Languages of Sumba: State of the field

Joseph LOVESTRAND

SOAS University of London

This paper reviews the state of documentation and description of the languages of Sumba, an island located in the province of Nusa Tenggara Timur (NTT) in Indonesia. Sumba is home to a population of over 800,000 people. The languages of Sumba can be divided into eight or more distinct languages. Language documentation and description in Sumba began in the 19th century with the arrival of the Dutch. The most well-described language of Sumba is Kambera, spoken by 150,000 or more people occupying the eastern half of the island (Klamer 1998a). There is much less documentation and grammatical description of the languages of West Sumba. with the exception of two recent MA theses. There are also many word lists, as well as orthographically transcribed ritual texts included in publications by anthropologists interested in this aspect of the culture. Video documentation of the languages of Sumba is almost completely lacking. Language vitality in Sumba remains relatively high, however, there are signs of a possible shift to Indonesian. This increases the urgency of the need for language documentation and description, as well as investment in programs that facilitate ongoing local language use.

1. Background

This paper provides a concise summary of the current state of published, academic knowledge of the languages of Sumba.[1] Figure 1 shows the location of the island of Sumba in the province of Nusa Tenggara Timur (NTT), Indonesia. Sumba is south of Flores, and accessed by flight via Bali (from the west) or Timor (from the east). Its area of about 11,000 km² (twice the size of Bali) is home to a relatively sparse population of over 800,000 people (BPS 2015).[2] Most of the islands of NTT are part of the active Sunda-Banda volcanic arc, but the island of Sumba is older, and its tectonic origins have been the subject of some debate (Rutherford, Burke & Lytwyn 2001:454).

The eastern and northern areas of Sumba have dry climates with hilly grassland and savannah landscapes. Because there is only a single rainy season lasting a few months, crops are mostly limited to cassava and corn (maize). In the wetter climates of the central and western parts of the island, farmers use buffalo to cultivate rice fields (Hoskins 1989a; Vel 1994:23–24; Twikromo 2008:42–50). Pigs are another common livestock. In modern economic terms, the average person living in Sumba is relatively poor, but this does not take into account regular exchange of material goods and services (Vel 1994). These reciprocal transactions are typically cashless, and represent a vibrant local economic system that is difficult to quantify.

Dutch traders (VOC, Dutch East India Company) made contact with the people of Sumba as early as 1750, but the Dutch were relatively uninvolved in Sumbanese society until the first Dutch missionary arrived in 1881 (Wellem 2004; Aritonang & Steenbrink 2008:319–320). The people of Sumba were resistant to the new religion and did not begin converting

[1] Thanks to the following researchers for their input on this article: Leif Asplund, Misraini Balle, Yustinus Ghanggo Ate, Janet Hoskins, Webb Keane, Marian Klamer, David Mitchell and Allahverdi Verdizade. Thanks also to two anonymous reviewers for their helpful comments. Naturally, any remaining errors or omissions are my own.

[2] In 2015, BPS (*Badan Pusat Statistik*) projected the 2020 population of Sumba would be 817,000 based on the 2010 census in which the population was 689,000.

LOVESTRAND, Joseph. 2021. 'Languages of Sumba: State of the field'. NUSA 70: 39-60. Permanent URL: http://hdl.handle.net/10108/100089. doi: https://doi.org/10.15026/100089

in large numbers until the beginning of the 20th century, coinciding with an increasing presence of formal education (run by missionaries) and a national political structure (Vel 2008:35–41). A few decades ago, the majority of the people of Sumba still identified the traditional religion of the island, known as Marapu, as their predominant religious affiliation.[3] Today, around 90 percent of the people of Sumba identify as Christian (Protestant or Catholic), although that statistic may not give a realistic sense of the continued importance of Marapu in Sumba.[4]

Figure 1. Map of Sumba in Indonesia[5]

The Marapu religion and its associated ritual and social practices have been the topic of many anthropological studies. Marapu practitioners are well-known for their elaborate funerals centered around megalithic (large stone) gravesites (e.g., Hoskins 1986; Adams 2007) and animal sacrifice (e.g., Hoskins 1993; Keane 2018). They also practice a number of seasonal rituals (e.g., Hoskins 1994), including the annual Pasola festival in which a ceremonial (yet dangerous) spear-throwing battle is played out on horseback (Adams 2007:121–125). The people of Sumba are also world-renown for their traditional *ikat*—hand-crafted fabrics woven and dyed in intricate patterns (e.g., Adams 1969; 1980; Hoskins 1989b; Geirnaert-Martin 1992; Forshee 2001) and their traditional architecture (e.g., Keane 1995a). Sumbanese traditional religious and sociopolitical events are infused with ritual speech. Ritual speech is a poetic form that differs from everyday language

[3] Thirty-five years ago, Hoskins (1985:148) noted that 80% of Kodi people "remain faithful to the worship of their ancestors (called *Marapu*)". Keane (1995b:291) cites a 1986 census publication stating that 63.6% of the people of Sumba were not associated with any of the five official religions of Indonesia.

[4] Data on religious affiliation are currently available on individual webpages for each of the four *kabupaten* of Sumba (sumbatengahkab.bps.go.id, sumbabaratkab.bps.go.id, sumbatimurkab.bps.go.id and sumbabaratdayakab.bps.go.id [accessed 15 July 2020]). The data are presumably based on the 2010 census, and projected to various different years. About 42% of the population are estimated to identify as Protestant and 48% as Catholic. Only two *kabupaten* report people affiliating with an "other" category, which in the case of Sumba Timur is explicitly identified as Marapu. The "other" category make up about 5% of the population of the island, and those who identify as Muslim make up another 5%.

[5] © Gunkarta / CC-BY-SA-3.0 Wikimedia Commons, labels and circle marking Sumba added.

primarily in its extensive use of parallelism and conventionalized metaphors (e.g., Fox 1988). Ritual speech has been of particular interest to anthropologists. Anthropological publications include orthographic transcriptions of ritual or poetic speech in most languages of Sumba (Section 3).

The grammatical structures of the languages of Sumba have not been as well-described as their poetic style and its social functions. The remainder of this paper reviews what literature is available to date on the grammatical description and documentation of the languages of Sumba. As an overview of the linguistic diversity of the island, Section 2 reviews various perspectives on how many languages there are in Sumba, or how the various linguistic communities of the island can most sensibly be organized into categories of languages and dialects. Section 3 reviews the state of grammatical description and multipurpose documentation of the languages of Sumba. Based on the limited studies available, Section 4 outlines a few grammatical features of the languages of Sumba highlighting potential topics for future cross-linguistic research. Section 5 is a discussion of language vitality and language endangerment in Sumba. Section 6 is a brief summary and a consideration of possible next steps for linguistic research in Sumba.

2. Language and dialect boundaries

The languages of the Sumba group can be traced to migrants who arrived on the island around 3500 years ago (Lansing et al. 2007).[6] The Austronesian migrants were not the first inhabitants of the island, but as they mixed with the local population, their languages dominated and wiped out nearly all traces of whatever languages were previously spoken in Sumba. The modern-day languages of Sumba, together with the Hawu language of the neighboring Savu islands, form a sub-group in the family of Malayo-Polynesian (Austronesian) languages (Blust 2008; Klamer 2009).[7]

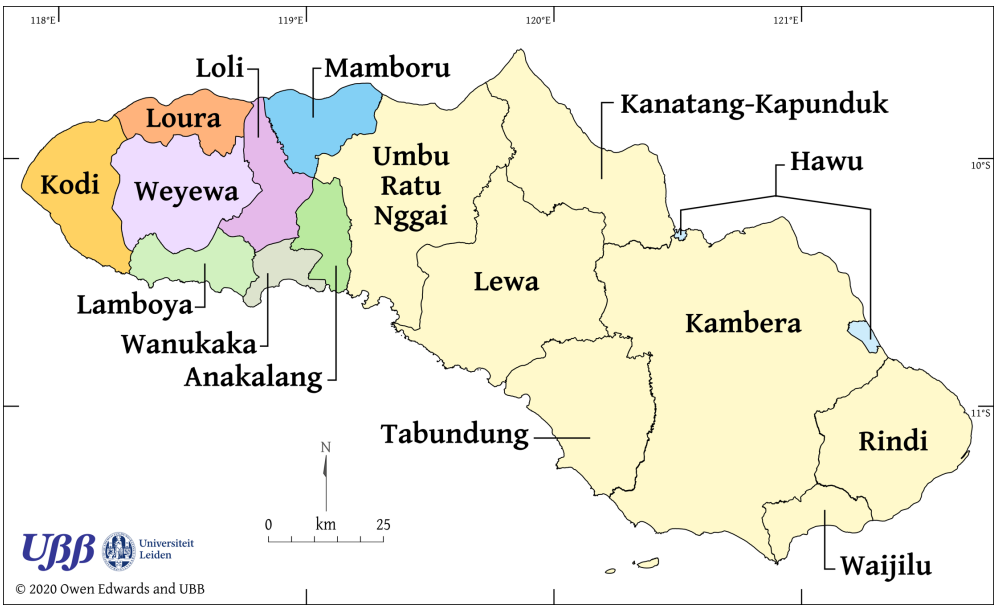

Figure 2. UBB Sumba language map[8]

[6] See also Lansing's (2008) reply to a critique by Donohue (2008).

