

New Directions in Field Linguistics: Training Strategies for Language Documentation in Indonesia

Margaret Florey and Nikolaus P. Himmelmann

7.1 Introduction

Language documentation has only been conceptualized as a sub-field of linguistics following the publication of Himmelmann's (1998) foundational paper, and in the ensuing ten years has become an area of intensive development in linguistics. A stimulating discussion is being generated in various forums about what it means to undertake language documentation and how this can be achieved, focusing on methods, technologies, and tools (e.g. Austin 2003; Bird and Simons 2003; Barwick and Thieberger 2006; Gippert, Himmelmann, and Mosel 2006). In part, the genesis of this new field has been triggered by the issue of language endangerment, and the growing awareness that rich documentation is *the* means of ensuring that a lasting multipurpose record of a language will be accessible to community members, researchers from a range of disciplines, policymakers, educators, and other stakeholders. Crucially for endangered languages, such a record also becomes a vital resource for language revitalization and maintenance activities.

The authors gratefully acknowledge the assistance received from the Volkswagen Endangered Languages Program for the two training workshops which were run under the auspices of the project 'Capacity building on a local and national level: Documenting Totoli, Central Sulawesi, Indonesia'. We also thank all of the linguists who volunteered their time to come and teach in Bali, and Dr Wayan Pastika, Udayana University, Bali, for his generous support with local organization of the workshops. We are grateful to Heidi Johnson and Nick Thieberger for information about current archiving activities. Florey's research and training workshops have been supported by a Major Documentation Project grant (MDP0009) from the Hans Rausing Endangered Languages Programme, SOAS, UK ('Documentation of four moribund Moluccan languages: Eastern Indonesia and the Dutch diaspora') and by an Australian Research Council Discovery Project grant (DP0343379) ('Cross-linguistic study of endangered Maluku languages: Eastern Indonesia and the Dutch diaspora').

Concurrent with the development of this new field, a shift in the ethics and politics of linguistic fieldwork has been taking place, which Florey (2008) calls the 'new linguistics'. Linguistic fieldworkers and language activists are unifying in promoting a research paradigm that prioritizes the leadership and involvement of community members in linguistic research, and sees linguists and community members working more and more in partnership. This stance has been articulated, for example, by Cameron *et al.* (1993); Craig (1992); Grinevald (2003); and Smith (1991, 1999), who argue the need for academic linguists to reassess their working relationships with speakers and their communities, particularly in regard to community participation in decision-making vis-à-vis the research process. It is apparent that the greatest progress towards documenting a substantial number of the world's languages will be made through the activities of well-trained linguists working in their own region. There is also growing recognition that language renewal and long-term language maintenance are not sustainable if they are dependent on external actors, and fundamentally require the frontline involvement of Indigenous language activists.

Language documentation, revitalization, and maintenance all require a complex set of skills, and, in order for community members and local linguists to participate more fully in such endeavours, the need for training is evident (e.g. the papers in Austin 2004). The authors of a report commissioned by the Netherlands Organization for Scientific Research remarked on the lack of fieldwork courses at universities in countries where endangered languages are spoken, and called for the creation of training programmes both for young researchers and for more established researchers who to date have not been involved in language documentation (NWO 2000: 1). This perspective has wider currency and a range of training and capacity building programs have been founded internationally both within and external to academe.

The best-known and longest-standing training models are those established in the Americas, such as the Guatemalan *Proyecto Lingüístico Francisco Marroquín* (PLFM) and *Oxlajuuj Keej Maya' Ajtz'uib'* (OKMA) (England 2003; 2007), and, in North America, Advocates for Indigenous California Language Survival (Hinton 1994; 2002; Hinton and Hale 2001) and the Indigenous Language Institute (New Mexico). The American Indian Language Development Institute (Arizona) is a highly successful model (McCarty *et al.* 1997; 2001; McCarty and Watahomigie 1999; McCarty and Zepeda 1998) which has led to the birth of offspring institutes in various locations in North America, including Oregon (Northwest Indian Language Institute), Oklahoma (Oklahoma Native American Language Institute) and Alberta, Canada (Canadian Indigenous Languages and Literacy Development Institute). In Australia, relevant courses include those at the Institute for Aboriginal Development, Alice Springs, and the Centre for Australian Languages and Linguistics (CALL) at Batchelor Institute of Indigenous Tertiary Education. Pilbara TAFE College in Western Australia offers an Indigenous Language Worker Program via the Certificate in Aboriginal Language Work.

Educational programmes dedicated to the training of Indigenous linguists have more recently been launched within a number of universities in Australia, North America, and the Pacific. These programmes are broadening the scope of tertiary education in linguistics and endeavouring to better meet the educational needs of Indigenous peoples. Examples include the First Nations Languages Program (University of British Columbia, Canada), Certificate in Aboriginal Language Revitalization (the University of Victoria, Canada), the Endangered Languages Academic Program, (SOAS London), the MA in Language Documentation and Conservation (the University of Hawai'i), The Center for Indigenous Languages of Latin America (the University of Texas, USA) (Woodbury and England 2004), the Pacific Languages Unit (the University of the South Pacific), and the Alaska Native Language Program (the University of Alaska Fairbanks).

