

Efforts in language documentation in a linguistically diverse country: Building up collaborations of various stakeholders

Yanti (Atma Jaya Catholic University of Indonesia)
Asako Shiohara (ILCAA, Tokyo University of Foreign Studies)

Presented at NINJAL International Symposium
Approaches to Endangered Languages in Japan and Northeast Asia: Description, Documentation and Revitalization
Tachikawa, Tokyo, August 6 – 8, 2018

Introduction

- Indonesia is well-known for its linguistic and cultural diversity.
- Many languages remain under-described and are also seeing drastic reductions in speaker numbers as communities shift to the national language, Indonesian, or more vigorous regional languages (Jukes, Shiohara, and Yanti 2017).
- Documenting these languages before they disappear are necessary.
- In this presentation, we share our efforts in building up collaborations for documenting endangered indigenous languages of Indonesia, especially those spoken the Nusa Tenggara Timur Province (hereafter NTT).

Structure of this talk

1. Sociolinguistic situation in Indonesia and NTT
2. Early efforts for language documentation in Indonesia
 - Phase 1: Establishing network with local institutes
 - Phase 2: Working with native speaker students in NTT
3. Current efforts for language documentation in NTT
 - Phase 3: Setting a more concrete goal and collaborating with more parties.
4. Conclusion and future plans

Language diversity in Indonesia

- Indonesia:
 - Located in South East Asia
 - Ethnologue: Indonesia contains more than 700 languages – nearly 10% of the world's total (Simons, Gary F. and Charles D. Fennig 2018)
 - Linguistically the world's second most diverse country after Papua New Guinea (Arka, 2013)
 - National language: (Standard) Indonesian