[7] For higher levels of subgrouping, see Gasser (2014) and Fricke (2020).

[8] © 2020 Owen Edwards and UBB, used by permission.

This paper is limited to discussion of the Sumba language group, excluding the communities of Hawu speakers who have lived in Sumba for more than a century, as well as various other migrants such as the Makassarese traders who came via Flores (Hoskins 2002:799–800). In addition to these local languages, the national language (*bahasa Indonesia*) is widely spoken across the island, and there are some efforts to increase the use of English as a language of tourism.[9]

There are minimally seven speech varieties (or lects) that every source includes in their list of languages of Sumba. These seven are the only ones included as languages in the Glottolog (Hammarström et al. 2020). Figure 3 shows these seven languages written in bold in a single row. Below this row are further subdivisions which some sources treat as distinct languages as well.[10] One area of difference is in regards to the Weyewa cluster. While Glottolog considers this cluster to form a single language, the Ethnologue recognizes Loura as a language distinct from Weyewa, bringing the total number of languages of Sumba to eight (Eberhard, Simons & Fennig 2020).[11] The UBB map (Figure 2) breaks out another lect, Loli, recognizing it as a separate language as well, resulting in a total of nine distinct languages in Sumba (Edwards & UBB 2018).[12]

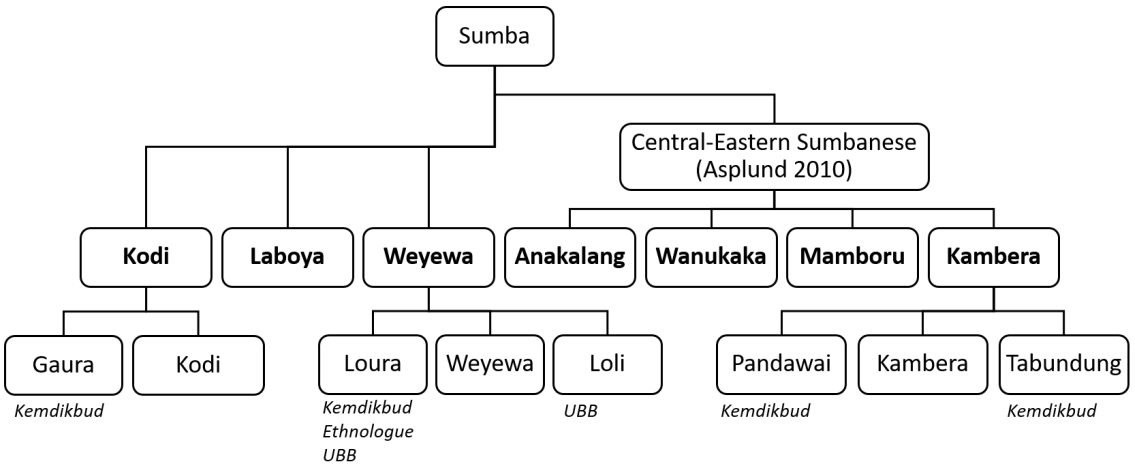

Figure 3. Four different classifications of the languages of Sumba

The official government language map from the Indonesian Ministry of Education and Culture (*Kementerian Pendidikan dan Kebudayaan*), also known as Kemdikbud, further expands the number of languages of Sumba by splitting Kambera in East Sumba into three languages (Kemdikbud 2019). The name Kambera was originally the name of just one ethnolinguistic community in East Sumba, the one studied extensively by Klamer (Section 3.1), but the name also groups together all of the lects (excluding Hawu) in East Sumba (Sumba Timur) and some in Central Sumba (Sumba Tengah) as a single

[9] For example, The Sumba Foundation recruits volunteers to teach English to children in West Sumba (https://sumbafoundation.org/education/english-classes/ [accessed July 5, 2020]).

[10] Note that Figure 3 does not include all lects of Sumba, but only those that are labeled as a language in the publications under discussion.

[11] The Glottolog cites van der Velden (1900) as a source for considering Loura a dialect of Weyewa. Note that Loura is often called "Laura" but the latter is considered offensive.

[12] Although Loli may be closely related to Weyewa linguistically, there has historically been conflict between the groups, including violent confrontation as recently as 1998 (Vel 2001; Kuipers 2009; Gunawan Mitchell 2019).

language.[13] The UBB language map (Figure 2) divides up the linguistic communities of East Sumba into seven dialects.[14] This appears to be based on a map provided by Onvlee (1984) who further divides the Kambera area in Figure 2 to include two smaller groups, "Massoe Karera" and "Melolo". Onvlee's (1984) dictionary entries give cognates (or semantic equivalents) for Lewa and Mangili (Rindi), as well as for other languages, but no criteria for determining whether these lects should be considered part of the Kambera language or not. Asplund (2010) treats them as one language based on a lexical similarity of over 80%, but Gasser (2014) places Kambera and Lewa at a greater distance from each other than the distance between Anakalang and Wanukaka. Further documentation of variation and of the linguistic identities held by various groups of East Sumba is needed to clarify whether these lects should be subsumed under one language name or treated as multiple languages, as done by Kemdikbud.

The Kemdikbud map also identifies Gaura as a language of West Sumba (*kecamatan Laboya Barat*), bringing their total number of languages of Sumba to eleven.[15] The Ethnologue and the Glottolog treat "Nggaura" as a dialect of Kodi. Lexically, Gaura appears to be relatively close to Kodi (Asplund 2010; Gasser 2014). However, in personal communication with speakers of Kodi, Gaura and Laboya, the unanimous view is that Gaura is a distinct language, and not a dialect of Kodi or Laboya (also confirmed by Hoskins, personal communication). Therefore, a reasonable working assumption is that Gaura is a distinct language of West Sumba.

Another group that may have been overlooked in previous linguistic work is Mbukambero (Needham 1980:25–26). Mbukambero is considered (at least by some) to be a dialect of Loura/Weyewa.[16] Asplund (2010) also considers Ponduk and Baliledu as lects that might be best classified as languages, closely related to Weyewa and Anakalang (see also Gasser 2014). Given what data is available, it seems reasonable to say that Sumba has at least eight languages (including Gaura), but that there are potentially several more communities with an exclusive linguistic identity that would be best described as a language rather than a dialect. Any real insight into the questions of interintelligibility and linguistic identity among various groups in Sumba will require more research.

3. Documentation and description

This section reviews published works on the languages of Sumba from the perspective of documentary and descriptive linguistics (Himmelmann 1998). Descriptive linguistics is concerned with the analysis of the phonological, morphological, syntactic and discourse patterns found in languages, as well as their interplay and their intersection with other

---

[13] The autonym for the predominant language of East Sumba is *Humba* (reflecting a phonological shift from /s/ to /h/), and the language has also been called Sumbanese. However, since the name Sumba has been adopted as the name of the island, speakers of other languages also view themselves as Sumbanese. To avoid confusion, Klamer (1998a:3) and others use the dialect name, Kambera, to refer to the language as a whole. Alternatively, the language has also been called East Sumbanese (although there is no "West Sumbanese" counterpart).

[14] The Kambera dialects or ethnolinguistic communities, sometimes called "domains", as listed on the UBB map in Figure 2 are: Umbu Ratu Nggai, Lewa, Kanatang-Kapunduk, Tabundung, Kambera, Waijilu and Rindi. The name *Hawu* on the map refers to the language of the nearby Savu islands whose speakers have lived in Sumba for over a century (see the second paragraph of Section 2).

[15] Note that Kemdikbud recognizes Loura, but not Loli (Figure 3).

[16] The Glottolog cites Wielenga (1917) as a source for classifying Mbukambero as a dialect of Loura.

areas of research. Documentary linguistics is concerned creating "a lasting, multipurpose record of a language" (Himmelmann 2006:1). These aims have not been a primary focus of previous studies of the languages of Sumba. This should not be taken as a criticism of past work, which has been valuable from other perspectives, such as anthropological linguistics and historical linguistics.

Published studies of Sumba began in the 19th century with the arrival of the Dutch. Goh (1991) is an extensive bibliography that includes these early writings, but most of them have little or nothing to say about the languages of Sumba with the exception of a few word lists. A focus on linguistic description and documentation in Sumba begins in the 20th century with documentation of Kambera in East Sumba, the language that remains the most well-described to date (Section 3.1). Early linguists also documented the vocabulary of some languages of West Sumba, but the linguistic diversity of the western part of the island remains severely underdocumented (Section 3.2).

