Language documentation and the multi-dimensionality of capacity building:

framing research diversity in an Indonesian ethno-ecological context

I Wayan Arka

Australian National University/Udayana University

1 Table of Contents

Abstract

This paper describes language documentation and maintenance practices in an Indonesian context with attention given to the complexities of minority ethnolinguistic speech communities. Unlike the traditional linguistic research conducted in Indonesia in the 19th to 20th centuries, fieldwork research in Modern Indonesia must now contend with multiple dimensions, such as the impact of language policy on minority speech communities, the effects of economic developments and social mobility on previously isolated communities, and the role of motivation or capacity building in intergenerational language transmission. These multi-dimensional aspects are explored within the context of Indonesia's superdiverse ethnolinguistic landscape with evidence provided based on my fieldwork sites across Indonesia, including in Papua, central Flores, and eastern and western Indonesia. My experiences across these fieldwork projects emphasise the need for long-term capacity building and local leadership practices, and the importance for linguists to approach language holistically by considering its socio-linguistic, cultural-political, ecological and cognitive dimensions.

Keywords: ethnolinguistics, capacity building, language documentation, endangered languages, language policy, multilingualism, diglossia

Language documentation and the multi-dimensionality of capacity building:

framing research diversity in an Indonesian ethno-ecological context

1. Introduction

This paper discusses language documentation and capacity building in Indonesia, focusing on challenges and opportunities offered by Indonesia's rich resources in ethnolinguistic research. First, I contextualise the discussion by providing some background on Indonesia's ethnolinguistic superdiversity. Then, I examine language documentation and linguistic studies in Indonesia during the past two centuries, broadly classifying them into classic and modern documentation research. Classic research covers studies undertaken in the early/mid-19th to 20th centuries in Indonesia, primarily by foreign missionaries, while modern language documentation research, including capacity-building activities, covers three periods in Indonesia: late 20th Century, early 21st Century (pre-COVID-19), and 2020 until today (including COVID-19). Finally, based on my personal experiences, I reflect on the research opportunities in Indonesia and the challenges of collaborative research in a local community context. I also highlight the opportunities and challenges of enhancing local capacity building and engagement activities specifically in the context of COVID-19 and its long-term local impact on linguistic research in Indonesia.

2   Ethnolinguistic diversity and language documentation in Indonesia

Indonesia is one of the world's most ethnolinguistically superdiverse regions and is home to approximately 700 languages. Genealogically, these languages belong to two major groups: Austronesian and non-Austronesian (Papuan), with the former language family predominantly covering central and western Indonesia, and the latter encountered only in eastern Indonesia, in mainland New Guinea and the adjacent regions of Alor-Pantar and north

Maluku/Halmahera. The precise subgrouping of Indonesian languages is subject to ongoing research (cf. Blust, 2009; Donohue and Grimes, 2008). In eastern Indonesia, such as in North Halmahera, Timor-Alor-Pantar, and Bird's Head, Austronesian and Papuan languages have been in contact with each other for millennia. Consequently, contact-induced changes, combined with other kinds of internal diversifications, have led to linguistic complexity in the region. Similarly, in central and western Indonesia, Austronesian languages have undergone gradual diversification, resulting from extensive dialectal variation in dialect chains, forming a linkage (cf. Ross, 1988: 9–11). The Austronesian languages in these regions are also subject to contact, making it difficult to identify discrete proto-language(s) – consider, for instance, the debate on the existence of Central Malayo-Polynesian (Blust, 1993, 2009; Donohue and Grimes, 2008; Klamer, 2002a, 2002b).

The rich ethnolinguistic superdiversity of Indonesia serves as a living laboratory for research in linguistics and related fields. Linguistically, the varying properties across different structures in their internal grammatical systems (phonology, morphology and syntax), and their interface with discourse-pragmatics, are of interest to typological and theoretical linguists, and will be discussed later in this chapter (see section 4.2). Socio-culturally, the Austronesian and Papuan communities have rich histories and traditions reflecting a blend of local/indigenous and non-indigenous practices (e.g., related to Hinduism and Islam) with language playing an essential role and therefore serving as a window into the richness of such practices.

