

## Training communities in documentation and technology: The Language Documentation Training Center model

Jennifer Sou

*University of Hawai'i at Mānoa*

Leah Pappas

*University of Hawai'i at Mānoa*

Khairunnisa

*National Research and Innovation Agency*

Gary Holton

*University of Hawai'i at Mānoa*

Language documentation is increasingly seen as a collaborative process, engaging community members as active participants. Collaborative research produces better documentation that is valuable for both the academic community and the speakers. However, in many communities, speakers and language advocates lack the skills necessary to fully engage in collaborative projects. One way to overcome this barrier is to provide language documentation training to community members. Such training should teach participants how to ethically and comprehensively complete every stage of the documentation process while offering opportunity for theoretical discussion and practical application. In this paper, we offer one possible model for community-based training in language documentation and conservation that focuses on bidirectional learning and capacity building. We describe a training workshop that was held in 2018 in Kupang, the capital of Indonesia's Nusa Tenggara Timur (NTT) province. A collaboration between the University of Hawai'i, Leiden University, and Artha Wacana Christian University, this workshop implemented a model based on the practices of the Language Documentation Training Center (LDTC), an organization devoted to training speakers to document their own languages. We detail the NTT workshop itself, summarize post-workshop feedback, and offer suggestions to others looking to provide similar training in speaker communities.

**1. Introduction<sup>1</sup>** Language documentation is best conceived as a collaborative process, engaging community members as active participants (Czaykowska-Higgins 2009). Summarizing the current state of the field, Grenoble notes that “[t]here is wide-spread agreement among linguists engaged in language documentation today that they must engage in collaborative work with the communities of native speakers whose languages they document” (2010: 295). The moral argument for engaging local participants in the documentation process is clear. Engaging speakers in the documentation of their languages represents a form of decolonization, allowing speakers to take control of the process of creating a lasting record of their language. Most of the literature on collaborative documentation has focused on the Indigenous languages of North America and Australia, and we must be careful not to impose North American research paradigms in non-North American contexts (cf. Holton 2009; Crippen & Robinson 2013). Nonetheless, in many parts of the world, lack of access to language documentation training remains a significant barrier to overcoming extractive, outsider-led research models. The need for greater diversity of participation in language documentation projects is particularly acute in regions such as Indonesia, where access to documentation training has been limited (Arka 2018).

Beyond this moral argument, there is good evidence that collaborative research simply produces better documentation. Truly collaborative projects benefit from diverse expertise and result in documentation that is valuable for both the academic community as well as the speakers (Leonard & Haynes 2010). The resulting work can more accurately reflect community knowledge as well as patterns of language use (Olko 2018). Native speaker linguists also tend to have greater opportunity to spend time “in the field.” While outside linguists may be able to spend a few months at a time documenting languages in Indonesia, duties toward work, study, and family require that they eventually return to their home countries. In contrast, native speaker linguists live and work in their field areas, providing ample opportunity for documentation work. Finally, given the enormous amount of work to be done as soon as possible, there is simply no way to effectively complete language documentation tasks by relying solely on outside linguists. As such, it is necessary and desirable to train a cadre of local linguists and language advocates who can carry out this work and sustain documentation efforts into the future.

One such way to achieve collaborative documentation is to provide language documentation training to community members. In this respect, training should be intended to teach participants how to ethically and comprehensively complete every stage of the documentation process while offering opportunity for theoretical discus-

---

<sup>1</sup> The Documenting Minority Languages in NTT Workshop was supported by National Science Foundation grant BCS-1761223, the Netherlands Organization for Scientific Research, and Artha Wacana Christian University, and its success depended on the collective efforts of many people involved. The authors would especially like to acknowledge: A.L. Blake, Jacob Hakim, and Trent Ukasick at the University of Hawai‘i at Mānoa; Marian Klamer, Hannah Fricke, George Saad, and Yunis Sulistyono at Leiden University; and June Jacob at Artha Wacana Christian University. Thanks are especially due to the numerous colleagues and staff at Artha Wacana Christian University who assisted with logistics in Kupang. Of course, we would like to thank all the participants and, last but not least, the community members who welcomed us into their homes and allowed us to document their languages.

sion and practical application. In this paper, we offer one possible model for community-based training in language documentation and conservation. We describe a training workshop that was held in 2018 in Kupang, the capital of Indonesia's Nusa Tenggara Timur (NTT) province. A collaboration between the University of Hawai'i, Leiden University, and Artha Wacana Christian University, this workshop implemented a model based on the practices of the Language Documentation Training Center (LDTC), an organization devoted to training speakers to document their own languages. We begin in §2 with a brief overview of the language situation in NTT and the need for training. In §3, we describe the LDTC training model, which served as a basis for the workshop design. In §4, we discuss the ways in which the LDTC model was adapted to the local context in NTT and summarize the implemented workshop. In §5, we evaluate the effectiveness of the workshop, drawing on feedback from workshop participants and leaders. Finally, in §6, we conclude with a generalization of the impacts of the workshop and its broader applicability to other contexts.

## 2. Languages and language documentation in NTT

**2.1 Linguistic diversity** With more than 700 languages, Indonesia is among the most linguistically diverse areas of the world. This diversity is particularly evident in NTT, a meeting ground between the vast Austronesian family and the non-Austronesian (or Papuan) languages of the Timor-Alor-Pantar family. The Austronesians settled in NTT perhaps 4,000 years ago (Grimes et al. 1997), but there is evidence of human occupation dating back to more than 40,000 years (O'Connor et al. 2011), and linguistic relationships in the region remain the subject of much speculation. Although the Timor-Alor-Pantar languages have long been assumed to belong to the vast Trans-New Guinea group, recent work finds little evidence to support the relationship of the Timor-Alor-Pantar languages to other non-Austronesian languages of the New Guinea mainland (Holton & Klamer 2017). The relationships within the Austronesian languages of the region are also poorly understood. Most belong to the so-called Central Malayo-Polynesian group, but the status of this group remains the subject of much debate, and few lexical items can be reconstructed at this level (Blust 1993; Donohue & Grimes 2008). The general picture that emerges is one of an extended period of contact between Austronesian and non-Austronesian languages, resulting in significant lateral transfer (cf. Saad et al. 2019). This picture is consistent with the genetic evidence, which unexpectedly reveals that speakers of Austronesian languages in NTT share more genetic similarities with speakers of non-Austronesian languages in New Guinea than do the speakers of Timor-Alor-Pantar languages (Mona et al. 2009).

As a result of this history of contact between Austronesian and non-Austronesian peoples, today about fifty-eight languages are spoken in NTT (Klamer & Ewing 2010). Their distribution is shown in Figure 1. Despite the evident diversity in the region, the languages of NTT have similar features. They are known for their general lack of morphology and little to no voice system (Arka 2002). More importantly, NTT is included in the so-called linguistic area of East Nusantara, which consists of approximately 400 languages (Klamer & Ewing 2010). Thomason (2001) defines

a linguistic area as a geographical area where a group of three or more languages exist side by side and share some structural features as a result of contact instead of by accident or inheritance of a common ancestor. Thus, in the case of languages in NTT, it is expected that some, if not all, of the languages contain shared features, which have diffused between Austronesian and non-Austronesian languages. Klamer and Ewing (2010) report that this phenomenon is found in Alor and Pantar, where the Austronesian language Alor consists of a base Austronesian lexicon but applies morphosyntactic features associated with (substrate) Papuan languages.