3.1. East Sumba (Kambera)

Klamer (1998a:3–4) reviews the early literature on Kambera. The earliest works are short lists of words (e.g., Heijmering 1846; Roos 1872; Vermast 1895). Wielenga (1909) published the first grammatical description of the language. The descriptive work was expanded by Onvlee who published a grammar with some transcribed texts (Onvlee 1925) and a dictionary (Onvlee 1984), as well as some pedagogical works and a New Testament translation. Onvlee's collaborator, Kapita, published a Kambera grammar in Indonesian (Kapita 1983) and a Kambera-Indonesian dictionary (Kapita 1982), along with other texts. The description of the grammar of Kambera was continued by Klamer (1998a; as well as 1994; 1996; 1997; 1998b; 1999; 2000a; 2000b; 2001; 2002a; 2002b; 2004; 2005) in what remains the most complete description of any language of Sumba.[17]

The earliest documentation of Sumbanese texts did not include any interlinearized annotation, and so "these texts can only be analyzed with the help of an informant or by someone who has studied the language first" (Klamer 1998a:4). The same issue holds of transcriptions of ritual speech in Kambera and other languages (e.g., Forth 1988). Exceptions to this generalization are the six Kambera texts appended to Klamer (1998a:355–394) and the ritual texts in Asplund (2017). Both sources provide morpheme-by-morpheme interlinearized glossing.

By current standards, the documentation of Kambera is relatively meager in its size, and it does not include any video documentation—"the ideal recording device" (Himmelmann 1998:168). Thirty audio recordings along with additional transcriptions and field notes from Klamer's research in Sumba are archived by The Language Archive at the Max Planck Institute for Psycholinguistics in Nijmegen.[18] The documentation of Kambera could also be expanded in its dialectical coverage. The majority of the description and documentation available focuses on the eponymous Kambera dialect.

---

[17] A Kambera word list of over 600 items compiled from published sources (Onvlee 1984; Klamer 1994) is available online from LexiRumah (Kaiping, Edwards & Klamer 2019): https://lexirumah.model-ling.eu/languages/kamb1299 [accessed 8 Dec 2020].

[18] Klamer's collection can be freely accessed online via The Language Archive (Max Planck Institute): https://archive.mpi.nl/tla/islandora/object/tla%3A1839_00_0000_0000_001E_A655_8 [accessed 8 Dec 2020].

3.2. West Sumba

Compared to Kambera, there is much less material available for each of the languages of West Sumba. The available documentation of these languages has predominately been of two kinds. First, there is a significant number of word lists that have been used in lexical comparisons (Djawa 2000; Budasi 2007; Lansing et al. 2007; Putra 2007; Blust 2008; Greenhill, Blust & Gray 2008; Asplund 2010). Along the same lines, the Ministry of Education and Culture has published several extended word lists (*kamus*) for languages of West Sumba (Anakalang: Kemdikbud 2003a; Kodi: Kemdikbud 2003b; Weyewa: Kemdikbud 2003c; Loli: Ori 2010; Laboya: Rina & Kabba 2011),[19] and Mitchell (2008) produced a dictionary for Wanukaka. Second, there are orthographically-transcribed ritual texts from at least five West Sumba languages included in publications by anthropologists interested in this aspect of the culture (Kodi: Hoskins 1985; 1986; 1987a; 1987b; 1988a; 1988b; 1989a; 1989b; 1989c; 1994; 1996; 1998; 2004; 2005; Wanukaka: Mitchell 1988; Weyewa: Kuipers 1986; 1988; 1990; 1998; Renard-Clamagirand 1988; Laboya: Keller 1989; Anakalang: Keane 1991; 1995a; 1997; Loli: Rothe 2004). Some anthropological studies also reference other linguistic topics, such as kinship and other vocabulary (e.g., Needham 1980; Forth 1982; 1983). These lists of words and transcribed texts generally lack phonetic detail, and original recordings are not available. In most cases, only excerpts of texts are published. Such orthographic transcriptions are an important aspect of the documentation of the languages of Sumba, but they do not provide enough detail for phonological and grammatical description.

In the last few years there have been three different efforts to document languages of West Sumba, Kodi and Weyewa in particular, but also Wanukaka and Laboya. In the process of analyzing the phonology of Kodi, a 790-item list (although many items are repeated) was recorded and archived at PARADISEC (Balle & Lovestrand 2019a). Shorter word lists of around 300 items in Wanukaka and Laboya were also recorded and archived.[20] Weyewa audio recordings are included in a corpus that was created as part of a collaborative program pairing graduate students from the USA with local linguists (Yanti & Shiohara 2018). The Weyewa corpus contains 49 recordings totaling 8 hours and 15 minutes of audio and one 3-minute video recording (Bibbs et al. 2018).[21] More video documentation can be found in the collection of Sou (2019).[22] The Sou collection includes five Weyewa (Wee Lurri) videos totaling 1 hour and 45 minutes, of which 35 minutes includes time-aligned transcription and translation. The collection also has a 20-minute Kodi video with transcription and translation. More Kodi video documentation can be expected in the near future, since a Kodi speaker and linguist, Ghanggo Ate, received a grant from The Endangered Language Fund in 2020 to purchase equipment for

[19] See also the Loli word list appended to Rothe (2004) and the Kodi word list appended to Twikromo (2008).

[20] The Kodi, Wanukaka and Laboya word lists and audio recordings are accessible online at: https://catalog.paradisec.org.au/collections/JL4 [accessed 7 July 2020]. Registration is required for access. A partial version of the Kodi word list is also available online via LexiRumah (Kaiping, Edwards & Klamer 2019) (https://lexirumah.model-ling.eu/sources/lovestrandballe18 [accessed 8 Dec 2020]), as is an extended list of over 600 items in Laboya based on the field notes of Allahverdi Verdizade (https://lexirumah.model-ling.eu/sources/verdizade19 [accessed 8 Dec 2020]).

[21] The Weyewa corpus is accessible online at: https://catalog.paradisec.org.au/collections/WEW2018/ [accessed 8 July 2020]. Registration is required for access.

[22] Sou's videos can be accessed at: https://scholarspace.manoa.hawaii.edu/handle/10125/64694 [accessed 8 July 2020].

video documentation of Kodi.[23] There are other videos containing West Sumba languages, but these are edited for a commercial audience and not amenable to linguistic analysis (Hoskins & Scheerer Whitney 1988; 1991; Larsson 2013).

Given that language documentation has only recently begun for the languages of West Sumba, it is not a surprise that the description of any of the languages is also fairly limited. A short grammar sketch of Loura was published by van der Velden (1900). Onvlee published very little about the grammar of Weyewa, considering the fact that he worked on a New Testament in the language (Kuipers 1990: xxii, 56).[24] In the last two decades, several Indonesian students have shown an interest in describing particular aspects of languages of West Sumba (e.g., Limy 2003; Walangara 2003; Kasni 2012; Ngongo 2013; 2015; Sukerti 2013; Sukerti & Ghanggo Ate 2016; Bulu 2018; Bili 2020). These studies offer valuable insights, but are generally narrow in scope and limited in data. More systematic and holistic analyses of the phonology, morphology and syntax of each language are still needed.

Recently, two languages, Kodi and Laboya, have been the focus of more comprehensive and detailed description. The MA thesis of Ghanggo Ate (2018) focuses on reduplication in Kodi, but also outlines some basic phonological and morphosyntactic patterns. More phonological research on Kodi is in progress (Balle & Lovestrand 2019b; Lovestrand & Balle 2019). Another recent MA thesis, by Verdizade (2019), provides a sketch of aspects of Laboya grammar. Both Ghanggo Ate and Verdizade intend to continue to pursue their studies in these areas, so there is a possibility that more descriptions of Kodi and Laboya grammar are forthcoming.

4. Linguistic features

With only limited documentation and description, a general typology of the languages of Sumba is not possible, but from the data available, a few patterns emerge that are worth noting. These patterns are primarily related to phonology and morphology. It should be emphasized that the findings are preliminary and subject to further investigation.

4.1. Phonological inventories

Like most Austronesian languages, the languages of Sumba minimally have a set of five basic phonemic vowel segments: /a, e, i, o, u/. Diphthongs are common, but do not necessarily have the same status in every language. Kambera has the diphthongs /ai/ and /au/ which Klamer (1998a:13–14) analyses as the long forms of /e/ and /o/ respectively. Verdizade (2019:18–19) identifies the same diphthongs in Laboya. Kodi does not treat diphthongs as segments, but does form diphthongs through metathesis (Section 4.3).

It appears that all the lects of Sumba have the same eight sonorant consonants (although the phonemic status of the palatal nasal is not always clear):

- two glides: /j/ and /w/
- two liquids: /l/ and /r/
- four nasals: /m/, /n/, /ɲ/ and /ŋ/

[23] For more information on The Endangered Language Fund see: http://www.endangeredlanguagefund.org/ [accessed 6 July 2020].