Unfortunately, Indonesia's ethnolinguistic superdiversity is not yet comprehensively documented from the perspective of modern language documentation[1]. This is not to deny the existence of classic language documentation[2] undertaken, for example, by Dutch/German

---

[1] Himmelmann's(2006:v) idea of language documentation to produce 'a lasting, multipurpose record of a language' and 'a comprehensive record of the linguistic practices characteristic of a given speech community'

[2] 'Classic' or 'traditional' documentation refers to the type of documentation typically done by means of pen and pencil, without rich multimedia datasets and without using much digital technology. This is typically

Christian missionary linguists and anthropologists, who had already started their research in Indonesia in the 19th Century. In western Indonesia, the famous H.N. Van Der Tuuk (1824–1894) worked on a grammar and bilingual dictionary of Batak. In Flores, central Indonesia, J.A. Verheijen (1908–1997) worked on the ethnobiology-linguistic documentation of Manggarai and published a bilingual Manggarai dictionary, text collections, and a book on Manggarai ethnography (Verheijen 1967, 1977c, Verheijen 1977b, Verheijen 1977a, Verheijen 1982, Verheijen 1990, Verheijen 1991). In eastern Indonesia, Petrus Drabbe (1887–1970) worked on the languages of Tanimbar and Southern New Guinea, and published grammatical descriptions and texts (Drabbe 1950, 1954, 1955).

Traditional linguistic research in Indonesia has also been undertaken by the Summer Institute of Linguistics (SIL) International. They conducted extensive linguistic research across Indonesia for more than four decades before significantly reducing their work by the early 2000s.[3] The Indonesian government through *Badan Pengembangan dan Pembinaan Bahasa* (The Language Development and Fostering Agency) has also undertaken language documentation. Primarily politically motivated by nation building for Indonesia's unity, most of their efforts in the period 1975–2007 were devoted to the research and development of Bahasa Indonesian and Malay varieties (69.7%) and only a third of them (30.3%) related to local vernacular languages (Arka, 2013:89).[4] In addition, the distribution of sites of their documentation is rather skewed, with 69% of their publications concerning languages of

---

the type done in earlier periods of documentation. It is unclear at this stage whether this type of documentation is still practiced as in the old days without being supplemented by some modern documentation techniques (i.e., without using any digital technology such as a modern digital recorder). The main deficiency of such documentation is that its non-digital outcomes (e.g., textual corpora) are not usable/processable using current techniques in the quantitative approach to corpus linguistics.

[3] SIL's research is undertaken in conjunction with bible translation and Christian missionary activities. SIL therefore walked a fine or fuzzy line of research, open to the possibility of distortive interpretative labelling, e.g., carrying out activities classified by the Indonesian authority as a 'threat' to the established dominant religious grouping. This led to SIL being evicted from south Sulawesi (Arka 2013), and its operational permit was not renewed in other parts of Indonesia.

[4] These figures are based on my earlier research (Arka, 2013) and were current at the time it was written. Present developments show that more attention has been given to local languages, but determining exact figures for the different categories is a matter of further investigation,

western Indonesia even though western Indonesia is home to only 31% of the total number of languages in Indonesia.

There are also studies on local languages undertaken by Indonesian academic staff and students across different universities in the country. They include research reports and theses completed as part of their degrees in linguistics and language studies. However, these are largely unpublished and are often not publicly available online (as they were typically submitted only in hard copy to the linguistics/language departments without further cataloguing); their precise number is, therefore, unknown.

While classic language documentation has been done mainly by Indonesians, modern language documentation in Indonesia has been undertaken mostly by foreign linguists (Sawaki and Arka 2018). Modern documentation in Indonesia in the last two decades includes four projects funded by the Volkswagen Foundation's *Dokumentation bedrohter Sprachen* (DOBES) in Totoli, Ujir, Wooi, and Central Papuan Languages; projects funded by the Endangered Languages Documentation Programme (ELDP) in Kata Kolok Bali, Ratahan, Tonsawang, Rongga, Alor-Pantar, Moor, Marori, Smerky, and Semeuleu; and two projects funded by the AHRC (Arts and Humanities Research Council) on Enggano. These documentation projects have resulted in rich multimedia corpora, available publicly through online archives including the DOBES Archive[5], the Pacific and Regional Archive for Digital Sources in Endangered Cultures (PARADISEC)[6] and the Endangered Languages Archive (ELAR)[7].