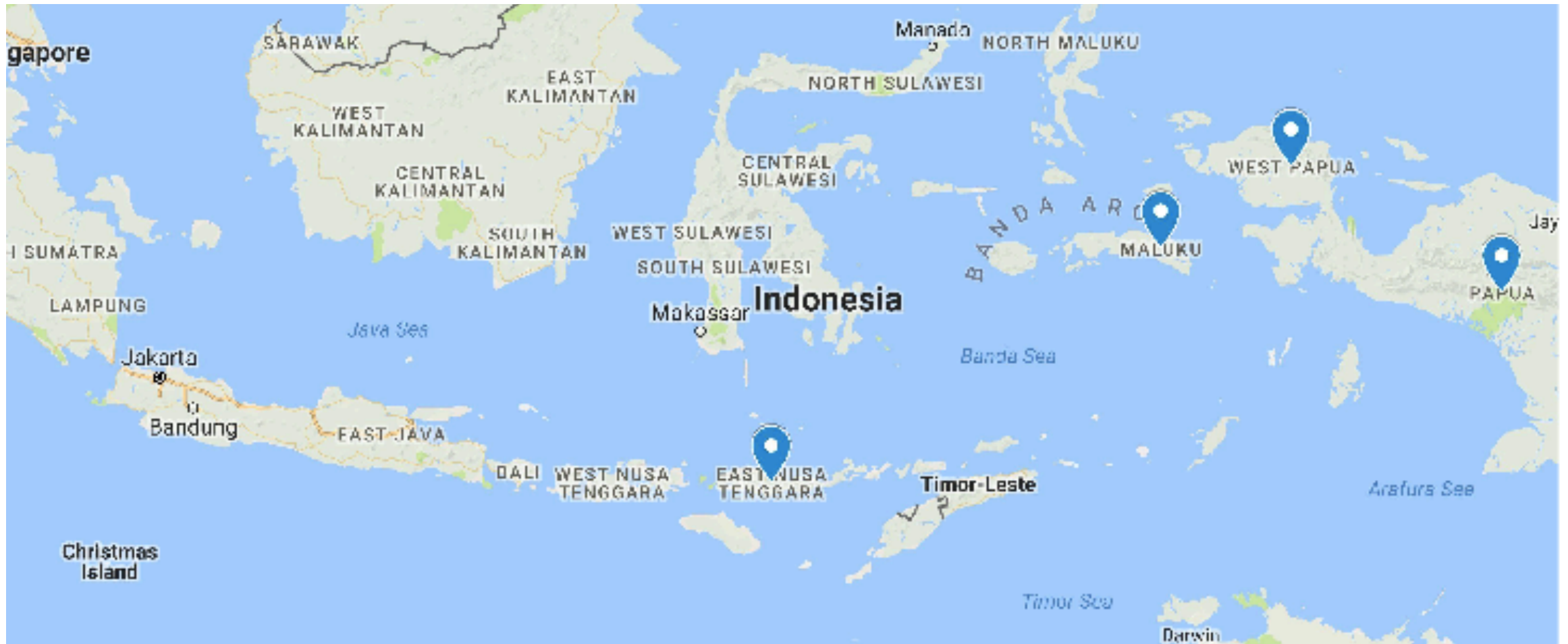
Map of Indonesia



Language diversity in Indonesia

- Languages spoken in Indonesia belong to both Austronesian languages (+/- 64%) and Non-Austronesian languages (+/-36%) (Arka, 2013).
- Most of these languages are found in the eastern part of Indonesia, especially in the provinces of Papua and West Papua in the island of New Guinea, as well as in the provinces of Maluku and East Nusa Tenggara, including many small islands in eastern Indonesia.

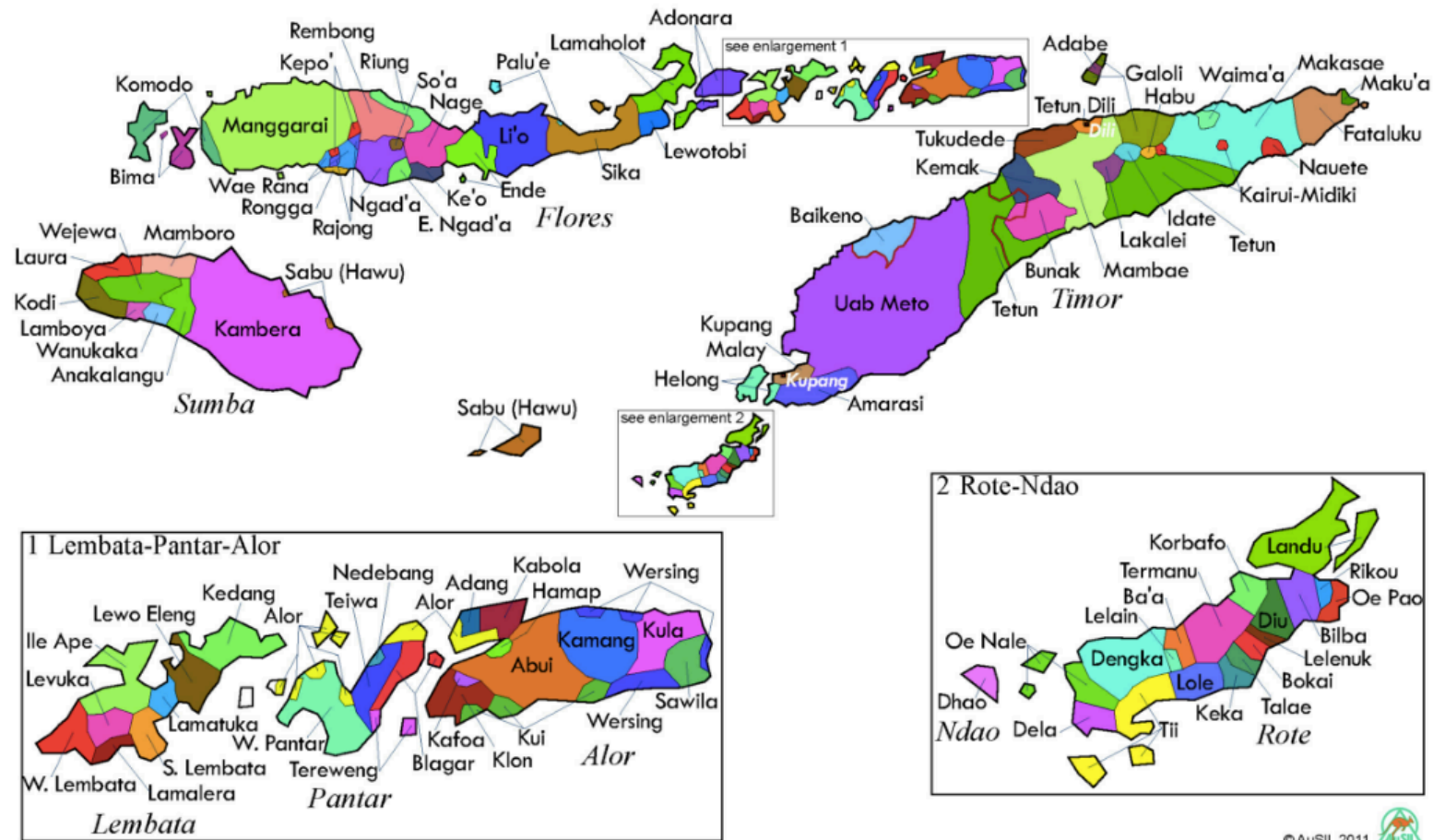
Locations where high linguistic diversity is observed



Languages of Nusa Tenggara Timur Province

- Approximately 70 languages are spoken in the Nusa Tenggara Timur Province, that corresponds to 10% of the total languages spoken in Indonesia.
- Languages spoken in this province belong to two language families: Austronesian languages and non-Austronesian languages (also called Papuan languages) which are mainly spoken in Alor Island and Pantar Island.