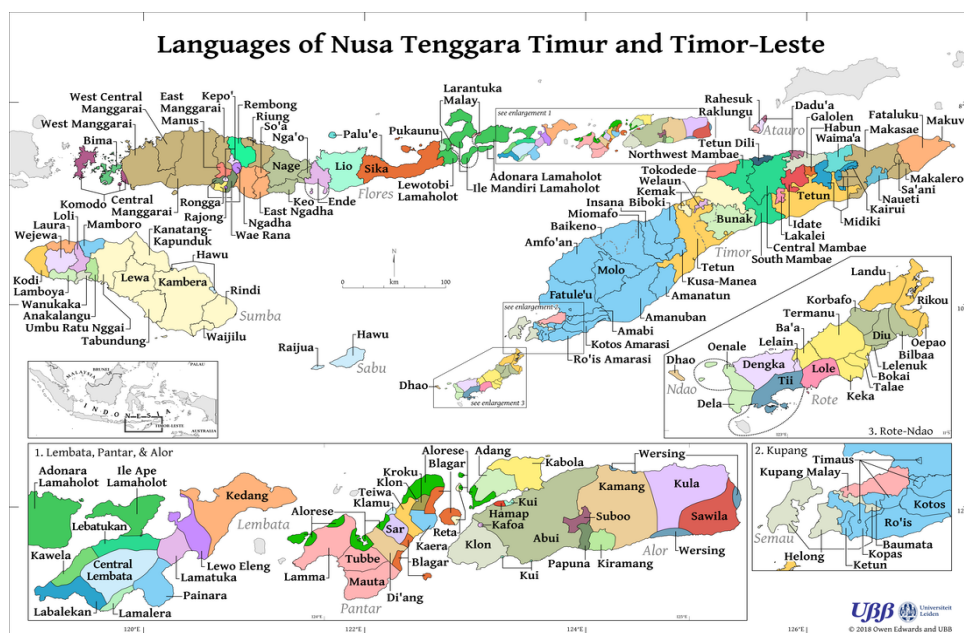


Figure 1. Language varieties of Nusa Tenggara Timur and Timor-Leste (Edwards & UBB 2018). Distinct languages are marked by distinct colors; dialects share the same color.

In spite of a number of recent documentation efforts, including a multiyear international collaboration under the auspices of the European Science Foundation, many of the languages of NTT remain only cursorily documented. Several grammatical descriptions and dictionaries exist, but few large-scale documentary corpora are available in language archives. Many of the languages continue to be used in village contexts, but a shift to Malay is occurring at an accelerated rate in conjunction with rapid urbanization (Ewing 2014). Nonetheless, the assumption among practitioners has been that for the most part, even where shift has occurred in urban environments, most languages remain “safe” in their home village environments. However,

recent sociolinguistic research suggests this assumption may not be valid. Saad's (2016) work with Abui, one of the largest and most vibrant languages of the region, reveals a much higher rate of shift to Malay and a lower knowledge of Abui among youth in village settings. This suggests that the well-known phenomenon of urban language shift in the region is now taking place in rural environments as well. Even when the local language is used, younger speakers in particular are adapting more and more Malay grammar to be used in the local language (Saad 2020).<sup>2</sup> Language shift has often led to wholesale loss of morphology, even among adult speakers, as reported for the Austronesian language Alorese (Moro 2019). In no areas of NTT are local languages used as a medium of education or for any other formal discourse.

Moreover, it is well established that specialized domains of knowledge tend to become endangered well before the language itself (Si 2011). Recent ethnobotanical research in Timor suggests that knowledge of indigenous plant names is greatly reduced in regions with increased access to modern medical facilities (Collins et al. 2006). Documentation of NTT languages needs to be prioritized now, while speakers with detailed knowledge of grammar and semantic systems still remain.

## 2.2 Challenges and the need for access to training in documentary linguistics

The Indonesian government provides inadequate financial support for linguistic research and language documentation and conservation efforts. The bureaucracy in the country can be complicated, and oftentimes, funding for language documentation projects is unlikely to be awarded to individuals acting as independent researchers. Instead, a memorandum of understanding between institutions and the government is needed for the funding to be granted. Local linguists do have the option to compete for international grant funding, which is more competitive than national funding. However, only a few Indonesians have tried, and even fewer have succeeded. As a result, most language documentation in Indonesia has been done by foreign linguists (Arka 2018).

Even if funding is successfully secured, meaningful engagement of native speakers in collaborative language documentation research is not easy, as it requires community members to have some knowledge of each step of the language documentation process. In order to substantially contribute to the project, speakers should ideally have some basic linguistic knowledge, such as the ability to differentiate orthography from individual sounds and to identify words and morphemes. Moreover, they need to be familiar with digital tools such as audio- and video-recording techniques, data management protocols, and a range of software such as ELAN (Sloetjes & Wittenburg 2008), WeSay (Albright & Hatton 2007), and Fieldworks Language Explorer (FLEX) (SIL International 2021), among others, to transcribe, translate, and gloss data. Additionally, it is necessary for them to have an arrangement with an archive with which they can deposit data. As a result, native speakers without

---

<sup>2</sup> We use the term *Malay* as a cover term for what is essentially a continuum of language practice ranging from regional Malay varieties to formal or standard Indonesian. For additional discussion of the sociolinguistics of regional Malay varieties in Eastern Indonesia, see Jacob & Grimes (2006), Paauw (2008), and the references therein.

training are not able to participate in documentation projects as full partners. Often, their participation may be limited to serving as guides or language consultants during fieldwork or working with researchers to transcribe and translate data.

With few natural resources and a limited economy, NTT today remains the least developed of Indonesia's twenty-seven provinces, with a regional per capita income just one third of the national average (Badan Pusat Statistik 2021). As a result, residents are rarely able to travel outside the region; hence, NTT residents have few opportunities for linguistic training. Further, despite the urgent need for community-led language documentation in NTT, let alone in Indonesia as a whole, efforts to provide local training are not yet adequate.

The most readily available options in linguistic training in NTT are through one of the local universities and through the local Christian church. While none of the local universities offer a degree program in linguistics, both the public Nusa Cendana University and the private Artha Wacana Christian University offer training in English language and English language teaching. Many of the faculty in these programs do have formal linguistics training, and students in the programs are often encouraged to complete theses comparing the structure of their native languages to that of English. While these programs have been extremely successful in promoting greater awareness of local languages, students receive little training in the theory or techniques of documentary linguistics. Moreover, university-level training is accessible only to degree-seeking students enrolled in a university.

A second locally available option for linguistics training is through the Language and Culture Unit (UBB) of the synod of the Evangelical Church of Timor (*Gereja Masehi Injili di Timor*). The primary mission of UBB is to translate religious materials into the local languages of NTT. In fulfillment of this mission, UBB provides training to volunteer and professional staff, drawing on trainers from both within and outside Indonesia. Training may include basic linguistic analysis, as well as the use of software tools such as FLE<sub>x</sub> and WeSay. Again, this training is available only to UBB affiliates and hence not broadly accessible to all residents of NTT.

With greater and broader access to training, native speakers can act as agents in the documentation process. When trained, they are equipped to set up and execute recording sessions in an ethical and culturally appropriate manner; they can store and share data safely, and they can independently transcribe and translate data. A growing body of research shows that local involvement produces work that advances the field of linguistics by asking questions that are relevant to the community and providing a perspective that differs from that of academia (cf. Leonard & Haynes 2010; Roche et al. 2010; Fitzgerald & Linn 2013; Leonard 2017). In other words, training in documentary linguistics surely can enhance native speakers' involvement, and the resulting collaborative research approaches can benefit all involved (Fitzgerald 2018).