[24] Leif Asplund (personal communication) points out that there is a brief description of Weyewa grammar in Onvlee (1973:165–196).

Most lects use the glottal fricative /h/, which is a historical reflex of an alveolar fricative /s/ (Klamer 1998a:12), but some lects appear to have retained /s/, including Loli (Ori 2010), Weekombaka (Weyewa) (Bulu 2018), Mamboru, Baliledu and Ponduk (Asplund 2010). A few lects have an additional voiced fricative. At least one Weyewa lect has a voiced alveolar fricative /z/ (Kuipers 1998:xvii). Kodi and Tana Righu (Weyewa) have a voiced velar fricative /ɣ/.[25] Loura has both a voiced velar fricative and a voiced interdental fricative /ð/ (Ghanggo Ate, personal communication).

The languages of Sumba have three series of stops by manner of articulation, as shown in Table 1. They all have voiceless stops and implosives stops, but they vary in whether the third series of stops is prenasalized (voiced) stops or (non-nasalized) voiced stops. In Kodi, Weyewa and Kambera the third series is prenasalized (voiced) stops. In Anakalang, Laboya and Wanukaka, the third series is (non-nasalized) voiced stops (Keane 1997:xxviii; Verdizade 2019:21; Mitchell n.d.).[26] Verdizade (2019:21) points to evidence that the denasalization was a more recent innovation that several neighboring lects underwent, distinguishing themselves phonologically from other languages in Sumba. Other lects that lack prenasalized stops are Loli (Ori 2010), Baliledu and Ponduk (Asplund 2010).

Table 1. Stops and fricatives inventories of four languages of Sumba[27]

Kodi	p	t	(c)	k	mb	nd	(ɲʄ)	ŋg	ɓ	ɗ				ɣ	h
Weyewa	p	t	c	k	mb	nd		ŋg	ɓ	ɗ		ʔ	z		(h)
Laboya	p	t		k	**b**	**d**	**ɟ**	**g**	ɓ	ɗ	ʃ	ʔ			h
Kambera	p	t		k	mb	nd	ɲʄ	ŋg	ɓ	ɗ	ʃ				h

There are other minor differences in the inventories of stops, particularly in regards to palatals. Kodi and Weyewa do not have a palatal implosive that other languages of Sumba have (although it is not a common sound in any language). Weyewa has no voiced palatal stop, but does have a voiceless palatal stop which seems to be absent from other languages (although the sound exists in Kodi as a palatalized form of /t/). Kodi and Kambera do not have a contrastive glottal stop phoneme that is found in other languages of Sumba. In Laboya, the distribution of the glottal stop appears to be limited, and its phonemic role in the language remains uncertain (Verdizade 2019:23).

---

[25] For Tana Righu, see Asplund's word list on the Austronesian Basic Vocabulary Database where <gh> represents a voiced velar fricative (https://abvd.shh.mpg.de/austronesian/language.php?id=326 [accessed 16 July 2020]).

[26] Note that Laboya is also referred to as Lamboya, most likely reflecting a version of the name that was recorded before the loss of prenasalized stops, or a misinterpretation of the sound by speakers of languages with prenasalized stops.

[27] Sources are as follows: Kodi (Lovestrand & Balle 2019), Weyewa (Kuipers 1998:xvii), Laboya (Verdizade 2019:20; author's own fieldwork), Kambera (Klamer 1998a:10). Klamer actually describes Kambera has having alveolar affricates, not palatal stops. These are presented as palatal stops here assuming that whatever the precise phonetic articulation, these sounds are phonemically comparable to palatal stops. Wanukaka (Mitchell n.d.; author's own fieldwork) and Anakalang (Keane 1997:xxviii) appear to have the same consonant inventory as Laboya.

4.2. Contrastive length

There are at least two types of languages of Sumba in regards to the function of vowel and consonant length. Weyewa and Kodi have contrastive long consonants, and non-contrastive lengthening of stressed vowels—typically the penultimate syllable (Kuipers 1998:xvii; Lovestrand & Balle 2019). In Kodi (and presumably in Weyewa also), the lengthening of a stressed vowel is blocked when it is followed by a long consonant, as shown in the minimal pairs in examples (1) and (2). This can be analysed in a straightforward way as a syllable weight constraint, assuming that the long consonant occupies the coda of the preceding syllable adding a mora to its syllabic weight.

(1)	a.	/poto/	[ˈpoːto]	‘four’	**Kodi**
	b.	/potto/	[ˈpotːo]	‘type of bamboo’	
(2)	a.	/lod̪o/	[ˈloːd̪o]	‘sing’	
	b.	/lod̪d̪o/	[ˈlod̪ːo]	‘day’	

The other type of language is represented by Kambera and Laboya. These languages do not have contrastive consonant length, but they have a corresponding phenomenon of contrastive vowel length, which is limited to stressed syllables (Klamer 1998a:13–14; Verdizade 2019:18; author’s field research).[28] This pattern can be seen in the Laboya minimal pair in example (3).

(3)	a.	/mŏro/	[ˈmoɾo]	‘medicine’	**Laboya**
	b.	/moro/	[ˈmoːɾo]	‘green’	

The restriction on the distribution of vowel length contrast can be analyzed as an innovation based on the pattern seen in Weyewa and Kodi. In Weyewa and Kodi, there are two possible phonetic cues for the presence of a long consonant: the consonant length itself, and the lack of lengthening in a preceding stressed vowel (left side of Figure 4). It appears that Laboya and Kambera have discarded the use of consonant length as a contrastive feature, and rely solely on the variation in vowel length (right side of Figure 4), resulting in an asymmetric distribution of vowel length as a contrastive feature.

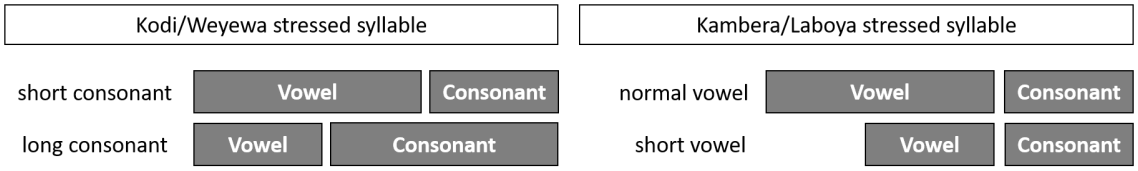

Figure 4. Schematic representation of length contrasts in stressed syllables

4.3. Morphology

The languages of Sumba have varying types of morphophonological alternations. Kambera uses word-final epenthesis (“paragogic” vowel) to preserve canonical phonotactic structures (Klamer 1998a:18–20). In contrast, Laboya uses apocope (final-vowel deletion) which results in consonant-final syllables (Verdizade 2019:26–27). Kodi uses both apocope and metathesis (example 4) to shorten word forms, in addition to a palatalization process that changes the pronunciation of labial and alveolar stops (Lovestrand & Balle 2019). These alternations are known to be context-sensitive, but the

[28] A short vowel in Kambera can also be phonetically realized as a lax vowel (Klamer 1998a:13–14).

contexts they occur in cannot be adequately described without further documentation and analysis of natural speech.

(4) **Metathesis in Kodi**

a.	/watu/	[ˈwaːtu]	~	[waut]	‘stone’
b.	/api/	[ˈaːpi]	~	[aip]	‘fire’
c.	/marapu/	[maˈɾaːpu]	~	[maˈɾaup]	‘spirits’

Based on the description of Kambera, and the more limited descriptions of other languages, it can be assumed that the languages of Sumba typically have a large number of pronominal and tense-aspect clitics that phonologically attach to the verb root. Another morphological feature, reduplication, has been described for Kambera and Kodi. Reduplication has various functions including intensification and pluractionality, as in example (5). Other than these basic observations, there are not yet enough descriptive studies available for comparative studies of the morphology and syntax of the languages of Sumba.

(5) **Reduplication in Kodi**

A	*loka=nggu*	*na=**inu~inu***	*kopi.*
DEF	uncle=1SG.POSS	3SG.SBJ=RED~drink	coffee

‘My uncle drank coffee abundantly (more and more).’ (Ghanggo Ate 2018:45)

5. Language vitality and maintenance

Around the world, a language goes out of use about once every three months (Campbell et al. 2013; Simons & Lewis 2013). In Indonesia, about half of the approximately 700 languages are declining in language vitality, and a quarter are likely to no longer be spoken within two or three generations (Anderbeck 2015). It is unclear how much these trends have impacted Sumba. Impressionistically, it is relatively common for children who grow up in one of the major towns (Waingapu, Waikabubak, Waitabula) to only learn to speak the national language, and not a local language. According to the 2010 census, 104,000 people in Sumba primarily spoke Indonesian in the home out of a total of 685,000 respondents (~15%), and that number has likely increased significantly.

