Detecting Papuan Loanwords in Alorese: Combining Quantitative and Qualitative Methods

Francesca R. Moro, Yunus Sulistyono and Gereon A. Kaiping

Introduction

In many parts of eastern Indonesia and Melanesia, speech communities often lack archaeological data and historical written sources, meaning that linguistic data is the only means by which to reconstruct past social interactions of speech communities (Ross 2013; Klamer 2015). Alorese, a language spoken in eastern Indonesia in a small-scale bi-/multilingual setting, is one such community. To reconstruct the sociolinguistic past of the Alorese, this paper analyses quantitative and qualitative patterns of lexical borrowing between Alorese and its neighboring languages.

Alorese is the only Austronesian language spoken on the coasts of the Alor and Pantar archipelago. On current accounts, it consists of 13 dialects or varieties corresponding to the main coastal villages where Alorese is spoken (see Figure 7.3 in §2). The other languages spoken on those islands are the Alor-Pantar languages[1] (henceforth AP), which belong to the (Papuan) Timor-Alor-Pantar family (henceforth TAP see §1).[2] As a point of contact between Austronesian surrounded by non-Austronesian languages, Alorese constitutes a special 'natural laboratory' for language contact studies. Since their arrival on the archipelago about 600 years ago, Alorese varieties have been in contact with the local AP languages. This long-term contact has affected the Alorese grammar, resulting in morphological simplification and a few structural borrowings (Klamer 2011; Moro 2018, 2019; Moro & Fricke 2020).

Interestingly, the two earlier publications on the topic (Klamer 2011; Robinson 2015) seem to indicate that Alorese lexicon is less affected by the long-term contact than the grammar. Both of these studies focus on a small part

---

1   Note that despite the name, Alorese itself is not a Timor-Alor-Pantar language.

2   The following abbreviations are used: AP = Alor-Pantar, PAL = Proto Alorese, PAP = Proto Alor Pantar, PFL = Proto Flores-Lembata, PMP = Proto Malayo-Polynesian, PTAP = Proto Timor-Alor-Pantar, PWL = Proto Western Lamholot, TAP = Timor-Alor-Pantar.

of the basic vocabulary (a Swadesh list), which might be more resistant to borrowing that the lexicon overall. A section in the short grammar of Alorese by Klamer (2011: 104–107) indicates an estimated percentage of 5.2%, while an article by Robinson (2015) discussing Austronesian borrowings into AP languages and AP borrowings into Austronesian languages finds about 3.8% AP loans in Alorese. These numbers are surprisingly small, considering the length of contact.

In this paper, we research whether the observation applies beyond the core vocabulary by extending the data to a 596-concept list, including all 13 Alorese dialects. Unlike other studies investigating Austronesian-Papuan borrowings (see among others Klamer's chapter in this volume), we did not pre-select the semantic domains to study, but investigated the entire dataset, and got the semantic domains of the loanwords inductively. In order to detect borrowing events, an algorithm was used to sifts loanwords out of a huge lexical pool: ~600 words × 13 Alorese dialects, × 55 Austronesian languages, × 42 TAP language varieties = approximately 66,000 word forms (see § 2). This pool is much larger than the dataset used in the previous research on AP borrowing in Alorese.

The present chapter, thus, illustrates an innovative methodological approach to the study of loanwords which uses an algorithm for automatic lexical similarity detection to study loans across two linguistic families. In this chapter, we describe the two-step procedure that was employed and how the results compare to work that has done this manually, to answer questions such as: does the size of a dataset make a difference when we investigate relative amount of borrowing? And does the percentage of borrowings increase when we investigate a large dataset, including highly borrowable concepts, compared to when we investigate a Swadesh list? Another innovative aspect of the chapter is that this is the only study in which 13 dialects of a minority language of Indonesia are compared. Comparing dialects on the patterns of lexical borrowing allows us to answer questions such as, do dialects of a language show differences in terms of their patterns of borrowing? Can this difference be related to their geographical location, their neighbours, or to the individual histories of the dialect communities?

A preliminary version of this research has been published in Chapter 6 of the PhD dissertation of Sulistyono (2022), in which lexical borrowings from and into Alorese and various languages including AP, Malay, Dutch and Portuguese are discussed. The present chapter has reconsidered the loan status of some words, excluding one concept, 'finished', and including four concepts 'dolphin', 'gravel', 'to breathe', and 'to hide'. Additionally, we provide an explanation to account for the limited lexical influence, and place our findings in
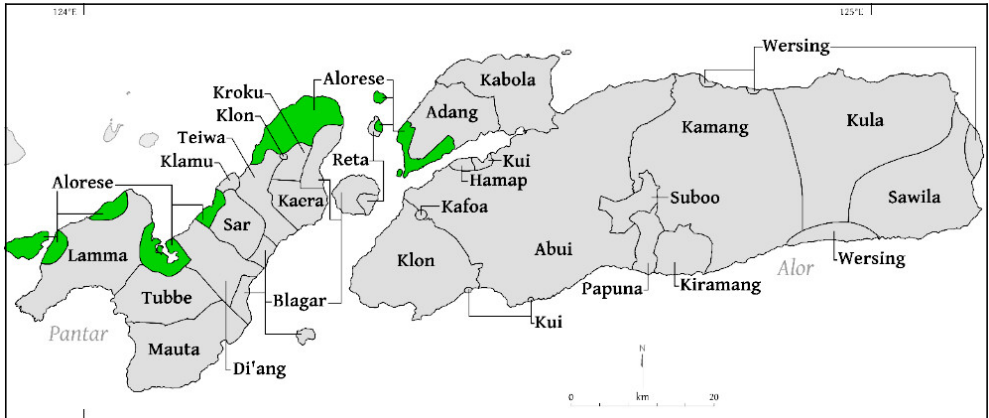
FIGURE 7.1   Alorese spoken on Alor and Pantar

a broader geographical perspective, relating our results to those of other studies in the present volume.

This chapter is organised as follows. As a background to this study, we begin by providing some basic information on Alorese and AP languages in §1; this is followed by §2 illustrating the research questions, the dataset, and the methodology of the present study. §3 presents the main findings, while §4 discusses the findings and gives some concluding remarks.

1   Alorese and the AP Languages

Alorese has approximately 25,000 speakers (Eberhard, Simons, & Fennig 2019). It is spoken along the coasts of Alor and Pantar, and on two small islands in the Alor-Pantar Strait in the Indonesian province of Nusa Tenggara Timur (see green areas on Figure 7.1 above). Besides Indonesian and the local Malay variety, Alorese is the only Austronesian language and is indigenous in the area.

The other languages spoken on those islands are roughly 25 Papuan languages of the Alor-Pantar (AP) subgroup, which belongs to the Timor-Alor-Pantar (TAP) family (Schapper, Huber, & van Engelenhoven 2017). There is evidence that the AP languages are spoken on the Alor archipelago since ~3,000 BP (Klamer 2017: 10), thus long before the arrival of the Alorese.

On Alor, Alorese is only spoken on the northern peninsula, alongside Adang; on Pantar, it is spoken alongside Kroku, Teiwa, and Nedebang (Klamu), among others. The historical situation of Alorese as Austronesian language spoken amid a mosaic of AP languages continues to the present day.
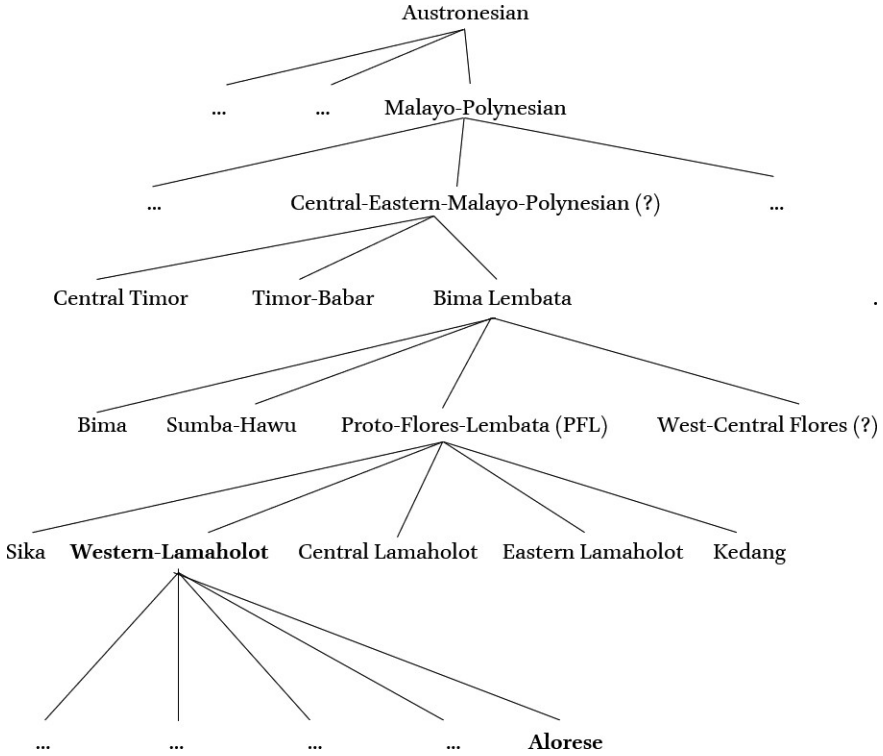
FIGURE 7.2   Genealogical classification of Alorese

Historically, Alorese speakers are descendants of groups migrating east-wards from the neighbouring island of Flores and its offshore islands (Klamer 2011: 8–15; Wellfelt 2016: 248–249; Sulistyono 2022). Historical linguistics indic-ate that the language spoken by these migrating groups was a western Lama-holot variety that later developed into what we today call 'Alorese'. Therefore, from a genealogical perspective, the closest relatives of Alorese are western Lamaholot varieties (Doyle 2010: 30; Elias 2017; Fricke 2019; Sulistyono 2022). Alorese and western Lamaholot varieties belong to the Flores-Lembata sub-group of Malayo-Polynesian languages, which also includes the eastern and central Lamaholot varieties, Sika, and Kedang (Fernandez 1996; Fricke 2019). Figure 7.2 above shows the genealogical classification of Alorese (Sulistyono, 2022:144; Fricke, 2019:20).

According to Anonymous (1914: 77), the first Alorese settlers arrived "5 to 600 years ago", meaning that they arrived around 1300–1400. Local oral his-tory suggests that the northeastern Pantar area, in particular today's villages of Pandai (see Figure 7.3 on the next page), was the first area settled by the Alorese in the 14th century. It was followed by the expansion to the Alor pen-

FIGURE 7.3 The 13 Alorese varieties spoken on Alor and Pantar, labeled according to village name

insula in the 16th century, and later expansions to the west and to the Strait starting in the 18th century (see Sulistyono 2022 for a detailed account on oral histories in this region). We will see in § 3.1 that this scenario is supported by the patterns of AP lexical borrowings in Alorese investigated in this paper. As a result, from a geographical and historical perspective, there are 13 Alorese varieties grouped in four main clusters: Northeast Pantar, Alor Peninsula, Strait, and West Pantar (from oldest to most recent). This geographical grouping is useful when determining the spread of loanwords in the varieties.

Traditionally, the Alorese practise exogamy with the neighbouring AP communities. In the past, exchanging women was a necessity for the Alorese, because their settlements only numbered about 200–300 people (Anonymous 1914: 89–90). Today, exogamy is still practised; however, the percentage of Papuan women has dropped considerably, as the Alorese settlements have become larger (approximately 1,500–2,000 inhabitants) and it has become easier to find a spouse within the same settlement. The settlement patterns are patri-virilocal, and the women generally move to the husband's village and are expected to learn Alorese (cf. also DuBois 1944: 85). Exogamy and patri-virilocal culture are inevitably linked to specific language acquisition patterns. In the past, Alorese villages must have been home to a continuous and considerable influx

of Papuan women who learned Alorese as a second language (L2), as well as bilingual children growing up learning both Alorese and an AP language (from their mother). These acquisition patterns are currently changing, as the local Malay variety and Indonesian are both gaining more ground.

Turning to the issue of language equality, at some point in the history of the Alorese, their language started to enjoy slightly more prestige over the AP languages, due to its role as lingua franca in the area of the Alor-Pantar Strait before Indonesian was introduced in the 1960s (Stokhof 1975: 8; DuBois 1944: 16). The status of Alorese as lingua franca arose due to the involvement of the Alorese in a Chinese-Muslim trade network bringing goods and slaves to Alor. Furthermore, during colonial times, the Alorese rulers acted as intermediaries between the inland Papuan population and the colonial governments (Stokhof 1984: 111). This situation must have led to the increase of asymmetric multilingualism, with Papuan speakers learning Alorese, but Alorese speakers remaining mostly monolingual.

2   The Present Study

This paper investigates Alor-Pantar (AP) loanwords in Alorese looking at a large lexical dataset. Using a two-step combination of automatic pre-screening and qualitative checks, we classify as candidate loanwords in Alorese all forms that are not inherited from the ancestor language (Proto Flores-Lembata), but that are formally similar to their semantic equivalents in one or more AP language, and check them individually.

