2010:778)
smoke	haʔi masun	haʔi masun	*ma-qasu
stone/rock	fatu	fatu	*batu
house	uma	uma	*Rumaq, =T.
field	likal	likal	
road	salan	salan	*zalan
machete	sakulu	sakulu	
knife	tudik	tudik	=T.
spoon	kaneun	kaneu	
slingshot	karetupel	karetupel	Malay <i>kartupel</i>
rope	tali	tali	*talih, =T.
needle	daum	daun	< *zaRum (irr. *z > d)
mortar	kosun	kosu	**ŋəsun < *ləsun
pestle	alu	alu	*qahəlu, =T.
canoe/boat	bero	bero	=T.
big	bein	bein	
small	disi	disi anoan	
good	loil	loil	cf. Kemak <i>məloi</i>
bad	daat	ta loil	loan? < *zaqat
fat	boʔur	boʔur	cf. Tetun <i>kbokur</i>
dry	mara	mara-aan	<T. < *maja
far	laisoo	laisoo	*zauq, first element perhaps <i>lai</i> 'land'
near	besik	besik	=T.
new	founaan	fouʔaan	*baqəRu
old/former (house)	uma tuan-aan	tuan-aan	*tuqah, =T.
young (person)	anu kulaʔaan	kulaʔaan	*ŋuda
old (person)	anu matas-aan	anu matuan(-aan)	*ma-tuqah
hot	fanas	fanas	*ma-panas
cold	suman	kamar	
short	badak	para kanoan	*pandak, =T. <i>badak</i>
long	naruk	naruk	<T. < *anaduq
straight	lotar	lotar	
blind	mata-t daat	mata-n daat	
deaf	lika-t bere	lika-n bere	cf. Kemak <i>liga-r pereŋ</i>
thirsty	sarota	sarota	
hungry	salaen	salaen	cf. Kemak <i>slaen</i>
all	hotu-hotu	hotu-hotu	=T.
many	waʔin	waʔin	=T.
round	kabual	kabual	cf. Tetun <i>kabuar</i>
full	honu	honu oli	*pənuq
empty	maman	tanoo, maman	<i>tanoo</i> from *taq + <i>noo</i> 'exist'
stink, rotten	foon	foon	*bahu, =T.
white	baras	baras-aan	?*balaR
black	metan	metan	*ma-qitəm
yellow	nakat	unir	*kunj

English	Oele'u	Mahein	Etymology
red	meak	meak-aan	*ma-iRaq
green	mosok	nata-aan	cf. Tetun <i>modok</i>
blue	biru	nakat(-aan)	<Malay <i>biru</i>
here	laʔa koo	laʔa koo	
there	bete nia	nia kolaʔa	
there (further)	beta nia laa	nia kolaʔa	
underneath	bet lata-n	nia lata kolaʔa	
on top of	bete alain	nia alain kolaʔa	
edge	bete tehen	tehen, ninin	cf. Hawu <i>təbi</i>
behind	bete kaliʔu-n	kaliʔu-n	*likud
in front of	laʔ luam-aan	luam-aan	
outside	laʔ kaliʔu-n	nia luak-aan	luak-aan < *luqaR
inside	laʔa lamaraan	bete lamaraan	< **laram < **ralam < *daləm
from	n-osi	n-osi, hosi	=T. <i>hosi</i>
east	timur	timur	< Malay <i>timur</i>
west	barat	barat	< Malay <i>barat</i>
one	isa	isa	cf. Tetun <i>ida</i> , ?*əsa
two	rua	rua	<T. *duha
three	tolu	tolu	*təlu, =T.
four	hoat	hoat	*əpat
five	lima	lima	*lima, =T.
six	inam	inan, inam	*ənəm
seven	hitu	hitu	*pitu, =T.
eight	walu	walu	*walu, =T.
nine	siwi	siwi	*siwa, =T. (Fehan <i>siwi</i>)
ten	sakulu	sakulu	*saŋapuluq
twenty	rua kulu	rua kulu	
one hundred	atus isa	atus isa	*Ratus
one thousand	rihun isa	riun isa	loan? < *Ribun
speak	dale	dale	=T. (Belu 'relate, narrate')
sing	ananu	ananu	=T. <i>hananu</i>
cry	soo	soo	
laugh	mali	mali	*malip (PCEMP)
laugh at s.o.	mali anu	mali anu	
hear	loka	loka	*dəŋəR
see	okin	ania, anai, okin	
know	alaan	k-alaan	
forget	baluan	baluan	
remember	anoin	k-anoin	=T. <i>hanoin</i> 'think, remember'
eat	aan	k-aan	*kaən, cf. Tetun <i>haan</i>
drink	enu	enu	*inum
bite	anisi	haʔu k-anisi	<T. <i>hanisi</i> < *ŋisi 'grin, show the teeth'
fall	aleal	haʔu mout	
drop	a-mout	haʔu k-a-mout	cf. S. Mambae <i>mou</i> 'fall'
burn (intr.)	sunu	haʔu tunu	*tunu, cf. Tetun <i>sunu</i>
pound	baʔat	haʔu bai	<i>bai</i> < *bayu, =T.
die	mate	haʔu mate	*matay, =T.
dry in sun	awai	k-awai	*waRi
plait/braid	hanan	hanan	
fly	nele	(haʔu) nele	