Intergenerational transmission is higher in villages, but there are some signs that even in the villages there is a possibility of a future shift to Indonesian (Simanjuntak 2018). The influence of Indonesian is growing in the villages through education and media via television and smart phones (Twikromo 2008:82–23; Vel 2008:51). Metaphorical ritual speech is falling out of use faster than other domains. Kuipers (1998:2) describes a multiple choice test that was administered to Weyewa 12-year-olds in which the task is to select the correct interpretation (in Indonesian) of a traditional Weyewa couplet, such as that in example (6).[29] In 1979, 60% of students understood the meaning of these couplets. By 1994, that number for the same test in the same school had dropped to just 25%.

[29] The correct answer for example (6) is c.

(6) ***a ndita we'e*** 'who fills the water cisterns' **Weyewa**
a powi api 'who keeps the fires burning'

a) *orang yang menyanyi* 'one who cries'
b) *orang yang rasa dingin* 'one who feels cold'
c) *orang yang mendiami rumah adat* 'one who cares for the traditional home'

Another factor to consider is the overall number of speakers of each language, as shown in Table 2. The Ethnologue reports numbers of speakers, but does not offer sources, and in the three cases where a date is given it shows that the data is more than twenty years old (Eberhard, Simons & Fennig 2020). More accurate and up-to-date information was captured by the 2010 national census.[30] This data shows that there are three languages that have around 100,000 speakers or more: Kambera, Kodi and Weyewa. The other languages are relatively small, with 20,000 speakers or less.

Table 2. Estimated numbers of speakers for the languages of Sumba

Name (as listed in census)	2010 Census	Ethnologue
Anakalang	19,100	16,000
Kambera / Sumbanese	151,000	240,000
Kodi	93,000	20,000
Lamboya	17,400	25,000 (1997)
Loura	15,500	10,000 (1997)
Mamboru	8,400	10,000
Weyewa / Sumba Barat	130,000	55,000
Wanukaka	1,200	10,000 (1981)
Bahasa Indonesia	103,000	
Unlisted/No response	104,000	

There is some uncertainty about the accuracy of the 2010 census data. 104,000 (~15%) of the respondents did not give an answer to the language question. In some cases, this could be due to them using a different name to refer to their language than the one listed by the census taker. For example, perhaps Wanukaka speakers from Rua reported Rua as their language name, resulting in an underestimated total number of Wanukaka speakers. More uncertainty is introduced by the use of the name Sumba Barat as an equivalent of Weyewa. Perhaps speakers who did not recognize any of the language names as their own chose this option, thus inflating the total number of Weyewa speakers. Note that for

[30] Both sources intend to identify a single language as a primary language, but in a multilingual context like Sumba it is not always clear how frequently or in how many domains a person must use a language for it to be considered their primary language. In the 2010 census, respondents were specifically asked which language they use daily in their home (Ananta et al. 2015:275).

unrecognized linguistic groups like Gaura and Loli there is no data on their number of speakers. Nonetheless, this data represents the best information available on the number of speakers of the languages of Sumba.

There have been several efforts over the years to develop materials to encourage expanding the use of local languages into the domain of literacy, potentially strengthening their vitality. For Kambera and Weyewa, this tradition goes back to Onvlee's work (Klamer 2009:259), but there is no evidence that literacy in these languages has become common practice. Pedagogical works have been published for other languages as well (e.g., Kodi: Labu Djuli, Ratu Koreh & Tanda Kawi 2005; Laboya: Djawa 2013), but these works are not known to be in use. More recently, the Indonesian foundation Suluh Insan Lestari received a government grant (INOVASI) to partner with Kodi speakers in developing a school-based mother-tongue literacy program. The program ran for one year, with some moderate signs of success, but local school administrators declined to continue the funding beyond the one-year pilot project (Balle 2020). Suluh continues to work with local Protestant churches (Gereja Kristen Sumba) on informal literacy programs with the goal of producing translation of Christian scriptures in local languages. Along similar lines, Lembaga Alkitab Indonesia recently sponsored the production of a book of Bible stories in Kodi (LAI 2018).

Part of the success of these literacy efforts will depend on issues of orthography design which requires accurate phonological and grammatical descriptions, as well as local organizational structures for building consensus and promoting broadly comprehensible approaches to writing local languages (Balle & Lovestrand 2019b; Ghanggo Ate 2019; Lovestrand 2020). In this sense, there is a connection between documentation and description of languages and their ongoing use. There is also a need for a higher degree of local investment in all aspects of language research and literacy programs in Sumba.

6. Summary and prospectus

As is the case in many places in Indonesia, Sumba is an island rich in linguistic diversity and cultural heritage. Great interest has been shown in understanding the more traditional aspects of Sumbanese culture, including the use of ritual language in religious and social events. However, from the perspective of documentary and descriptive linguistics, these studies have only treated the languages superficially. The first step to more grammatical description of the languages of Sumba is more phonological analysis, along the lines of Klamer's (1998a:9–44) description of Kambera, and ongoing work on Laboya (Verdizade 2019) and Kodi (Lovestrand & Balle 2019). Since it is known that some morphophonological alternations are context sensitive (Section 4.3), phonological description requires the recording, transcription and translation of natural speech. The documentation, in turn, requires a phonological analysis to guide the transcription process, so the description and documentation are mutually dependent.

Primary linguistic data should be documented (annotated and archived) in an accessible manner, and should be seen as an essential method for extending the impact of a research project. Given the advances in technology in the last decade, it is now possible for virtually any researcher to include video documentation in their data collection, and this should be included as a goal in every linguistic project. Documentation not only grounds descriptive analyses and provides data for further analysis, it also has the function of preserving aspects of endangered linguistic practices. It should be a priority to create a video documentation corpus of ritual speech in languages across Sumba before it goes out of use. Documentation and description are also valuable as a basis for engaging the linguistic communities of Sumba in activities that promote continued use of local

languages. Orthographies, natural reading materials, dictionaries and digital content in local languages are all improved by language documentation and description.

A major challenge for linguists working in Sumba is to balance the need for in-depth documentation and description with the task of capturing the breadth of linguistic diversity in Sumba. There are three obvious areas where there is a need for further documentation and sociolinguistic investigation of particular linguistic communities: 1) the dialectical variation of Kambera, 2) the status of Gaura as a separate language, and 3) the relationships between Weyewa, Loura, Loli and other closely-related lects.

Abbreviations

Abbreviations are as follows:

1	first person	3	third person
DEF	definite	red	reduplication
SG	singular	SBJ	subject

References

Adams, Marie Jeanne. 1969. *System and meaning in East Sumba textile design: a study in traditional Indonesian art* (Cultural Report Series 16). New Haven: Yale University Southeast Asia Studies.

Adams, Marie Jeanne. 1980. Structural aspects of East Sumbanese art. In James J. Fox (ed.), *The flow of life: essays on Eastern Indonesia*, 208–220. Cambridge: Harvard University Press.

Adams, Ron L. 2007. *The megalithic tradition of West Sumba, Indonesia: an ethnoarchaelological investigation of megalith construction*. Simon Fraser University.

Ananta, Aris, Evi Nurvidya Arifin, M. Sairi Hasbullah, Nur Budi Handayani & Agus Pramono. 2015. *Demography of Indonesia's ethnicity*. Singapore: Institute of Southeast Asian Studies.

Anderbeck, Karl. 2015. Portraits of language vitality in the languages of Indonesia. In I. Wayan Arka, Ni Luh Nyoman Seri Malini & Ida Ayu Made Puspani (eds.), *Language documentation and cultural practices in the Austronesian world: papers from 12-ICAL. Volume 4*, 19–47. Canberra: Asia-Pacific Linguistics, College of Asia and the Pacific, Australian National University.

Aritonang, Jan S. & Karel A. Steenbrink (eds.). 2008. *A history of Christianity in Indonesia* (Studies in Christian Mission v. 35). Leiden: Brill.

Asplund, Leif. 2010. The languages of Sumba. Presented at the East Nusantara Conference, Kupang, NTT.

Asplund, Leif. 2017. *Some formal characteristics of parallel speech in Kambera*. Stockholm University MA thesis. https://www.diva-portal.org/smash/get/diva2:1108705/FULLTEXT01.pdf (8 December, 2020).

Balle, Misriani. 2020. Models of language maintenance or revitalization in Sumba. Presented at the conference on Indonesian Languages and Linguistics: State of the Field, Atma Jaya University, Jakarta, Indonesia.

Balle, Misriani & Joseph Lovestrand. 2019a. Sumba wordlists and phrases. PARADISEC. http://catalog.paradisec.org.au/collections/JL4 (7 July, 2020).

Balle, Misriani & Joseph Lovestrand. 2019b. How to write implosives in Kodi. In Yanti (ed.), *Konferensi Linguistik Tahunan Atma Jaya (KOLITA) 17*. Universitas Atma Jaya, Jakarta.