Map of Languages of Nusa Tenggara Timur (source: <http://e-ubb.org/peta-bahasa-ntt/>)



Languages in Nusa Tenggara Islands

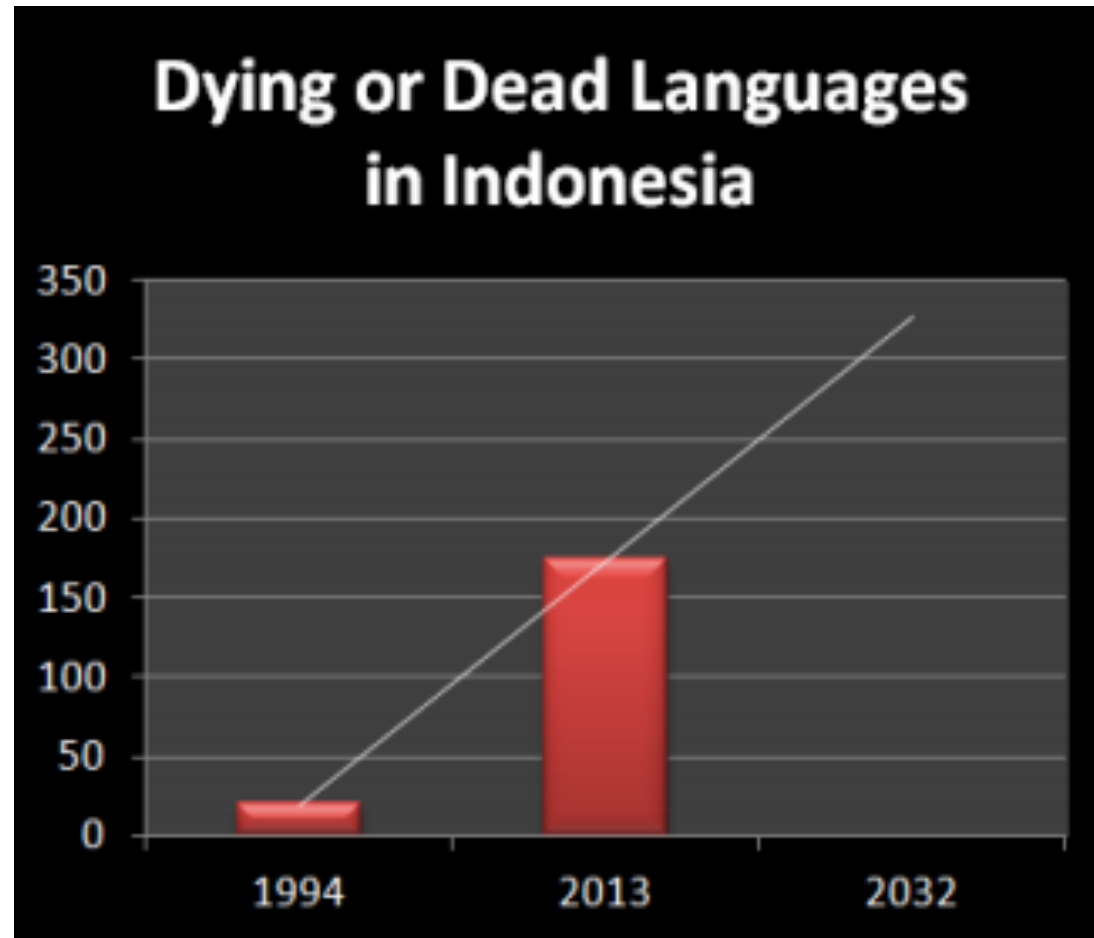
- The languages in this area, as a whole, are still understudied.
- Recent studies, however, have revealed many interesting features of these languages, such as numeral systems, metathesis, head-marking morpho-syntax, and SVCs (e.g. Arka, 2016, Edwards, 2016).
- These features attract many linguists to work on them in recent years.
- We have been working on two languages in this area:
 - Helong, an Austronesian language spoken in the eastern part of Timor Island (an on-going collaborative project with Yanti)
Helong Words Website:
<http://helong-bolok.aa-ken.jp/helong-bolok.html>
 - Kui, a non-Austronesian language spoken in Alor Island. (Windschuttel and Shiohara 2017)

Recent status of minority languages in Indonesia

- Many of the languages - even some with seemingly quite large numbers of speakers - have become endangered.
- Anderbeck (2015:21): “We are witnessing an definite weakening of the local languages in Indonesia.”

