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)

Indonesia is well-known for its linguistic and cultural diversity. However, many languages spoken there remain under-described and there are drastic reductions in speaker numbers as communities shift to the national language, *Bahasa Indonesia*, or more vigorous regional languages (e.g., Jukes, Shiohara, and Yanti 2017).

This presentation will introduce our efforts in documenting endangered indigenous languages in Indonesia, especially in the province of East Nusa Tenggara (hereafter NTT Province). We shall first briefly review the sociolinguistic situation in Indonesia and the NTT Province, and then report our activities funded by the LingDy project of the Research Institute for Languages and Cultures of Asia and Africa from 2014 to the present (2018). Our attempts are described in three phases below.

Phase 1: Establishing networks with local institutes

Our team's activities began with a workshop on language documentation at Udayana University in Bali in 2013. During the workshop, we gave lectures on the theory of documentary linguistics and training on practical issues useful for documenting languages such as recordings, data management, and annotation using ELAN software. Furthermore, we also conducted similar workshops from 2014 to 2016 in other cities in Indonesia and Malaysia, i.e., Samarinda, Manado, Jambi, and Kota Kinabalu.

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

The series of workshops mentioned above might have attracted attention to language endangerment in Indonesia and enabled us to establish international networks with local institutes. Unfortunately, these workshops did not lead to actual, substantial actions for language documentation. 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.

The situation was different when we conducted the workshop in Kupang in August 2015. As the capital of the NTT Province, one of the most linguistically diverse provinces in Indonesia, most of the participants were speakers of indigenous languages. They realized that their languages were disappearing and were distinctly motivated to document them. They were determined about making recordings and annotating the data during the practical session. As a result, we archived some data from nine indigenous languages to PARADISEC (http://www.paradisec.org.au/home.html), an international language archive.

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

After the success of the workshop in Kupang, we determined to focus our activities in this region hereafter. In addition, we also set a more concrete goal in our efforts in language documentation i.e., to archive substantial linguistic data of the languages in the region by expanding our collaboration. We realized that linguistic data would not be sufficiently valuable and useful unless data had been annotated based on appropriate knowledge of the language and general linguistics. Students who are interested and motivated enough to document their own language are invaluable resources as they are experts in their mother tongues. Nevertheless, while they are generally very talented and have a good command of English, their linguistics knowledge is still limited.

In an attempt to solve this issue, Peter Cole and Gabrielle Hermon of the University of Delaware, who had been working with the first author of this talk in their research in Indonesia, launched the "EAGER: Collaborative Approach to the Documentation of Endangered Languages in Linguistically Diverse Locales" program and offered to collaborate with us. In this program, we will form teams of one American graduate student interested in fieldwork and documentary linguistics, and two Indonesian students interested in documenting their own language. These teams will receive an intensive training course in linguistic fieldwork and language documentation and spend a few weeks in the field practicing what they have learned by collecting and annotating data on the Indonesian team members' language.

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