These are exciting developments which are significantly strengthening the participation of Indigenous peoples in language documentation and maintenance. However, in other parts of the world, particularly in resource-poor developing countries in Africa and Asia, there have been few responses from universities, governments or language planning agencies, and similar capacity building ventures have been slow to emerge. The need for training in modern linguistic techniques to facilitate documentation is nowhere more apparent than in Indonesia, where some 737 Austronesian and Papuan languages are spoken (Gordon 2005). Ethnolinguistic groups vary widely in their size across this island nation, and, whilst the western Indonesian languages number hundreds of thousands or even millions of speakers, those in the east are typically much smaller, ranging from several hundred to tens of thousands (Tryon 1995). The eastern Indonesian languages are also amongst the most endangered languages of the Austronesian region (NWO 2000: 28; Florey 2005a). With an estimated 6912 languages in the world (Gordon 2005), Indonesia contains approximately 10.7 per cent of the world's linguistic resources, yet we estimate that fewer than 10 per cent (and possibly as few as 5 per cent) of the languages of Indonesia have been the subject of modern linguistic documentation. Himmelmann *et al.* (2005) noted that 'Documentation work proper—i.e. the recording, transcription and annotation of communicative events—has no tradition in Indonesian linguistics and to date has not been practised by Indonesian linguists'. The research which has been undertaken in Indonesia since the development of new documentation methods and technologies has overwhelmingly been led by foreign researchers funded by international grants. For example, a review of the projects funded by three of the major international agencies between 2002 and 2007 indicates that grants were awarded to seven projects documenting Indonesian languages (ELDP 5; Volkswagen 1; NWO 1) and an Indonesian linguist led just one of those projects (I Wayan Arka, whose training was undertaken in Australia).

Faced with large numbers of un(der)-documented languages, high levels of endangerment, and a severe shortage of Indonesian linguists with the requisite training, linguists have begun to ask how linguistic fieldwork practice could contribute to

capacity building for language documentation and maintenance in Indonesia. The NWO Advisory Committee advocated a strategy of cooperation, suggesting that 'Among the countries of the so-called Third World, the amount of local expertise, and hence of potential for local collaboration, varies widely. . . Where feasible, western researchers should cooperate with linguists in the country where the research is carried out. Researchers should be encouraged to train local linguists where such a need exists, and to establish or join a local network in which the research can be continued, and language revival activities initiated' (NWO 2000: 25). The philosophy of empowerment and the desire to expedite linguistic research and maintenance activities within Indonesia has seen the expansion of grassroots training of community members alongside linguists undertaking fieldwork in Indonesia. However, the scope of the problem requires larger-scale action and this persuaded Himmelmann *et al.* (2005) to seek funding from the Volkswagen Endangered Languages Program to train a new generation of Indonesian linguists and language activists, and to build documentation expertise at both the local and national level. This chapter discusses the training programme which resulted from that proposal. It explores the goals and methods of the program, analyses the outcomes to date, and considers the sustainability of this training model.

7.2 Developing a Training Model

The goals of the proposed training program were (1) to build the capacity for Indonesian linguists and language activists to initiate locally developed and implemented language documentation and maintenance projects, and (2) to facilitate a flow-on effect from this training to other people and institutions in the country. The Volkswagen Foundation funded Himmelmann's team to run two Training Workshops on Language Documentation. An intensive, residential model was planned for the workshops, which would be held over ten days in 2006 and six days in 2007. In accord with the training goals, the principal objectives of the two workshops were that, upon completion, participants would have

- (1) a basic understanding of the theory and principles of language documentation and language maintenance,
- (2) developed the ability to begin applying language documentation and maintenance methods and technologies in the field,
- (3) acquired knowledge of funding agencies and skills in the preparation of grant proposals, and
- (4) gained familiarity with pedagogical methods to support the transfer of skills more widely in universities and communities in Indonesia.

The structure and content of the workshops drew on various training experiences which provided a more realistic perspective of what might be achieved within the

available time frame, including Himmelmann's involvement in organizing three summer schools on field methods and language documentation in Germany in 1993, 1999, and 2004, and Florey's training work with Wangka Maya Pilbara Aboriginal Language Centre in Western Australia in the early 90s. The model also built extensively on the more localized training experiences of Himmelmann for Waima'a in East Timor, and Florey and colleagues in eastern Indonesia and in the Moluccan diaspora in the Netherlands (Florey 2001c; 2004; 2009; Florey and Ewing in Press; Florey and van Engelenhoven 2001).

We acknowledged in the original funding application that the short duration of the workshops would constrain the number of topics and the depth of the instruction which we would be able to provide. However, this model was necessary from a practical perspective because the organizers, tutors and most of the participants have full-time jobs, which precluded longer workshops. Further, as this particular kind of workshop had not been undertaken in Indonesia before, we considered it important to test and evaluate the training model before seeking support for a further set of workshops. Pedagogically, the intensive model provided the opportunity for immersion in the workshop topics. The experience which Florey and colleagues had gained in a six-day intensive workshop held in Maluku in 2004 (*Mari belajar bahasa tanah 'Let's study indigenous languages'*), also demonstrated the benefits of a residential mode which distanced participants from the business of their daily lives.

7.2.1 Logistics

In the Indonesian context, we needed to take into account several organizational issues which might not play such an important planning role in other parts of the world. Some students lived in quite remote locations and had little or no previous experience of air travel. About one third of the participants had no email or telephone access, and bringing together such a diverse group required logistical support from the tutors with regard to the participants they proposed inviting. Recent religious conflicts in Maluku and Sulawesi meant that extra care had to be taken in organizing the accommodation and meals, and in ensuring free time on Fridays and Sundays for Islamic and Christian religious services. After some discussion, Bali was selected as the site for the workshops as it best filled the criteria of access (good flight schedules and reasonably central for all participants) and infrastructure (hotels and local transport). The students and tutors lived and worked together in a hotel for the duration of the workshops. This arrangement also supported the informal exchange of experiences and aspirations, and helped to build a network among the participants, which we hoped might continue into the future to support their ongoing linguistic research.