English	Oele'u	Mahein	Etymology
swim	nani	nani	<T. < *naŋuy
bathe	diis	diis	*diRus (irr. *d = d, irr. *u > i)
bathe (a child)	diis wali oan	diis wali oan	
kill	hesa	k-esa	
give	fee	fee	*bəRay
cough	boo	boo	cf. Meto <i>n-boho</i>
vomit	muta	muta	*mutaq, =T.
spit	anilu	k-anilu	cf. Tetun <i>taniru</i>
itch	katal	katal	*gatəl
go	laʔa	laʔa	*lakaw
come	mai	mai (oli)	*ma(R)i
run	alai	alai	*laRiw, =T. <i>halai</i>
walk	kai lai	kai lai	
stay	hein	(oo) m-ein, afii	hein =T. 'wait, stop and wait'
stand	afii	afii	
stand sth. up	fafiin	tau nafii	
sit	tulan	tulan	*tudaŋ
sleep	buta	buta	
lie down	buta akiʔis	buta akiis	
sleepy	akati	matan duʔur	
dream	meʔi	meʔi	<T. (Fehan <i>meʔi</i>) < *mahipi
wake up	mata	mata	
wake s.o. up	fakun anu	oo m-atana	*baŋun
build/erect (house)	fafiin uma	safiin uma	
pregnant	niti	niti	
be born	moli	moli	*maqudip
pull	doʔe	doʔe	
read	baʔa	ania, lee	Malay <i>baca</i> , Portuguese <i>ler</i>
play	baʔan	baʔan	cf. Meto <i>na-ʔbaʔe</i>
take	ola	ola	cf. Tetun <i>hola</i> , ?*alap
indicate	atudu	atudu	< T. <i>hatudu</i> < *tuzuq
send, order	donu	donu	cf. Amarasi (Meto) <i>n-renu</i>
not	ta	lale	*taq, cf. Tetun <i>lale</i>
yes	iya	soʔin	
don't	muin	ta bele	
where?	beta bee	beta bee	
who?	sese	see	*sai, =T.
how much?	hira	hira	< Tetun < *pija
how?	tau bee, duma bee	tau bee	
what?	saa	saa	=T. <i>saa</i>
why?	tau saa	duma bee	
when?	wain hira	wain hira	*waRi, = T. <i>wai hira</i>

REFERENCES

- de Almeida, António. 1976. Da Origem Lendária e Mitológica dos Povos do Timor Português. *Memórias da Academia das Ciências de Lisboa, Classe de Ciências* 19:575–607.