Anderbeck 2015

Trend line for dying or dead languages in Indonesia, p. 33



Recent status of minority languages in Indonesia

- These languages are endangered as a result of language shift, generally to the national language, Indonesian, or to other regional varieties of Malay, such as Papuan Malay, Kupang Malay or Manado Malay. (e.g. Musgrave 2014)

Documentary linguistics: why and how

- Language endangerment is so rapid in Indonesia, especially in its eastern part in which the size of each language community is small.
- What language to use to the next generation is the decision made by the community. We, as outsiders, cannot do anything when parents decide not to pass on their mother tongue to their children.
- Language shift often occurs as an inevitable and necessary consequence of the migration or intermarriage, without a clear and conscious decision made by the individual or community concerned.

Documentary linguistics: why and how

- However, we realize that the loss of a language means the loss of the culture and the local wisdom as well as the traditional knowledge, and it is important to document the languages before they are gone forever.
- It is in this context that we started working on our collaboration for documenting indigenous languages in Indonesia, which is supported by Linguistic Dynamic Science (LingDy) Project.

Efforts for language documentation in Indonesia

- ILCAA started Linguistic Dynamic Science Project in 2007; funded by the Japanese government (P.I. Prof. Toshihide Nakayama).
- The project aimed at building mutually beneficial relationships with minority language communities and local research institutes through active collaboration with them.
- One of the project activities is to host language documentation workshops in Indonesia, which started in 2013.

Previous workshops in Indonesia

- In 2006 and 2007 Nikolaus Himmelmann and Margaret Florey: workshops on language documentation in Ubud, Bali.
- The Center for Endangered Languages Documentation (CELD) at Universitas Negeri Papua (UNIPA) in Manokwari has organized occasional training sessions for local team members since 2009.
- The Indonesian Academy of Sciences (LIPI): in-house training from 2012-2013 for their documentation projects, assisted by Dr. John Bowden (the Jakarta Field Station of the MPA EVA)

Efforts for language documentation in Indonesia

- The efforts can be divided into three phases:
 - Phase 1: Establishing network with local institutes
 - Phase 2: Working with native speaker students in the NTT province
 - Phase 3: Setting a more concrete goal and collaborating with more parties.

Phase 1: Establishing networks with local institutes

- A total of 6 workshops were carried out in Indonesia:
 - 2013
 - Udayana University, Bali, Indonesia: 5-6 August 2013
 - 2014:
 - Udayana University, Bali, Indonesia: 15-16 August 2014
 - Mulawarman University, Samarinda, Indonesia: 19-21 August 2014
 - 2015
 - Seloko Institute, Politeknik Jambi, Indonesia: 15-16 June 2015
 - Manado State University, Manado, Indonesia: 6-7 August 2015
 - Balai Bahasa Denpasar, Bali, Indonesia: 15-16 August 2015

Map of workshop locations



Phase 1: Establishing networks with local institutes

- We worked with local institutes to host workshops.
- The primary aims were to introduce the idea of language documentation and to give a basic training for documenting languages.
- The local institutes selected participants (ideally young researchers and students) to take part in the workshops.
- Each workshop lasted for two days.
- Each workshop was taught by international linguists and one or two local linguists.

Phase 1: Establishing networks with local institutes

- Workshop program: two main parts
 - Introductory lectures on theories and methods of language documentation
 - Practical sessions for making recordings and annotating using ELAN.

Picture of participants and instructors from the 2013 workshop



Phase 1: Establishing networks with local institutes

- Technical aspects
 - We introduced high quality equipment (digital recorders, video cameras, microphones) and headphones from Tokyo, especially in the first few workshops
 - In early stage of this phase, the lectures, slides, and trainings were mostly conducted in English.

High quality equipment



Phase 1: Establishing networks with local institutes

- Evaluation of phase 1:
 - More time was allocated to lectures than to practice
 - During practical sessions, especially ELAN session, issues such as unfamiliarity with computers and data organizing hindered the practice.
 - Providing lecturers, instructions and slides in English were less effective than doing that in Indonesian; we shifted the language of the workshops from English to Indonesian in the later workshops.