In summary, there has been considerable traditional language documentation in Indonesia since the 19th century, predominantly in the areas of core grammar and skewed towards the languages of western Indonesia. In modern research, we have seen increasing

---

[5] https://archive.mpi.nl/tla/islandora/object/tla%3A1839_00_0000_0000_0001_305B_C
[6] https://paradisec.org.au
[7] https://www.elararchive.org/

interest in modern language documentation on the minority and endangered languages of Indonesia, mostly led by foreign linguists via internationally funded projects, including DOBES, ELDP and AHRC. Given that there remain many minority languages, typically marginalised and endangered, and often in remote locations in Indonesia which are not yet properly documented, additional documentation efforts are certainly needed.

This need for additional documentation raises the issue of local capacity building and whether such efforts should be led by local agents, and ideally, members of the speech communities under investigation. In the next section, I address these issues based on my field experience in the Enggano Project[8]. I first discuss the dynamic nature of modern Indonesia from the perspective of minority speech communities, and the complex challenges they face. Then, in the context of local capacity building from an academic perspective, I highlight how recent COVID-19-related developments provide a momentum towards socially responsible research for long-term mutual benefits.

3   Challenges for minority speech communities in Modern Indonesia

Indonesia has undergone significant changes across different ecological dimensions (socio-political-cultural-economic) since the late 19th to early 20th centuries, which was when Dutch/German missionaries started their language documentation. In particular, since Soeharto's New Order Era (i.e., late 1960s), there have been unprecedented changes affecting the wellbeing of minority speech communities. I will illustrate these changes based on my fieldwork with minority speech communities in three regions across Indonesia: Flores (central Indonesia), Merauke (eastern Indonesia) and Enggano (western Indonesia).

The first significant change relates to language policy. Indonesian language policies have been politically driven by the need to maintain and enforce the unity of a nation-state,

---

[8] https://enggano.ling-phil.ox.ac.uk/

and this 'Indonesianisation' was particularly harsh during Soeharto's authoritarian regime. The government adopted a centralised language management approach (cf. Moeliono, 1994196). Its implementation[9], combined with other ecological factors (discussed below), negatively impacted the wellbeing of minority languages, leading to their accelerated marginalisation and endangerment in recent decades. Given the history of Indonesian politics and the current status of Indonesian as the official, formal and unifying language (*bahasa persatuan*), the use of Indonesian will remain an important part of the strategy of the central government in its nation building effort, and in controlling its territorial integrity in years to come (see Arka (2013) for further discussion, including the 'cognitive filter' and strategic issues in language management in Indonesia).

A second, unprecedented change in modern Indonesia that has impacted physical and socio-cultural-linguistic ecologies has been caused by economic developments. In this context two elements have been prioritised by the government: modern education (including literacy) and infrastructure programs. These two priorities have had far-reaching consequences on the wellbeing of minority speech communities. Literacy education throughout Indonesia, which includes English in an increasingly globalised world, has resulted in unstable multilingualism with a shift towards Indonesian at both an individual and societal level. Younger generations, particularly in urban areas, have become increasingly multilingual with Indonesian becoming more dominant over time—an unhealthy sociolinguistic situation for the intergenerational transmission of local languages. This even applies for more predominant languages, like Javanese, in urban settings (e.g., see Ravindranath & Cohn, (2014).

---

[9] Oftentimes, implementation was oppressive at the grassroots level. For example, local Rongga children were physically punished when they used their local language in the classroom (Arka, 2005).

Development (or *pembangunan*) was the focus and slogan of Soeharto's three-decade government, formulated in the so-called Pelita (*Pembangunan Lima Tahun*, or Five-year Development plans). The massive investment across sectors resulted in a booming economy and improved prosperity. In fact, this massive investment has continued under the current president, Jokowi. His election campaign, which focussed on infrastructure development, brought him to power in two free elections in Indonesia in 2014 and 2019. Under the development banner, new roads and ports (i.e., sea, air) have been built, opening up and strengthening the interconnectivity of remote places across the archipelago. This infrastructure development has indeed led to new business opportunities and therefore triggered the accelerated mobility of people around the country.

Relevant to our discussion here is the mobility of people to previously remote and isolated areas (e.g., villages in Merauke close to the border with PNG). Access to these places has improved due to developments in land, sea and air transportation. Travel has also become cheaper due to healthy competition in modern Indonesia's free economy. Consequently, there has been an increase in contact situations between local minority speech communities and outsiders who are typically part of a dominant ethnic group, such as Javanese and Buginese/Makassarese. This intensified contact, as attested from my fieldwork sites in central Flores, Merauke (Papua) and Enggano (western Indonesia), is responsible for the marginalisation and endangerment of the local languages. Of course, the negative linguistic impacts of outsider contact is not unique to Indonesia—the opening up of previously isolated communities, which has resulted in language endangerment, has also been reported in Thailand, for instance in the case of Ugong (Bradley 1989).