**3. The Language Documentation Training Center** An essential component to the success of the NTT workshop was the LDTC model and participation of its members. LDTC is a student-run organization at the University of Hawai'i at Mānoa Department of Linguistics. Its mission as stated on its website is "[...] to train communi-

ty members of underdocumented languages in the skills of language documentation, while spreading awareness about linguistic diversity and language endangerment.”<sup>3</sup> Since its founding in 2003, LDTC has worked with speakers of over one hundred languages, and the results of that documentation have been deposited with the Kaipuleohone Digital Archive. In addition, numerous web portals have been created, which provide language vignettes and help promote awareness of endangered languages.<sup>4</sup>

LDTC’s workshop series are held biannually on the University of Hawai‘i at Mānoa campus. Each of the eight two-hour workshops is held over the course of nine weeks during the fall and spring semesters. At the end of the workshop series, LDTC hosts a banquet where participants can share their language and the work they completed with the general public. Each workshop is led by an experienced LDTC member, the workshop leader. Workshops are intended to cover the fundamentals of language documentation theory and practice so that participants emerge from the workshop with the ability to document languages independently. Workshop themes therefore cover theoretical topics including language endangerment, traditional knowledge, and revitalization; linguistics fundamentals including phonology, orthography, syntax, and morphology; and technological skills including audio and video recording, data management, archiving, transcription, and subtitling. A typical LDTC workshop schedule is shown in Table 1. Each two-hour workshop session consists of a lesson, which takes place during the first hour, followed by hands-on implementation of the material during the second hour. Food and snacks are provided as a small incentive and give participants a chance to socialize.

Workshop participants, who are all voluntary, may fit one of two categories: Language Experts and Language Mentors. Language Experts are those who speak a minority language and are interested in learning more about linguistics and language documentation. Language Mentors are anyone who is interested in language documentation that does not speak an endangered language; this category includes both students and community members. Usually, both Experts and Mentors are beginners to language documentation, and they are paired up to achieve personalized synergistic bidirectional learning. The relationship between Expert and Mentor is a crucial aspect of LDTC (Rehg 2007). In these reciprocal relationships, both Mentors and Experts learn from each other as they work together throughout the workshops to ultimately create language materials – such as a website or a subtitled video – that provide information about the Expert’s language. Through this classroom simulation of community-led language documentation (Czaykowska-Higgins 2009), LDTC is not just a language endangerment awareness program but also an experiential learning opportunity for participants (Ajo et al. 2010; Butler 2011). LDTC members were intent on maintaining this aspect of LDTC during the workshop in Kupang.

---

<sup>3</sup> [ldtc.org](http://ldtc.org) (Accessed 2020-09-22.)

<sup>4</sup> [ldtc.org/languages](http://ldtc.org/languages) (Accessed 2020-09-22.)

**Table 1.** Spring 2019 LDTC workshop schedule indicating topics and activities covered during each workshop

Date	Topic	Activity
1 Jan 30	Language Attitudes, Endangerment, Traditional Knowledge	Language Endangerment Index evaluation
2 Feb 6	Phonology & Orthography	Modified Swadesh wordlist
3 Feb 13	Archiving & Audio Recording	Audio recording Swadesh wordlist
4 Feb 20	Morphology & Syntax	Word/sentence elicitation to identify common morphemes and word order
5 Feb 27	Video Recording	Video and audio record a story, recipe, song, etc. in your language
6 Mar 6	Biodiversity Walk	Video and audio record traditional knowledge about the use of plants or animals
7 Mar 13	Subtitling & YouTube	Practice ELAN transcription of video and upload to LDTC YouTube site
8 Mar 27	Revitalization and Finishing Websites	Outline a revitalization plan appropriate for your language
9 Apr 3	Closing Banquet and Presentations	-----

**4. Adapting the LDTC model to Nusa Tenggara Timur** Thus, the participation of LDTC was critical to the success of the Documenting Minority Languages of NTT workshop. The involvement of experienced LDTC members ensured that the collaborative approach to learning – in which the Mentors share their linguistic and technical knowledge while Experts share information about their languages and cultures – were met. LDTC members’ experience with curriculum planning and training beginners in language documentation crucially contributed to the development of the workshop schedule and modules. However, adjustments on many fronts were needed in order to successfully adapt the LDTC training model to a one-week workshop in Kupang with participants of various levels of knowledge about linguistics and technology. Some of the most significant changes and considerations are discussed presently.

**4.1 Workshop organization** To conduct the workshop so that it was broadly accessible to many participants, we had to identify an appropriate location, assure availability of equipment, and choose a language of delivery. The location of the workshop and having a local affiliate were forefront in shaping how everything could be



organized and run. In order to make this multiday workshop widely available to residents of NTT, it was hosted in Kupang, the capital of the province. As with the LDTC workshops, participation was entirely voluntary, and food was provided during the workshop sessions. As such, Kupang was logistically feasible because people from all over the region regularly travel to Kupang for work or school, and as a result, a diverse pool of interested participants were able to join the workshop. Kupang was also ideal in terms of access to utilities such as electricity and the internet, facilities needed to host a large group of people, paved roads, and multiple modes of transport. The local organizing institution, Artha Wacana Christian University, is located in Kupang, so its campus served as an accessible location for participants. Accommodation for participants was simplified in the capital city, since many participants already resided in Kupang or had family who could host them. Therefore, housing and travel costs were greatly diminished, and the workshop was able to provide travel stipends to those coming from neighboring islands. Additionally, Kupang has high linguistic diversity within the city and the surrounding areas. This allowed us to incorporate a field practicum as part of the workshop, without requiring participants to travel great distances.

To ensure that all participants had access to software and equipment, participants were requested to bring their own laptop computers. The workshop leaders provided audio and video equipment from their home institutions. Each participant was given a USB stick that had a variety of applications and workshop materials predownloaded as a “Language Documentation Starter Kit,” which included ELAN, SayMore, Audacity, OpenOffice, FFmpeg, Handbrake, a metadata sheet, and presentation slides. The internet, while accessible, was not reliable, so preparing the relevant software before the workshop avoided unnecessary technical issues with regard to downloading programs.

Finally, all workshop presentations were conducted either only in Malay or in English with Malay translation. Most participants had English abilities, but efforts to use Malay as the language of instruction aided in the understanding of concepts.

**4.2 Roles** The LDTC model of Expert–Mentor pairs had to be adjusted to suit the Kupang circumstances. In Kupang, we expanded upon the LDTC model to train participants to document the languages of the greater community. Participants were still welcome to document their own language, but this was not a requirement. This adjustment made it possible for several participants – all speakers of different languages – to work together on a documentation project.

As a result, there were no Expert–Mentor pairings; instead, there were workshop leaders and participants. Workshop leaders were professors and graduate students from the University of Hawai‘i at Mānoa and Leiden University (including two Indonesian nationals), who were the equivalent to LDTC Mentors and Leaders. These leaders served as workshop presenters and also worked with individual participants and teams to provide personalized mentoring and answer questions that arose during the presentations.