Phase 1: Establishing networks with local institutes

Evaluation of phase 1:

- Although the participants of the workshops, as linguists, were interested in the general sociolinguistic situation in Indonesia, almost none of them were “stakeholders.”
- In big cities where we conducted the workshops, students and lecturers belonged to larger language communities such as various Malay varieties (e.g., Manado Malay and Jambi Malay) or Balinese, one of the dominant indigenous languages, and therefore, did not belong to “a party concerned” about the loss of languages and cultures.
- For that reason, the workshops did not lead to collaboration for substantial work of language documentation.

Phase 2: Working with native speaker students in the NTT province

- Like in previous workshops, we worked with a local institute to host the workshop
- The workshop was carried out at Nusa Cendana University in Kupang, Indonesia on August 11-12, 2015, preceded by a half-day seminar on August 10, 2015.
 - General lectures for a wider audience were given in the seminar, and then a restricted number of participants attended the workshop, where they could focus on intensive documenting work.
 - In the workshop, allocating more time for the participants to work on their projects for concrete outcomes, including the session on data management and practical session for annotation using ELAN software.
 - Simpler equipment was used: Zoom H-1.

Map of Kupang







Phase 2: Working with native speaker students in the NTT province

Evaluation of phase 2:

- As the capital of the NTT province, which is one of the most linguistically divergent provinces in Indonesia, almost all the participants of the workshop are speakers of indigenous languages, that is, people concerned in the issue.
- They have a strong motivation to document their language and leave the data for the next generations. They seriously participated in making recordings and data annotating during the practical session.

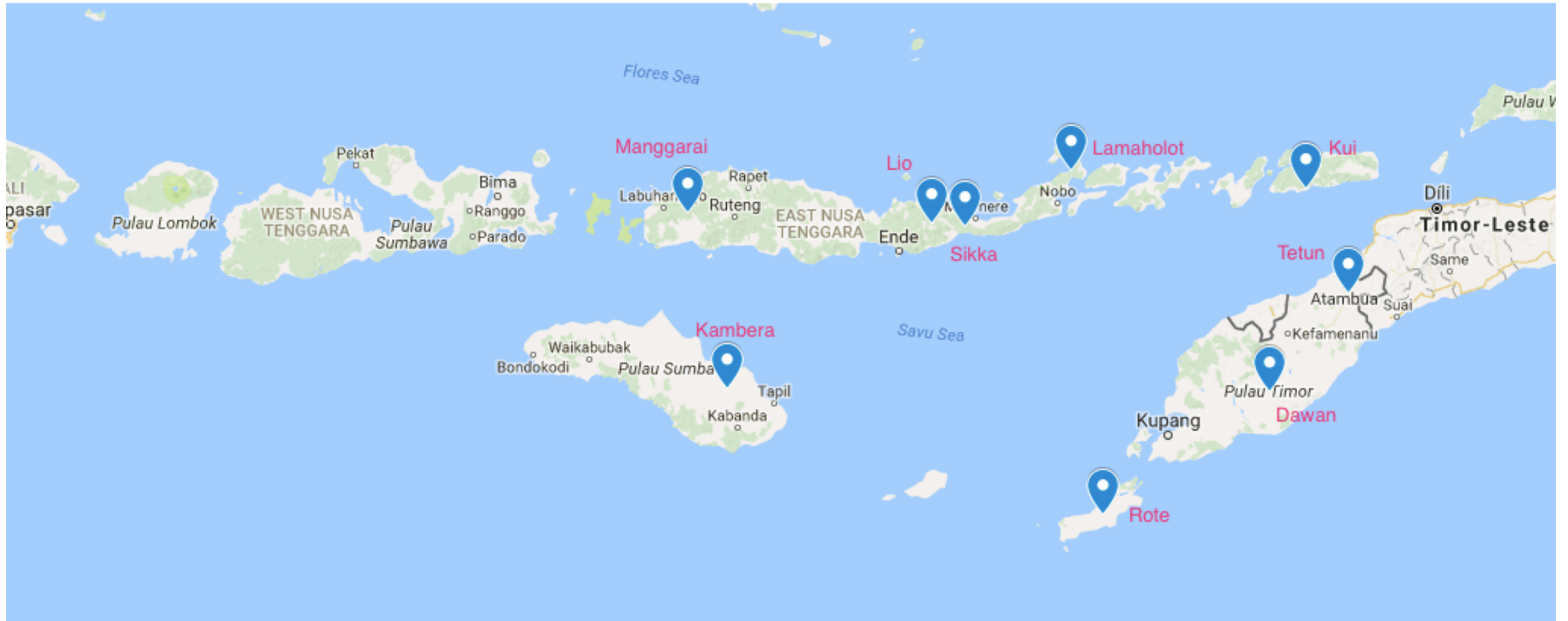
Phase 2: Working with native speaker students in the NTT province