The challenges faced by minority speech communities are huge, and language endangerment is only one of the side effects of the two aforementioned changes. Life in modern Indonesia is increasingly competitive across different dimensions with day-to-day

basic economic needs being the immediate concern of locals. Addressing language-related issues, such as proper language documentation and developing language maintenance and/or revitalisation programs, is not a priority for local communities. Based on my field research experience in eastern Indonesia, there are at least four common issues and challenges to undertaking a language documentation program for language maintenance/revitalisation in the region (Arka, 2013):

1. *Content issues*: the creation of descriptive materials on the language, such as grammars and other pedagogical resources.

2. *Participation issues*: encouraging and ensuring active participation of speech community members in language maintenance programs.

3. *Support issues*: providing long-term institutional and/or organisational and financial support and incentives for language maintenance programs.

4. *Capacity building and leadership issues*: recruiting and training local leaders and community members so that they can do language maintenance and/or revitalisation programs themselves, including finding external support as necessary.

Linguists are typically trained to address the content issue in (1) and not the remaining ones in (2)–(4), since they can be difficult to deal with and require collaboration with different stakeholders (see the related discussion on support issues in Arka, 2013). In what follows, I discuss issue (4), *Capacity building and leadership,* and frame it in relation to issue (1), C*ontent issues,* via modern (ethno)linguistic research, including in a (pre-)COVID-19 context.

4   Ethnolinguistic research and capacity building

In connecting (ethno)linguistic research with capacity building, we are faced with the following questions: (a) what strategy is needed to enable community participation in language documentation and overcome any challenges of continuing to do it in modern Indonesia in the long-term? (b) what area of (ethnolinguistic) research in language documentation is to be undertaken for the purpose of (a), particularly in the context of COVID-19. I address these questions below with reference to my language documentation projects in Flores, Merauke and Enggano.

4.1   Capacity building and local leadership.

Long-term active participation requires local agents in speech communities to drive language documentation activities, especially follow-up maintenance and revitalisation programs, if any. Ideally, local agents are enthusiastic and are young language activists with a good educational background, since this enables fast hands-on training in complex language documentation tasks and local leadership activities. Initial training requires building expert knowledge and skills, including a capacity to document speech events using a modern video recorder, as well as data processing activities on a laptop/computer, such as metadata entry or transcription activities, especially via specialised software like ELAN[10]. In my previous three documentation projects in Indonesia, I have been fortunate in sourcing a young, local assistant on each fieldwork site who was enthusiastic and willing to receive documentation training. As part of capacity building efforts that reach a wider Indonesian audience, I have organised documentary linguistic workshops in Indonesia with other linguists that were funded by DOBES and ELDP and attended by linguistics students and language activists.

---

[10] https://archive.mpi.nl/tla/elan

However, the training required for long-term capacity building—that is, enabling local community members to handle the four challenges mentioned in the previous section—involves knowledge, skills and experience beyond the documentary linguistic modules of video camera handling and/or data

Figure 1. Ecotourism in the Marori territory of Wasur, Merauke, West Papua, Indonesia.

collection and processing. While these skills are essential, leadership and management qualities—such as stakeholder engagement and communication, or end-to-end project management, including from project conception and grant application through to project execution and completion—are not incorporated in existing training workshop programs in Indonesia, at least not extensively[11]. Nonetheless, based on my work with local communities, such leadership and management skills are nurtured through a continuous mentoring process—and this takes time. For example, my two Marori documentation projects in west Papua began in 2009, but the fruits of local leadership training and mentoring only became visible in 2019—ten years after project commencement. Building on linguistic-ethnobiological documentation, the local Research Assistant (RA), Agustinus Mahuze, used his knowledge, skills, networks and team-work experience to establish *Mahuze Mandiri*, a local community-based non-governmental organisation (NGO), which continues the documentation work and also runs an ecotourism project as an innovative way of making use of the documentation materials. These projects include a recreation centre with a swimming facility, and walking tracks featuring *Musamus*, or termite mounds (**Error! Reference source not found.**), and an ethnobotanical garden. His NGO also successfully applied for a grant to

---

[11] Grant writing was part of the 2007 Ubud training.

have further cultural documentation (e.g., endangered dances with dancing workshops). The local regency government of Merauke has also provided support for the development of eco-tourism in Kampung Wasur and has used it as a model for similar community-based projects in other tribes in the wider region of Merauke.