Bibbs, Richard, Christian Doda, Marlince Lenda & Yanti. 2018. Recordings of various texts in Wejewa, a language spoken in Sumba Island, East Nusa Tenggara, Indonesia. PARADISEC. http://catalog.paradisec.org.au/collections/WEW2018 (8 July, 2020).

Bili, Angelina Yulianti. 2020. *Morfosintaksis bahasa Wewewa dialek Loura: kajian awal*. Tambolaka, Sumba Barat Daya: Sekolah Tinggi Keguruan dan Ilmu Pendidikan (STKIP) Weetebula.

Blust, Robert. 2008. Is there a Bima-Sumba subgroup? *Oceanic Linguistics* 47. 45–113.

BPS. 2015. *Proyeksi Penduduk Kabupaten/Kota Provinsi Nusa Tenggara Timur 2010–2020*. Jakarta: Badan Pusat Statistik.

Budasi, I Gede. 2007. *Kekerabatan bahasa-bahasa di Sumba (Suatu kajian linguistik historis komparatif)*. Ph.D. Thesis. Yogyakarta: Universitas Gadjah Mada.

Bulu, Yuliana. 2018. *Fonologi bahasa Wewewa di desa Weekombaka, kecamatan Wewewa Barat, Sumba Barat Daya, NTT: sebuah kajian awal*. Tambolaka, Sumba Barat Daya: Sekolah Tinggi Keguruan dan Ilmu Pendidikan (STKIP) Weetebula.

Campbell, Lyle, Nala Huiying Lee, Eve Okura, Sean Simpson, Kaori Ueki & John Van Way. 2013. New knowledge: Findings from the Catalogue of Endangered Languages (ELCat). Presented at the 3rd International Conference on Language Documentation and Conservation, February 28 – March 3, 2013, University of Hawai‘i at Mānoa. http://hdl.handle.net/10125/26145 (8 December, 2020).

Djawa, Alex. 2000. *Rekonstruksi protobahasa Kambera – Loli – Kodi – Lamboya di Sumba, Provinsi Nusa Tenggara Timur*. MA Thesis. Denpasar, Bali: Universitas Udayana.

Djawa, Alex. 2013. *Buku pelajaran hanewe Laboya, kabupaten Sumba Barat [Textbook in Laboyan Language, West Sumba Regency]*. Kupang: UPT Bahasa dan Sastra Indonesia dan Daerah Dinas Kebudayaan dan Pariwisata Provinsi Nusa Tenggara Timur.

Donohue, M. 2008. How many guests, what sort of hosts? *Proceedings of the National Academy of Sciences* 105. E1–E1.

Eberhard, David M., Gary F. Simons & Charles D. Fennig. 2020. *Ethnologue: languages of the world*. Twenty-third edition. Dallas, Texas: SIL International. http://www.ethnologue.com (8 December, 2020).

Edwards, Owen & UBB. 2018. *Languages of Sumba and Sabu*. Kupang: Language & Culture Unit (UBB) and Leiden University. http://vici.marianklamer.org/uploads/1/2/4/7/124768088/sumba_sabu_languages.pdf (8 December, 2020).

Forshee, Jill. 2001. *Between the folds: stories of cloth, lives, and travels from Sumba*. Honolulu: University of Hawai’i Press.

Forth, Gregory. 1982. Sumbanese finger names: Some comparative remarks. https://repository.ubn.ru.nl/bitstream/handle/2066/73424/73424.pdf.

Forth, Gregory. 1983. Blood, milk and coconuts: a study of intra-cultural ritual variation. *Man* 18(4). 654–668.

Forth, Gregory. 1988. Fashioned speech, full communication: aspects of eastern Sumbanese ritual language. In James J. Fox (ed.), *To speak in pairs: essays on the ritual languages of eastern Indonesia*, 129–160. Cambridge: Cambridge University Press.

Fox, James J. (ed.). 1988. *To speak in pairs: essays on the ritual languages of eastern Indonesia*. Cambridge: Cambridge University Press.

Fricke, Hanna. 2020. New subgrouping evidence for the Bima-Lembata languages (Austronesian, east Indonesia). Presented at the 12th International Austronesian and Papuan Languages and Linguistics Conference (APLL12), University of Oslo (online). https://www.hf.uio.no/iln/forskning/aktuelt/arrangementer/konferanser/2020/APLL12/program/fricke.pdf (8 December, 2020).

Gasser, Emily. 2014. Subgrouping in Nusa Tenggara: the case of Bima-Sumba. In Jeffrey Connor-Linton & Luke Wander Amoroso (eds.), *Measured Language: Quantitative Studies of Acquisition, Assessment, and Variation*. 63–78. Washington, DC: Georgetown University Press.

Geirnaert-Martin, Danielle C. 1992. *The woven land of Laboya: Socio-cosmic ideas and values in West Sumba, Eastern Indonesia*. Leiden: Centre of Non-Western Studies, Leiden University.

Ghanggo Ate, Yustinus. 2018. *Reduplication in Kodi*. MA thesis. Australian National University.

Ghanggo Ate, Yustinus. 2019. Pengajaran literasi berbasis bahasa ibu: sebuah alternatif mengurai isu literasi di Sumba. Presented at the Makalah Orasi Ilmiah disampaikan dalam rangkah Wisuda Perdana, STKIP Weetebula, Sumba-NTT. https://www.academia.edu/40538459/Ghanggo_Ate_Y._2019._Pengajaran_literasi_berbasis_bahasa_ibu_Sebuah_alternatif_mengurai_isu_literasi_di_Sumba._Makalah_Orasi_Ilmiah_disampaikan_dalam_rangkah_Wisuda_Perdana_STKIP_Weetebula_Sumba-NTT_28_Januari_2019 (19 July, 2020).

Goh, Taro. 1991. *Sumba bibliography*. Canberra, ACT, Australia: Dept. of Anthropology, Research School of Pacific Studies, Australian National University.

Greenhill, S. J., R. Blust & R. D. Gray. 2008. The Austronesian Basic Vocabulary Database: from bioinformatics to lexomics. *Evolutionary Bioinformatics* 4. 271–283.

Gunawan Mitchell, Istutiah. 2019. Changing times and changing patterns of collective violence in Sumba. In *The 7th International Symposium of Journal Antropologi Indonesia (ISJAI) Proceedings*, 219–223. Universitas Gadjah Mada, Yogyakarta: Departemen Antropologi Universitas Indonesia. https://simposiumjai.ui.ac.id/wp-content/uploads/20/2020/03/p219.pdf (8 December, 2020).

Hammarström, Harald, Robert Forkel, Martin Haspelmath & Sebastian Bank. 2020. Glottolog 4.3. https://glottolog.org/ (8 December, 2020).

Heijmering, Geerlof. 1846. Bijdrage tot de kennis van de taal der ZW Eilanden, benevens een proeve van vergelijking derzelve met acht andere inlandsche talen. *Tijdschrift voor Nederlandsch-Indië* 8. 1–81.

Himmelmann, Nikolaus P. 1998. Documentary and descriptive linguistics. *Linguistics* 36. 161–195.

Himmelmann, Nikolaus P. 2006. Language documentation: What is it and what is it good for. In Jost Gippert, Nikolaus P. Himmelmann & Ulrike Mosel (eds.), *Essentials of language documentation*. 1–30. Berlin/New York: Mouton de Gruyter.

Hoskins, Janet Alison. 1985. A life history from both aides: the changing poetics of personal experience. *Journal of Anthropological Research* 41. 147–169.

Hoskins, Janet. 1986. So my name shall live: stone-dragging and grave-building in Kodi, West Sumba. *Bijdragen tot de Taal-, Land- en Volkenkunde* 142. 31–51.

Hoskins, Janet. 1987a. The headhunter as hero: Local traditions and their reinterpretation in national history. *American Ethnologist* 14. 605–622.

Hoskins, Janet. 1987b. Complementarity in this world and the next: gender and agency in Kodi mortuary ceremonies. In Marilyn Strathern (ed.), *Dealing with inequality: analysing gender relations in Melanesia and beyond*. 174–206. Cambridge: Cambridge University Press.

Hoskins, Janet. 1988a. Etiquette in Kodi spirit communication: the lips told to pronounce, the mouth told to speak. In James J. Fox (ed.), *To speak in pairs: essays on the ritual languages of eastern Indonesia*. 29–63. Cambridge: Cambridge University Press.

Hoskins, Janet. 1988b. The drum is the shaman, the spear guides his voice. *Social Science & Medicine* (Special Issue Techniques of Healing in Southeast Asia) 27. 819–828.

Hoskins, Janet. 1989a. Burned paddy and lost souls. *Bijdragen tot de Taal-, Land- en Volkenkunde* 145. 430–444.