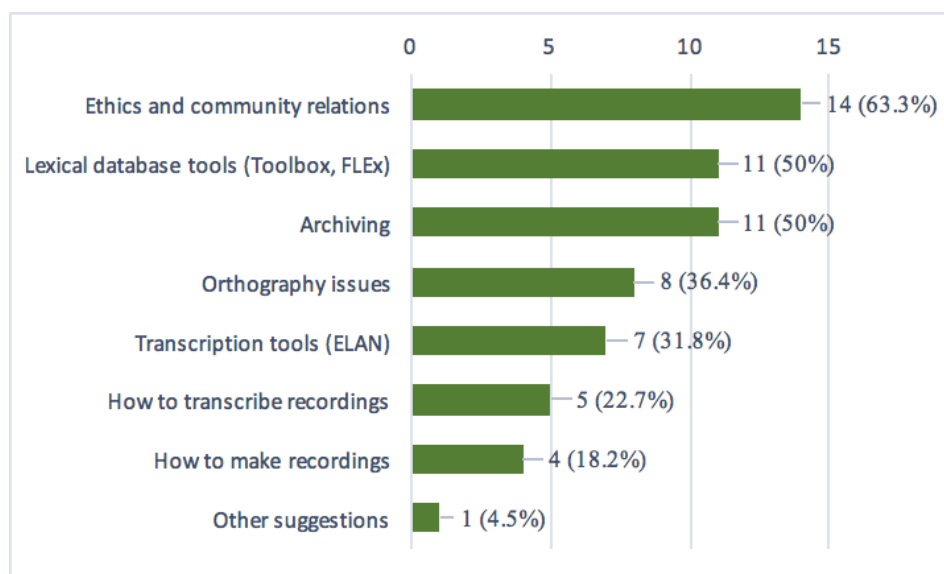
Workshop participants attended the workshop to be trained in documentation methods and served as cultural experts during the field practicum. The workshop

was open to anyone, though active recruitment targeted people who expressed interest in language documentation work through Artha Wacana Christian University and people who were already engaged in language documentation. A pre-workshop online survey was conducted, and of the twenty-two people who responded, one (4.5%) was a high school student, nine (40.9%) were current undergraduate students, five (22.7%) recently graduated from university, one (4.5%) was a master's student, one (4.5%) was a PhD student, one (4.5%) was a teacher, and four (18.2%) had a high school education but were not currently students. Seven (31.8%) noted that they had some experience in language documentation, while the remaining 15 (68.2%) did not. This range of educational backgrounds and knowledge of language documentation enriched the workshop and elevated the experience for everyone involved.

**4.3 Fieldwork practicum** As mentioned previously, holding the workshop in Kupang allowed for a fieldwork practicum component because of the great linguistic diversity of the area. This aspect differs greatly from the LDTC model, where there is continuous Expert–Mentor training and collaboration throughout the documentation process. By changing that dynamic to the leader and participant roles, a three-day fieldwork experience would emulate the same role of bidirectional learning and utilize the language documentation techniques learned throughout the training sessions, which would be conducted prior to going into the field.

The speaker communities were selected and prepared in advance by the affiliated local university faculty. Groups of three to five participants were paired with one or two workshop leaders and traveled to nearby areas to work with native speakers of several different NTT languages. Local relationships were key in allowing outside linguists and workshop participants to be accepted in the various communities. The process of building trust and establishing community relationships can take considerable time. By leveraging existing relationships between university faculty and NTT communities, we were able to create a field experience for the workshop participants, which would not normally be possible within the short time period available. All financial aspects of the practicum, from travel to host gifts and payment, as well as lodging and food, were covered by the workshop.

**4.4 Topics and schedule** Since the NTT workshop focused heavily on the fieldwork practicum and time was limited – with just two days of classroom instruction as opposed to eight two-hour individual sessions – the LDTC workshop topics had to be adjusted. A pre-workshop online survey was used to assess the needs of the workshop participants. Figure 2 displays the responses to the multiple-choice question about what topics participants would like covered in the workshop.



**Figure 2.** Pre-workshop survey responses for “What topics would you like to see covered in this workshop?” Multiple selections were allowed, and not all who participated in the workshop responded to this survey ( $n = 22$ ).

With this input from the participants themselves, and taking from the experience LDTC’s model provided, workshop leaders considered which themes would be most useful for the group. They ultimately decided to concentrate on the technological aspects of language documentation rather than linguistic concepts, which were well aligned with community needs. There were several reasons for this decision. First, the goal was to train participants in modern language documentation techniques, and use of recording equipment and software programs are core to that concept. For many people, utilizing audio and video equipment can be intimidating, especially if they have never used such tools before. Allowing participants to have adequate time and the opportunity to “play” with the equipment and get comfortable with them prior to having to do so in the field, relieves some of the stress and possibly embarrassment experienced when recording community members and respected elders. Moreover, the creation of tangible materials is an aspect of language documentation that people care about. Recordings and annotations can be shared; they are long-lasting digital materials for speakers and linguists, and they can give prestige to a language. While linguistic analysis is important, it can be argued that training in the proper methods of recording and annotating minority languages is the first step to community collaboration. In-depth linguistic analysis can occur after recordings are made. We do note, however, that linguistic analysis is not completely separate from documentation. For instance, lessons on basic concepts of phonetics and phonology were included in the classroom instruction specifically to prepare participants to an-

notate recordings while in the field. If more time was available or another workshop possible, topics such as morphological and syntactic analysis could be included as a next step.

While the LDTC model relies on a one-hour lesson and then one hour of hands-on practice of the material, this was not possible in Kupang. With only two days of training, the modules were adjusted so that an hour for hands-on practice was only included for the technical portions: audio recording, video recording, and transcription. Hands-on training in file management, metadata, and archiving instead occurred throughout the lecture and was reinforced during the practicum. Each module built upon the previous lesson. Participants practiced ethical recording in their audio- and video-recording practice. The resulting recordings were used as files to practice metadata management and file-naming conventions. Then these properly managed files were also used for practice with ELAN transcription. Once in the field, participants could employ this workflow to efficiently navigate ethical expectations, use equipment to set up recording sessions, manage data, and transcribe and translate recordings with members of the language communities. The full workshop schedule – consisting of two days of classroom instruction, followed by three days of fieldwork, and a final two days in the classroom for data processing and debriefing – is shown in Table 2.

**Table 2.** Schedule for the 2018 Documenting Minority Languages of NTT Workshop

Date	Schedule
Mon, May 7	Introductions Ethics & community relations Audio recording + 1 hour of hands-on training Video recording + 1 hour of hands-on training
Tues, May 8	Metadata, archiving, & file management Phonetics & Phonology ELAN + 2 hours of hands-on training
Wed–Fri, May 9–11	Field Work Practicum: Data collection, transcription, & translation
Sat, May 12	Orthography (Guest lecture with Charles Grimes) Grant writing
Sun, May 13	Final project preparations Certificate ceremony