Another effective strategy is to connect language documentation outcomes with community-based entrepreneurships, which have clear local economic benefits. This comes with its own unique set of challenges since it involves excellent planning, management skills and substantial financial investment. Collaborations with, and support from, relevant departments in the local government are also essential, in addition to the involvement and input from locals themselves.

4.2   Linguistic research and language documentation

Linguists applying for grants to do fieldwork-based research typically frame their documentary research in the context of interdisciplinary issues, questions and debates. In this subsection, I discuss embedding linguistic research in documentation projects for the benefit of both academic and local communities, with illustrations based on my Indonesian documentation projects.[12] I conceive documentation research to necessarily have linguistic strands (with interconnected descriptive, theoretical and comparative explorations) and also bundled with clear practical implications, such as policy making and local capacity building. Pure language documentation alone—that is, textual or multimedia data collection and processing—is not competitive in the context of grant applications even for funding agencies specifically devoted to language documentation, like ELDP.

---

[12] Ideally in a language documentation project, there is a process of gathering goals from community members and talking to foreign funding bodies to accommodate community goals. However, in my experience, minority communities typically have no awareness of language endangerment. Even when they do, they tend to have no clear idea how to maintain/revive their languages, and what assistance is needed.

Therefore, linguistic research in language documentation and endangerment must be multi-dimensional and should approach language as a means of communication in the social and cultural space of human interaction and existence across time. I identify at least four intertwined dimensions that academic research should consider while doing language documentation: (a) internal structural grammatical systems; (b) socio-cultural dynamics in a contemporary context; (c) history and contact in extended language ecology; and (d) motivation. Each dimension is discussed in turn below.

4.2.1   The grammatical and lexical dimension

Traditionally, this dimension relates to structural linguistics across the core modules of phonology, morphology and syntax. The nature of this research is typically mixed and may involve descriptive, typological and/or theoretical strands, with different weightings depending on the focus of a given project, and the research team's interests. All my documentation projects have been designed with the grammatical and lexical/lexicographic dimensions in mind. For instance, the Enggano Project consisted of both a historical and typological strand in its investigation of shared and unique linguistic properties of the language within the broader Austronesian family. We have won another 2-year project to focus on the Enggano lexicon, with the outcomes including a learner's dictionary. The Rongga documentation project, on the other hand, was more descriptive in its linguistic outcomes, and included a grammar of Rongga (Arka 2016) and papers on phonetics and phonology (Suparsa and Arka 2006, Suparsa 2009). Other descriptive outcomes include sociolinguistic descriptions (Sumitri and Arka 2019, Arka 2011a, 2010) and dictionaries (Arka 2012, 2011b).

As a general principle, the descriptive linguistic strand is prioritised for good reason. Descriptive work is logically the first entry point, and a necessary prerequisite, before complex comparative and theoretical linguistic analyses can be performed. In addition, the

outcome of descriptive work, such as grammar sketches, extended word lists or short dictionaries, is typically more accessible to wider audiences, including local speech communities.[13] Further, the outcome of descriptive work directly and easily feeds into practical and applied documentation strands. For example, descriptive phonetics-phonology is essential for the development of practical orthographies, which may be urgently needed for developing teaching materials. Likewise, basic grammatical properties, such as verbal voice morphology or preposition usage, are essential components of teaching materials for local school children. In short, linguistic descriptive works allow for a quick and wide research impact on the community.

4.2.2   *The sociolinguistic and cultural-political dimension*

The sociolinguistic and cultural-political dimension straddles multiple domains including sociolinguistics, ethnolinguistics and linguistic anthropology. The issues here are often intertwined with history, such as past contact situations with other ethnolinguistic groups, which is discussed further below. A critical sociolinguistic concern in Indonesia relates to unstable diglossic or multiglossic multilingualism involving prolonged inter-/intra-group communication, which results in language shift and language endangerment. I have discussed these issues in the context of Rongga in Central Flores (Arka (2005, 2011a); Suparsa & Arka (2009)), and in the context of Marori (Arka 2013). Commonalities across these regions include the cultural and political dominance of regional ethnic groups (e.g., the Manggarai in western Flores, and the Marind in Merauke), which has adversely affected the vitality of minority languages in those areas. Such regional dominance is ongoing and has been historically present since before the official formation of modern Indonesia in 1945.

