



Universiteit
Leiden
The Netherlands

Traces of language contact: The Flores-Lembata languages in eastern Indonesia

Fricke, H.L.A.

Citation

Fricke, H. L. A. (2019, November 13). *Traces of language contact: The Flores-Lembata languages in eastern Indonesia*. LOT dissertation series. LOT, Amsterdam. Retrieved from <https://hdl.handle.net/1887/80399>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/80399>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/80399> holds various files of this Leiden University dissertation.

Author: Fricke, H.L.A.

Title: Traces of language contact: The Flores-Lembata languages in eastern Indonesia

Issue Date: 2019-11-13

Part I

Grammatical description

CHAPTER 2

Introduction to Part I

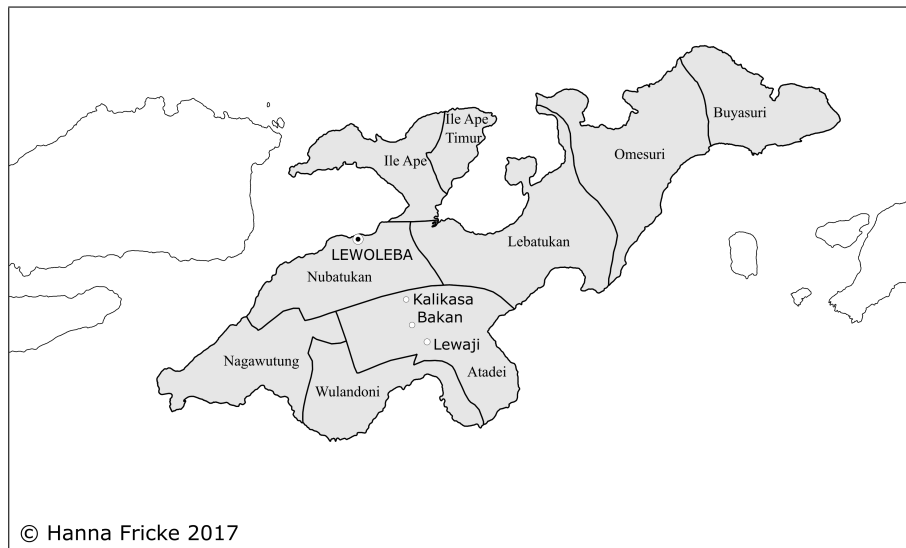
2.1 Location and speakers

Central Lembata is spoken by about 3,000 people in six villages in the central mountains on the island of Lembata: Kalikasa (Katakeja), Bakan (Ile Kerbau), Lewaji (Dori Pewut), Kolilerek (Tubuk Rajan), Waimuda (Nuba Boli) and Mudalerek (Nogo Doni).¹ All six Central Lembata villages are part of the Atadei district (Indonesian: *kecamatan*) in the Lembata regency (Indonesian: *kabupaten*).² The number of speakers is calculated according to recent census data (Badan Pusat Statistik Kabupaten Lembata 2018). The map in Figure 2.1 shows the three biggest villages of the Central Lembata speaking area. The map also contains district names and boundaries in the Lembata regency. Kalikasa is the administrative capital of the Atadei district.

¹ The names in brackets are the official village names only used in very formal or written contexts, while the other names are the original names that have been in use since long before these villages became officially administrative units of *desa* (Indonesian: ‘village’). The official *desa* names were created by village elders and are not used in daily conversation.

² Lembata has been a single regency since 1999. Previously, it was part of Flores Timur which is located towards the west of Lembata, encompassing the eastern tip of Flores and the islands Adonara and Solor.

Figure 2.1: Fieldwork locations on Lembata



In the Atadei district, infrastructure is still developing. During the time of my fieldwork trips, electricity was produced by generators but only for a few hours at night. Kalikasa had mobile coverage whereas the other two villages, Bakan and Lewaji, did not. The inhabitants of the Atadei district are mainly farmers. There are only a few places of employment like in schools or the medical center in Kalikasa. This is why many people leave the area to work in the regency's capital Lewoleba or even further away in Kalimantan, Java or Malaysia. Roads are available reaching all villages in the district with daily running trucks to transport passengers and goods. However the roads are generally in bad states which makes travelling slow and uncomfortable.