2.1   *Dataset and Methodology*
In order to understand the patterns of loanwords in Alorese, we worked with word list data collected from field work and published sources aggregated in the online lexical database LexiRumah (Kaiping, Edwards, & Klamer 2019). We use version 1.0.0 of the database.[3] The sources of the individual word lists and forms used here can be found on LexiRumah. The dataset contains between 104 and 756 forms (counting all synonyms, and counting polysemous words once for every meaning) associated to a list of 596 concepts. The concept list contains pronouns and numerals, and nouns and verbs relating to both basic human activities (e. g., 'knife', 'to pull', 'to work', 'fireplace ash'), as well as the

---

3   A more recent version 3.0.1 includes an expanded set of languages, which are mostly more
    distant Austronesian or other Papuan languages, and thus not relevant for the Alorese lex-
    icon.

natural and cultural world of the region (e.g., 'sun', 'island', 'mountain', 'dolphin', 'rice ear bug', 'chicken', 'to plant yam', 'to clear land by burning').

The language dataset contains 13 Alorese varieties or dialects (one given in two sources), each of which displays between 450 and 756 forms for those concepts, 55 other Austronesian languages or varieties, and 42 TAP languages or varieties. For the Austronesian and TAP languages, the dataset contains between 104 and 756 forms (counting all synonyms, and counting polysemous words once for every meaning) associated to the list of concepts.

The first step we took was a data-mining process to discover potential loanword patterns in such a large dataset. So, we investigated lexical data from a quantitative perspective, by applying automatic lexical similarity detection. In the second step, we conducted a qualitative fine-grained analysis on the similarity sets whose patterns of distribution were compatible with borrowing event between Alorese and AP languages.

Borrowing between Alorese and AP languages would be visible in forms that are similar between Alorese and AP languages, and not explained otherwise: If forms in AP and Flores-Lembata languages are also similar, other explanations are assumed (e. g., borrowing before the genesis of Alorese, or widespread borrowing from Indonesian). Borrowed forms may have different meanings from the form in the donor language due to an originally general term being applied to a more specific foreign concept or due to subsequent semantic shift (Winter-Froemel 2013). In wordlist data, semantically different borrowed forms are hard to detect (List & Forkel 2021), and thus beyond the scope of our study. We thus focus on etymologically related forms within each concept.

In order to find candidates of etymologically related forms shared between AP languages and Alorese, we applied the automatic lexical similarity detection tool LexStat (List 2012), implemented in LingPy 2.6.5 (List et al. 2019). The LexStat algorithm uses a simplified 'sound class' representation (List 2012) of the forms in each language. Forms are matched with each other, and their sound class sequences are aligned with each other, giving a score that describes how many sounds in a form need to be changed to generate the corresponding form in a different language. Using stochastic methods, LexStat extracts the information whether the correspondence between different sounds is systematic or sporadic. LexStat's cognacy score then describes how many effective changes, discounting systematic differences, are needed to transform one form into another—lower scores mean that two forms are likely cognate, higher scores point to a lack of etymological relation. All pairs forms that have a cognacy score more similar than a set threshold of 0.55 are then connected into a network. The resulting network of forms is then split into discrete cognate classes using a graph partition algorithm, such as Infomap (Rosvall, Axelsson

TABLE 7.1   Relevant patterns of distribution of lexically similar forms in languages of the region, and the
corresponding borrowing or inheritance history of such a form

Hypo-thesis	AP lan-guages	Alorese	Flores-Lembata	Indo-nesian	Other Aus-tronesian	Explanation
1	Present	Absent	Absent	Absent	Absent	Inherited TAP vocabulary
2	Present	Present	Absent	Absent	Absent	Loan from AP into Alorese or vice versa, to be further inspected
3	Present	Present	(likely present)	Present	(likely present)	Indonesian loan into local languages
4	Present	Present	Present	Absent	Absent	Likely Alorese loan (inherited from PFL) into AP

& Bergstrom 2009). The algorithm creates classes which are strongly connected internally but have only weak connections between different classes.

While LexStat has been designed, and tested (Rama et al., 2018) for identifying cognate forms under systematic sound correspondence, the underlying similarity scoring is also promising for loan detection. While borrowing from one language into another does not follow systematic diachronic sound laws, phonological adaptation from the donor language to the recipient language may nonetheless introduce systematic changes (Uffmann 2015) and the general surface similarity should be picked up by LexStat's sequence alignment algorithm.

In order for an item to be an indication of borrowing between Alorese and an AP language, the lexically similar forms must be present in at least one Alorese dialect and at least one AP language. Different patterns of distribution of such forms outside Alorese and the AP languages indicate different hypotheses about the history of the word. The most important such hypotheses are summarized in Table 7.1. In this table, "present" means that the lexical similarity set contains a form for at least one language/dialect of that group, "absent" means that no language in that group has an attested form in that lexical similarity set.

As illustrated by hypothesis 2 in the previous table, an AP loan candidate in Alorese must be present in at least one Alorese variety, in at least one AP language, but in no other Austronesian language. To illustrate the automatic loan detection, an example is presented in Table 7.2, which shows the lexical similarity set for the concept 'to breathe'.

TABLE 7.2   Examples of a lexical similarity set associated with the concept
‘to breathe’ generated using automatic comparison

Concept	Language	Alignment	Form
to_breathe	Alorese-Munaseli	h ɔ - p a ŋ	hopang
to_breathe	Blagar-Bama	s o - p a ŋ	sopang
to_breathe	Blagar-Kulijai	h o - p a ŋ	hopang
to_breathe	Blagar-Nule	h o - p a ŋ	hopang
to_breathe	Blagar-Pura	h o - p a ŋ	hopang
to_breathe	Deing	- o - p a ŋ	opang
to_breathe	Kaera	s u ʔ p a ŋ	su'pang
to_breathe	Western Pantar-Tubbe	h o - p a ŋ	hopang
to_breathe	Reta-Pura	h oː - p a ŋ	hoopang
to_breathe	Reta-Ternate	h u - p a ŋ	hupang

The automatic comparison recognized that one Alorese variety, i.e., Alorese-Munaseli has the word *hopaŋ* ‘to breathe’ which is similar to forms attested in several AP languages. The AP forms are related and follow semi-regular sound changes (PAP initial *s > Kaera *s*, Blagar *h*, see Holton & Robinson 2017: 56). Therefore, this set potentially indicates a loanword from AP languages into Alorese (the Munaseli variety).

From the 596 concepts, the automatic detection filtered 167 sets of loan candidates, such as the one in Table 7.2 above. The resulting lexical similarity sets were inspected according to their pattern of distribution in the languages and dialects of the region. We manually checked the 167 loan candidates in more detail, to see whether the etymological relationship between the forms as hypothesized by LexStat makes sense beyond only the word lists. Of the potential 167 lexical similarity sets, 74 turned out to be erroneous, leaving us with 93 loan candidates. The erroneous cases include meaning mismatches, whereby, due to the different word order in Alorese and AP languages, the two aligned forms have formal similarity but are not semantically related. An example of an error due to a meaning mismatch is given in Table 7.3.

In Alorese, which is verb medial, the form for ‘bite’ is *gaki*, while *ata* means ‘person’ (*gaki ata* ‘to bite someone’). In Blagar-Pura, which is verb final, the form for ‘bite’ is *adaŋ*, while *jabar* means ‘dog’. The algorithm aligned Alorese *ata* ‘person’ with Blagar *adaŋ* ‘bite’ on the basis of formal similarity, but semantically the two forms are not related.

TABLE 7.3   Examples of a meaning mismatch due to word order

Concept	Language	Alignment	Form
to_bite	Alorese-Ternate	a t a -	gaki ata
to_bite	Blagar-Pura	a d a ŋ	jabar ing adang

The potential 93 loanwords were inspected more carefully to establish whether they are indeed AP loanwords in Alorese. Out of the 93 candidates, 28 turned out to be AP loanwords, the others were loanwords in the other direction (from Alorese into AP), or from Malay, or loanwords of unclear direction, or the resemblance was due to chance. In the following section, we present the 28 AP loanwords.

3   AP Loanwords in Alorese

In this section we present the AP loanwords organized by semantic fields, a choice shared by other contributions in this volume (e.g., Klamer, Edwards, and Schapper & Huber), to gain an additional perspective on the type of contact between the Alorese and the AP speakers. We assigned the AP loanwords to five semantic fields from the most prone to borrowing to the more resistant to borrowing: *Basic actions and technology* (§3.1), *Social and political relations* (§3.2), *Agriculture and vegetation* (§3.3), *The physical world* and *Animals* (§3.4), and a miscellaneous field including (*quantity, emotions, motion, kinship, the body, spatial relations, sense perception*) (§3.5). The semantic fields are those of Tadmor et al. (2012), but where slightly modified to be consistent to those of Edwards (this volume). *Basic actions and Technology* thus includes *Tools* as well as *Weapons*, and *The house*. The *Law* and *Religion and belief* were combined with *the Social and political relations* field. Unlike Edwards, we also combined *Animals* and *The Physical world*. Approximately half of the AP loanwords occur in the three most borrowable semantic fields (*Basic actions and technology*, *Social and political relations*, and *Agriculture and vegetation*).

In §3.6 we will draw generalizations regarding their distribution among Alorese varieties, and their donor languages. All comparisons presented in this section were made with the tool EDICTOR, (etymological dictionary editor) at https://digling.org/edictor/. EDICTOR visualizes and allows to edit the cognate judgements in a lexical database. The tool also aligns similar sounds within the sets which helps to discover sound correspondences.

3.1 *Basic Actions and Technology*

3.1.1 'Fish trap'

The Alorese and AP forms for the concept 'fish trap' are presented in Table 7.4.

TABLE 7.4 Lexical similarity set associated with the concept 'fish trap'

Language	Alignment
Alorese-Pandai	k ɛ r
Blagar-Bakalang	v e r
Blagar-Bama	w e r
Blagar-Nule	k e r
Blagar-Tuntuli	v e r
Kula-Lantoka	g a r
Nedebang	tʃ a r
Teiwa	k ɛː r
Wersing-Maritaing	- a r

The Alorese-Pandai word *ker* 'fish trap' is an innovation, different from the inherited form *pukoʔ* < PFL **pukət* 'fish trap' (Fricke 2019: 240) present in the other Alorese varieties. The source for this innovation are AP languages, which present similar forms. The sound changes among AP languages are semi-regular, because initial stops are usually retained among AP languages, so the change *k* > *v*/*w* in some Blagar varieties remains unexplained; final *-r is retained unchanged in all the AP languages, except Klamu. In Klamu, the PAP final *-r is expected to be lost (Holton et al. 2012: 94), but it is possibly irregular because a retention of *-r is also attested in PAP **dur* > Klamu *dur* 'rat'. Due to the geographical spread of the word among AP languages, we consider this a loanword from AP languages, most likely Blagar-Nule or Teiwa, into Alorese-Pandai.

3.1.2 'Bed'

The Alorese and AP forms for the concept 'bed' are presented in Table 7.5.

Alorese-Pandai, Alorese-Munaseli and Alorese-Alor Besar have innovated the form *deki* for 'bed; raised platform'. This word is likely to be a loanword from AP languages which have similar forms that are related and reflect regular sound changes. Initial PAP **d* is retained in all the languages. Medial *k is retained in Blagar, reflected as *ʔ* in Adang (Holton & Robinson 2017: 56). Among

the AP languages, Blagar, Reta, Kaera, and Western Pantar have the most similar form because the medial *k* is retained. From a geographical perspective, however, Blagar or Reta are likely the donor language(s) as this form is found in Alorese varieties spoken around the Alor-Pantar Strait.

TABLE 7.5   Lexical similarity set associated
             with the concept 'bed'

Language	Alignment
Alorese-Pandai	d ɛ k i
Alorese-Munaseli	d e k i
Alorese-Alor Besar	d e k i
Western Pantar-Lamma	d e k i
Blagar	d e k i
Reta-Pura	d ɛ k i
Teiwa	d e k
Adang	d e ʔ

3.1.3   'To fold'

The Alorese and AP forms for the concept 'to fold' are presented in Table 7.6.

TABLE 7.6   Lexical similarity set associ-
             ated with the concept 'to fold'

Language	Alignment
Alorese-Pandai	l a k u k -
Kula-Lantoka	l a k u p -
Sawila	l a k u p i
Blagar	p i l i k u

For the concept 'to fold', the Alorese-Pandai variety in northeast Pantar uses both the inherited form *lepe* and an innovation *lakuk*. Among the AP languages, the most similar forms are *lakup*(*i*) in Kula and Sawila, *pəliku*/*piliku* in Blagar. It is unclear whether the AP words for 'to fold' presented in Table 7.6 are all related. In Sawila, *kupi* means 'to fold' (Kratochvíl 2014: 408), but the additional syllable *la-* is of unclear origin. In Blagar, *pi-* is an inalienable possessor for first

person plural inclusive (Steinhauer 2014: 182). Since at least some of the AP word seem to form a cognate set, while Alorese-Pandai is the only variety to use this form, we consider this a loanword from AP languages, and most likely Blagar into Alorese-Pandai. The change of the vowel from *i* to *a* (*liku* to *lakuk*) is also attested in other Blagar loanwords, such as Alorese *kalaki* 'angry' from Blagar *kilikil*, Alorese *reha* 'monitor lizard' from Blagar *rihi*, and Alorese *tera* 'to close' from Blagar *teriŋ*.