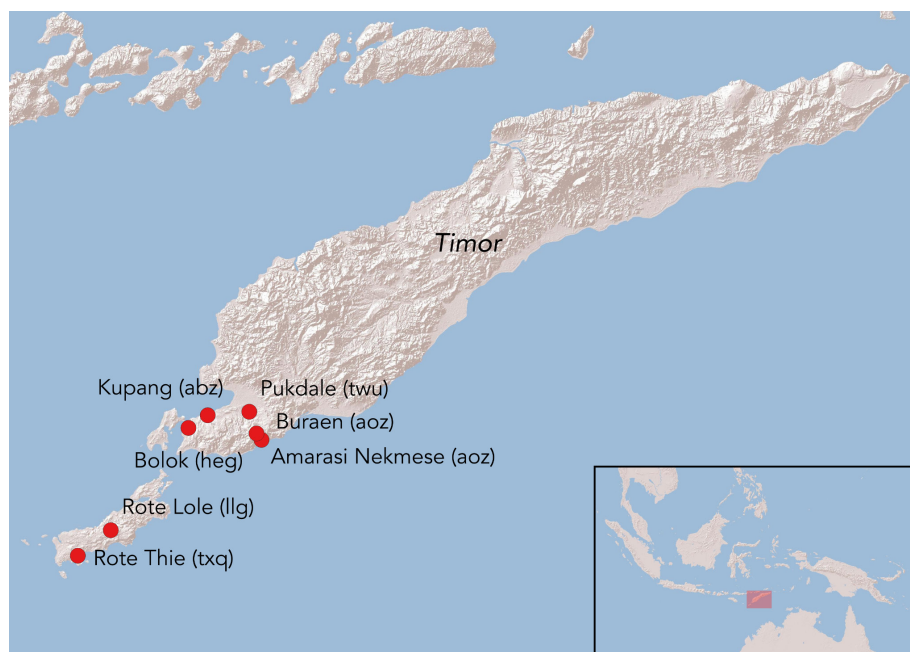
**4.5 Summary of the Documenting Minority Languages in NTT workshop** The Documenting Minority Languages in NTT workshop in Kupang took place over the course of one week from May 7–13, 2018. It was a collaborative effort across three universities: the University of Hawai‘i, Leiden University, and Artha Wacana Christian University. The professor from the local university organized everything that needed to be done at a local level, from the recruitment of many of the participants to ordering the catering for the week-long workshop. The university teams from abroad each included a professor who had conducted research in NTT for several decades, six linguistics graduate students from the University of Hawai‘i, and three from Leiden University. Some of the students had worked in Indonesia prior, while others had extensive experience with LDTC. The diverse experience of the incoming group of organizing participants was key in pre-planning and in executing the workshop.

Community participants consisted of thirty-four Indonesians and one participant from Singapore. Roughly 50% of the Indonesian community members were residents of Kupang at the time of the workshop (though only a handful of them were native to Kupang), and 50% resided on a neighboring island and had to travel to attend the workshop. Most participants were students or recent graduates and had no language documentation experience, but a few had experience up to the graduate level. Participants represented over twenty different Austronesian and Papuan languages.

In total, seven groups of workshop participants documented six regional minority languages. These included four Austronesian languages – Helong (ISO 639-3 *heg*), Termanu (*twu*), Rote Thie (*txq*), and Rote Lole (*llg*) – as well as two varieties of Uab Meto (*aoz*), as spoken in Buraen and Amarasi Nekomese. A seventh group worked with diaspora speakers of the Alor-Pantar language Abui (*abz*) in metropolitan Kupang.<sup>5</sup> The locations of these host communities are shown in Figure 3. Groups were tasked with collecting three types of data – wordlists, folk stories, and the Surrey Video Stimuli elicitation experiment (Fedden et al. 2013) – and using ELAN to transcribe the story. This amounted to approximately two days of data collection and one day of data transcription.

---

<sup>5</sup> In addition, Sar (glottocode *sarr1247*) speaker Henrik Lambolang traveled from the island of Pantar to join the workshop in Kupang, where he worked with Amor Sir and Marian Klamer to document his language. However, no workshop participants received fieldwork training using the Sar language.



**Figure 3.** The locations of the field practicum portion of the workshop. Groups worked with Helong (ISO 639-3 heg) at Bolok, Abui (abz) within the Kupang diaspora, Uab Meto at Buraen (aoz), Uab Meto at Amarasi Nekkese (aoz), Termanu at Pukdale (twu), Rote Thie (txq), and Rote Lole (llg).

The classroom time following the field practicum was not only a time to debrief and wrap things up but also to introduce participants to ways in which they could find funding for their language projects, which is essential knowledge for the success and sustainability of any community-based approach. As noted in §2.2, because insufficient funding for language documentation is available from the Indonesian government, participants must be made aware of the international funding available to them and the techniques to write a successful application. The final day provided time for preparing data to be archived in PARADISEC and finishing a professionally printed and bound book that presented a compilation of the data and transcriptions from the field practicum (Jacob et al. 2018).<sup>6</sup> Copies of the book were given to the speakers we worked with and workshop participants. The production of a tangible product in printed form was important both symbolically and practically for the local participants and the involved speaker communities, serving as a physical rep-

<sup>6</sup> Workshop data and recordings can be accessed at <https://catalog.paradisec.org.au/collections/NTT> (accessed 2020-10-18). PARADISEC, a founding member of the Digital Endangered and Languages and Music Archiving Network (DELAMAN), was chosen as an appropriate repository because of its long-standing commitment to endangered and minority language communities in the region.

resentation of the accomplishments of the workshop while at the same time bringing a level of prestige and awareness to both the language and the field of language documentation.

**5. Assessment of workshop outcomes** In this section, we assess the effectiveness of the workshop. We discuss the positive outcomes (§5.1) and the challenges (§5.2). When possible, we draw on responses from a Google Forms survey that was sent to all participants after the workshop's conclusion.

The survey contained twenty-three questions divided into four categories: demographics, pre- and post-workshop experiences, workshop satisfaction, and general feedback about the workshop. A majority of the questions were open-ended to allow participants to voice their opinions, but to measure satisfaction with each workshop module or activity, a three-point Likert scale was used.<sup>7</sup> All questions were asked in Malay, and 18 of the 35 participants (51%) responded to the survey. The survey provided an option for anonymous response, but all respondents chose to reveal their names. While the lack of anonymity may have biased the responses to the more evaluative questions in the survey, it is unlikely to have affected the arguably more useful questions about workshop content.

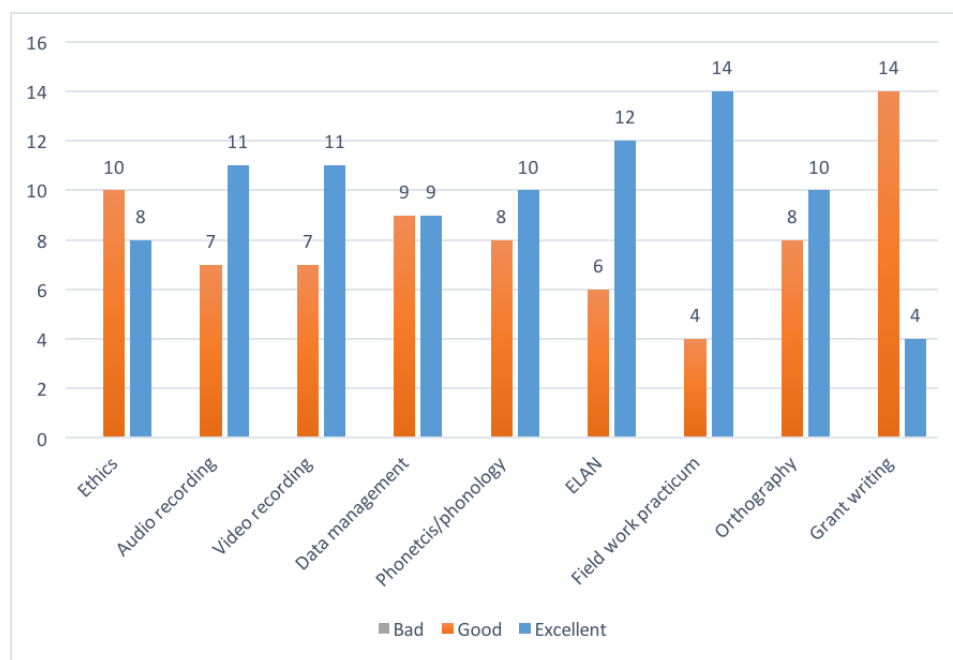
**5.1 Positive outcomes** Participants viewed the workshop as a valuable cultural exchange. For instance, one wrote that the workshop allowed them to “establish good relations with citizens at home and abroad and learn valuable lessons.” Because the participants were so diverse and workshop leaders were from different universities, participants and workshop leaders alike had access to different perspectives and approaches to theoretical and technical topics. Leaders taught participants in the classroom, whereas the participants taught the leaders during the field practicum. They guided leaders in understanding and following social norms, allowing leaders to gain a more nuanced understanding of language documentation techniques and their cross-cultural applications. As a result, the workshop served to train both speakers of endangered languages and graduate students, and this reciprocity encouraged a collaborative atmosphere that allowed for cultural, personal, and academic exchange.