Kalikasa is officially named Katakeja 'I am the one in the middle' in the language of Central Lembata. This refers to its central position on the island of Lembata. The word Kalikasa denotes a cactus plant which is found in the area and which used to be planted around the village for protection. The official name of the village Bakan is Ile Kerbau which means 'buffalo hill'. This name was chosen because the hills around Bakan used to be full of buffaloes (L2:117). According to oral history, the word *bakan* has three meanings (i) 'plain', (ii) 'a kind of grass' and (iii) 'older siblings' (cf. L2, L3). The village Lewaji is officially named Dori Pewut meaning 'joining the bees'. This goes back to the origin story of the village Lewaji which tells about the Pewut

‘bee’ clan as the only native clan of the village that emerged out of the earth and other immigrant clans that fled from a tsunami that had destroyed their island Lapan Batan and joined the Pewut ‘bee’ clan (L2:36-48, §1.2.4.3). The name Lewaji is short for *lewu wajinu* ‘the village of the younger siblings’. This goes back to an event in oral history that caused the inhabitants of Lewaji to be seen as younger siblings, thus subordinate to those of Bakan, which they had to address from then on as ‘older sibling’ (L2:110). It becomes clear that the official names are the newer names of the villages because their meaning is transparent in the local language, whereas the meanings of the older names are already more obscure and only known by certain people, mainly elders. Sometimes the older names have several possible meanings, as in the case of Bakan.

Based on oral history (L3:398-408), the oldest village among the Central Lembata villages is Bakan and people have come from there and founded the other surrounding villages of Kolilerek, Kalikasa, Waimuda and Mudalerek. However, for Lewaji, it has been said that there were people already living there with which the Bakan people started to live together (L2:37-39).

The Atadei district reaches further southeast and includes about 16 villages where the neighbouring language Atadei Painara (also: Eastern Atadei, or South Lembata) is spoken.³ Based on information from a sociolinguistic questionnaire archived in the Central Lembata Corpus (Fricke 2019) and on information given in Keraf (1978a), the variety of Central Lembata is also spoken in a few more villages towards the west. Among those are Boto (Laba Limut) and Udek in the Nagawutung district and Lewuka (Belobao) in the Wulandoni district. Nowadays, for administrative and infrastructural reasons, there is not much contact between the Central Lembata people from Atadei and those from the other districts.

In previous publications, the Central Lembata area was considered to be part of the Lamaholot dialect chain (Arndt 1937; Keraf 1978a; Grangé 2015b), although very little information was available on the languages of that area. During my fieldwork, it became evident that the speakers of Central Lembata do not consider their language to be Lamaholot. Lamaholot for them is the language spoken in the northern part of Lembata and on the other is-

³ According to district records the Atadei Painara speaking villages are: Karangora (Ile Kimok), Waipei, Atalojo, Bauraja, Waiwejak, Lewokurang, Paulolo (Leba Ata), Lewokoba, Watuwawer (Ata Kore), Atawolo (Lusi Lame), Brenai, Lerek, Lamanuna, Duli, Tobilolong and Baulolong. Not all of these are administrative units of *desa*.

lands to the west, Adonara, Solor and Eastern Flores, thus the language that has been linguistically classified as Western Lamaholot (cf. §4.1). I chose the language name Central Lembata as a geographical label to cover the six Atadei villages described above. The speakers themselves do not have a specific name for their language variety. They would refer to it as *bahasa daerah* (Indonesian: ‘local language’), or combine the Indonesian term *bahasa* ‘language’ with a village name, such as *Bahasa Bakan* ‘Bakan language’.

2.2 Data collection and processing

The fieldwork to collect data for the grammar of Central Lembata in Chapter 3 was conducted in three villages: Kalikasa (Katakeja), Bakan (Ile Kerbau) and Lewaji (Dori Pewut) in the Central Lembata speaking area (cf. §2.1). The central part of Lembata was chosen as field site as it shows the highest linguistic diversity of the Lamaholot area, and the varieties spoken there had not been studied before (cf. §1.1 and 4.1.4). The data collection in the field was carried out for a total of about 9 months distributed over the years of 2015, 2016 and 2018. Out of these 9 months, two months were spent working with native speakers of Central Lembata in Yogyakarta, a student town on Java.

Recordings of Central Lembata speech were stored in ‘wav’ format for audio data and in ‘mp4’ format for video data. All audio recordings, in most cases with corresponding video, were transcribed using the linguistic annotation software ELAN produced by the Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands, accessible at <https://tla.mpi.nl/tools/tla-tools/elan> (Wittenburg et al. 2006).⁴ Transcription was done by me together with native speakers of Central Lembata. While transcribing, I made notes on word meanings. ELAN produced ‘eaf’ files with time-aligned annotations and these were exported as ‘flextext’ files and then imported into a FLEX project that contains a lexical database. FLEX is free-ware provided by SIL at <https://software.sil.org/fieldworks>. The program allows linguists to build a text corpus with glossings and translations that are connected to a lexical database that can be built and expanded while glossing. In the FLEX corpus, the texts were glossed and translated into Eng-

⁴ One recording of a conversation could not be transcribed due to bad quality (cf. C4).

lish and in most cases also Indonesian. This was also done with help of native speakers when needed. Finally, the glossed and translated transcriptions were re-imported into ELAN to link them to the recordings. These eaf files are archived together with their corresponding recordings in the Central Lembata Corpus (cf. §2.3). For some of the transcriptions, the exported glossings are still work in progress.