3.1.4   'To pull'

The Alorese and AP forms for the concept 'to pull' are presented in Table 7.7.

TABLE 7.7   Lexical similarity set associ-
ated with the concept 'to pull'

Language	Alignment
Alorese-Pandai	- w a k
Blagar-Kulijahi	a w a k
Blagar-Nule	a v a k
Abui	h a f i k
Klon	g ə b ɪ k
Adang	ʔ a b i ʔ i ŋ
Kabola	a p i ʔ i ŋ

For the concept 'to pull', Alorese-Pandai uses both the inherited form *vider* 'to pull' and an innovation *vak* 'to pull'. For this concept, the majority of Alorese varieties use a Malay loan *tarek* 'to pull'. The Alorese-Pandai word *wak* 'to pull' is possibly a Blagar loan, because a similar form *awak/avak* 'to pull' is attested in Blagar-Kulijahi and Blagar-Nule. The initial vowel *a-* is a prefix in Blagar indicating causative (Steinhauer 2014: 160, 194). This Blagar word seems to be related to the other AP words listed in the table. The sound changes are semiregular, as initial *b is reflected as *f* in Abui, and can be reflected as *v* if in intervocalic position (after the addition of a prefix) in Teiwa and Nedebang (Holton & Robinson 2017: 56), and in this case also in Blagar. The vowel *a* in Blagar remains difficult to explain, although Edictor found one other correspondence of Blagar *a* and, for instance, Klon *i*: Blagar-Nule *havaʔ* 'house' ~ Klon-Hopter *ʔəwi* 'house'. The Blagar forms are formally the most similar, hence we identify Blagar as the donor language for this loan.

3.1.5 'To wash'

The Alorese and AP forms for the concept 'to wash' are presented in Table 7.8.

TABLE 7.8   Lexical similarity set associated
             with the concept 'to wash'

Language	Alignment
Alorese-Baranusa	- - l a m i ŋ
Alorese-Munaseli	- - l a m i ŋ
Alorese-Pandai	- - l a m i ŋ
Adang-Lawahing	- - l a: m - -
Adang-Otvai	- - l a m - -
Reta	- - l a: m i ŋ
Deing	- - l a n a ŋ
Hamap	n a l a m - -
Kabola	- - l a m - -
Kafoa	- u l ɑ m - -
Western Pantar-Lamma	- - l a m i ŋ

For the concept 'to wash', some Alorese varieties on Pantar, have innovated the form *lamiŋ*, next to the inherited *bema* ('to wash' for clothes) and *hue* ('to wash' for dishes). The Alorese word *lamiŋ* 'to wash' in the varieties of Baranusa, Pandai, and Munaseli appears to be a loanword from an AP source. This form is an inherited AP form, with related forms in several AP languages, as can be seen in Table 7.8. Reta (*laamiŋ*) and Western Pantar-Lamma (*lamiŋ*) have the most similar forms to Alorese and both these languages are in contact with Alorese on Pantar; Reta is close to Munaseli and Pandai, while Western Pantar is close to Baranusa. Therefore, these are the most likely donor languages. This AP loanword is also mentioned by Klamer (2011: 105) and by Robinson (2015: 28), both pointing to Western Pantar as the donor language.

3.2 *Social and Political Relations*

3.2.1 'To pray'

The Alorese and AP forms for the concept 'to pray' are presented in Table 7.9.

TABLE 7.9   Lexical similarity set associated
with the concept 'to pray'

Language	Alignment
Alorese-Munaseli	- g a m a r
Kaera	a ʔ - m u r
Western Pantar-Lamma	- h a m u r
Reta	- - a m u r
Teiwa	- h a m a r

The Alorese-Munaseli variety in northeast Pantar innovated *gamar apa* for 'to pray', while in the other Alorese varieties the more widely used term for 'to pray' is *səbeaŋ* (< Malay loan *sambayang* [sambaʰiaŋ] 'to pray; to worship God'). The Munaseli form *gamar apa* comprises *gamar* (external origin) and *apa* (Alorese 'something'). In the set for the concept 'to pray', it seems that Alorese-Munaseli has borrowed *gamar* from a neighboring AP language, such as Teiwa, which has *hamar* for 'pray'. Since cognates of the Teiwa form for 'to pray' are attested across several AP languages, it is likely that this is an inherited AP form. Teiwa is very likely to be the donor because the vowels are identical to the Alorese-Munaseli word *gamar*. The initial *g* in the Alorese-Munaseli word may come from the Teiwa form *ga-hamar* 'pray for someone', whereby *ga-* is a third person singular pronoun (Klamer 2010: 55).

3.2.2 'Adultery'

The Alorese and AP forms for the concept 'adultery' are presented in Table 7.10.

TABLE 7.10   Lexical similarity set associated
with the concept 'adultery'

Language	Alignment
Alorese-Alor Besar	b u h a
Alorese-Munaseli	b u h a

TABLE 7.10   Lexical similarity set associ-
              ated with the concept 'adultery'
              (*cont.*)

Language	Alignment
Kaera	b u s -
Blagar-Pura	b u h a
Reta	b u h a
Teiwa	b uː s -

No similar forms to the Alorese-Alor Besar and Alorese-Munaseli word *buha* 'adultery' are attested in the near-by Flores-Lembata languages and no proto forms are available for this concept. Conversely, the AP forms are historically related and reflect regular sound changes. The initial PAP *b is expected to be retained unchanged in all the languages; the final PAP *s is expected to be retained regularly as *s* in Teiwa, Kaera, and Sar, and changed into *h* in Blagar and in Reta (Holton & Robinson 2017: 56). This sound change is attested in several Blagar words, such as PAP *mis > Blagar *mihi* 'sit' and PAP *bis > Blagar *bihi* 'mat', where an epenthetic vowel is added after the weakening of *s. Given the presence of the glottal fricative *h* and the vowel in Alorese varieties, we conclude that Alorese-Alor Besar and Alorese-Munaseli borrowed *buha* from either Blagar or Reta.

3.3   *Agriculture and Vegetation*

3.3.1   'Digging stick'

The Alorese and AP forms for the concept 'digging stick' are presented in Table 7.11.

TABLE 7.11   Lexical similarity set associated
              with the concept 'digging stick'

Language	Alignment
Alorese-Bana	- - n o r u ʔ
Alorese-Helangdohi	- - n ɔ r u ʔ
Blagar-Nule	- - n o r u k
Reta	h a n oː r u k
Western Pantar	- - s o r u -

The table shows that the Alorese-Bana and Alorese-Helangdohi varieties in northeast Pantar innovated *noruʔ* 'digging stick'. The other Alorese varieties use the word *kuaŋ* inherited from PWL (*nuaŋ*, Sulistyono 2022: 255).

Among the AP languages Blagar-Nule (*noruk*), Reta (*hanoruk*), and Western Pantar (*soru*) show related forms for the concept 'digging stick', these forms seem to go back to a form like #sVnoru(k) 'digging stick'. It seems that a semantic change occurred in Western Pantar to the relatively close concept 'stick; pole'. The sound changes are semiregular. The initial *s is regularly reflected as *h* in Reta and retained unchanged in Blagar and Western Pantar (Holton & Robinson, 2017: 56). The intervocalic *n is expected to be retained unchanged in all languages, however the sequence -*Vn*- is lost in Western Pantar (#sVnoru (k) > *soru*). The intervocalic *r shows irregular reflexes in Western Pantar because it is expected to be retained as *l*. Finally, the final *k is expected to be lost in Blagar and retained in Western Pantar (Holton & Robinson 2017: 56), but here we see the opposite pattern. Even though the sound correspondences among the AP languages are semi-regular, we consider the form inherited in AP languages. We, therefore, consider that the Alorese word *noruʔ* is a loanword and that the most likely donor for this concept is Blagar-Nule which has the form *noruk*, most similar to the Alorese form *noruʔ*.

3.3.2   'Garden'

The Alorese and AP forms for the concept 'garden' are presented in Table 7.12.

TABLE 7.12   Lexical similarity set associated with the concept 'garden'

Language	Alignment
Alorese-Munaseli	b u t a ʔ
Adang-Lawahing	b u t u -
Adang-Otvai	b u t - -
Blagar-Warsalelang	b u t a x
Blagar-Tuntuli	b u t a q
Blagar-Pura	b u t a
Kabola	b u t u ʔ

The table shows that Alorese-Munaseli in northeast Pantar innovated *ekaŋ buta?* for 'garden'. For the concept 'garden', the general Alorese term that goes back to PWL is *ekaŋ* 'garden' (Sulistyono 2022:255). However, the Munaseli vari-

ety uses a compound *ekaŋ butaʔ*, which comprises an inherited form *ekaŋ* (<
PWL *eka 'garden') and the new form *butaʔ*, which is of external origin.

The AP words meaning 'garden' are clearly related and attested in AP lan-
guages spoken around the Alor-Pantar Strait. Possibly, a Proto Nuclear Alor-
Pantar form *butVq 'garden' could be reconstructed based on this cognate set
(Kaiping & Klamer 2019: 35). Alorese-Munaseli has borrowed the form *butaʔ*
to form a compound *ekaŋ butaʔ* 'garden'. The donor language is most likely a
Blagar variety because they have the most similar forms and are geographically
close to Munaseli. The lenition of final stop *x/q* in Blagar into a glottal stop in
Munaseli is expected because Alorese does not allow final *x/q*.

3.3.3 'Rattan'
The Alorese and AP forms for the concept 'rattan' are presented in Table 7.13.

TABLE 7.13   Lexical similarity set associ-
ated with the concept 'rattan'

Language	Alignment
Alorese-Munaseli	l u - a
Alorese-Pandai	l u - a
Blagar-Kulijahi	l i - a
Blagar-Nule	l i j a
Blagar-Pura	l i - a
Blagar-Bama	l e g
Blagar-Tuntuli	l e - g
Blagar-Warsalelang	l eː - g
Reta	l i - a g
Reta	l i j a g
Kaera-Abangiwang	l ɛː - g
Kabola-Monbang	l o j o ʔ
Blagar-Bakalang	l i j a
Adang-Otvai	ˈl e
Teiwa-Lebang	l i j a g
Kui-Labaing	l e
Adang-Lawahing	l ɛ ʔ
Deing	l i a x
Sar-Adiabang	l i j a h
Sar-Nule	l i j a g

We assume that the form *lua* 'rattan' in Alorese-Pandai and Alorese-Munaseli comes from an external source. This form is used alongside a different form *ue/uwe*, inherited from PFL **uay* 'rattan' (Fricke 2019: 477). For this concept, regular sound correspondences can be seen among AP languages, namely PAP initial **l* is retained unchanged in all languages, as expected. There is no PAP sound that is reflected as final *g*, but the synchronic correspondences are fairly regular: for instance, Teiwa *bag* ~ Deing *bax* 'seed', Teiwa *og* ~ Deing *ox* 'hot'. Interestingly, no AP language shows the vowel combination *ua* found in the Alorese form *lua*, so we hypothesize that Alorese borrowed the form *lia* from Blagar, but it changed the diphthong from *ia* to *ua*, to be reminiscent of the inherited PFL form **uay* 'rattan'.

3.3.4   'Root'
The Alorese and AP forms for the concept 'root' are presented in Table 7.14.

TABLE 7.14   Lexical similarity set associated with the concept 'root'

Language	Alignment
Alorese-Alor Kecil	- a l i - - ŋ
Alorese-Dulolong	- a l i - - ŋ
Adang-Lawahing	- a l i ʔ i ŋ
Adang-Otvai	- a l i ʔ a ŋ
Hamap	- a l i - a ŋ
Kabola	h a l i ʔ i ŋ
Kafoa	- ɪ l i: k ɑ ŋ
Kamang	- a l i: - - - - -
Abui	- a i - - - - -

For the concept 'root', the majority of the Alorese varieties use the inherited form *ramuʔ* (< PMP **Ramut* 'root', see Sulistyono 2022: 265). The form *aliŋ* is an innovation in Alorese-Alor Kecil and Alorese-Dulolong in the Alor Peninsula and it is borrowed from AP languages. The PAP intervocalic **-l-* is retained unchanged in all modern-day AP languages (Holton et al. 2012: 94), except in Abui where it is lost. It seems that Alorese borrowed the word *aliŋ* 'root' from Adang, Kabola, or Hamap. These languages are located close to the Alor Peninsula varieties. Among these languages, Adang is most likely to be the donor. The Adang-Lawahing word *aliʔiŋ* is the most similar to Alorese *aliŋ* 'root'. It is

likely that Alorese borrowed the word from Adang-Lawahing and dropped the ʔ in the process, as Alorese varieties spoken on the Alor Peninsula and in the Strait do not have word medial glottal stop.