7.2.2 Tutors

Our earlier training experiences highlighted the importance of a low student-teacher ratio to facilitate the achievement of our objectives, and to cope with the challenge

of the topics we would present and the different educational backgrounds the participants would bring to the workshops. Nine linguists with research based in Indonesia volunteered to assist Himmelmann and Florey as lecturers and tutors for the first workshop, and six of the nine were able to return for the second workshop. We thus had a total teaching staff of eleven people in 2006 and eight in 2007, who contributed an impressive range of backgrounds and expertise. The Volkswagen grant supported the participation of Himmelmann and Florey, as well as Claudia Leto (Bochum University, Germany) and Jani Kuhnt-Saptodewo (Museum of Ethnology, Vienna) from Himmelmann's research team, and I Wayan Arka (Australian National University) (see Plate 10) and Betty Litamahuputty (Max Planck Institute Jakarta Field Station). A further five tutors were supported from their own or other sources - Anthony Jukes (SOAS, London), Michael Ewing (University of Melbourne), Simon Musgrave (Monash University, Melbourne), Antonia Soriente (Max Planck Institute Jakarta Field Station), and Jan Wohlgemut (Max Planck Institute Leipzig). The contributions of these linguists were a major factor in the overall success of the two workshops.

7.2.3 Students

The financial and teaching constraints necessitated limiting the number of participants in Workshop 1 to twenty-five (see Plate 9). The tutors were asked to suggest potential participants, looking in particular for people with a background of interest in indigenous languages. We sought participants across several categories: those who were relatively early in their careers and would have the greatest opportunity to utilize the training they would receive, those who were further into their careers and in leadership roles in teaching and thus would have the opportunity to introduce documentation-related topics into their curricula, and community language workers. The student body was heterogeneous in region, culture, education, gender and religious affiliation. Amongst the nine women and sixteen men in Workshop 1, about half were MA or PhD students in linguistics in various Indonesian universities. The other half were university lecturers with overseas PhDs, people working in relevant allied disciplines or organizations (e.g. in a regional archive or museum), or language activists who have been involved in internationally funded research projects. All major regions of Indonesia were represented, from Sumatra to Papua, with a deliberate slight overrepresentation of eastern Indonesia due to the higher levels of diversity and endangerment in Sulawesi and further east. In a predominantly Muslim country, this overrepresentation also meant that about 50 per cent of the participants came from the more Christian areas in the east (Nusa Tenggara, Maluku and Papua). The indigenous languages represented amongst the participants of Workshop 1 as first or second languages included Adang (Alor), Alune, Balinese, Dani, Gorontalo, Haruku, Maybrat, Javanese, Malay, Minang, Mentawai, Punan, Rotinese, Selayar, Sundanese, Taa, Tombulu, and Totoli. All of the participants were speakers of Indonesian and/or

a Malay variety (Ambonese Malay, Brunei Malay, Melayu-Papua) and for some (a minority) of the participants that is their only language.

Approximately 45 per cent of this first group returned to participate in Workshop 2 in 2007. In order to qualify for the second workshop, the 2006 students were required to fulfil three criteria by February 2007:

- preparation of a five-page research proposal outlining a documentation or language maintenance project for an indigenous language of Indonesia, and following the grant preparation guidelines taught in Workshop 1;
- a one page summary of the proposal;
- a recording and transcription of at least 15 minutes of linguistic data.

A total of sixteen of the 25 participants (64 per cent) fulfilled the criteria, which exceeded our expectations and further underlined the success of Workshop 1. Eleven participants (four women and seven men) eventually took part in Workshop 2 as several participants encountered various obstacles (primarily health problems, but also workplace exigencies making it impossible to attend).¹

7.3 Workshop Structure and Content

Careful planning of the workshop structure and content was essential to optimize this training opportunity. With tutors spread around the world or in the field in Indonesia, most of the workshop organization took place by email and phone calls. In both years, the tutors first met together the day before the workshop was to begin. It was only at this point that we could finalize the program and discuss the strengths and weaknesses of the participants vis-à-vis education, language background, research experience, and familiarity with technologies.

The schedule for each workshop was structured in four daily sessions, which began at 8.30 am and concluded around 5.45 pm. In Workshop 1, Sessions 1 and 3 were 45-minute lectures, while sessions 2 and 4 were one and a half or two hour tutorials. In Workshop 2, much more time was provided for tutorial work and there was an average of one lecture per day. In both workshops, a tutors' meeting was held at the end of each teaching day to give the staff an opportunity to report on the challenges and successes of the day, and to review the activities for the following day. This thirty to sixty-minute session was essential, in part because of the tutors' varied backgrounds, interests and expertise, and also because of our shared desire to remain flexible and able to respond to any issues and concerns which might have arisen.

¹ In this regard we thankfully note the good fortune that in Workshop 1 all participants arrived on time, no-one became sick, and everyone returned home safely.

7.3.1 Lectures

The lectures drew on the expertise of all of the tutors. All lectures were given in Indonesian, which was the lingua franca of the workshop. Most lectures included PowerPoint presentations, and students were given handouts with the PowerPoint slides and other lecture notes in Indonesian. At the conclusion of each workshop, a CD with the presentations, notes and photos was given to all students and tutors.

The primary goal of Workshop 1 was to build technical skills in language documentation methods, and fourteen lecture topics (listed below) were developed to work towards Objectives 1–3. As these topics indicate, the students were introduced to a number of computer programs used in language documentation, including Audacity (for the capture of audio data from analogue or DAT recorders), ELAN (used to segment, time-align, and transcribe audio and video files), and Toolbox (used to interlinearize, gloss, and analyse data, and to build a lexicon) (see Plate 11).

- Introduction to language documentation (Himmelmann)
- Recording technologies and techniques (Jukes)
- Speakers and speech communities (Florey)
- Organizing metadata (Musgrave)
- Capture of audio and video (Leto and Himmelmann)
- Basic orthography issues (Wohlgemut and Soriente)
- Using software for transcription (Arka)
- Some principles for segmenting discourse (Himmelmann)
- Notes on dealing with conversation (Ewing)
- Notes on dealing with ritual language (Kuhnt-Saptodewo)
- Preparing grant applications (Florey)
- Toolbox (Arka, Jukes, Musgrave)
- Commenting on meaning (Arka, Jukes, Musgrave)
- Commenting on grammar (Arka, Jukes, Musgrave)

In addition, two summarizing lectures were given, one on day 7 and one on the last day, which highlighted major features of the previous lectures and related them to recurring problems encountered in the practical tutorials (e.g. the issue of metadata and the need to employ a proper folder structure for the data stored on a computer).