The final product, which was a printed book of the work completed during the workshop, was unanimously lauded as a positive substantial result, though a few participants expressed that they did not receive copies themselves. Another open-ended survey question asked about their satisfaction with having the workshop presentations in both Malay and English, to which all respondents reported that the use of Malay was helpful and they enjoyed learning new vocabulary in English; however, they were most satisfied when presenters used their native language. Figure 4 shows the satisfaction levels for each component of the workshop. The answer options in Malay were *buruk*, *bagus*, and *luar biasa*, which can be translated as ‘bad’, ‘good’,

---

<sup>7</sup> A three-point Likert scale was used instead of a five-point Likert scale for simplicity, as it made the survey easier to complete. The survey was not intended for statistical analysis, so we did not aim for fine-grained responses.

and ‘excellent’, respectively.<sup>8</sup> No one rated a module as “bad,” so only “good” and “excellent” are reported in the figure.



**Figure 4.** Participant satisfaction levels for each module or activity ( $n = 18$ )

The practicum was the highlight of the workshop, with the majority of participants (76%) giving it the highest rating. Other activities that were most highly regarded were the ELAN training and the audio- and video-recording sessions. This feedback supports our decision to focus on the technical aspects of language documentation in order to prepare participants for the field, especially since the practicum was such a success.

As a direct result of the workshop, many respondents reported an increased awareness of the language endangerment issues in their region, and they had a better understanding of cultural and linguistic preservation, which, in turn, led to greater engagement of community members in language documentation projects throughout NTT. Post-workshop, 72% of respondents were engaged in language-related activities. Six participants continue to work as translators or transcribers. One is conducting this work with UBB, the local language activist program. Two others have served as participants in language documentation projects. Additionally, two participants

<sup>8</sup> These three options were the most suitable as they are commonly used together in Indonesia.



were continuing their language documentation training with other workshops or with SIL International. One participant, a teacher, took the initiative to incorporate local languages in the local content class at a public vocational high school in Alor. The local content class in Indonesia aims to educate and raise awareness of cultural heritage, including language varieties in the region. Therefore, given the many varieties spoken in Alor, the workshop participant was inspired to document stories, poems, and personal narratives in Abui, his native language, and use them as teaching materials to stimulate his students to speak their own mother tongues, which include varieties related to Abui, such as Mataru, Kafel, Luwenwi, and Kafoi.

A majority (61%) of respondents requested yearly training programs or were eager for future workshops to be held nearer to their homes, which would allow different linguistic groups to access training and extend training beyond a university context. This desire indicates that with increased awareness of language endangerment, speakers become more aware of the need for and value of training.

**5.2 Challenges** There were three main challenges to organizing and conducting this workshop: technology, time, and travel. Technology was a consistent challenge for both planning and conducting the workshop. Since LDTC workshops are always held on the University of Hawai'i campus, participants have easy access to the internet and are often computer literate. However, the Kupang participants had varying levels of computer skills and possessed laptops of varying degrees of capabilities. While the preprogrammed USBs did help to reduce software installation times, many people still faced difficulties with computer use, and teaching the software had unforeseen problems. For instance, the ELAN workshop took more than a half a day of instruction, much longer than anticipated. Even with a step-by-step tutorial, participants required extensive one-on-one aid as they learned to navigate the program. As such, we had to adapt our schedule and forego teaching SayMore, which had originally been scheduled for the end of the second workshop day. However, participants' satisfaction with the technological modules suggests that accepting and overcoming these challenges is worthwhile.

Another big challenge was the limited amount of time for the workshop and the ability to train each module fully. While the total number of contact hours at the NTT workshop was comparable to or exceeded that of the typical LDTC workshop series, compressing the workshop into just one week rather than eight allowed much less time for participants to develop and refine their skills. While workshop leaders were able to cover a wide variety of topics in just seven days, the short timeframe limited the amount of practice that participants received, and it required some important topics to be sacrificed (e.g., SayMore, basic morphology and syntax). A longer workshop would not only allow a broader array of topics to be addressed, but it would crucially give participants more time to become familiar with software and recording equipment with the help of an experienced mentor. Language documentation skills only improve through practice and trial and error, so a longer workshop would have led to more experienced and confident participants.

Based on feedback from participants, some modules may have deserved more time while others could have taken less. For instance, when participants were asked

about what they found to be the most interesting and most difficult topics, the top two responses for both were ELAN training and the phonetics and phonology session. In particular, many participants noted that the whole transcription process was challenging. This overlap possibly suggests that these two topics could have deserved additional training time. In contrast, the grant-writing session, though still rated “good,” was the least favorite module. This is also reflected in the fact that the workshop did not generate funding proposals; no participants sought grants to support their own work, even though people so often mention lack of funds as a barrier to language work. These results, coupled with the long and intense training that is usually required to prepare graduate students to search for and write successful grants, suggest that more intensive training and individual mentoring is required before participants are comfortable with seeking funding to conduct their own documentation projects. We do believe, however, that it is useful for participants to have a better understanding of the process and the kinds of funding available.

Lastly, because of its favorable outcomes, the practicum required considerable preparation and coordination. Organizing transportation and lodging for so many groups with such a tight schedule was not an easy task. Some groups that traveled farther away via ferry experienced transportation issues on their return due to weather, and they struggled to complete everything within the three allotted days because of increased time spent traveling. Those participants returned to Kupang by flight, which added an unexpected cost. The field practicum was also challenging for workshop leaders, who needed to serve as experts and mentors in unfamiliar language communities. Fortunately, though, speakers within the communities were overwhelmingly generous and accommodating, largely thanks to pre-established connections with the workshop organizers. Similarly, two participants from neighboring islands noted in the survey that the cost of traveling between their home islands and Kupang was a drawback to the workshop. While a stipend to travel via ferry was provided, participants’ travel situations were each unique, which may have led to budgetary challenges. Complications with budgeting and money are a common theme in language documentation and training projects since it is near-impossible to anticipate all issues that could arise. Therefore, though we strove to avoid such problems, we were pleased to hear that these logistical challenges were our only critiques. Ultimately, these logistical challenges served to introduce participants to some of the hurdles commonly encountered in language documentation fieldwork.