3.3.5 'Taro'
The Alorese and AP forms for the concept 'taro' are presented in Table 7.15.

TABLE 7.15   Lexical similarity set associated with the concept taro

Language	Alignment
Alorese-Alor Besar	a - g ɔ l
Alorese-Alor Kecil	g o l o -
Alorese-Dulolong	g ɔ l ɔ -
Alorese-Baranusa	g o l o -
Blagar-Pura	a u
Blagar-Bakalang	a w g o l
Blagar-Kulijahi	a w g o l
Blagar-Nule	a w g o l
Blagar-Pura	g o l - -
Adang-Otvai	a - g ɔ l
Hamap	a - k o l
Reta	a i g o l

The Alorese word for 'taro' is formed by two elements, the word *au/ai* is possibly from PFL *kayu 'tree' (see Fricke 2019: 521) and the word *gol/golo* which is likely an innovation borrowed from AP languages. Alorese-Dulolong has *kadʐo golo* with the form *kadʐo* 'tree; wood', also from PFL *kayu 'tree'.

Among the AP languages, the words *awgol* (Blagar-Bakalang), *au gol* (Blagar-Pura), *agɔl* (Adang-Otvai), and *ai gol* (Reta) are most similar to the Alorese forms. The AP forms are also formed by an element *au/ai/a* meaning 'tree' or 'tuber' and an element *gol/go/hol/ho* meaning 'taro'. This pattern is found also to refer to other tubers, as for instance Blagar-Pura *au benu* 'cassava', and *au kasi* 'sweet potato'. Languages where these form for 'taro' are used are located close to the Alor Peninsula, and the most likely donor seems to be Blagar or Adang. The fact that Alorese-Baranusa also uses a similar form, *au golo*, for 'taro' may indicate that this word was borrowed early on, although the fact that is absent in the northeastern Pantar varieties may also indicate a

later borrowing from Alorese varieties in the Alor Peninsula to the Baranusa variety.

3.4 Animals and Physical World

3.4.1 'Coral rock'

The Alorese and AP forms for the concept 'coral rock' are presented in Table 7.16.

TABLE 7.16 Lexical similarity set associated with the concept 'coral rock'

Language	Alignment
Alorese-Munaseli	k ɔ - k a -
Blagar-Bakalang	k o - k a -
Blagar-Bama	k o - q a s
Kaera	q o ʔ q i s
Teiwa	q o - q a s
Adang-Otvai	ʔ o ʔ - o i
Kabola	k o ʔ - o i

The Alorese-Munaseli innovation *koka* for the concept 'coral rock' is a loanword from AP languages, in which the forms for this concept are cognates. The sound changes are semi-regular: PAP initial and intervocalic *q is reflected as *k* in Blagar, *q* in Teiwa, and *ʔ* in Adang (Holton & Robinson 2017: 56); Kaera is expected to have *x* but has *q*, and Adang is expected to have zero in intervocalic position, but here may have a glottal stop to avoid a sequence of two identical vowels. The correspondence of Teiwa *q* and Kabola *k* is regular, and attested in other words such as Teiwa *qab* ~ Kabola *kaba* 'spear' and Teiwa *qarnuk* ~ Kabola *karnu* 'ten'. The PAP final *s is retained in Teiwa and Kaera, and reflected as *h* in Blagar and Adang (Holton & Robinson 2017: 56), here however, Adang and Kabola have final *i* and Blagar-Bama has retained the final *s*. The correspondence *a* ~ *i* in Teiwa and Kaera is attested also in other forms such as Teiwa *saxaʔ* ~ Kaera *siʔaq* 'chicken', Teiwa *hasak* ~ Kaera *isʔik* 'empty'. The most likely donor for this concept is Blagar-Bakalang which has the form *koka*, identical to the Alorese one.

3.4.2   'Mud'

The Alorese and AP forms for the concept 'mud' are presented in Table 7.17.

TABLE 7.17   Lexical similarity set associ-
ated with the concept 'mud'

Language	Alignment
Alorese-Alor Kecil	b a n a -
Alorese-Dulolong	b a n a -
Alorese-Ternate	b a n a -
Blagar-Pura	b a n a k u ŋ
Nedebang	b a n a q a
Reta	b a n a k u ŋ
Sar	b e n a: q
Abui	f a n a q

The Alorese word *bana* 'mud'[4] has no similar forms in the neighboring Flores-
Lembata languages or in the proto languages. This innovation is likely to come
from AP languages. For the concept 'mud', it looks like that the AP forms go
back to a form like #banak or #banaq. The sound changes are regular, as PAP
initial *b is retained in all languages, but Abui where it is reflected as *f* (Holton
& Robinson 2017: 56). The addition of a final syllable *-uŋ* in Blagar and Reta
remains unclear, although similar additions of final syllables are found in the
Strait Alorese varieties, where the final syllables *-uŋ*, *-iŋ*, and *-aŋ* are added to
some words.[5] Alorese varieties on Alor Peninsula and in the Strait apparently
borrowed *bana* 'mud' from either Blagar or Reta, as these languages have the
most similar form.

---

4   Note that the word *bana* in Alorese also means 'forest', from PMP *banua 'inhabited land, ter-
     ritory supporting the life of a community'. Robinson (2015: 22) considers the word for 'forest'
     an Alorese loan into AP languages: Alorese *banna* 'forest' (cf., Lamaholot (Ile Ape) *bəənawa*
     'forest') > Retta *vana*, Adang *bana*, Kula *banan* 'forest'.
5   The Alorese varieties Ternate, Buaya, Alor Besar, Alor Kecil, and Dulolong, which are spoken
     in the Alor-Pantar Strait area form a subgroup that is based on the exclusively shared sound
     change of PAL *w > f in all positions, PAL *ai > ei in word-final position, and the addition of
     the syllables *-uŋ*, *-iŋ* and *-aŋ* in final position in some words (see Sulistyono 2022: 216).

3.4.3 'Gravel'

The Alorese and AP forms for the concept 'gravel' are presented in Table 7.18.

TABLE 7.18   Lexical similarity set associated
with the concept 'gravel'

Language	Alignment
Alorese-Alor Besar	b a l ɔ f a -
Alorese-Alor Kecil	b - l o f a -
Alorese-Dulolong	b a l o f a -
Alorese-Munaseli	g - l o w a r
Blagar-Bakalang	g ə d o b a r
Blagar-Kulijahi	g ə d o w a r
Blagar-Nule	g ə n o v a r
Blagar-Warsalelang	d o l o w a r
Teiwa	d a l a w a r
Adang-Otvai	d a r o f e
Sar	- - - - - - - w a r
Deing	d a l a w i r

For the concept 'gravel' Alorese-Alor Besar, Alorese-Alor Kecil, and Alorese-Dulolong on the Alor Peninsula innovated the forms *balofa* and *blofa* respectively, while Alorese-Munaseli in northeast Pantar innovated *gelovar*. The other Alorese varieties have an inherited form similar to *vato kar:ik* 'gravel' which is constituted of *vato* (< PMP *batu 'stone') and *kar:ik* (< PMP *kədi 'small', Blust & Trussel 2020).

A number of similar forms for the concept 'gravel' are found in several AP languages. Although it is not clear whether all the AP forms are cognates, most of them are related and are likely to be inherited. The forms *dobar/dowar/nowar/lowar/lawar* are preceded by the syllable *gə* in some Blagar varieties, and by the syllable *do/da* in Blagar-Bama and Blagar-Warsalelang, and in Adang, Teiwa, Sar, and Deing. In Blagar, *do-* is a deictic morpheme that means 'up there' (Steinhauer 2014: 159). The correspondence of *d* and *n* in the two words *gedobar* (Blagar-Bakalang) and *genowar* (Blagar-Nule) is also seen in *kədumu* (Blagar-Bakalang) and *kənumu* (Blagar-Nule) 'to suck'.

The form *gelovar* 'gravel' in Alorese-Munaseli is possibly borrowed from several different sources or is a mixed of different forms. The initial syllable *ge-* is similar to Blagar-Bakalang, Kulijai and Nule; the part *-lowar* is similar to Blagar-Warsalelang *dolowar*. The word *balofa* on the Alor Peninsula is also similar to

the form in Blagar-Warsalelang *dolowar* 'gravel'. The sound change *w* > *f* is regular in the varieties of the Alor Peninsula and attested also in other loanwords, such as *safa* 'rice field' from Malay *sawah*. It is unclear why Alorese Alor Besar and Alor Kecil added the initial syllable *ba-*, one possible explanation may be that *ba-* is a shortening of the Malay word *batu* 'stone'. The most likely donor seems to be Blagar-Warsalelang with the form *dolowar* 'gravel', although the differences in forms may point to independent borrowing events in the Alorese varieties.

3.4.4   'Dolphin'

The Alorese and AP forms for the concept 'dolphin' are presented in Table 7.19.

TABLE 7.19   Lexical similarity set associated
with the concept 'dolphin'

Language	Alignment
Alorese-Alor Besar	k u dʒ a - e
Alorese-Alor Kecil	k u dʒ a - i
Alorese-Bana	k u j a - -
Alorese-Buaya	- u dʒ a - e
Alorese-Dulolong	k u dʒ a h i
Alorese-Helangdohi	k u dʒ a - -
Alorese-Munaseli	k u dʒ a - -
Alorese-Pandai	- u dʒ a - -
Alorese-Ternate	k u dʒ a - e
Alorese-Wailawar	k u dʒ a - -
Adang-Lawahing	- u s a h a
Adang-Otvai	- u s a h -
Blagar-Bakalang	k u dʒ - - a
Blagar-Bama	k u dʒ - - a
Blagar-Kulijahi	k u dʒ - - a
Blagar-Nule	k u dʒ - - a
Blagar-Tuntuli	k u dʒ a h -
Blagar-Warsalelang	k u dʒ - - a
Deing	k u - i - -
Kaera	x u j a - -
Blagar-Pura	k u ɟ a - -
Sar-Adiabang	k u j a - -
Teiwa	k u j a ʔ -

The word for 'dolphin' is widely attested both in Alorese varieties and in the AP languages. In other Austronesian languages to the far north of Alor, similar forms are attested; *uas* in Geser-Gorom (south Maluku), and *kuraf* in Uruangnirin (spoken on west Papua), and a Proto Oceanic form *kuriap 'dolphin' has been reconstructed (Blust & Trussel 2020). However, no similar form is attested in the closest relatives of Alorese, the near-by Flores-Lembata languages, and no PMP forms are available for this concept. For this reason, we consider this to be an innovation in Alorese, possibly a loan from AP languages.

In the AP languages, there are similar forms showing regular sound correspondences, which indicate shared ancestry. The sound correspondences enable the reconstruction of a possible early AP form *kujasi 'dolphin'. The initial *k is retained in all AP languages, except Karea where it is reflected as *x* (*xuja*), and Adang where it is usually reflected as glottal stop, but here it is lost (*usaha*). The approximant *j is retained in Teiwa, Kaera, Sar and Blagar-Pura, but lost or changed in others, such as Adang where we find *s*. Medial *s is retained, except in Adang where it is regularly reflected as *h* (*usaha*).

In Blagar, the approximant [j] only occurs in the interjection *jo* 'yes' and a few borrowings, such as the recently adopted Christian name *Yohan* [johan] and the word *rayat* [rajat], borrowed from Indonesian *rakyat* 'the people'. With this evidence, the most likely scenario for this concept seems to be that PAP, the ancestor of AP languages, borrowed the form from an Austronesian donor and when the Alorese arrived in the Alor archipelago, they re-borrowed the form from AP languages. The similarity between the Alorese *kuʤae* and the Blagar word *kuʤa* 'dolphin' may also indicate recent contact, with Blagar then re-borrowing the Alorese form more recently.

3.4.5 'Monitor lizard'

The Alorese and AP forms for the concept 'monitor lizard' are presented in Table 7.20.

TABLE 7.20   Lexical similarity set associated with the concept 'monitor lizard'

Language	Alignment
Alorese-Alor Besar	- - - - - r e h a
Alorese-Alor Kecil	- - - - - r e h a
Alorese-Dulolong	- - - - - r e h a

TABLE 7.20   Lexical similarity set associated
             with the concept 'monitor lizard'
             (*cont.*)

Language	Alignment
Blagar-Bakalang	- - - - - r i h i
Blagar-Bama	- i - r i s -
Blagar-Kulijahi	- - - - - r i h i
Blagar-Nule	- - - - - r i - -
Blagar-Tuntuli	- i - r i s -
Blagar-Warsalelang	- i - r i s -
Blagar-Pura	- a - r i - -
Deing	j e - r i s -
Kaera	- i ʔ r i s -
Kaera	t ɛ ʔ r ɛ s -
Klon-Hopter	w ə - r i h -
Kui	- - - - - r o s -
Nedebang	- - - - - l i s i
Sar-Adiabang	j i - r i s -
Sar-Nule	- - - - - r i s -
Teiwa	- - - - - r i s -

The Alorese varieties on the Alor Peninsula and in the Strait have innovated
the word *reha* 'monitor lizard'. The Alorese varieties on Pantar retain the inher-
ited form *eto/teto damar* (< PWL *eto 'monitor lizard' see Sulistyono 2022: 260;
damar is of unknown origin).