The content for Workshop 2 was driven by a number of factors. The curriculum built and expanded on the skills taught in 2006. It focused heavily on Objective 3 (skills in proposal writing) and introduced Objective 4 (gain familiarity with pedagogical methods to support the transfer of skills more widely in universities and communities in Indonesia) which was not taught in Workshop 1. The curriculum also responded to feedback from the students and incorporated a number of topics which they requested, including dictionary making and an overview of current documentation projects in Indonesia. The eight lecture topics covered:

- Summary comments on proposals (Himmelmann)
- Documentation projects in Indonesia (Leto)
- Using Edirol digital recorder (Litamahuputty)
- A note on digital archives (Himmelmann)
- Transferring skills in language documentation (Florey)
- Developing local language and culture centres (Florey and Kuhnt-Saptodewo)
- Toolbox: summary so far (Arka)
- Dictionary-making (Arka)

7.3.2 Tutorials

In tutorials, the students carried out practical exercises which formed the core of the workshops. The Workshop 1 students were placed in three tutorial groups with eight students in Group 1 (tutored by Arka, Himmelmann, Kuhnt-Saptodewo, and Leto) and Group 2 (tutored by Florey, Musgrave, and Soriente), and nine students in Group 3 (tutored by Ewing, Jukes, Litamahuputty and Wohlgemut). Group 1 included most of the younger students, and Group 2 included the community language activists who had the least familiarity with technologies. Within each group, participants usually formed subgroups of two each, often for the practical reason of sharing equipment (see Plate 12). In Workshop 2, all eleven students worked in one tutorial group which was taught in alternation by two teams of four people: Team A (Himmelmann, Kuhnt-Saptodewo, Leto, and Soriente), Team B (Florey, Arka, Ewing, and Litamahuputty). In both cases, there was a ratio of one tutor to two or three students, which ensured that someone was always on hand to meet the specific needs of a particular participant or participant duo.

Tutorial exercises in Workshop 1 encompassed all basic aspects of documentation. The students practised recording techniques, and the use of Audacity, ELAN, and Toolbox. They learned to prepare basic metadata using Excel or Word, and were given a demonstration of IMDI, a program which supports more detailed recording of metadata. Time constraints limited the opportunity to work with video data, but some time was given to this task in Workshop 1 and to learning to use Adobe Premiere Pro and iMovie.

Prior to and during Workshop 1 there was extensive discussion amongst the tutors about whether all topics would be accessible to or useful for all participants, particularly those who did not have access to computer facilities in their home location and, at least in the immediate future, would not be able to use the computational tools they were learning. Some alternate tutorial exercises were offered, focusing on techniques for transcription and metadata using pencil and notebooks, and on simple methods for producing resource materials for language maintenance activities. We found, however, that all students were keen to gain an understanding through participation and practice of the technologies which are used in modern language documentation. Even the students who came to the workshop with minimal computing experience demonstrated

high motivation to master these tools, and students often continued working late into the night.

In Workshop 2, students worked individually to practise their technical skills on their own data, which they brought to the workshop. Through this process, by the end of the workshop most of the students had acquired the ability to capture and process their data in Audacity and ELAN, to export ELAN files into Toolbox, and to produce an MDF lexicon from Toolbox (see Plate 13). In each session they prepared appropriate metadata, and gained a better understanding of the requirements and practices of depositing data in an archive.

7.3.3 Preparation of grant proposals

In both workshops, considerable time was devoted to skills in the preparation of grant proposals and acquiring knowledge of funding agencies (Objective 3), which we considered crucial in optimizing the possibilities for Indonesian linguists and language activists to work independently on their own documentation projects. We first needed to redress the inequity that the funding guidelines for most international agencies are available only in English and thus are not accessible to non-English speakers, which includes many Indonesians. The guidelines and application materials for the Hans Rausing Endangered Languages Project, the Foundation for Endangered Languages and the Endangered Languages Fund were translated into Indonesian by Lita-mahuputty and distributed as handouts to the participants.² In Workshop 1, students were given an introductory lecture on proposal writing and they were then placed into six smaller groups of four to five members each and given several days in which to meet together, come up with a project idea and draft a proposal. Each group met at least once with their tutors to get some initial feedback on the basic idea and the draft outline. On the second last day of the workshop, the groups distributed a one-page summary handout and made a fifteen-minute presentation of their proposals, with ensuing critical discussion and feedback from the tutors and participants.

In Workshop 2, each participant was allocated two tutors (one from each tutoring team) who worked closely with the student to bring their proposal to a standard acceptable for submission to a national or international funding agency. The participants redeveloped and improved their proposals both in tutorials and during breaks. Revised drafts were given to the tutors every second day, and the tutors then provided individual feedback to the two or three students they were mentoring. Students responded very positively to this strategy both throughout the workshop and in the evaluation process (see Section 7.4 below). Many commented on the benefits of working one-to-one with skilled researchers and learning from their experiences. One participant noted that although she had worked for sixteen years in a major regional institution and had previously been involved in research with a large international organisation, this was the first time that she had been trusted to prepare her

² The translated materials were also provided to these three funding bodies for wider distribution.

own research proposal. The mentoring process continues beyond the duration of the workshop, and tutors will assist students in finding an appropriate funding body and submitting their application. We hope that this method will result in the funding and execution of a number of documentation projects.

7.4 Evaluation

An anonymous evaluation procedure was used on the final day of each workshop to elicit feedback and to assess the outcomes (see Appendix). The evaluation procedure consisted first of rating aspects of the workshop on a five-point scale from 'very satisfactory' to 'not at all satisfactory'. Some of the questions in Part I differed in each evaluation. Part II consisted of short written comments on a number of topics such as 'What did you hope for from the workshop before it began? Were your hopes fulfilled?'