- The recordings with transcription and translation have been archived in an international language archive: PARADISEC.
- The languages are Dawan, Kambera, Kui, Lamaholot, Lio, Manggarai, Rote, Sika, Tetun (Thanks to the workshop participants and Nick Thieberger)

<http://catalog.paradisec.org.au/collections/NTT2015>

- Following this success, we decided that we would focus our language documentation efforts in the Nusa Tenggara Timur Province.

Map of the locations of the languages





PARADISEC Catalog

Asako Shiohara | [Sign out](#)

- [Home](#)
- [Dashboard](#)
- [Collections](#)
- [Items](#)
- [Contact](#)

[Return To Results](#)

[Edit collection](#)

Collection details

Ready for metadata export Private: Hide metadata from all users

Collection ID	NTT2015
Title	Recordings of various texts in several languages spoken in Nusa Tenggara Timur
Description	results of language documentation project hosted in 2016 at Nusa Cendara University
Archive link	http://catalog.paradisec.org.au/repository/NTT2015
Collector	Asako Shiohara Find similar
Operator	Nick Ward
Originating university	Tokyo University of Foreign Studies
Countries	Indonesia - ID <i>To view related information on a country, click its name</i>
Languages	Uab Meto - aoz Kuii - kvrl

Items in Collection (9)

[Add new item](#)

Item ▲▼	Title ▲▼	Digitised	Files	Actions
dawan20150811 open	a love song		4	View Edit
kambara20150811 open	a folkstory		4	View Edit
kui20150811 open	a description of a village		4	View Edit
lamaholot20150812 open	an advice		4	View Edit
lio20150811 open	a folkstory		4	View Edit
manggarai20150812 open	a folkstory		4	View Edit
rote20150811 open	a folkstory		4	View Edit
sika20150811 open	a folkstory		4	View Edit
tetun20150811 open	a song		4	View Edit

- [Show 10](#)
- [Show 50](#)
- [Show 100](#)
- [Show all 9](#)

Depositing information

Phase 3: Setting a more concrete goal and collaborating with more parties

- In phases 1 and 2, we were successful in building networks with local institutes and researchers as well as native speaker students and in providing language documentation trainings.
- Having worked with native speaker students in Kupang, we decided to go to the next step of the documentation by setting a more concrete goal.
- The new goal is to collect substantial amount of data with native speakers through our programs.

Phase 3: Setting a more concrete goal and collaborating with more parties

- The linguistic data we obtained from language documentation would not be sufficiently valuable and useful if they are not annotated based on appropriate knowledge of the language and general linguistics
 - the language documentation should be conducted side by side with language description in order to utilize the data in a better way.
- Local institutes are generally limited in terms of financial and human resources.
- Native speaker students are great resources, but they have not had adequate linguistics knowledge.
- Two-day workshops are not adequate to trigger substantial language documentation works.
- Thus, to achieve our more concrete goal, we began collaborating with with more parties.

Phase 3: Setting a more concrete goal and collaborating with more parties

- We started to collaborate with Prof. Peter Cole and Gabriella Hermon of the University of Delaware
- “EAGER: Collaborative Approach to the Documentation of Endangered Languages in Linguistically Diverse Locales” program
- A two-year program funded by the National Science Foundation (Award # 1747801)
- In this program:
 - Teams of one American graduate student interested in fieldwork and documentary linguistics and two Indonesian students interested in documenting their own language
 - Intensive training course in linguistic fieldwork and language documentation for one week in Kupang
 - Following the training, they spend a few weeks in the field practicing what they have learned by collecting and annotating data on the Indonesian team members’ language.
 - Data from the training and practice will be archived in PARADISEC.

Phase 3: Setting a more concrete goal and collaborating with more parties

- This training program is expected to benefit several stakeholders:
 - Local students who want to document their language as well as develop their academic career through language documentation
 - American graduate students who want to gain experience documenting under-described languages through fieldwork
 - Language communities and linguists who wish to utilize the linguistic data for education, research, and so on.