2.3 The Central Lembata Corpus

The data collected is archived at <https://hdl.handle.net/1839/ed1aa6f0-93da-4a44-8ddb-66221cbc01ce> in the Central Lembata Corpus at the MPI Language Archive in Nijmegen, the Netherlands. Table 2.1 gives an overview of the records in the corpus. It contains video and audio recordings of speakers native to Kalikasa, Bakan, Lewaji and surrounding villages. The time indications refer to the length of the whole recordings. The time of recorded Central Lembata free speech amounts to approximately 7,5 hours in form of narratives (~ 4,5 hours) and conversations (~ 3 hours). The total recordings are much longer because parts of the recordings are in Indonesian and there are also recordings that do not contain free speech but other contents, such as wordlists or elicited sentences.

The main elicitation tasks used were the Surrey Stimuli (Fedden et al. 2010) and the Event and Position List (Moro and Fricke submitted). Both elicitation lists are collections of short video clips and some pictures that the speaker had to describe using Central Lembata. The wordlist and the cultural questionnaire contributed to the LexiRumah database (Kaiping et al. 2019) and the CultureRumah database (Kaiping and Klamer to appear).

Apart from recordings, a lexical database of about 1,800 Central Lembata lexical entries is part of the corpus. A community dictionary based on this lexical database will be published with a local publisher in Indonesia.

Table 2.1: Overview of records in the Central Lembata Corpus

Genre	Recorded time	Speakers
Conversation	4 hours	11 females; 6 males 6 <30 years; 3 <60 years; 2 >60 years; 6 unknown
Narratives	5 hours	4 females; 10 males 2 <30 years; 8 <60 years; 4 >60 years
Surrey Stimuli	2 hours	3 females; 5 males 3 <30 years; 5 <60 years
Event and Position List	2,5 hours	2 females; 4 males 2 <30 years; 4 <60 years
Other elicitations	4,5 hours	1 female; 2 males 2 <30 years; 1 <60 years
Wordlist (600 items)	2 hours	1 male 1 <60 years
Sociolinguistic questionnaire	1,5 hours	2 males 2 <60 years
Cultural questionnaire	40 minutes	1 male 2 <60 years
Total	22 h 10 min	

2.4 Citation codes

Each recording session in my data collection has a citation code which is deduced from the original file name of the associated recording. A list of citation codes is found in the list of abbreviations on page xxii. These codes are used when citing examples from the corpus in this dissertation and elsewhere. An example of a citation would be L1:50. This means that the example sentence or phrase is found in the recording with the code L1 in line number 50. The line corresponds to the segment number in the corresponding ‘eaf’ file that can be opened with the program ELAN, as well as to the line in the flextext file readable by FLEX (cf. §2.2). In the FLEX corpus, as well as in the archive, the titles of the texts in the FLEX corpus all start with

the citation code. In the electronic version of this dissertation, the citation codes are linked to the Language Archive where the recordings are publically accessible (cf. §2.3).

A few examples in this dissertation are taken from the Hewa Collection collected by me in 2013 in the village Hewa, in the Wulanggitang district of the Flores Timur regency, Indonesia. This collection is archived at the MPI Language Archive in Nijmegen, the Netherlands (Fricke et al. 2019). The coding system for examples from this corpus follows the same strategy, with the difference that HC (=Hewa Collection) is the beginning of the code and that the line number does not correspond to an ELAN annotation but to a line in the corresponding Toolbox text file. An example is the code HC_SR:31 which refers to line 31 in the text SR (=Spatial relations) in the Hewa Collection.

2.5 Transcription and glossing conventions

All linguistic examples in the grammar of Central Lembata are either transcribed using the International Phonetic Alphabet (IPA) or a practical orthography with each character representing one phoneme. This orthography is laid out in §7.3. IPA transcriptions, in phonemic (/ /), are used in the section on the Central Lembata phonology, while the practical orthography, in italic font, is used in all other sections of the grammar.

For abbreviations in glosses, I follow the Leipzig glossing rules with my own additions whenever necessary (Max Planck Institute for Evolutionary Anthropology 2015). A list of all glosses and their meanings is given in the list of abbreviations on page xix.