7.4.1 Participants

The responses to the questions in Part I were tabulated for each of the 25 participants in Workshop 1 and 10 participants in Workshop 2.³ The average per Workshop 1 participant was 4.3/5 with scores ranging from 3.6 to 4.9, and per Workshop 2 participant was 4.4/5 with scores ranging from 3.6 to 5. In both workshops, the results for the three questions which addressed the value of lectures, tutorials, and materials yielded an average of 4.4/5. The overall impression of Workshop 1 (Question 8) rated 4.7/5. Students in Workshop 2 expressed a very high level of satisfaction with the two tutorial teams (Question 5) with an average response of 4.6, and with the individual guidance on proposal writing (Question 6), averaging 4.8.

These high scoring results are more meaningful in an Indonesian context than they might be in some other settings as cultural requirements of politeness and fitting into one's group and level increase the tendency to avoid the extremes and strongly favour choices in the middle regions of such scales. In Workshop 1, only Question 4, which pertained to the availability of equipment, attracted lower responses (with an average of 3.8/5), and this was also the point most repeatedly commented upon negatively in Part II. Participants had been asked, if at all possible, to bring their own recording equipment, headsets and laptops. While most were able to bring recording equipment and headsets, almost half did not have their own laptops, which meant that they had to share during tutorial exercises. Furthermore, those without their own laptops could not continue working by themselves during the afternoon break and in the evenings as many wished to do. In Workshop 2, we addressed this issue through asking tutors to bring extra laptops to ensure that each participant had full access to a computer for the duration of the workshop. The success of this strategy is reflected in the rise of the evaluation score on this question to 4.6/5.

³ One participant in Workshop 2 chose not to complete an evaluation.

In Part I of the Workshop 2 evaluation, only one question had an average response lower than 4: Question 8, which asked about the opportunity to finish the assigned tasks. Eight of the participants felt that the six-day workshop was too short and that some topics were not allocated sufficient time. Students wrote in Part II that in particular they would have liked more time for dictionary-making.

In the written section of the evaluation of Workshop 1, almost all participants commented favourably upon the practical tutorials, considering these the highlight of the workshop. In the Workshop 2 written evaluation, students reported that they had come to the workshop hoping to improve their skills in documentation, computational tools, dictionary preparation and proposal-writing, and indicated that their hopes were met in almost all areas. The participants consistently responded that the sessions which they most liked (Question 12) were those focused on preparation of their proposal and on dictionary-making. There were many favourable reports on the opportunity to work with experts and the friendliness and dedication of the tutors. When asked what they thought was the least satisfactory aspect of the workshops, nearly half of the participants in Workshop 1 and 80 per cent in Workshop 2 answered that there was none.

One issue which we had anticipated was that seniority and status might impact on the working relationships within the group as academic events in Indonesia are heavily dominated by considerations of status and protocol. Overall, our concerns in this regard were not realized and the lecture sessions and tutorials took on an unusually egalitarian atmosphere. Issues of status appeared only in the context of the grant proposal groups in Workshop 1. Here, the more senior members tended to set the agenda and to distribute workloads among the more junior members. This concern may need to be addressed more directly in future workshops.

7.4.2 Tutors' responses

The tutors did not participate in a formal evaluation process. The daily tutors' meetings gave us an opportunity to receive feedback and to learn how the tutors felt the workshop was progressing and where they saw challenges, both personally and for the participants. As with the students, the tutors' response to both workshops was overwhelmingly positive. The excitement and enthusiasm of the students was contagious, and, for the tutors, it was an opportunity to meet the next generation of linguists, to learn about the language-related activities which are happening around Indonesia and those which are planned for the future, and to share their expertise and the training experience with colleagues.

An unanticipated outcome of Workshop 1 for the tutors was the opportunity to learn alongside the participants. This is perhaps best exemplified with the case of ELAN. Quite a number of the tutors had never used ELAN and this meant that those tutors were challenged to keep at least one step ahead of the students in their tutorial group so that they could continue to support their learning. Similarly, not

all tutors had used IMDI or the video processing programs, though these were only peripherally introduced in Workshop 1. Again, the large number of tutors proved useful because at least one tutor in each group was an experienced user of the various programs. Most tutors explicitly noted that they profited directly from this workshop for their own work as it gave them the opportunity to learn the preliminaries of ELAN and other programs and they took away new skills which they intend to continue using in their fieldwork and research. Indonesian language skills varied widely across the pool of tutors and the less fluent speakers found these opportunities to lecture, tutor and interact informally in Indonesian greatly benefited their language skills.

7.5 The Question of Sustainability

At this point it is useful to return to the title of this chapter—'New directions in field linguistics: Training strategies for language documentation in Indonesia'—which reflects the strategy that we have utilized thus far to strengthen the capacity for language documentation in Indonesia: that is, to embed training within the fieldwork for our own documentation projects, and those of our linguist colleagues. This strategy has been born not from imagining the most appropriate training model we might develop under ideal circumstances, but rather from necessity, taking into account the dual limitations of our funding resources and the time we have available to dedicate to training. Funding remains a major challenge faced by those seeking to provide training to local linguists and language activists. There has been a laudably significant increase in funding available for the documentation of endangered languages, and some funding agencies, such as the Volkswagen Foundation, the US National Science Foundation and Hans Rausing Endangered Languages Project, permit and indeed encourage training to be embedded within documentation projects. However, there are very few grant sources which specifically address the capacity building needs of developing countries vis-à-vis language documentation and language maintenance. The Indigenous Language Worker Program at Pilbara TAFE College in Western Australia states as two of its aims:

- To provide Indigenous people with the necessary skills to organize and run language programs in their own communities; and
- To assist Indigenous people to make informed decisions about their own language and take control of their own language programs (Pilbara TAFE 2007).