Phase 3: Setting a more concrete goal and collaborating with more parties

- Assisted by the American Institute for Indonesian Studies (AIFIS):
- <http://sites.udel.edu/fieldwork-training-indonesia/>

Phase 3: Setting a more concrete goal and collaborating with more parties

- The first year program has (just almost) finished
- Participants:
 - 4 American graduate students
 - 8 Indonesian undergraduate students and fresh graduates
- Instructors:
 - Peter Cole (University of Delaware)
 - Gabriella Hermon (University of Delaware)
 - Yanti (Atma Jaya Catholic University of Indonesia)
 - Asako Shiohara (ILCAA, Tokyo University of Foreign Studies)
 - Jermy Balukh (STIBA Cakrawala Nusantara, Kupang)
- Invited speaker:
 - June Jacob (Arta Wacana Christian University)

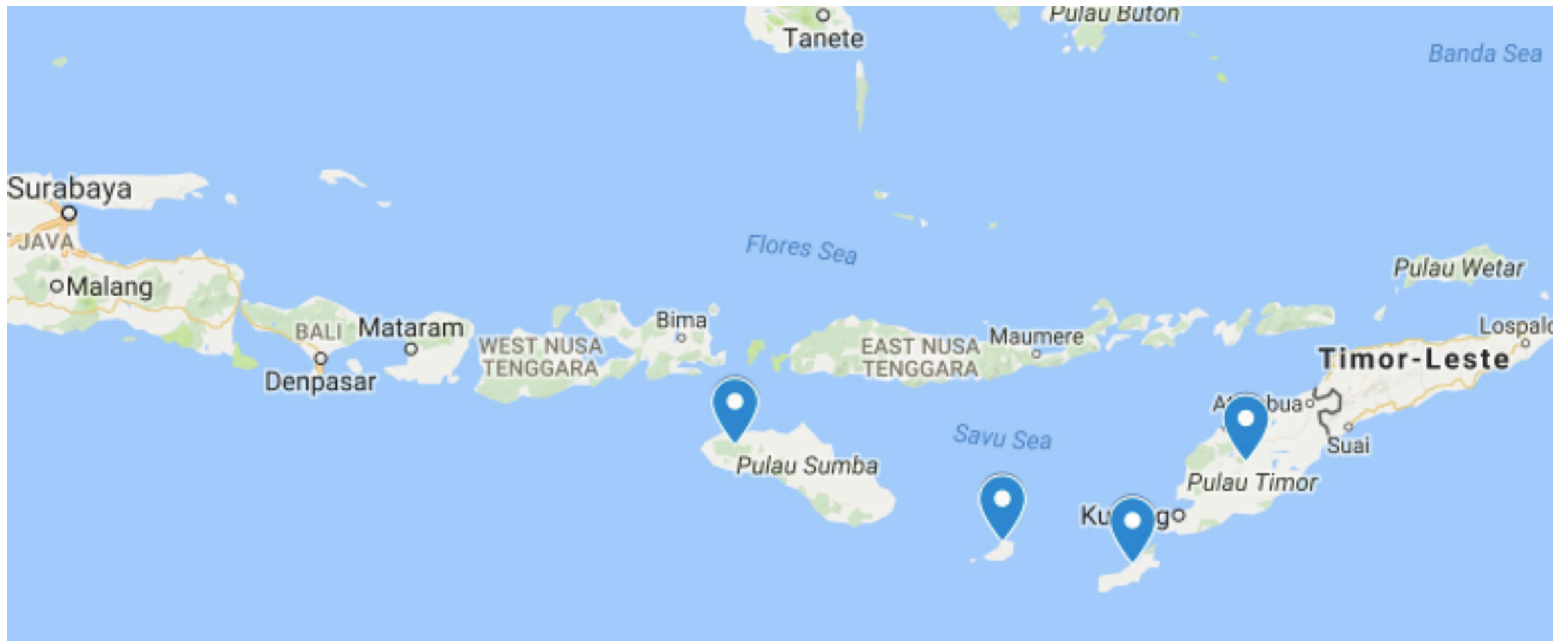
Phase 3: Setting a more concrete goal and collaborating with more parties

- Program:
 - 5 day training in Kupang (July 2 – 6, 2018)
 - 3 weeks fieldwork practice in the field (July 7- July 27, 2018)
 - 1 week data consolidation in Kupang (July 29 – August 3, 2018)

Phase 3: Setting a more concrete goal and collaborating with more parties

- Languages and locations:
 - Mollo dialect of Uab Meto in Bijaepunu, Molo Utara, Timor Island
 - Termanu dialect of Rotenese in Nggodimeda, Rote Island
 - Seba dialect of Hawu in Seba, Sabu Island
 - Wejewa language in Weekokora, Sumba Island

Map: Languages in the first year program



First year program

- Each group succeeded to make +/- 8 hours of recordings of various texts.
- Some of the recordings were transcribed, glossed and translated.
- The participants both the Americans and Indonesians were generally satisfied with the program and got really interested to continue to work.
- More detailed evaluation of the first year is not done yet.