From the distribution and the regular sound changes, it is evident that the
forms found among AP languages are a cognate set and go back to a proto form
(Robinson reconstructed PAP *lVsi 'monitor lizard', 2015: 29)

Based on the evidence presented in Table 7.20, we conclude that the donor
language for the Alorese word *reha* is probably Blagar (Bakalang and Kulijahi)
because it has the form *rihi*, which is the most similar to Alorese *reha*. Reasons
for the change of the non-final vowel *i* to *e* remain unclear, but the Alorese final
a from Blagar *i* in loanwords seems regular, as seen earlier in the Alorese *tera*
'to close' from Blagar *teriŋ* 'to close', and Alorese *lakuk* 'to fold' from Blagar *liku*
'to close'.

3.5 *Miscellaneous: Quantity, Emotions, Motion, Kinship, the Body, Spatial Relations, Sense Perception*

3.5.1 'Ten'

The Alorese and AP forms for the concept 'ten' are presented in Table 7.21.

TABLE 7.21 Lexical similarity set associated with the concept 'ten'

Language	Alignment
Alorese-Alor Besar	k a - r - t o u -
Alorese-Alor Kecil	k a - r - t o u -
Alorese-Bana	k a - r - t o u -
Alorese-Baranusa	k a - r - t o u -
Alorese-Beang Onong	k a - r - t o u -
Alorese-Buaya	k a - r - t o u -
Alorese-Dulolong	k a - r - t o u -
Alorese-Helangdohi	k a - r - t o u -
Alorese-Kayang	k a - r - t o u -
Alorese-Munaseli	k ə - r - t o u -
Alorese-Pandai	k ə - r - t o u -
Alorese-Ternate	k a - r - t o u -
Alorese-Wailawar	k a - r - t o u -
Abui-Fuimelang	k a - r - - - - - - - - -
Abui-Petleng	k a - r - n u k u
Abui-Takalelang	k a - r - n u k u
Abui-Ulaga	k a - r - n u k u
Adang-Lawahing	- a i r - n u - -
Adang-Otvai	ʔ e - r - n u - -
Abui-Atimelang	k a - r - - - - - - - - -
Blagar-Bakalang	- a - r - n u - -
Blagar-Bama	q a - r - n u k u
Blagar-Kulijahi	- a - r - - - - - - - - -
Blagar-Nule	- a - r - n u - -
Blagar-Tuntuli	q a - r - n u k -
Blagar-Warsalelang	x a - r - - - - - - - - -
Blagar-Pura	- a - r i n u - -
Deing	q a - r - n u k -
Hamap-Moru	- a i r - n u - -
Kabola	k a - r - n u - -

TABLE 7.21 Lexical similarity set associated with the concept 'ten' (*cont.*)

Language	Alignment
Kaera	x a - r - - - - - - - - -
Kafoa	k ɑ - r - n u k u
Kamang	k ɑ - r - n ɔ k -
Klon-Bring	k ɑ - r ə n ʊ k -
Klon-Hopter	k a - r - n u k -
Kiramang	k a - r - n u k u
Kui	k a - r - n u k u
Western Pantar-Lamma	k e - - a n u k u
Nedebang	k a - - - - - - - - - - -
Reta	k a - r a n u - -
Sar-Adiabang	q a - r - n u k -
Sar-Nule	q a - r - n u k -
Teiwa	q aː - r - - - - - - - - -
Proto Alor-Pantar	q a - r - - - - - - - - -
Proto Alor-Pantar	q a - r - - - - - - - - -
Proto Timor-Alor-Pantar	q a - r - - - - - - - - -

The Alorese numeral *kartou* 'ten' is formed combining the decimal base *kar* 'tens' and the numeral *tou* 'one'. The form for the decimal base *kar* is a borrowing from AP languages, while the numeral 'one' *tou* is inherited (< PWL *tou) 'one'. Besides the phonological material, Alorese also borrowed the pattern of forming 'ten' as 'ten-one' from AP languages (see Schapper & Klamer 2017 for an extensive description of numerals in Alor-Pantar languages). This is an innovation only present in Alorese, absent from the other Flores-Lembata languages, which all preserve reflexes of the Proto Austronesian form *puluq for 'ten' (Schapper & Klamer 2017: 320 ff.). This loan has also been discussed in Klamer (2011), Robinson (2015), and Moro (2018).

The PAP word *qar- 'tens' has been reconstructed by Holton et al. (2012: 115). As described above, it seems that Alorese only borrowed the part of the numeral that marks tens, *kar-*, but retained the PWL form *tou 'one' (Sulistyono 2022: 428). Since the form is present in all Alorese varieties, it is likely to be an old loan (see § 3.1). The donor is likely to be one which has initial *k* (and most likely one which has the exact syllable *kar*) because Alorese varieties also have initial *kar-*. Among the AP languages that have *kar*, the donor is most likely

one which is spoken close to the coast or located around the Alor-Pantar Strait, such as Klon or Reta.

3.5.2 'Angry'
The Alorese and AP forms for the concept 'angry' are presented in Table 7.22.

TABLE 7.22 Lexical similarity set associated with the concept 'angry'

Language	Alignment
Alorese-Alor Besar	k a - l a k i -
Alorese-Alor Kecil	k a - l a k i -
Alorese-Baranusa	k - - l i k i l
Alorese-Buaya	k - - l e kː i -
Alorese-Dulolong	k ə - l a k i -
Alorese-Kayang	k - - l i k i -
Alorese-Munaseli	k - - l i k i l
Alorese-Pandai	k - - l i k i l
Alorese-Ternate	k a - l a k i -
Alorese-Wailawar	k - - l i k i l
Blagar-Warsalelang	k i - l i k i l
Blagar-Manatang	- - a l i ʔ i l
Blagar Kulijahi	- - - - - - - - - l i l
Kaera	k e ʔ l i k i l
Klon-Hopter	k ə - l i k - -
Western Pantar-Lamma	k - - - - - a k i ŋ
Sar-Adiabang	k - - - - - a k a -
Teiwa	k ə - l e x e l

No similar forms to the Alorese word for 'angry' are attested in the near-by Flores-Lembata languages and no proto forms are available for this concept. In some Alorese varieties, the concept 'angry' is a compound consisting of an inherited root *onoŋ* 'inside' (< PFL *una 'house; inside; hole', Fricke 2019: 464) and the AP loanword *kelikil*. This word is likely a loan from AP languages which present similar forms, for the same concept, and which reflect semi-regular sound changes: PAP initial and medial *l is retained in all languages, with Western Pantar as an exception; PAP medial *k is retained unchanged, but reflected as *x* in Teiwa (Holton & Robinson 2017: 56). The correspondence of Blagar *k*

and Teiwa *x* is regular because it is also seen in other words, such as Blagar *tekil* ~ Teiwa *taxal* 'thin' and Blagar *sokil* ~ Teiwa *soxai* 'to dance'. Some varieties of Blagar have weakened and eventually lost the intervocalic *k and have *aliʔil*, as in Blagar-Manatang, and *lil* as in Blagar Kulijahi (possibly from a form like *kilikil* as in Blagar-Warsalelang). Weakening of intervocalic *k is found also in Kabola, for instance the Blagar *k* and Kabola *ʔ* correspondence is regular as seen in other words, such as Blagar *trukinuk* ~ Kabola *tiʔinu* 'nine' and Blagar *tətoku* ~ Kabola *atoʔo* 'stomach; belly'.

Given that the AP lexemes seem to form a historically related set, and that there are no similar forms attested in the other Flores-Lembata languages, we conclude that the Alorese varieties borrowed this form from AP languages, most likely from Blagar or Kaera. The Alorese varieties spoken on and around the Alorese peninsula (Alor Besar, Alor Kecil, Dulolong, and Ternate) have the form *kalaki*, whereby the vowels *i* have been changed into *a*. The change of the vowel from *i* to *a* is also attested in other Blagar loanwords, such as Alorese *tera* 'to close' from Blagar *teriŋ*, and Alorese *reha* 'monitor lizard' from Blagar *rihi*, and Alorese *lakuk* 'to fold' from Blagar *liku*.

3.5.3   'Road'
The Alorese and AP forms for the concept 'road' are presented in Table 7.23.

TABLE 7.23   Lexical similarity set associated with the concept road

Language	Alignment
Alorese-Baranusa	- - t ɔː r
Alorese-Munaseli	- - t ɔ r
Alorese-Beang Onong	- - t ɔ r
Deing	w u t o r
Deing	- - t o r
Kaera	- - t o r
Teiwa-Lebang	y i t a r
Western Pantar	y a t o r
Kafoa	y a - - - - -
Kui	y a - - - - -
Abui-Takalelang	j a - - - - -

In some Alorese varieties on Pantar, Alorese-Baranusa, Beang Onong and Munaseli, we find the AP loan *tor* 'road'. Klamer (2011: 105) and Robinson (2015: 28) list this form as an AP loanword into Alorese, from Western Pantar *ya tor* 'road'. The form *ya tor* is widespread among AP languages. The part *ya* is the AP word for 'road' and *tor/tar* is found in Kaera, Deing, Teiwa, and Western Pantar. The *tor/tar* element is semantically related to the word for 'tail' in AP languages (PAP *ora 'tail', see Holton & Robinson 2017: 78), such as Teiwa *t-or* 'tail; tailbone' and Klon *t-or* 'bone', both with the possessive prefix *t-*. We suggest that a semantic shift from 'tail' to 'main road' has taken place in some languages, probably due to the fact that a road with curves does resemble an animal's tail. Western Pantar is the only language where the compound is still complete. The other languages have either lost the *ya* part or the *tor* part. However, it is also possible that the varieties that only have *ya*, like Kafoa, Abui and Kui, might never have had the compound *ya tor*. In Abui-Takalelang, *foqa* means 'big'; thus, *ja foqa* means 'big road; highway'. Western Pantar and Deing are the most likely donor for this loanword.

3.5.4 'Younger sibling'

The Alorese and AP forms for the concept 'younger sibling' are presented in Table 7.24.

TABLE 7.24 Lexical similarity set associated with the concept 'younger sibling'

Language	Alignment
Alorese-Bana	- - k a u -
Alorese-Buaya	- - k a u -
Alorese-Munaseli	- - k a u -
Alorese-Pandai	- - k a u -
Blagar-Kulijahi	- - k a - w
Blagar-Warsalelang	- - k a - w
Blagar-Bakalang	- - k a - w
Blagar-Nule	n e k a - w
Blagar-Tuntuli	p i k a - w
Blagar-Pura	- e k a k u
Sawila	n i k a k u
Kula-Lantoka	ŋ a k a k u
Teiwa	n a k a ʔ a w
Wersing-Maritaing	n ɛ k a u k

TABLE 7.24   Lexical similarity set associated
with the concept 'younger sibling'
(*cont.*)

Language	Alignment
Wersing-Taramana	n e k a k u
Reta-Ternate	g a k a k u
Reta-Pura	- - k a k u
Western Pantar-Lamma	- i a k u

As for the concept of 'younger sibling', the form *kau* is quite widespread among the Alorese varieties. In Alorese-Kayang and Alorese-Wailawar, a medial glottal stop has been inserted. Another term for this concept in Alorese is *aring* 'younger sibling', which is related to the Lamholot forms *aring/arik* (Fricke 2019: 529).

The Alorese *kau* form shows similarities with several AP languages, in which the forms go back to PTAP *kaku 'younger relative' (see Schapper & Huber, this volume). In some AP languages, the form is presented with a possessive prefix. Blagar, Kaera, and Teiwa are all possible donors.

Interestingly, it seems that this form is highly borrowable, as it listed by Schapper and Huber (this volume), among the TAP etyma into the Austronesian languages of Timor. Unlike on Alor and Pantar, where the form has been borrowed together with its original meaning, on Timor the form has undergone a semantic shift from PTAP *kaku 'younger relative' to Makasae and Makalero 'small'.

3.5.5   'To bury'

The Alorese and AP forms for the concept 'to bury' are presented in Table 7.25.