- Badan, Pusat Statistik. 2012. *Penduduk Indonesia menurut desa hasil sensus penduduk 2010*. https://web.archive.org/web/20140327051059/http://bps.go.id/download_file/Penduduk_Indonesia_menurut_desa_SP2010.pdf [accessed 30/05/2018]
- Blust, Robert, and Stephen Trussel. Ongoing. Austronesian comparative dictionary. (Online open access site at: www.trussel2.com/ACD/.)
- Cristo Rei, João Maria, and Charles E. Grimes. 2011. *Disionariu Galolen: Galolen—Português—Tetun Dili; Português—Galolen; Tetun Dili—Galolen; Galolen—English; English—Galolen*. Dili: Konsehu Lian Inan Galolen and Australian Society for Indigenous Languages.
- Dyren, Isidore. 1965. *A lexicostatistical classification of the Austronesian languages*. Indiana University Publications in Anthropology and Linguistics, and Memoir 19 of the International Journal of American Linguistics. Baltimore: The Waverly Press.
- Direcção-Geral de, Estatística. 2016. *Timor-Leste Population and housing census 2015 population distribution by administrative area volume 2 (language)*. <http://www.statistics.gov.tl/category/publications/census-publications/2015-census-publications/volume-2-population-distribution-by-administrative/4-language/> [accessed 30/05/2018]
- Edwards, Owen. 2018. Top-down historical phonology of Rote-Meto. *Journal of the Southeast Asian Linguistics Society* 11:63–90. <http://hdl.handle.net/10524/52421>
- Fogaça, Helem Andressa de Oliveira. 2017. *O ecossistema fundamental da língua Mambae: aspectos endoecológicos e exoecológicos de uma língua austronésia de Timor-Leste*. PhD thesis, Universidade de Brasília. <http://repositorio.unb.br/handle/10482/31396>
- Fox, James J. 1991. The heritage of traditional agriculture in eastern Indonesia: Lexical evidence and the indications of rituals from the outer arc of the Lesser Sundas. *Bulletin of the Indo-Pacific Prehistory Association* 10:248–262.
- Grimes, Charles E. and Barbara D. Grimes. 1987. *Languages of South Sulawesi*. Canberra: Pacific Linguistics.
- Grimes, Charles E., Carlos Marçal, and Paolino Viana Ferreira. 2014. *Introductory Dictionary of Mambae (Same): Mambae—English, English—Mambae, Mambae—Indonesia—Tetun Dili, Indonesia—Mambae, Tetun Dili—Mambae*. Darwin: Australian Society for Indigenous Languages.
- Grimes, Joseph E. 1988. Correlations between vocabulary similarity and intelligibility. *Notes on Linguistics* 41:19–33 [reprinted in Eugene H. Casad (ed.). 1992. *Windows on Bilingualism*, 17–32. Dallas: SIL and UTA.]
- . 1995. *Language survey reference guide*. Dallas: SIL International.
- Hammarström, Harald, Sebastian Bank, Robert Forkel and Martin Haspelmath. 2018. *Glottolog* 3.2. Jena: Max Planck Institute for the Science of Human History. (Available online at <http://glottolog.org>, Accessed on 2018-05-27.)
- Hull, Geoffrey. 1998. The basic lexical affinities of Timor's Austronesian languages: a preliminary investigation. *Studies in Languages and Cultures of East Timor* 1:97–202.
- . 2003. Lia-Bekais nia Fatin iha Dalen-Lubun Timór. *Studies in Languages and Cultures of East Timor* 5:55–78.
- Klamer, Marian. 2002a. Field notes on Lakalei. In *LexiRumah*, ed. by Gereon Kaiping, Owen Edwards and Marian Klamer. Leiden: Leiden University Centre for Linguistics. (Available online at <https://lexirumah.model-ling.eu/lexirumah/>. Accessed on 2018-06-07.)
- . 2002b. Field notes on Tokodede. In *LexiRumah*, ed. by Gereon Kaiping, Owen Edwards and Marian Klamer. Leiden: Leiden University Centre for Linguistics. (Available online at <https://lexirumah.model-ling.eu/lexirumah/>. Accessed on 2018-06-07.)

- van Klinken, Catharina Lumien. 1999. *A grammar of the Fehan dialect of Tetun: an Austronesian Language of West Timor*. Canberra: Pacific Linguistics.
- Morris, Cliff. 1984. *Tetun-English dictionary*. Canberra: Pacific Linguistics.
- da Silva, Eng. Guilherme 'Puru-Berliku'. 2012. *Disionáriu Wekais-Tetun*. Dili: Secretaria de Estado da Cultura.
- Simons, Gary F. 1977. Recognizing patterns of divergence, and convergence in a matrix of lexicostatistic relations. In *Language variation and survey techniques: Working papers in Papua New Guinea Languages 21*, ed. by Richard Loving and Gary F. Simons, 107–134.
- Simons, Gary F. and Charles D. Fennig (eds.). 2018. *Ethnologue: Languages of the World, Twenty-first edition*. Dallas, Texas: SIL International. (Online version: <http://www.ethnologue.com>.)
- Smith, Kenneth D. 1984. The languages of Sabah: A tentative lexicostatistical classification. In *Languages of Sabah: A survey report*, ed. by Julie K. King and John Wayne King, 1–49. Canberra: Pacific Linguistics.
- Williams-van Klinken, Catharina. 2015. *Tetun-English interactive dictionary*. 2nd edition. Dili: Dili Institute of Technology. [<http://www.tetundit.tl/>]
- . 2015. *Mapping the mother tongue in Timor-Leste: Who spoke what where in 2010?* Dili: Dili Institute of Technology. [http://www.tetundit.tl/Publications/Timor-Leste languages 2010.pdf](http://www.tetundit.tl/Publications/Timor-Leste%20languages%202010.pdf). [Accessed April 30, 2015.]
- Wolff, John U. 2010. *Proto-Austronesian phonology with glossary*. Ithaca, NY: Cornell Southeast Asia Program Publications.