Hoskins, Janet. 1989b. Why do ladies sing the blues? Indigo dyeing, cloth production, and gender symbolism in Kodi. In Annette B Weiner & Jane Schneider (eds.), *Cloth and human experience*. 141–173. Washington: Smithsonian Institution Press.

Hoskins, Janet. 1989c. On losing and getting a head: warfare, exchange, and alliance in a changing Sumba, 1888–1988. *American Ethnologist* 16. 419–440.

Hoskins, Janet. 1993. Violence, sacrifice, and divination: giving and taking life in eastern Indonesia. *American Ethnologist* 20(1). 159–178.

Hoskins, Janet. 1994. *The play of time: Kodi perspectives on calendars, history, and exchange*. Berkeley: University of California Press.

Hoskins, Janet. 1996. From diagnosis to performance: medical practice and the politics of exchange in Kodi, West Sumba. In Carol Laderman & Marina Roseman (eds.), *The performance of healing*. 271–290. New York: Routledge.

Hoskins, Janet. 1998. *Biographical objects: how things tell the stories of people's lives*. Psychology Press.

Hoskins, Janet. 2002. Predatory voyeurs: tourists and "tribal violence" in remote Indonesia. *American Ethnologist* 29. 797–828.

Hoskins, Janet. 2004. Slaves, brides and other 'gifts': Resistance, marriage and rank in eastern Indonesia. *Slavery & Abolition* 25. 90–107.

Hoskins, Janet. 2005. Calling on the ancestors to fight crime: Reinvented ritual in an age of intermittent violence. In Pamela Stewart & Andrew Strathern (eds.), *Expressive culture and historical change*. 131–153. London: Ashgate Publishers.

Hoskins, Janet & Laura Scheerer Whitney. 1988. *The feast in dream village*. Documentary Educational Resources. https://store.der.org/the-feast-in-dream-village-p846.aspx (2 January, 2019).

Hoskins, Janet & Laura Scheerer Whitney. 1991. *Horses of life and death*. Documentary Educational Resources. https://store.der.org/the-feast-in-dream-village-p846.aspx (2 January, 2019).

Kaiping, Gereon A., Owen Edwards & Marian Klamer (eds.). 2019. *LexiRumah 3.0.0*. Leiden: Leiden University Centre for Linguistic. https://lexirumah.model-ling.eu/ (8 December, 2020).

Kapita, Oe H. 1982. *Kamus Sumba/Kambera-Indonesia*. Waingapu: Gereja Kristen Sumba.

Kapita, UH. 1983. *Tatabahasa Sumba Timur dalam dialek Kambera*. Ende, Flores: Arnoldus.

Kasni, Ni Wayan. 2012. *Strategi penggabungan klausa bahasa Sumba dialek Waijewa*. Ph.D. thesis. Denpasar: Universitas Udayana.

Keane, Webb. 1991. Delegated voice: ritual speech, risk, and the making of marriage alliances in Anakalang. *American Ethnologist* 18. 311–330.

Keane, Webb. 1995a. The spoken house: text, act, and object in eastern Indonesia. *American Ethnologist* 22. 102–124.

Keane, Webb. 1995b. Religious change and historical reflection in Anakalang, West Sumba, Indonesia. *Journal of Southeast Asian Studies* 26. 289–306.

Keane, Webb. 1997. *Signs of recognition: powers and hazards of representation in an Indonesian society*. Berkeley: University of California Press.

Keane, Webb. 2018. Killing animals: on the violence of sacrifice, the hunt and the butcher. *Anthropology of this Century* 22. http://aotcpress.com/articles/killing-animals-violence-sacrifice-hunt-butcher/ (20 July, 2020).

Keller, Edgar. 1989. Fashioned words of dispair: an autobiograhical speech of a dispossessed in Laboya, West Sumba. In W. Marschall (ed.), *Texts from the islands; oral and written traditions of Indonesia and the Malay world* (Proceedings of the 7th European Colloquium on Indonesian and Malay Studies). 265–309. Berne: The University of Bern, Institute of Ethnology.

Kemdikbud. 2003a. *Kamus bahasa Anakalang-Indonesia*. Waikabubak: Dinas Pendidikan dan Kebudayaan Kabupaten Sumba Barat, Sub Dinas Kebudayaan.

Kemdikbud. 2003b. *Kamus bahasa Kodi - Indonesia*. Dinas Pendidikan dan Kebudayaan, Kabupaten Sumba Barat, Sub Dinas Kebudayaan.

Kemdikbud. 2003c. *Kamus bahasa Wewewa-Indonesia*. Waikabubak: Dinas Pendidikan dan Kebudayaan Kabupaten Sumba Barat, Sub Dinas Kebudayaan.

Kemdikbud. 2019. Peta bahasa di Indonesia: Provinsi Nusa Tenggara Timur. Badan Pengembangan Bahasa Dan Perbukuan, Kementerian Pendidikan dan Kebudayaan. https://petabahasa.kemdikbud.go.id/ (6 July, 2020).

Klamer, M. A. F. 1999. Grammaticalized quotations in Kambera, Buru and Tukang Besi. In E. Zeitoun & P. Jen-kuei Li (eds.), *Selected Papers from the Eight International Conference on Austronesian Linguistics.* 579–598. Taipei: Academia Sinica.

Klamer, Marian. 1994. *Kambera: a language of eastern Indonesia*. Ph.D. thesis. The Hague: Vrije Universiteit Amsterdam.

Klamer, Marian. 1996. Kambera has no passive. In Marian Klamer (ed.), *Voice in Austronesian* (NUSA 39). 12–30. Jakarta: Universitas Atma Jaya.

Klamer, Marian. 1997. Spelling out clitics in Kambera. *Linguistics* 35. 895–928.

Klamer, Marian. 1998a. *A grammar of Kambera*. Berlin: Mouton de Gruyter.

Klamer, Marian. 1998b. Kambera intransitive argument linking. *Studia Linguistica* 52. 77–111.

Klamer, Marian. 2000a. Continuative aspect and the dative clitic in Kambera. In Ileana Paul, Vivianne Phillips & Lisa Travis (eds.), *Formal issues in Austronesian linguistics*. 49–63. Dordrecht: Kluwer.

Klamer, Marian. 2000b. How report verbs become quote markers and complementisers. *Lingua* 110. 69–98.

Klamer, Marian. 2001. Phrasal emotion predicates in three languages of eastern Indonesia. In Geert Booij & Jaap Van Marle (eds.), *Yearbook of Morphology 2000*. 97–121. Dordrecht: Springer Netherlands.

Klamer, Marian. 2002a. Report constructions in Kambera (Austronesian). In Tom Güldemann & Manfred Roncador (eds.), *Reported discourse. a meeting ground for different linguistic domains*. 323–340. Amsterdam: John Benjamins Publishing Company.

Klamer, Marian. 2002b. Semantically motivated lexical patterns: a study of Dutch and Kambera expressives. *Language* 78. 258–286.

Klamer, Marian. 2004. Multi-categorial items as underspecified lexical entries: The case of Kambera wàngu. In O. Fischer, Muriel Norde & Harry Perridon (eds.), *Up and down the cline: the nature of grammaticalization*. 299–323. Amsterdam: John Benjamins Publishing.

Klamer, Marian. 2005. Kambera. In Alexander Adelaar & Nikolaus Himmelmann (eds.), *The Austronesian languages of Asia and Madagascar* (Routledge Language Family Series). 709–734. London & New York: Routledge.

Klamer, Marian. 2009. The use of language data in comparative research: a note on Blust (2008) and Onvlee (1984). *Oceanic Linguistics* 48. 250–263.

Kuipers, Joel C. 1986. Talking about troubles: gender differences in Weyéwa speech use. *American Ethnologist* 13. 448–462.

Kuipers, Joel C. 1988. The pattern of prayer in Weyéwa. In James J. Fox (ed.), *To speak in pairs: essays on the ritual languages of eastern Indonesia*. 104–128. Cambridge: Cambridge University Press.

Kuipers, Joel. 1990. *Power in performance: the creation of textual authority in Weyewa ritual speech*. Philadelphia: University of Pennsylvania Press.

Kuipers, Joel C. 1998. *Language, identity, and marginality in Indonesia: the changing nature of ritual speech on the island of Sumba*. Cambridge: Cambridge University Press.

Kuipers, Joel C. 2009. Unjuk rasa ("expression of feeling") in Sumba. In Gunter Senft & Ellen B. Basso (eds.), *Ritual communication*. 223–242. Bloomsbury Academic.

Labu Djuli, M. Hum., Adriana Ratu Koreh & Matheos Tanda Kawi. 2005. *Bahasa Kodi: untuk murid kelas III SD*. UPTD Bahasa, Dinas Pendidikan dan Kebudayaan, Provinsi Nusa Tenggara Timur.