Conclusion

- We have been collaborating with local institutes for documenting endangered/ minority languages in Indonesia since 2013.
- We have adjusted the program according to the situation we encountered in locations we visited.
- Initially, our activities focused on hosting seminars and workshops to share the value of indigenous minority languages and the significance of documenting them, and establishing a network of researchers interested in language documentation

Conclusion

- Then, we started to do substantial work of documentation by collaborating with native speakers.
- Building collaborations with native speakers and local academic communities is crucial for the documentary work.
- Language documentation should be conducted side by side with language description in order to utilize the data in a better way. For that purpose we also need to invite linguists from outside of the community.
- The attempts of “EAGER” program that invites the US graduate students is attaining certain results.

Future plans

- Potential collaborators would be researchers who are interested in
 - having experience of fieldwork
 - working on understudied languages
 - Phonological or grammatical features characteristic of the languages in this area; (e.g., implosives, metathesis, head-marking morphosyntax, SVCs)

Future plans

- Native speaker students who are motivated to document their own language are invaluable resources.
- In collaborating with them, however, we always need to seek ways to reward for their work; helping the students to develop their academic career through language documentation may be one of the ways.
- Some of local universities encourage their students to write their thesis about their languages and cultures. We are trying to collaborate with the universities to support their attempts.

Acknowledgements

- We thank our collaborators: Peter Cole and Gabriella Hermon and Jermy Balukh
- We thank our colleagues who have involved in one or more workshops: John Bowden, Antonia Soriente, Atsuko Utsumi, Tim McKinnon, Kazuya Inagaki, Sri Satyawati, June Jacob, Nazaruddin, Anthony Jukes, Antoinette Schapper.
- We thank our local collaborators and local institutions for their generous help in hosting the workshops and coordinating the participants.

References

- Anderbeck, Karl. 2015. Portraits of language vitality in the languages of Indonesia. In I Wayan Arka, Ni Luh Nyoman Seri Malini, Ida yu Made Puspani (eds.), *Language documentation and cultural practices in the Austronesian world*, Papers from ICAL 12 (A-PL.019) 4:19-7. Canberra: Pacific Linguistic.
- Arka, I Wayan 2013, 'Language management and minority language maintenance in (eastern) Indonesia: strategic issues', *Language Documentation and Conservation*, vol. 7, no. 2013, pp. 74-105.
- Arka, I Wayan. 2016. *Bahasa Rongga: descripsi, tipologi dan teori*. Jakarta: Penerbit Universitas Atma Jaya.
- Austin, Peter. 2012. Language Documentation. Oxford Bibliographies. LAST MODIFIED: 19 DECEMBER 2012. DOI: 10.1093/OBO/9780199772810-0075.
- Edwards, Owen. 2017. Metathesis and unmetathesis: Parallelism and complementarity in Amarasi, Timor. Doctoral dissertation, The Australian National University
- Evans, Nicholas. 2009. *Dying words: endangered languages and what they have to tell us*. Oxford: Wiley-Blackwell.
- Jukes, Anthony, Asako Shiohara, and Yanti. 2017. Collaborative project for documenting minority languages in Indonesia and Malaysia, AALL11.
- Klamer Marian, Antoinette Schapper, Greville Corbett, Gary Holton, František Kratochvíl & Laura C. Robinson. 2014. Numeral words and arithmetic operations in the Alor-Pantar languages. In Marian Klamer (ed.), *The Alor-Pantar Languages: History and typology*, 337–373. Berlin: Language Science Press.
- Musgrave, S. 2014. Language shift and language maintenance in Indonesia. In P.Sercombe and R. Tupas (eds) *Language, Education and Nation-building : Assimilation and Shift in Southeast Asia*. Basingstoke: Palgrave Macmillan, pp87-105.
- Simons, Gary F. and Charles D. Fennig (eds.). 2017. *Ethnologue: Languages of the World*, Twentieth edition. Dallas, Texas: SIL International. Online version: <http://www.ethnologue.com>.
- Windschuttel , Glenn and Asako) Shiohara. 2017. “Kui”, Antoinette Schapper (ed.) *The Papuan Languages of Timor, Alor and Pantar: sketch grammars*, vol.2. De Gruyter Mouton, Berlin.