**6. Conclusion: Generalizing the model for other contexts** The achievements of the NTT workshop show that it is possible for a training effort of limited duration to be comprehensive and have a significant impact on capacity building. Adapting the LDTC model – a language documentation training workshop originally developed for an audience in urban United States – to succeed in a provincial Indonesian context, suggests that the LDTC workshop format is broadly applicable and potentially portable. We therefore encourage others to adapt this model. The following are some considerations that we believe should be taken into account when doing so.

- **Local partnerships:** It cannot be stressed enough that bringing this model into another context requires local support. This is critical not only for logistical planning but also for community collaboration and relationship building. It would have been nearly impossible to achieve what we did without having a local partner coordinating so many details and connecting us with speaker communities.
- **Modules and schedule:** At minimum, every workshop should incorporate ethics, phonology and orthography, audio recording, data management, and archiving, culminating in an experience in the field. With one full week to complete the training of these topics, inclusive of hands-on practice and a debriefing period, we do not recommend a shorter amount of time.
- **Language of instruction and in the field:** Aim to have training materials available in a common national/local language, and have some, if not all, workshop leaders be able to use that common language for instruction and for the field practicum.
- **Location and the infrastructure:** Carefully consider the location so that it is easily accessible by all participants. This may mean a larger urban locale, like our choice of Kupang, if the workshop is intended for speakers of diverse languages, or it could be a more localized place if the workshop is concentrating on a specific language or language group. But access to infrastructure, such as electricity, internet, roads, facilities, accommodations, and modes of transport, is crucial to the logistics of a workshop. Many areas of linguistic diversity may not have easy access to electricity; and the modules we suggest require electronics and recording equipment. Therefore, researchers looking to host a workshop in a location with little access to electricity should budget to bring a portable electricity source to charge equipment.
- **Computer literacy:** The LDTC model is based on participants being computer literate. As such, we did not need to factor in time to teach basic computer skills. However, in many parts of the world, this training may be necessary; participants must first learn to navigate a computer before using it to annotate and edit linguistic data. Alternatively, one could adjust our model from a software/technology-focused workshop and instead teach more traditional methods of language documentation using handwritten notes and transcriptions. This choice should be based on the prior knowledge and capabilities of workshop participants, and workshop leaders should aim to teach methods that will maximize participants' ability to apply the workshop content in their personal language documentation endeavors.
- **Group documentation:** LDTC uses an Expert–Mentor dynamic, while in Kupang a small group dynamic was utilized. In both cases, the documentation was done in small teams and fostered personalized hands-on assistance by workshop leaders. This also allowed participants learning new skills to rely on each other and not be afraid to make mistakes whether in the classroom or in the field. Furthermore, no single person held all the responsibility to fulfill the tasks during the field practicum.

- **Tangible finished product and archiving:** Plan and budget for some kind of product that can represent the work that was accomplished through training to document the language(s). A booklet, recipe book, poster, or vocabulary flashcards are some possibilities. Producing enough to share with all participants and language speakers is advisable. Additionally, have an archiving plan in place so that the materials of the workshop can seamlessly be deposited.

While local social, economic, and educational factors may limit the effectiveness of certain aspects of the workshop, we believe that the basic LDTC model is generalizable to contexts across the world. As a holistic approach to language documentation training, the LDTC model addresses issues of language endangerment, documentation, and ethics in the field. Workshop participants learn how to conduct every step of language documentation – including deciding what to record, setting up the recordings, transcribing and translating recordings, and managing and archiving data – so that they may independently continue the work in their own communities. Thus, we believe that every successful workshop should aim to be comprehensive, personalized, and hands-on. This is the essence of the LDTC model.

On a disciplinary level, this workshop is further evidence that collaborative efforts in language documentation have the potential to yield high returns and to avoid the pitfalls of the limitations associated with outside linguists solely conducting documentation work. In order for language documentation to be sustainable, language community members must be empowered as agents in the documentation process. Collaborative documentation engages community members, prepares native speakers to document their own or neighboring languages, has the potential to increase language prestige as a result of heightened interest in the language, and improves the cultural relevance and quality of the documentation, but effective collaboration requires access to training to ensure that all participants have the skills necessary to engage as active partners. The success of the Documenting Minority Languages of NTT workshop demonstrates the potential for the LDTC model to serve as a basis for developing community-based language documentation training. It is our sincere hope that other communities, researchers, and language advocates will continue to adapt the LDTC model to local contexts and, in doing so, increase the capacity for community-based language documentation across the world.

## References


- Ajo, Frances, Valérie Guérin, Ryoko Hattori, & Laura C. Robinson. 2010. Native speakers as documenters: A student initiative at the University of Hawai'i at Mānoa. In Grenoble, Lenore A. & N. Louanna Furbee (eds.), *Language documentation: Practice and values*, 275–286. Amsterdam: John Benjamins Publishing. doi:10.1075/z.158.26ajo
- Albright, Eric & John Hatton. 2007. WeSay: A tool for engaging native speakers in dictionary building. In Rau, D. Victoria & Margaret Florey (eds.), *Documenting and revitalizing Austronesian languages* (Language Documentation and Conservation Special Publication 1), 189–201. Honolulu: University of Hawai'i Press. (<http://hdl.handle.net/10125/1391>)
- Arka, I Wayan. 2002. Voice systems in the Austronesian languages of Nusantara: Typology, symmetry and Undergoer orientation. *Linguistik Indonesia* 21(1). 113–139.
- Arka, I Wayan. 2018. Reflections on the diversity of participation in language documentation. In McDonnell, Bradley, Andrea L. Berez-Kroeker, & Gary Holton (eds.), *Reflections on language documentation 20 years after Himmelmann 1998* (Language Documentation and Conservation Special Publication 15), 132–139. Honolulu: University of Hawai'i Press. (<http://hdl.handle.net/10125/24815>)
- Badan Pusat Statistik. 2021. *Statistik Indonesia: Statistical yearbook of Indonesia 2021*. (Publication No. 03200.2103). Jakarta: Badan Pusat Statistik Indonesia. (<https://www.bps.go.id/publication/2021/02/26/938316574c78772f27e9b477/statistik-indonesia-2021.html>) (Accessed 2021-11-30.)
- Blust, Robert. 1993. Central and Central-Eastern Malayo-Polynesian. *Oceanic Linguistics* 32(2). 241–293. doi:10.2307/3623195
- Butler, Katie. 2011. Community-based website building: The Language Documentation Training Center's approach to mentor-mentee partnership. (Paper presented at Sustainable data from digital research: Humanities perspectives on digital scholarship, Melbourne, 12–14 December). (<http://hdl.handle.net/2123/8112>)
- Collins, Sean, Xisto Martins, Andrew Mitchell, Awegechew Teshome, & John Arnaason. 2006. Quantitative ethnobotany of two east Timorese cultures. *Economic Botany* 60(4). 347–361.
- Crippen, James A. & Laura C. Robinson. 2013. In defense of the lone wolf: Collaboration in language documentation. *Language Documentation & Conservation* 7. 123–135. (<http://hdl.handle.net/10125/4577>)
- Czaykowska-Higgins, Ewa. 2009. Research models, community engagement, and linguistic fieldwork: Reflections on working within Canadian Indigenous communities. *Language Documentation & Conservation* 3(1). 15–50. (<http://hdl.handle.net/10125/4423>)
- Donohue, Mark & Charles E. Grimes. 2008. Yet more on the position of the languages of Eastern Indonesia and East Timor. *Oceanic Linguistics* 47(1). 114–158.