TABLE 7.25   Lexical similarity set associated
with the concept 'to bury'

Language	Alignment
Alorese-Alor Besar	t - - o - - u -
Alorese-Buaya	t - - o - - u -
Alorese-Dulolong	t - - u - h o
Alorese-Alor Kecil	t - - o - h u -

TABLE 7.25 Lexical similarity set associated with the concept 'to bury' (*cont.*)

Language	Alignment
Alorese-Ternate	t - - o - - u -
Blagar-Bakalang	t - - o - - - - - w
Blagar-Bama	t - r o - k u -
Blagar-Kulijahi	t - r o - - u -
Blagar-Nule	t ə r o - - - - - w
Blagar-Tuntuli	t o r o - k u -
Blagar-Warsalelang	t ə r o - k u -
Blagar-Pura	t a r o - - u -
Kaera	t - r a ʔ q o -
Makasae	t a r - - - - - u -
Teiwa	t a r a - x a ʔ
Kamang	f o ɹ u

The Alorese word *tou* 'to bury' is likely to be an AP loanword in Alorese varieties on the Alor Peninsula and in the Strait, because no similar forms are attested in the near-by Flores-Lembata languages and no proto forms are available for this concept. Conversely, the AP forms are historically related and reflect regular sound changes (Blagar-Bakalang *tou*, Blagar-Tuntuli *toroku*, Teiwa *taraħaʔ*, Kamang *foɹu*). Proto AP *tVroqu 'to bury' may be reconstructed because initial *t- is attested regularly in most of the AP languages and the intervocalic *-r- is also expected to appear unchanged in most of the languages. In Kamang, intervocalic *-r- is expected to change into *l*, but in one of the varieties in the Kamang cluster, namely Tiyei, it has changed into *ɹ*. As for the vowels, the correspondence Blagar *o* and Teiwa *a* is regular, and attested in other words such as Blagar-Tuntuli *bogori* 'yellow' ~ Teiwa *bahari* 'yellow'.

Among the AP languages that have a reflex of this form, Blagar-Bakalang has the most similar form to Alorese, suggesting that Alorese borrowed *tou* 'to bury' from this Blagar variety. In Alor Kecil, the addition of intervocalic *h*, as seen in *tohu* 'to bury' is also seen in other words, such as Alor Besar *tafeuŋ* ~ Alor Kecil *təfihuŋ* 'fog'.

3.5.6 'Heart'

The Alorese and AP forms for the concept 'heart' are presented in Table 7.26.

TABLE 7.26 Lexical similarity set associated with
the concept 'heart'

Language	Alignment
Alorese-Alor Besar	- - - - - k u b - a ŋ -
Alorese-Alor Kecil	- - - - - k u b - a ŋ -
Alorese-Bana	- - - - - k u b - a ŋ -
Alorese-Beang Onong	- - - - - k u b - a ŋ -
Alorese-Buaya	- - - - - - - u b - a ŋ -
Alorese-Dulolong	- - - - - k u b - a ŋ -
Alorese-Kayang	- - - - - k u b - a ŋ -
Alorese-Munaseli	t a p k u b - a ŋ -
Alorese-Pandai	- - - - - k u b - a ŋ -
Alorese-Ternate	- - - - - k u b - a ŋ -
Alorese-Wailawar	- - - - - - - u b - a ŋ -
Blagar-Bakalang	- - - - - k u b - a ŋ -
Blagar-Bama	- - - - - k u b - a ŋ -
Blagar-Kulijahi	- - - - - k u b - a ŋ -
Blagar-Nule	- - - - - k u b - a ŋ -
Blagar-Tuntuli	- - - - - k u b - a ŋ -
Blagar-Warsalelang	- - - - - k u b - a ŋ -
Kui	- - - - - k u b l a - i
Klon	- - - - - k u b
Reta-Pura	- - - - - k u m b a
Blagar-Pura	- - - - - k u b - a ŋ -
Wersing-Maritaing	- u - k a b - a ŋ -
Wersing-Taramana	g e u k a b - a ŋ -

For the concept 'heart', Alorese varieties innovated the form (*tapo/təpo*) *kubaŋ*. The part *tapo/təpo* means 'coconut' and is inherited (< Lamaholot-Kedang #tapu, see Samely 1991; Sulistyono 2022: 242), while the part *kubaŋ* is borrowed from AP languages. The form for 'heart' in AP languages is often given with a possessive prefix (*ge-* in Wersing, *ta-/eta-* in Klon and Reta)).

Based on the AP cognates presented in the table, a tentative PAP form*kVbaŋ 'heart' may be reconstructed. The PAP initial *k- is regularly retained as *k* in all

languages. Even though the PAP intervocalic *-b- is expected to be reflected as *p* in Wersing, a similar retention of intervocalic *-b- happens also in PAP *-lebur > Wersing *jebur* 'tongue' (Holton et al. 2012: 115).

The Alorese form (*tapo*/*təpo*) *kubaŋ* 'heart' is probably a loanword from Blagar, as this language has the form that is most similar to the Alorese form. As for the addition of the (*tapo*/*təpo*) 'coconut' part, a possibility may be that the Alorese have re-analyzed the first person plural inclusive or reciprocal prefix *t*V-, which is often attached to body parts, as the first syllable of the word *tapo* 'coconut', and hence have added this word to the concept for 'heart'.

Robinson (2015: 24) proposed the opposite pattern, namely that this is an Alorese loanword into Blagar and Wersing. This proposal was based on the similar form *taʔ kubaŋ* 'heart' found in Kedang. However, the collection of more AP forms, and the internal diversity among the AP languages, suggest that the form *kubaŋ* is likely of AP origin, while the Kedang word *taʔ kubaŋ* is a loanword from Alorese, or an AP loanword into Kedang via Alorese.

3.5.7 'To breathe'

The Alorese and AP forms for the concept 'to breathe' are presented in Table 7.27 (repeated from Table 7.2).

TABLE 7.27 Lexical similarity set associated with the concept 'to breathe'

Language	Alignment
Alorese-Munaseli	h ɔ - p a ŋ
Blagar-Bama	s o - p a ŋ
Blagar-Kulijai	h o - p a ŋ
Blagar-Nule	h o - p a ŋ
Blagar-Pura	h o - p a ŋ
Deing	- o - p a ŋ
Kaera	s u ʔ p a ŋ
Western Pantar-Tubbe	h o - p a ŋ
Reta-Pura	h oː - p a ŋ
Reta-Ternate	h u - p a ŋ

Alorese-Munaseli has innovated the word *hopang* 'to breathe' which is similar to forms attested in several AP languages, such as Blagar-Kulijahi, Nule,

Pura *hopang*, Western Pantar-Tubbe *hopang*, Kaera *supang*. Since, the AP forms are related and follow semi-regular sound changes (PAP initial *s > Kaera *s*, Blagar *h*, see Holton & Robinson 2017: 56), we consider this a loanword from AP languages into Alorese-Munaseli, with the most likely donor being a Blagar variety.

3.5.8 'Small'
The Alorese and AP forms for the concept 'small' are presented in Table 7.28.

TABLE 7.28   Lexical similarity set associated with the concept 'small'

Language	Alignment
Alorese-Alor Besar	k a e
Alorese-Alor Kecil	k a e
Alorese-Dulolong	k a e
Alorese-Ternate	k a e
Alorese-Buaya	k a e
Hamap	k a ʔ i - -
Kabola	k a ʔ a - i
Adang	k a ʔ a - i
Kaera	k i k i -
Blagar	k i k i -

The Alorese varieties on Pantar use three different inherited forms for the concept 'small'. There are two forms of PMP origin: *anaŋ* (< PMP *anak) 'small' and *kar:i* (< PMP *kədi 'small in size'), and one form which can only be traced back to PWL *kesi/*kisu 'small' > *kihu* 'small' (Sulistyono 2022: 264). The Alorese varieties on the Alor Peninsula and in the Strait, however, have innovated a new form *kae* which suggests an external source. We do not group *kae* 'small' together with *kari* 'small' because all the Peninsula and Strait varieties consistently use the form *kae*, in doing so they differ from the conservative Pandai variety which retains *kari* < PMP *kədi 'small in size'.

The innovative form *kae* 'small' may have been borrowed from Adang *kaʔai* 'small', with the loss of medial glottal stop, which Alorese varieties lack. Several AP languages have similar words. The change of *k into intervocalic *-ʔ-* in Adang is regular (Holton et al. 2012: 94). However, the change of *-in* to *-ai* in Adang and Kabola remains unexplained.

Robinson (2015: 23) holds a different view on *kae* 'small', which she considers a word of Austronesian origin due to the similarity of Alorese *kae* with Kedang *keke* and Tetun *kiʔik* 'small'.

The relationship between Alorese *kae* and Kedang *keke* (and Tetun *kiʔik*) is weak, and it is more likely that Alorese varieties borrowed *kae* from Adang *kaʔai* 'small'. About the origin of Kaera *kiki*, Blagar *kiki*, and Adang *kaʔai* 'small', we agree that the origin of the AP forms may ultimately be from an Austronesian language spoken in the area before the arrival of the Alorese.

3.5.9   'To close'
The Alorese and AP forms for the concept 'to close' are presented in Table 7.29.

TABLE 7.29   Lexical similarity set associated with the concept 'to close'

Language	Alignment
Alorese-Alor Kecil	- - - - - t e r a -
Alorese-Bana	- - - - - t e r a ʔ
Alorese-Baranusa	- - - - - t e r a -
Alorese-Beang Onong	- - - - - t e r a -
Alorese-Alor Besar	- - - - - t ɛ r a -
Alorese-Dulolong	- - - - - t e r a -
Alorese-Helangdohi	- - - - - t e r a ʔ
Alorese-Kayang	- - - - - t a r a ʔ
Alorese-Munaseli	- - - - - t ɛ r a ʔ
Alorese-Pandai	- - - - - t ɛ r a -
Alorese-Ternate	- - - - - f e r a -
Alorese-Wailawar	- - - - - t e r a -
Kula-Lantoka	- - - - - t i r a -
Deing	- - - - - t i a r
Kaera	w a n t ɛ r i ŋ
Tubbe	- - - - - t i a r I ŋ
Sawila	- l i ˈt i r a
Reta-Pura	u - - t i a l i
Adang-Lawahing	w a t ɛ l ɛ
Adang-Otvai	- - - - - uˈt e l
Blagar-Bakalang	v e n t e r i ŋ
Blagar-Bama	v e n t e r i ŋ
Blagar-Nule	v e n t e r i ŋ

TABLE 7.29   Lexical similarity set associated
              with the concept 'to close' (*cont.*)

Language	Alignment
Blagar-Tuntuli	v e n t e r i ŋ
Reta-Ternate	- - u t i e l i
Wersing-Maritaing	l ɛ t e r
Wersing-Taramana	l e t e r
Klon-Hopter	ʔ u ˈt ɛː r
Kiramang	- u t ɛ r
Kui-Labaing	- u t e r i
Kabola-Monbang	w h u ˈt e l e
Proto Alor-Pantar	- t i a r i n

For the concept 'to close', Alorese varieties innovated the form *tera(ʔ)* 'to close' which is different from the form #letuʔ found in Lamholot varieties (see Fricke, this volume, Table 5.9). It is unclear why Alorese-Ternate has *fera* with initial *f*. AP languages display similar forms, all reflexes of the PAP form *-tiari(n) (see Holton & Robinson 2017: 78). In many AP languages the root is preceded by an applicative prefix or by another verb: in Kaera *waŋ* is a verb which means 'be; exist' and occurs in serial verb constructions with various functions (Klamer 2014: 137). Considering that in AP languages the form is inherited, the form *tera(ʔ)* in Alorese looks like an AP loanword, whereby Alorese varieties have borrowed the root *teri* 'to close' from either Kui, Kaera or most likely Blagar. The change of the ultimate vowel from *i* to *a* in loans is also attested in other Blagar loanwords, such as Alorese *kalaki* 'angry' from Blagar *kilikil*, Alorese *reha* 'monitor lizard' from Blagar *rihi*, and Alorese *lakuk* 'to fold' from Blagar *liku*.

3.5.10   'To hide'
The Alorese and AP forms for the concept 'to hide' are presented in Table 7.30.

TABLE 7.30   Lexical similarity set associated
              with the concept 'to hide'

Language	Alignment
Alorese-Alor Besar	d a - f u - -
Alorese-Bana	d ə - w u - -

TABLE 7.30   Lexical similarity set associated
with the concept 'to hide' (*cont.*)

Language	Alignment
Alorese-Beang Onong	d a - wː u - -
Alorese-Beang Onong	d a - - u - -
Alorese-Buaya	d a - fː u - -
Alorese-Dulolong	d a - f u - -
Alorese-Dulolong	d a - - u - -
Alorese-Helangdohi	d ə - w u k -
Alorese-Kayang	d a - wː u - -
Alorese-Munaseli	d ə - w u k -
Alorese-Pandai	d a - w u - -
Alorese-Ternate	d a - f ː - u
Alorese-Wailawar	d e - w u - -
Abui-Takalelang	t a - b u - -
Adang	t a w u n i ŋ
Kabola	t ə w u n i
Reta-Pura	t a ɓ u n i ŋ
Western Pantar	- - - - - u n n i ŋ

For the concept 'to hide' all Alorese varieties display the form *dawu/ dəwuk* (in Pantar) or *dafu* (on the Alor Peninsula and in the Strait) that is not attested in the near-by Flores-Lembata languages. The change of *v* into *f* in Alor Peninsula and in the Strait varieties is regular (Sulistyono 2022: 214). This form is an innovation, possibly an old one, as it is found in all Alorese varieties. The source are AP languages, which present forms that are similar to the Alorese ones. According to Robinson (2015: 25), the AP forms were borrowed from Malay (*buɲyi* 'hide'), or from another Austronesian language going back to PMP *buni 'to hide'.