These aims highlight empowerment, participation and autonomy. Realistically, it will be hard for Indonesians to achieve Indigenous control when local linguists and language activists largely remain reliant on foreign linguists to initiate and bring research projects to their country. At this stage in reviewing our capacity building activities, we identify the critical questions as:

- (1) How far does this model go towards meeting the capacity building needs of Indonesia vis-à-vis language documentation and maintenance?
- (2) Is the model we utilized sustainable?
- (3) Is there a more suitable model which might better support the building of both capacity and autonomy?

7.5.1 Meeting capacity building needs

With regard to the first question, it is apparent that training only twenty-five people in a large and linguistically very diverse country will only meet a minuscule portion of the actual needs. However, there are very positive indications that even a short period of intensive training can trigger documentation and maintenance activities. The main goal of the two workshops we planned is to build the capacity for Indonesian linguists and language activists to initiate locally developed and carried out language documentation and maintenance projects. The outcomes suggest that the range of skills which were covered is already making a good start towards meeting needs, with reports of a number of encouraging developments immediately following Workshop 1 and more which developed later in the twelve months leading to Workshop 2.

Several participants immediately began to share their documentation skills with their colleagues, students and fellow community members. A course in linguistic research methods focusing on the documentation of minority Indonesian languages was introduced at the private university *Universitas Nasional* in Jakarta. A new upper level subject on language documentation will be taught from 2008 in the Arts Faculty at the State University of Papua in Manokwari. Co-applicant and participant I Wayan Pastika has begun to teach Audacity, ELAN, and Toolbox and archiving practices at Udayana University in Bali. The National Language Centre in Jakarta (*Pusat Bahasa*) invited Himmelmann to give a short course on language documentation in August 2006, which he did with the assistance of the four *Pusat Bahasa* participants in Workshop 1 (Luh Anik Mayani, Citra Aniendita Sari, Dira Hildayani, and Yayat Hendayana). The director and staff of *Pusat Bahasa* were very enthusiastic about putting language documentation on their main agenda and participant Luh Anik Mayani intends to run further short courses on the subject for her colleagues at *Pusat Bahasa*. More informally, participants in West Papua, Maluku, and Nusa Tenggara reported raising awareness of language endangerment amongst community members and training fellow university students in recording techniques and in the use of relevant documentation software.

Following Workshop 1, one participant, Jermy Balukh, remained in Bali for six weeks to work in apprenticeship with I Wayan Arka and native speakers of Rongga for Arka's Rongga documentation project. In Maluku, participant Ana Lewier took a class of students from Pattimura University in Ambon to Lohiasapalewa village on Seram Island, the home of fellow participant Johanis Soriale. Lewier, Soriale and the

Pattimura students spent a week doing a pilot survey of oral traditions and recording stories and songs in the local Alune language. These data will form the basis for an ongoing documentation collaboration between the two participants.

We noted in Section 7.3.3 above that the grant-writing sessions at the two workshops were a critical strategy for meeting capacity building needs and also building towards autonomy. Participants in Workshop 1 commented on the importance of these sessions and asked for more time to be devoted to this topic in Workshop 2. In the second workshop, substantial progress was made in developing participants' proposals to a stage at which they could actually be submitted to a funding body. One real test for the success of this workshop program will be the number of proposals which are actually submitted. At the time of writing, one participant, Jermy Balukh, had successfully applied to the Endangered Language Fund, and had received the Bill Bright Memorial award for his application. We are optimistic that a majority of the participants (perhaps eight of the eleven) will submit applications within the next twelve months (though of course the outcome of the applications remains uncertain). This number would represent approximately one-third of the original group, and this may seem to be a relatively small number. However, given that (a) most participants have never drafted a proposal before and that (b) prior to the workshops, there had only been a single successful application by an Indonesian scholar in the field of language documentation (Arka), it would still be considered a success.

On a more general level, the workshop experience has made it clear to us that regional linguists and language activists will be much better placed to initiate and develop their own research projects when they are able to prepare competitive grant applications. The preparation of translated grant application materials also works towards those dual goals. The Hans Rausing Project has made a good start towards the provision of grant information and application materials in languages other than English with application packages available in Indonesian, French, Spanish, and Russian. This task could be further supported if more linguists were able to supply translated materials in other languages. It would also assist applicants from non-English speaking backgrounds if international agencies were willing to accept translations of applications and annual reports, etc. A small translation item may have to be built into budgets for this work.

7.5.2 Sustainability

The unequivocal answer to the question of whether this model is sustainable is 'no, not in the present format'. In Section 1 we suggested that (a) the greatest progress towards documenting a substantial number of the world's languages will be made through the activities of well-trained linguists working in their own region, and that (b) language renewal and long-term language maintenance are not sustainable if they are dependent on external actors. If we are serious about these points, then

sustainability means building towards autonomy. It also means separating out the roles of grassroots activism and volunteerism from the kinds of activities which work towards 'indigenizing the academy' (Mihsuah *et al.* 2004) and handing over training to regional institutions, linguists, and language activists. The current model relies on linguists (1) gaining funding for documentation activities from national and international funding agencies, (2) having both the motivation and the skills to undertake capacity building, (3) scheduling training activities alongside of fieldwork, and (4) volunteering as tutors. These factors delimit the sustainability of training. As noted above, we made some headway in the two workshops towards our goal of training Indonesian linguists and language activists in documentation methods. However the methods and tools in our field are complex, and need to be reinforced and increased through a cycle of training and practical fieldwork experience, as made very apparent by the experiences in the second workshop. It is quite clear that perhaps with the exception of the two most talented and eager participants, most other participants would need at least two further workshops of ten to fourteen days in order to achieve a skill level that would allow them to carry on independently and to transfer their skills to other interested members in their communities. Indeed, participants themselves called for further training in use of the software, in skills such as video recording and processing, and in pedagogical methods for sharing their training with others.