LAI (trans.). 2018. *Paneghe ndahha mangnga nggambarona [Kodi Children's Bible]*. Jakarta: Lembaga Alkitab Indonesia.

Lansing, J. S. 2008. Reply: inference and Proto-Sumba. *Proceedings of the National Academy of Sciences* 105. E2–E2.

Lansing, J. Stephen, Murray P. Cox, Sean S. Downey, Brandon M. Gabler, Brian Hallmark, Tatiana M. Karafet, Peter Norquest, John W. Schoenfelder, Herawati Sudoyo & Joseph C. Watkins. 2007. Coevolution of languages and genes on the island of Sumba, eastern Indonesia. *Proceedings of the National Academy of Sciences* 104. 16022–16026.

Larsson, Milene. 2013. Blood sacrifice in Sumba. *Vice*. https://www.youtube.com/watch?v=iIu38uhrBFk (8 December, 2020).

Limy, Stafanus Dangga. 2003. *Sistem sapaan bahasa Wewewa dalam upacara perkawinan*. MA thesis. Kupang: The University of Nusa Cendana.

Lovestrand, Joseph. 2020. Comparing orthography design in Barayin (Chad) and Kodi (Indonesia). In *Proceedings of the First International Seminar on Languare, Literature, Culture and Education, ISLLCE, 15-16 November 2019, Kendari, Indonesia*. Kendari, Indonesia: EAI. https://doi.org/10.4108/eai.15-11-2019.2296216. http://eudl.eu/doi/10.4108/eai.15-11-2019.2296216 (22 July, 2020).

Lovestrand, Joseph & Misriani Balle. 2019. An initial analysis of Kodi phonology. Presented at the International Conference on the Austronesian and Papuan Worlds (ICAPaW), Denpasar, Bali.

Mitchell, David. n.d. Sounds and spellings in the Wanukaka language.

Mitchell, David. 1988. Method in the metaphor. In James J. Fox (ed.), *To speak in pairs: essays on the ritual languages of eastern Indonesia*. 64–86. Cambridge: Cambridge University Press.

Mitchell, David. 2008. *Hilu Wanukaka- a working dictionary of the Wanukaka language*. http://www.wanukaka.com/Home/hilu-wanukaka (8 December, 2020).

Needham, Rodney. 1980. Principles and variations in the structure of Sumbanese society. In James J. Fox (ed.), *The Flow of Life: Essays on Eastern Indonesia*, 21–47. Cambridge, MA: Harvard University Press.

Ngongo, Magdalena. 2013. *Teks Kette Katonga Wei Kawendo pada masyarakat adat Wewewa di pulau Sumba: analisis linguistik sistemik fungsional*. Ph.D. thesis. Denpasar: Universitas Udayana.

Ngongo, Magdalena. 2015. Analisis kohesi gramatikal dan leksikal text lisan peminangan dalam bahasa Waijewa. *Jurnal Tutur* 1. 189–202.

Onvlee, Louis. 1925. *Eenige Soembaasche Verteilingen*. Leiden: Brill.

Onvlee, Louis. 1973. *Cultuur als antwoord*. 's-Gravenhage: Martinus Nijhoff.

Onvlee, Louis. 1984. *Kamberaas (Oost-Soembaas)-Nederlands woordenboek*. Dordrecht: Foris Publications.

Ori, Lobu. 2010. *Kamus bahasa Lolina*. Waikabubak: Dinas Kebudayaan dan Pariwisata Kabupaten Sumba Barat.

Putra, Anak Agung Putu. 2007. *Segmentasi dialektal bahasa Sumba di pulau Sumba: Suatu kajian dialektologi*. Ph.D. thesis. Denpasar: Universitas Udayana.

Renard-Clamagirand, Brigitte. 1988. Li'i marapu: speech and ritual among the Wewewa of west Sumba. In James J. Fox (ed.), *To speak in pairs*. 87–103. Cambridge: Cambridge University Press.

Rina, A. Dj & John Lado B. Kabba. 2011. *Kamus bahasa Lamboya, kabupaten Sumba Barat*. Kupang: Dinas Kebudayaan dan Pariwisata Kabupaten Sumba Barat.

Roos, S. 1872. Bijdrage tot de kennis van taal- land en volk op het eiland Sumba. *Verhandelingen van het Bataviaasch Genotschap van Kunsten en Wetenschappen* 36. 1–125.

Rothe, Elvira. 2004. *Wulla Poddu: Bitterer Monat, Monat des Tabus, Monat des Heiligen, Monat des Neuen Jahres in Loli in der Siedlung Tarung-Waitabar, Amtsbezirk der Stadt Waikabubak in Loli, Regierungsbezirk Westsumba, Provinz Nusa Tenggara Timur, Indonesien*. Ph.D. thesis. München: Ludwig Maximilians-Universität.

Rutherford, E., K. Burke & J. Lytwyn. 2001. Tectonic history of Sumba Island, Indonesia, since the Late Cretaceous and its rapid escape into the forearc in the Miocene. *Journal of Asian Earth Sciences* 19. 453–479.

Simanjuntak, Tiar. 2018. Laporan hasil diskusi ketahanan bahasa Kodi [kod] yang dilakukan di SD INPRES Homba Tana, Waikadada, Kodi Bangedo, Sumba Barat Daya, NTT. Suluh Insan Lestari.

Simons, Gary F. & M. Paul Lewis. 2013. The world's languages in crisis: a 20-year update. In Elena Mihas, Bernard Perley, Gabriel Rei-Doval & Kathleen Wheatley (eds.), *Responses to language endangerment: in honor of Mickey Noonan. New directions in language documentation and language revitalization*. 3–20. Amsterdam: John Benjamins Publishing Company.

Sou, Jennifer. 2019. Sou collection - Indonesia. Kaipuleohone, University of Hawai'i Digital Language Archive. https://scholarspace.manoa.hawaii.edu/handle/10125/64694 (8 December, 2020).

Sukerti, Gusti Nyoman Ayu. 2013. *Relasi gramatikal bahasa Kodi: kajian tipologi sintaksis*. MA thesis. Denpassar, Bali: Universitas Udayana.

Sukerti, Gusti Nyoman Ayu & Yustinus Ghanggo Ate. 2016. Pola pemarkahan argumen bahasa Kodi. *Linguistik Indonesia* 34. 129–145.

Twikromo, Y. Argo. 2008. *The local elite and the appropriation of modernity: a case in East Sumba, Indonesia*. Ph.D. thesis. Nijmegen: Radboud Universiteit. http://webdoc.ubn.ru.nl/mono/t/twikromo_y/locaelant.pdf (21 July, 2020).

Vel, Jacqueline A. C. 1994. *The Uma-economy: indigenous economics and development work in Lawonda, Sumba (Eastern-Indonesia)*. Wageningen.

Vel, Jacqueline A. C. 2001. Tribal battle in a remote island: crisis and violence in Sumba (eastern Indonesia). *Indonesia* 72. 141.

Vel, Jacqueline V. A. 2008. *Uma politics: an ethnography of democratization in West Sumba, Indonesia, 1986-2006*. Leiden: KITLV Press.

Velden, J. H. van der. 1900. Proeve eener Spraakleer van de Laoraneesche Taal. *Tijdschrift voor Indische Taal-, Land- en Volkenkunde* XLII. 57–101.

Verdizade, Allahverdi. 2019. *Selected topics in the phonology and morphosyntax of Laboya*. MA thesis. Stockholm University.

Vermast, AM. 1895. Lijst van Soembaneesche woorden en uitdrukkingen, alphabetisch gerangschikt. *Veeartsenijkundige Bladen voor Nederlandsch-Indië* 9. 122–42.

Walangara, Trovina R. P. E. 2003. *Verba bahasa Anakalang*. MA thesis. Kupang: The University of Nusa Cendana.

Wellem, Frederiek Djara. 2004. *Injil dan marapu: suatu studi historis-teologis tentang perjumpaan Injil dengan masyarakat Sumba pada periode 1876-1990*. BPK Gunung Mulia.

Wielenga, D. K. 1909. *Schets van een Soembaneesche Spraakkunst naar 't dialekt van Kambera* (Verhandelingen van Het Koninklijk Bataviaasch Genootschap van Kunsten En Wetenschappen). Vol. XLVIII. Batavia: Landsdrukkerij.

Wielenga, D. K. 1917. *Vergelijkende woordenlijst der verschillende dialecten op het eiland Soemba en eenige Soembaneesche spreekwijzen* (Verhandelingen van Het Bataviaasch Genootschap van Kunsten En Wetenschappen). Vol. LXI:6. 's Gravenhage: M. Nijhoff.

Yanti & Asako Shiohara. 2018. Efforts in language documentation in a linguistically diverse country: building up collaborations of various stakeholders. Presented at the NINJAL International Symposium: Approaches to Endangered Languages in Japan and Northeast Asia: Description, Documentation and Revitalization, Tachikawa, Tokyo.