- Edwards, Owen & UBB. 2018. Languages of Nusa Tenggara Timur and Timor-Leste (map). Kupang: Language & Culture Unit (UBB) and Leiden University. (<https://vici.marianklamer.org/linguistic-maps.html>) (Accessed 2020-09-22.)
- Ewing, Michael C. 2014. Language endangerment in Indonesia. *International Journal of Education* 8(1). 12–24.
- Fedden, Sebastian, Dunstan Brown, Greville Corbett, Gary Holton, Marian Klamer, Laura C. Robinson, & Antoinette Schapper. 2013. Conditions on pronominal marking in the Alor-Pantar languages. *Linguistics* 51(1). 33–74. doi:10.1515/ling-2013-0002
- Fitzgerald, Colleen M. 2018. Reflections on language community training. In Bradley McDonnell, Andrea L. Berez-Kroeker, & Gary Holton (eds.), *Reflections on language documentation 20 years after Himmelmann 1998* (Language Documentation & Conservation Special Publication 15), 86–99. Honolulu: University of Hawai'i Press. (<http://hdl.handle.net/10125/24811>)
- Fitzgerald, Colleen M. & Mary S. Linn. 2013. Training communities, training graduate students: The 2012 Oklahoma Breath of Life Workshop. *Language Documentation & Conservation* 7. 185–206. (<http://hdl.handle.net/10125/4596>)
- Grenoble, Lenore A. 2010. Language documentation and field linguistics: The state of the field. In Grenoble, Lenore A. & N. Louanna Furbee (eds.), *Language documentation: Practice and values*, 289–310. Amsterdam: John Benjamins Publishing. doi:10.1075/z.158.28gre
- Grimes, Charles E., Tom Therik, Barbara Dix Grimes, & Max Jacob. 1997. *A guide to the people and languages of Nusa Tenggara*. Kupang, Indonesia: Universitas Kristen Artha Wacana and Alfa Omega Foundation.
- Holton, Gary. 2009. Relatively ethical: A comparison of linguistic research paradigms in Alaska and Indonesia. *Language Documentation & Conservation* 3(2). 161–175. (<http://hdl.handle.net/10125/4424>)
- Holton, Gary & Marian Klamer. 2017. The Papuan languages of East Nusantara and the Bird's Head. In Palmer, Bill (ed.), *The languages and linguistics of the New Guinea area: A comprehensive guide*, 569–640. Berlin: De Gruyter Mouton. (<https://hdl.handle.net/1887/3463908>)
- Jacob, June & Barbara Dix Grimes. 2006. Developing a role for Kupang Malay: The contemporary politics of an eastern Indonesian creole. (Paper presented at the 10th International Conference on Austronesian Languages, Puerto Princesa, Philippines, 17–20 January.) (<https://sil-philippines-languages.org/ical/papers.html>) (Accessed 2022-10-06.)
- Jacob, June, Marian Klamer, & Gary Holton (eds.). 2018. Workshop on documenting minority languages in Nusa Tenggara Timur Indonesia. Kupang: Universitas Kristen Artha Wacana. (<http://hdl.handle.net/10125/58880>)
- Klamer, Marian & Michael C. Ewing. 2010. The languages of East Nusantara: An introduction. In Ewing, Michael C. & Marian Klamer (eds.), *Typological and areal analyses: Contributions from East Nusantara*, 1–24. Canberra: Pacific Linguistics.

- Leonard, Wesley Y. 2017. Producing language reclamation by decolonising ‘language’. *Language Documentation and Description* 14. 15–36.
- Leonard, Wesley Y. & Erin Haynes. 2010. Making “collaboration” collaborative: An examination of perspectives that frame linguistic field research. *Language Documentation & Conservation* 5. 268–293. (<http://hdl.handle.net/10125/4482>)
- Mona, Stefano, Katharina E. Grunz, Silke Brauer, Brigitte Pakendorf, Loredana Castri, Herawati Sudoyo, Sangkot Marzuki, Robert H. Barnes, Jörg Schmidtke, Mark Stoneking, & Manfred Kayser. 2009. Genetic admixture history of Eastern Indonesia as revealed by y-chromosome and mitochondrial DNA analysis. *Molecular Biology and Evolution* 26(8). 1865–1877. doi:10.1093/molbev/msp097
- Moro, Francesca. 2019. Loss of morphology in Alorese (Austronesian): Simplification in adult language contact. *Journal of Language Contact* 12(2). 378–403. doi:10.1163/19552629-01202005
- O’Connor, Sue, Rintaro Ono, & Chris Clarkson. 2011. Pelagic fishing at 42,000 years before the present and the maritime skills of modern humans. *Science* 334(6059). 1117–1121.
- Olko, Justyna. 2018. Spaces for participatory research, decolonization and community empowerment: Working with speakers of Nahuatl in Mexico. *Language Documentation and Description* 16. 1–34.
- Paauw, Scott H. 2008. *The Malay contact varieties of Eastern Indonesia: A typological comparison*. Buffalo: State University of New York. (Doctoral dissertation.)
- Rehg, Kenneth L. 2007. The Language Documentation and Conservation Initiative at the University of Hawai’i at Mānoa. In Rau, D. Victoria & Margaret Florey (eds.), *Documenting and revitalizing Austronesian languages* (Language Documentation & Conservation Special Publication 1), 13–24. Honolulu: University of Hawai’i Press.
- Roche, Gerald, Ban+de mkhar, Bkra shis bzang po, G.yu lha, Snying dkar skyid, Tshe ring rnam gyal, Zla ba sgrol ma, & Charles Kevin Stuart. 2010. Participatory culture documentation on the Tibetan Plateau. *Language Documentation and Description* 8. 140–158.
- Saad, George. 2016. Coming of age: Diagnosing contact-induced change in Abui-Malay bilinguals. (Paper presented at the 8th Austronesian and Papuan Languages and Linguistics Conference, London, 13–14 May.)
- Saad, George. 2020. *Variation and change in Abui: The impact of Alor Malay on an indigenous language of Indonesia*. Leiden: Leiden University. (Doctoral dissertation.)
- Saad, George, Marian Klamer, & Francesca Moro. 2019. Identifying agents of change: Simplification of possessive marking in Abui-Malay bilinguals. *Glossa: A Journal of General Linguistics* 4(1). 1–29. doi:10.5334/gjgl.846
- Si, Aung. 2011. Biology in language documentation. *Language Documentation & Conservation* 5. 169–186. (<http://hdl.handle.net/10125/4497>)

- SIL International. 2021. Fieldworks Language Explorer (FLEX, version 9.0.17) [Computer software]. Dallas: SIL International. (<https://software.sil.org/fieldworks/download/>) (Accessed 2021-12-23.)
- Sloetjes, Han & Peter Wittenburg. 2008. Annotation by category: ELAN and ISO DCR. In *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008)*, 816–820.
- Thomason, Sarah G. 2001. *Language contact*. Edinburgh: Edinburgh University Press.


Jennifer Sou  
jsou@hawaii.edu

 [orcid.org/0000-0002-0972-2590](https://orcid.org/0000-0002-0972-2590)


Leah Pappas

 [orcid.org/0000-0002-3329-7686](https://orcid.org/0000-0002-3329-7686)

Khairunnisa

 [orcid.org/0000-0002-5676-2392](https://orcid.org/0000-0002-5676-2392)

Gary Holton

 [orcid.org/0000-0002-9346-1572](https://orcid.org/0000-0002-9346-1572)