Cycles of training workshops over an extended period of four to five years, however, are not possible under the present circumstances. Most documentation projects run for two to four years, and, if a research team does not apply for and receive further funding when its project concludes, the expertise of the team and the training momentum will be lost. Preparing and running training activities besides regular fieldwork also puts an extra burden on the tutors which will only be sustainable if this kind of activity receives adequate funding and is acknowledged as part of the 'regular duty' of a fieldworking linguist. In the current case, although all the tutors were very willing to give their time to trialling this kind of activity, participation posed some financial burden for several of the tutors. And despite the fact that both workshops did not last longer than ten days, it would be difficult to sustain such intensive training efforts over a longer period of time. It was particularly challenging given that the tutors were either heading home from fieldwork to teach in their regular jobs or were going to the field.

7.5.3 Alternative Models

Third, we pose the question of whether there is a more suitable model which might support the building of both capacity and autonomy. Section 1 noted that there are sustainable capacity building models which have been developed in other parts of the world. The *Oxlajuuj Keej Maya' Ajtz'uib'* (OKMA), for example, has run since the

1970s, and provides nine-month long courses in linguistics and documentary methods. The American Indian Language Development Institute (AILDI) has held six-week intensive summer institutes for thirty years, and offers a range of classes in linguistics and language maintenance methods. In Australia, Aboriginal Language Centres have provided training and support to local language communities since the mid- 1980s and are now being developed in parts of North America.⁴ The language centre model has been adapted this decade in the development of Indigenous Knowledge Centres (Taylor 2003; Taylor 2004; Queensland Government 2007). Could these models be adopted in Indonesia?

The language resource centre model appears well suited to large, diverse countries such as Australia and Indonesia. This idea is being explored in some parts of Indonesia. John Bowden has recently received funding from the ELDP to document Helong, spoken near Kupang in Timor. Bowden plans to incorporate training in his project and hopes to establish a language centre in the region in close cooperation with the workshop participants from Kupang (Haan, Balukh). Participant Yusuf Sawaki, a lecturer at the State University of Papua in Manokwari, plans to establish a language centre supporting local languages at that university. Margaret Florey and Michael Ewing have held preliminary discussions with the Siwalima Museum in Ambon about a Centre for Endangered Languages, Cultures, and Musics at that institution.

Language centres are established and managed regionally, and thus stand a better chance of understanding and meeting the needs of their local language communities. Some of the activities of a centre include coordinating local research projects, training staff in formal courses and through apprenticeship, hiring external linguists as necessary on short and longer-term contracts, acting as regional repositories and archives for data, and as literature production centres. The centres can start small and grow as the resources for both human and physical infrastructure become available. But the core issue of course is: Where would the financial, human, and technological resources come from for establishing such centres?

Both OKMA and AILDI are affiliated with universities and are able to draw on their infrastructure. In Indonesia as well, we believe, language centres may most easily be established through piggybacking onto appropriate existing institutions such as universities, libraries, museums or the *Pusat Bahasa*. The *Pusat Bahasa*, in fact, has recently begun a process of decentralization and plans to establish regional language centres (*Balai Bahasa*) in most of Indonesia's thirty-three provinces (usually in the capital of the province). It has also added language documentation to its brief. But it remains to be seen whether the *Balai Bahasa* or the other regional institutions just mentioned (universities, libraries, and museums) are willing to become cooperative partners and supporters for more local initiatives rather than just acting as regional representatives of their national headquarters.

⁴ A useful list of Australian language centres can be found at <http://www.fatsil.org/links/nild.htm>.

Language centres might also go some way towards addressing the shortage of computing equipment which is a chronic problem in many parts of Indonesia. In his report to Volkswagen, Himmelmann noted that, despite the enthusiasm of the *Pusat Bahasa* to include language documentation on their main agenda, the staff at that institution generally lack technical facilities and staff training in IT to support state of the art documentation projects. A language centre could house computing equipment and a digitization station which would be made available for use by regional linguists and language activists. Language centres could thus provide local archiving facilities and also increase access to archival material by the communities from which the data originates. Such localized initiatives are being encouraged and developed by some major preservation archives. For example, PARADISEC (Pacific and Regional Archive for Digital Sources in Endangered Cultures) is working with the Vanuatu Cultural Centre in Port Vila, the Institute of Papua New Guinea Studies in Port Moresby, and the Tjibaou Cultural Centre in New Caledonia to repatriate materials and providing data backup services. The Max Planck Institute is developing plans to establish local archive outlets in Latin America and elsewhere, linked to projects funded by the Volkswagen Foundation's DoBeS (Documentation of Endangered Languages) program. DELAMAN (Digital Endangered Languages and Musics Archive Network) hosted its annual meeting in late 2007 in Mexico City on the theme of global-local archive relationships, and one sub-theme focused on training in digital techniques for researchers, archivists, and communities.

Most importantly in the context of the current chapter perhaps, language resource centres could provide a focal point for training activities of linguists doing fieldwork in the area. Training would not depend exclusively on the initiative of a small dedicated group of fieldworkers but ideally could make use of all researchers working in the area, both foreign and local, thus distributing the training load across a broader range of individuals. It would also not be limited to the lifetime of a single project but could draw on the expertise and time of consecutive fieldwork projects, thus becoming more sustainable. One- or two-day workshops held outside a context which provides for regular training opportunities may be of little use, but a series of short workshops in such a centre may well achieve results which go beyond what can be achieved in longer one-off workshops. Last but not least, costs for training workshops could be drastically reduced as participants would not have to travel from all over Indonesia to a single location.