Some of the AP languages attached the reciprocal prefix *tV-* to the root and obtained forms like *taɓuniŋ* (Reta) or *təwuni* (Kabola). Since the Alorese forms *dawu/ dəwuk* are more similar to the AP forms (with the *tV-* prefix) than they are to PMP *buni, we conclude that Alorese borrowed this form from AP languages, rather than inheriting it from PMP.

3.5.11   'Dirty'
The Alorese and AP forms for the concept 'dirty' are presented in Table 7.31.

TABLE 7.31   Lexical similarity set associated with the concept 'dirty'

Language	Alignment
Alorese-Alor Besar	k a l i t a -
Alorese-Alor Kecil	k - l i t a -
Alorese-Bana	k - l i t a ʔ
Alorese-Baranusa	k - l i t a k
Alorese-Beang Onong	k a l i t a -
Alorese-Buaya	k a l i t a -
Alorese-Dulolong	k a l i t a -
Alorese-Helangdohi	k - l i t a ʔ
Alorese-Kayang	k - l i t a ʔ
Alorese-Munaseli	k - l i t a ʔ
Alorese-Pandai	k - l i t a -
Alorese-Ternate	k a l i t a -
Alorese-Wailawar	k - l i t a ʔ
Blagar-Bakalang	k - l i t a k
Blagar-Kulijahi	k ə l i t a h
Blagar-Nule	k ə r i t a k
Blagar-Pura	k a r i t a -
Reta	k a r i t a -
Teiwa	k - l i t a ʔ

For the concept 'dirty', all Alorese varieties innovated a form like #k(a)lita(ʔ/k), different from the remaining Flores-Lembata languages that use a form reconstructable to PWL *mila 'dirty' (Sulistyono 2022: 245, 401). Klamer (2011: 105) lists this form among the Alorese loanwords from AP languages. Robinson (2015: 24), on the contrary, assumes this to be an Alorese loanword into AP languages, because it has a potential cognate in nearby Austronesian languages: Alorese *kalita* (cf., Lamaholot (Ile Ape) *prita*).[6] We agree with Klamer and consider this an AP loanword into Alorese varieties for two reasons: (i) the AP forms share regular sound changes, (ii) the relationship between Alorese *kalita* and Lamaholot (Lamatuka) *prita* is weak.

---

6   According to LexiRumah 3.0.0, which reports data from Keraf 1978, it is Lamaholot Lamatuka which has *prita* for 'dirty', and not Ile Ape, which has *milan* 'dirty'.

The reflexes of AP forms show regular sound correspondences. In Blagar medial *l* corresponds to Reta medial *r* as seen in several other words, such as Blagar *bulaŋ* ~ Reta *buraŋ* 'sky' and Blagar *bulit* ~ Reta *kaburit* 'arrow'. A similar form *kilaʔɛ* 'dirty' is attested in Fataluku (a Papuan/Timor-Alor-Pantar language spoken in east Timor), which strengthens the proposal that this set is of AP origin. In addition to that, a similar cognate set with different semantic meaning, namely 'old; elderly (people)' is attested across AP languages, namely Abui *kalieta/kaleita*, Kafoa *kalta*, Kiramang *kaleta*, and Kui *kakaleta* 'old; elderly (people)'. A semantic change might have occurred within the AP languages from 'old; elderly' to 'dirty'. Finally, a comparison showed no correspondences between Alorese initial *k* and Lamatuka *p*. Given this evidence, we consider this a loan in Alorese from AP languages.

3.6 *Distribution of Loanwords*

Not all 28 AP loanwords occur in all 13 Alorese varieties. Some loanwords occur in all Alorese varieties, while others have a more limited geographical spread. Table 7.32 presents the distribution of loanwords in five groups: all Alorese varieties, only the varieties in northeast Pantar, only northeast Pantar and Alor Peninsula, only in the Alor Peninsula and in the Strait, and finally only in Pantar. The distribution of loanwords is informative about the relative age of loanwords, because loanwords attested in all Alorese varieties as regularly inherited forms were borrowed very early on before Alorese spread on the coastal areas of Alor and Pantar. The second group of loanwords are also possibly quite old, as those are found in northeast Pantar varieties, the area that is considered to be the homeland of the Alorese (see Sulistyono 2022).

TABLE 7.32 Distribution of loanwords based on geographic groups

Geographic groups	Concepts
All Alorese varieties	'heart', 'ten', 'younger sibling', 'angry', 'taro', 'to close', 'to hide', 'dolphin', 'dirty'
Northeast Pantar	'rattan', 'garden', 'digging stick', 'fish trap', 'coral rock', 'to breathe', 'to fold', 'to pull', 'to pray'
Northeast Pantar and Alor Peninsula	'bed', 'gravel', 'adultery'
Alor Peninsula and Strait	'monitor lizard', 'small', 'to bury', 'mud'
Pantar	'road', 'to wash'
Alor Peninsula	'root'

The semantic fields add a perspective on the type of contact. There is a difference between the early AP loans and the more recent loans attested in, for instance, the Alor Peninsula varieties. On the one hand, the early loans contain more basic vocabulary, such as numerals ('tens'), a kinship term ('younger sibling'), emotions ('angry') and body parts ('heart'). On the other hand, the more recent loanwords mainly concern nouns, particularly relating to the physical world, such as 'mud', 'monitor lizard', and 'root'.

Not surprisingly, since northeast Pantar is likely the homeland of the Alorese, the Alorese varieties more prone to borrowing are Alorese-Munaseli with 19 loanwords, and Alorese-Pandai with 15 loanwords on northeast Pantar, followed by the Alorese varieties on the Alor Peninsula (Alor Kecil, Dulolong and Alor Besar) (see Table 7.33), which are the second oldest group after the varieties of Munaseli and Pandai.

TABLE 7.33   Alorese varieties with their number of loanwords

Variety	Number of loanwords	Concepts
Munaseli	19	Adultery, angry, bed, coral rock, dirty, dolphin, garden, gravel, heart, rattan, road, small, ten, to breathe, to close, to hide, to pray, to wash, younger sibling
Pandai	15	Angry, bed, dirty, dolphin, fish trap, heart, rattan, small, ten, to bury, to close, to hide, to pull, to wash, younger sibling
Alor Kecil	13	Angry, dirty, dolphin, gravel, heart, monitor lizard, mud, root, taro, ten, to bury, to close, to hide
Dulolong	13	Angry, dirty, dolphin, gravel, heart, monitor lizard, mud, root, taro, ten, to bury, to close, to hide
Alor Besar	13	Angry, adultery, bed, dirty, dolphin, gravel, heart, monitor lizard, taro, ten, to bury, to close, to hide
Ternate	9	Angry, dirty, dolphin, heart, mud, ten, to bury, to close, to hide
Buaya	8	Angry, dirty, dophin, heart, ten, to bury, to hide, younger sibling
Bana	9	digging stick, dirty, dolphin, heart, small, ten, to close, to hide, younger sibling
Wailawar	8	Angry, dirty, dolphin, heart, small, ten, to close, to hide
Baranusa	8	Angry, dirty, road, taro, ten, to close, to hide, to wash
Helangdohi	7	Digging stick, dirty, dolphin, small, ten, to close, to hide

TABLE 7.33   Alorese varieties with their number of loanwords (*cont.*)

Variety	Number of loanwords	Concepts
Kayang	6	Angry, dirty, heart, ten, to close, to hide
Beang Onong	6	Dirty, heart, road, ten, to close, to hide

The varieties with the smallest number of loanwords are the most recent ones, such as Beang Onong, which was established in the early 1960's, and Bana and Wailawar established in 1966 and 1996 respectively (see §1).

As for the main donor language(s), the AP donor languages are mainly languages spoken around the Alor-Pantar Strait. Blagar had an important role as donor language. In fact, out of 28 loanwords, 20 are likely to come from Blagar (or at least have Blagar among the possible donors): 'fish trap', 'bed', 'to pull', 'to fold', 'adultery', 'digging stick', 'garden', 'rattan', 'taro', 'mud', 'coral rock', 'gravel', 'monitor lizard', 'dolphin', 'younger sibling', 'to close', 'to breathe', 'to bury', 'heart', and 'angry'. That Blagar is the dominant donor comes as no surprise, since Alorese and Blagar have a close, historical relationship. Both communities are bound in a century-old sociopolitical alliance, called *Galiyao Watang Lema* (see Sulistyono 2022: 15–16).

Other AP languages around the Alor-Pantar Strait that have also contributed AP loanwords to Alorese are Adang, Klon, and Kaera. The contribution of these languages varies according the Alorese subgroup in question. Adang is more likely to be the donor of loanwords found in the Alor Peninsula varieties, while Klon is more likely the donor for loanwords found both in northeast Pantar and Alor Peninsula varieties. Kaera probably had one of the earliest contacts with Alorese, because almost all loans from Kaera belong to the first group. Western Pantar and Deing are donors only for Alorese varieties spoken on Pantar.

4   Discussion

In the previous section, we have presented evidence for lexical borrowing from the AP languages into Alorese. After a close inspection, applying automatic lexical similarity detection, and subsequently a qualitative fine-grained analysis (see §2), we have detected 28 loanword events between Alorese and AP languages on a list of 596 items. The percentage of AP loanwords in Alorese is, thus, approximately 4.7%, confirming previous results of Klamer (2011) and Robin-

son (2015), which were based on smaller datasets. This result shows, on the one hand, that the percentage of AP loanwords in Alorese is indeed small, and on the other hand, that conducting loanword analysis on a Swadesh list, like in Klamer (2011), is likely to give a representative figure of the number of loanwords in a language. Having said this, the innovative use of automatic lexical similarity detection used in the present chapter looks promising and it deserves to be tested further in other studies, because it allows the screening of large datasets and a comparison across language families in a short amount of time. An obvious issue is that the first screening by distribution patterns can turn up false positives (forms marked as related which turn out not to be). These can be filtered out, but a small chance remains that few additional, actual loanwords are not found because they have a spurious similarity to e.g., other Austronesian forms, which makes them not pattern as loanwords. There might be other caveats that we are not aware of, which future studies using the same methodology may unravel.

The limited lexical influence from AP languages into Alorese is not so peculiar if seen in a broader geographical perspective. Similar findings are reported in two contributions of the present volume: Schapper and Huber (this volume), who focus on lexical borrowing from Papuan languages into Austronesian languages of Timor; and Klamer (this volume), who presents evidence for the opposite pattern, namely ancient Austronesian words attested in the lexicon of the TAP languages. However, the way these two studies compiled their dataset was very different from ours. An interesting result, that is shared by Schapper and Huber's, Klamer's and our contribution is that, despite the length of contact, the number of loanwords is relatively small: a dozen loanwords in Schapper and Huber, 14 ancient loanwords in Klamer, and 28 loanwords in our study. For Schapper and Huber (this volume), one possible explanation for the small number of loanwords is the lack of data from the Papuan languages of Timor, especially in the semantic field of plants and animals, which is a domain that attracts a considerable number of Papuan borrowings. According to Klamer (this volume), the limited and scattered lexical borrowing from Malayo-Polynesian languages into TAP languages points to a contact scenario involving relatively superficial contact in few socio-cultural domains such as trade or marriage, which does not require a community to be bilingual.

We now turn to the discussion of the 28 AP loanwords, that, despite the small number, can still be regarded as significant and informative about the type of influence that AP languages had on Alorese. AP lexical influence on Alorese is reflected by loans involving agriculture and vegetation (digging stick, garden, rattan, root, taro), the physical world (coral rock, mud, gravel), animals (dol-

phin, monitor lizard), and basic actions and technology (fish trap, bed, to fold, to pull, to wash). So, it seems that the AP languages mainly contributed with terms referring to the environment, or referring to tools and actions related to the environment. This is confirmed by the study of Schapper and Huber (this volume), who show that Papuan lexical influence on the Austronesian languages of Timor is mostly found in the domains of plants and animals. A similar result is also presented in Edwards (this volume) who found that possible loanwords or innovation in the regional and west Timor strata of the Rote-Meto lexicon are robustly attested in semantic spheres very prone to borrowing, such as *Tools* and *Vegetation*. This is especially interesting, if seen in contrast with the Austronesian lexical influence on the TAP languages (Klamer, this volume), which is reflected by loans involving textile technology (needle, to weave, to sew); societal structures (slave, king/ruler), subsistence and trade (salt, seed, maize, skin), and marriage (bride price).