Still, funding will continue to be an issue. But with support from funding bodies, linguists could build into grant applications an item to support their teaching as visiting lecturers. We can continue grassroots training through taking language centre staff to the field with us to strengthen their skills. Another possibility to make language centres more sustainable in the long-term is for them to partner with foreign universities. For example, the centres could serve as a host institution for in-country summer or semester-long documentation and field methods classes. This class could also involve a period spent volunteering as tutors at training workshops. Such arrangements will

be of two-way benefit in which foreign students and regional linguists and language activists can learn from and support each other.

In sum, we clearly see a role for training activities forming an integral and regular part of the new conceptualization of linguistic fieldwork which is emerging under the label of *language documentation*. And we strongly believe that given the right kind of contextualization or embedding (in the form of local or regional language centres), such training will contribute to the excitement of doing fieldwork and enrich our own experiences rather than becoming an additional burden. Ken Rehg (2007: 19) noted a 'stirring sense of camaraderie' born from the University of Hawai'i's documentation centre. That spirit was apparent at both the Bali workshops and is felt more widely among linguists and language activists involved in fieldwork, documentation, training and language maintenance who are working towards the shared goal of strengthening and supporting indigenous languages.

Appendix: Workshop Evaluations

In both workshops, a number of issues were evaluated using the following scale: Very satisfied (5), Satisfied (4), Reasonable (3), Dissatisfied (2), Very dissatisfied (1)

Table 7.1 Evaluation of Workshop 1

Q no.	Question	Average
PART 1 (n = 25)		
1	Organization of travel to Bali	4.5
2	Hotel and facilities (room, food, service, etc.)	4.5
3	Bali as workshop location	4.8
4	Availability of electronic tools (recording equipment, computers, etc.)	3.8
5	Lectures (content, delivery, etc.)	4.5
6	Tutorials (practicum, usefulness, etc.)	4.5
7	Handouts (clarity, helpfulness, etc)	4.3
8	Overall impression of the workshop	4.7
PART 2		
9	What did you hope for from the workshop before it began? Were your hopes fulfilled?	
10	Which part of the workshop did you like the most? Why?	
11	Which part of the workshop did you like the least? Why?	
12	Are you likely to use the knowledge or skills which you gained through this workshop? How? Give a short explanation.	
13	What other topics could or should be discussed in a workshop such as this?	
14	Please add any other suggestions or comments.	

Table 7.2 Evaluation of Workshop 2

Q no.	Question	Average
PART 1 (n = 10)		
1	Lectures (content, delivery, etc.)	4.5
2	Tutorials (practicum, usefulness, etc.)	4.6
3	Schedule and structure of the workshop (division of time between lectures and tutorials)	4.1
4	Study materials (handouts, PowerPoint presentations, etc)	4.2
5	Number of tutors for each group (2 teams with 4 tutors each)	4.6
6	Direction in developing a proposal (individual guidance from two tutors)	4.8
7	Availability of electronic tools (recording equipment, computers, etc.)	4.6
8	Opportunity to finish the assigned tasks	3.7
9	Hotel and facilities (room, food, service, etc.)	4.6
10	This workshop took place over 6 days. Should it have been longer, the same, or shorter?	8 2 0
PART 2		
11	What did you hope for from the workshop before it began? Were your hopes fulfilled?	
12	Which part of the workshop did you like the most? Why?	
13	Which part of the workshop did you like the least? Why?	
14	What other topics could or should be discussed in a workshop such as this?	
15	Please add any other suggestions or comments.	

8

Anxious Respect for Linguistic Data: The Pacific and Regional Archive for Digital Sources in Endangered Cultures (PARADISEC) and the Resource Network for Linguistic Diversity (RNLD)

Nick Thieberger

8.1 Introduction

It is clear that the wind of change blowing through the discipline of linguistics in the recent past is resulting in a heightened focus on fieldwork methodology. Leaving aside that part of our discipline that has never concerned itself with field recording, there is a change in the way that linguists can now conceive of linguistic data, in particular, but not solely, as it relates to endangered languages. Three main determinants of this change are: the realization of the urgency to record as much as possible before speakers stop speaking small languages; developments in technology¹ which allow us to work with data with an immediacy not previously possible; and a growing involvement in linguistic projects of speakers or their descendants who want access

This chapter was written while I was supported by the following Australian Research Council grants: SR0566965—Sharing access and analytical tools for ethnographic digital media using high speed networks, DP0450342—New methodologies for representing and accessing resources on endangered languages: a case study from South Efate, LEO453247—Digital archiving equipment for PARADISEC research archive of Asia-Pacific region audio recordings, LEO560711—PARADISEC, the Pacific and Regional Archive for Digital Sources in Endangered Cultures: Accessibility and Decentralization.

¹ 'New technologies' is used here to refer to digital recording, software, tools and methods that have become available in the past decade.

Endangered Languages of Austronesia

edited by
MARGARET FLOREY

OXFORD
UNIVERSITY PRESS

OXFORD

UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York

Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto

With offices in

Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam

Oxford is a registered trademark of Oxford University Press
in the UK and in certain other countries

Published in the United States
by Oxford University Press Inc., New York

© Editorial matter and organization: Margaret Florey 2010

© The chapters: their various authors 2010

The moral rights of the authors have been asserted
Database right Oxford University Press (maker)

First published by Oxford University Press 2010

All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
without the prior permission in writing of Oxford University Press,
or as expressly permitted by law, or under terms agreed with the appropriate
reprographics rights organization. Enquiries concerning reproduction
outside the scope of the above should be sent to the Rights Department,
Oxford University Press, at the address above

You must not circulate this book in any other binding or cover
and you must impose the same condition on any acquirer

British Library Cataloguing in Publication Data
Data available

Library of Congress Cataloguing in Publication Data
Data available

Typeset by SPI Publisher Services, Pondicherry, India
Printed in Great Britain
on acid-free paper by
CPI Antony Rowe, Chippenham, Wiltshire

ISBN 978-0-19-954454-7

1 3 5 7 9 10 8 6 4 2