This limited lexical influence does rule out an adult language-shift scenario in the Alor archipelago, because this is usually accompanied by the retention of (specialist) vocabulary from the heritage language (Ross 2013; see also Klamer, this volume).[7] The wholesale adoption of a good amount of lexical items is very frequent when there is an unequal relation between the languages, such that one community shifts to another language, and in doing so, it retains parts of the L1's vocabulary,[8] or one community adopts many words from a prestigious L2 (Muysken 2013). Neither of these scenarios applies in the case of Alor and Pantar. In the Alor archipelago, bilingualism involving Alorese and AP languages was long and stable, as is today, and never ended in a shift.

Evidence from contact-induced grammatical changes in Alorese show that Alorese was initially spoken in bilingual communities characterized by symmetric bilingualism, dense social networks, and low normativity, with many bilingual children who introduced new grammatical constructions in Alorese on the model of their AP languages (Moro 2018, Moro & Fricke 2020). After this period, which was relatively short, Alorese communities became larger, networks were looser, the language started to enjoy more prestige and became a lingua franca in the area (see §1). Consequently, Alorese was learned as an L2 by many adult AP speakers, and the outcome of this type of contact was severe simplification of morphology (Klamer 2012, 2020; Moro 2019). These acquisitional and socialisation patterns can still be observed today, as local people on Alor report that many Adang speakers can speak Alorese, but that

---

7   According to Ross (2013: 30), adult language shift appears to have been rare in Melanesia.
8   This shift scenario is hypothesized for Rote-Meto (Edwards, this volume).

Alorese people cannot speak Adang. Therefore, the asymmetric bilingual patterns that have started sometime in the past continue to the present day (see Moro 2021).

Two factors, thus, explain the relatively small amount of AP loanwords in Alorese. First, as discussed above, the bilingualism situation that led to grammatical borrowing did not last long enough, and when Alorese became more prestigious, the pattern became asymmetric. The fact that bilingualism in the AP language(s) was not reciprocated by the Alorese prevented the adoption of AP words in the Alorese language. Second, it is likely that in the exogamous Alorese community, the spouses came from different AP communities and thus spoke different AP languages, as we can still observe in Munaseli today. In a fieldwork trip conducted in 2016, Francesca Moro recorded 12 AP speakers who had married an Alorese spouse and had moved into the Alorese Munaseli community: they had six different L1s: Kroku (five speakers), Blagar (three speakers), Teiwa (one speaker), Sar (one speaker), Kaera (one speaker), Klamu (one speaker). So, a possible answer to the question "why did the Papuan mothers not introduce more of their native Papuan *lexicon* into the Alorese they used?" (cf. Klamer 2012: 104), is that the many different AP languages involved might have prevented heavy lexical borrowing from one specific AP language. A similar outcome is found in creoles, where the presence of several L1s interfering with each other prevents transfer from a single L1 (cf. Muysken 2013: 717). We can conclude that the bilingualism had more influence on the grammar of Alorese than on its lexicon, as the grammar usually falls below the threshold of consciousness, and the grammatical changes were either shared by almost all the L1s (presence of a plural word, converged give-constructions, see Moro 2018 and Moro & Fricke 2020), or they were simplification process independent of the L1s (loss of inflectional morphology, see Moro 2019).

Finally, Schapper and Huber (this volume) point out that "it is important not to exclude a lexeme as a possible loan candidate just because it has a known Austronesian etymology". We agree with this observation, as we also report cases, such as 'to hide' or 'dolphin', where lexemes coming from an Austronesian source were borrowed into AP languages, and then from there borrowed again into Alorese.

To conclude, we inspected the whole available lexicon (~600 words) of 13 Alorese varieties and found that, despite the length of contact between Alorese and AP speakers, the presence of AP loanwords is 'only' 4.7 %. The bilingualism scenario found in Alorese-AP communities had more influence on the grammar of Alorese than on its lexicon. This limited lexical influence is accounted for by the asymmetric bilingualism patterns and by the presence of several L1s interfering with each other. Yet, the AP loanwords can tell us that contact

between the Alorese and AP speakers revolved around agriculture and vegetation, the physical world, and basic actions and technology, and that Blagar had an important role as donor language, probably due to its position on Pantar and in the Strait.

Acknowledgements

References

Anonymous (1914). De eilanden Alor en Pantar, Residentie Timor en Onderhoorigheden. *Tijdschrift van het Koninklijk Nederlandsch Aardrijkskundig Genootschap* 31, 70–102.

Blust, Robert A. and Stephen Trussel (2020). *The Austronesian Comparative Dictionary, Web edition*. Available at: http://www.trussel2.com/ACD.

Doyle, Matthew (2010). *Internal divisions of the Flores-Lembata subgroup of Central Malayo-Polynesian*. Master thesis, Leiden University.

DuBois, Cora (1944). *The people of Alor: A social-psychological study of an East Indian Island*. Harvard University Press.

Eberhard, David M., Gary F. Simons, and Charles D. Fennig (2019). *Ethnologue: Languages of the world* (22nd ed.). SIL International. http://www.ethnologue.com

Edwards, Owen (this volume). Substrate Transference in Rote-Meto.

Elias, Alexander (2017). Subgrouping the Flores-Lembata languages using historical glottometry. Presentation on the *9th International Austronesian and Papuan languages and linguistics conference* (APLL-9) 21–23 June 2017, Paris.

Fernandez, Inyo Yos (1996). *Relasi historis kekerabatan Bahasa Flores*. Penerbit Nusa Indah.

Fricke, Hanna (2019). *Traces of language contact: The Flores-Lembata languages in eastern Indonesia*. Doctoral dissertation, Leiden University.

Holton, Gary, Marian Klamer, Frantisek Kratochvíl, Laura C. Robinson, and Antoinette Schapper (2012). The historical relations of the Papuan languages of Alor and Pantar. *Oceanic linguistics* 51(1): 86–122.

Holton, Gary and Laura C. Robinson (2017). The internal history of the Alor-Pantar language family. In Klamer, Marian (ed.), *The Alor-Pantar Languages: History and Typology*, pp. 49–91. Berlin: Language Science Press.

Kaiping, Gereon and Marian Klamer (2019). Subgroupings of the Timor-Alor-Pantar languages using systematic Bayesian inference. *SocArXiv*: 1–48.

Kaiping, Gereon, Owen Edwards, and Marian Klamer (eds.) (2019). *LexiRumah 3.0.1.* Leiden: Leiden University Centre for Linguistics. Available at: http://www.model-ling.eu/lexirumah/.

Klamer, Marian (2010). *A Grammar of Teiwa*. Berlin: Mouton De Gruyter.

Klamer, Marian (2011). *A Short Grammar of Alorese (Austronesian)*. München: Lincom Europa.

Klamer, Marian (2012). Papuan-Austronesian language contact: Alorese from an a real perspective. In Nicholas, Evans and Marian Klamer (eds.), *Melanesian languages on the Edge of Asia: Challenges for the 21st Century*, pp. 72–108. University of Hawai'i Press.

Klamer, Marian (2014). Kaera. In Schapper, Antoinette (ed.), *The Papuan Languages of Timor, Alor and Pantar: Volume 1*, pp. 97–146. Berlin: Mouton De Gruyter.

Klamer, Marian (2015). *Language as a time machine*. Inaugural lecture. https://scholarly publications.universiteitleiden.nl/handle/1887/32306

Klamer, Marian (2017). The Alor-Pantar languages: Linguistic context, history and typology. In Klamer, Marian (ed.), *The Alor-Pantar languages: History and typology*, pp. 1–49. Berlin: Language Science Press.

Klamer, Marian (2020). From Lamaholot to Alorese: Morphological loss in adult language contact. In Gil, David and Antoinette Schapper (eds.), *Austronesian Undressed: How and why languages become isolating*, pp. 339–368. Amsterdam: John Benjamins.

Klamer, Marian (this volume). Traces of pre-modern contacts between Timor-Alor-Pantar and Austronesian speakers.

Kratochvíl, Frantisek (2014). Sawila. In Schapper, Antoinette (ed.), *The Papuan Languages of Timor, Alor and Pantar: Volume 1*, pp. 351–438. Berlin: Mouton De Gruyter.

List, Johann-Mattis (2012). SCA: Phonetic Alignment based on sound classes. *New Directions in Logic, Language and Computation*, pp. 32–51. Springer.

List, Johann-Mattis, Simon J. Greenhill, Tiago Tresoldi and Robert Forkel (2019). *LingPy. A Python Library for Quantitative Tasks in Historical Linguistics*. Version 2.6.5. Zenodo. http://doi.org/10.5281/zenodo.3482794.

List, Johann-Mattis and Robert Forkel (2021). Automated identification of borrowings in multilingual wordlists [version 2]. *Open Research Europe* 2021, 1:79.

Moro, Francesca R. (2018). The plural word 'hire' in Alorese: contact-induced change from neighboring Alor-Pantar languages. *Oceanic linguistics* 57(1): 177–198.

Moro, Francesca R. (2019). Loss of morphology in Alorese (Austronesian): Simplification in adult language contact. *Journal of Language Contact* 12(1): 378–403.

Moro, Francesca R. (2021). Multilingualism in eastern Indonesia: Linguistic evidence of a shift from symmetric to asymmetric multilingualism. *International Journal of Bilingualism* 25(4):1102–1119.

Moro, Francesca R. and Hanna Fricke (2020). Contact-induced change in Alorese give-constructions. *Oceanic Linguistics* 59(1/2): 116–147.

Muysken, Pieter. (2013). Language contact outcomes as the result of bilingual optimization strategies. *Bilingualism: Language and Cognition 16*(4): 709–730.

Rama, Taraka, Johann-Mattis List, Johannes Wahle and Gerhard Jäger (2018). Are Automatic Methods for Cognate Detection Good Enough for Phylogenetic Reconstruction in Historical Linguistics? *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2* (*Short Papers*), pp. 393–400. New Orleans: Association for Computational Linguistics.

Robinson, Laura C. 2015. The Alor-Pantar (Papuan) languages and Austronesian contact in east Nusantara. In Ross, Malcom and I Wayan Arka (eds.), *The Asia-Pacific Linguistics: Language Change in Austronesian Languages*, pp. 19–33. Canberra, Australia: Pacific Linguistics.

Ross, Malcom (2013). Diagnosing contact processes from their outcomes: The importance of life stages. *Journal of Language Contact* 6(1):5–47.

Rosvall, Martin, Daniel Axelsson, and Carl T. Bergstrom (2009). The map equation. *The European Physical Journal Special Topics* 178(1): 13–23.

Samely, Ursula (1991). Wordlist on Kedang-Leuwayang. In Kaiping, Gereon, Owen Edwards, and Marian Klamer (eds.), *LexiRumah 3.0.1*. Leiden: Leiden University Centre for Linguistics. Available at: https://lexirumah.model-ling.eu/lexirumah/.

Schapper, Antoinette and Marian Klamer (2017). Numeral system in the Alor-Pantar languages. In Klamer, Marian (ed.), *The Alor Pantar Languages: History and Typology*, pp. 277–329. Berlin: Language Science Press.

Schapper, Antoinette, Juliette Huber, and Aone van Engelenhoven (2017). The relatedness of Timor-Kisar and Alor-Pantar languages: A preliminary demonstration. In Klamer, Marian (ed.), *The Alor-Pantar languages: History and typology*, pp. 91–146. Berlin: Language Science Press.

Schapper, Antoinette and Juliette Huber (this volume). Entwined histories: the lexicons of Kawaimina and Maka languages.

Steinhauer, Hein (2014). Blagar. In Schapper, Antoinette (ed.), *The Papuan Languages of Alor and Pantar: Volume 1*, pp. 148–218. Berlin: Mouton De Gruyter.

Stokhof, Willem A. (1975). *Preliminary Notes on the Alor and Pantar Languages* (*East Indonesia*). Canberra: Pacific Linguistics.

Stokhof, Willem A. (1984). *Annotations to a text in the Abui language* (*Alor*). *Bijdragen tot de taal-, land-en volkenkunde* 140(1): 106–162.

Sulistyono, Yunus (2022). *A history of Alorese (Austronesian): Combining linguistic and oral history*. Doctoral dissertation, Leiden University.

Tadmor, Uri, Martin Haspelmath, and Bradley Taylor. (2012). Borrowability and the notion of basic vocabulary. In Wichmann, Søren and Anthony P. Grant (eds.), *Quantitative approaches to linguistic diversity*, pp. 35–55. John Benjamins Publishing Company.

Uffmann, Christian (2015). Loanword Adaptation. In Honeybone, Patrick and Joseph Salmons (eds.), *The Oxford Handbook of Historical Phonology*, pp. 645–667. Oxford University Press.

Wellfelt, Emilie (2016). *Historyscapes in Alor: Approaching Indigenous Histories in Eastern Indonesia*. Doctoral dissertation, Linnaeus University.

Winter-Froemel, Esme (2013). Formal variance and semantic changes in borrowing: Integrating semasiology and onomasiology. In Zenner, Eline and Gitte Kristiansen (eds.), *New Perspectives on Lexical Borrowing: Onomasiological, Methodological and Phraseological Innovations*, pp. 65–100. Berlin, Boston: De Gruyter Mouton.