---

[13] While overseas funding bodies are often interested in selective linguistic aspects of theoretical/academic significance, they are also happy to support research activities for descriptive and practical outcomes provided that these outcomes are well integrated, and justified in their integration, with the academic goals of the projects.

Evidence of the negative impacts of cultural dominance on minority speech communities is clear, for example, in the domain of language and identity. In Merauke, minority communities like Marori adopt the identity, cultural values and related belief systems of the dominant group, the Marind, resulting in the shrinking, blurring and ultimate disappearance of certain aspects of a Marori's identity and belief system. This can be seen in Marind forms of contemporary address via the kinterms *namek* 'brother' and *namuk* 'sister', which are used widely in Merauke, including by the Marori people; the Marori terms of these are *mborumen* and *mbondombur* respectively.

The negative impact of regionally dominant groups has been compounded in contemporary Indonesia due to modern democracy. The increasing autonomy of local governments mean that minority ethnolinguistic groups, comprising three-to-four thousand people like the Rongga people in western Flores, are not typically in control of their local jurisdiction. They are disadvantaged, for example, in their local literacy development since resources are allocated to the dominant ethnolinguistic group, namely the Manggarai, thus jeopardising the ethnolinguistic vitality of the Rongga.

However, a community's ethnolinguistic vitality does not necessarily correlate with the number of speakers of the language. This is the case with Loloan Malay, spoken by an ethnic minority in Bali: the language has high ethnolinguistic vitality due to their distinctive language ideology (Sosiowati, Widiastuti, and Arka 2017, Sosiowati et al. 2019), despite being surrounded by, and in constant contact with the Balinese, which is further discussed below.

4.2.3   The ecological dimension

This dimension falls within environment and sustainability theory, and its application to language endangerment and maintenance has been discussed under the rubric of the Extended Ecology Hypothesis and 'situated ecologies' (Steffensen and Fill 2014). The impact

of change in a physical ecology on the wellbeing of minority languages is evident in all three languages targeted in my documentation projects. The common cause of this change is the opening up of previously isolated communities and/or the relocation of minority groups which results in intensified contact with more powerful ethnolinguistic groups. The Marori people for example, relocated to their current place in Kampong Wasur in the early-to-mid 20th century. In their original ecology, they lived in a forest. Their relocation to a village on the trans Papuan road gradually detached them from their original forest habitat. In addition, economic developments in modern Merauke have led to massive deforestation, resulting in the loss of rich forest-related cultural practices and indigenous ethnobotanical knowledge.

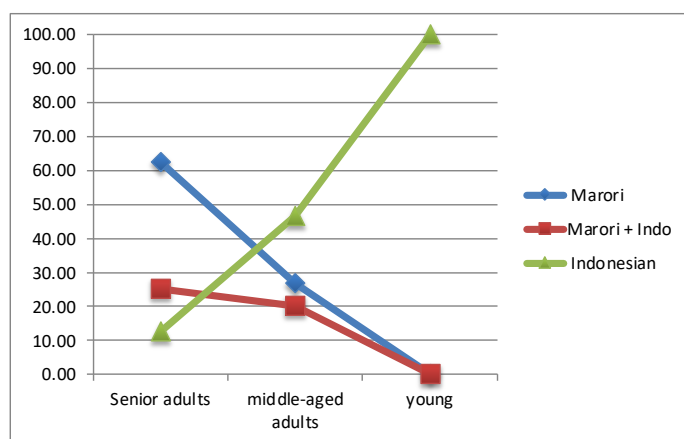New economic opportunities also attracted more non-Papuan outsiders (typically from Java, Sulawesi and Maluku) to live permanently with local Marori people in the same space of Kampong Wasur. Overtime, the number of outsiders in Kampung Wasur steadily increased, disturbing the local social networks. There are currently no longer the close-

Figure 2: Languages used by the Marori people (across different groups) when they speak to each other informally in the village.

knit social networks necessary for a healthy transmission of the Marori language in Kampung Wasur. Consequently, it is not surprising to see that there is a complete shift to Indonesian among the young generation as confirmed by my ethnographic work and current survey. As seen in Figure 2, Marori is still used by around 62.5% of senior adults (aged approximately 50 years old and above) who participated in the survey when they communicate with each other informally in the village, however Marori has been completely replaced by Indonesian in the same context among the young generation (those aged 20 years old and under). The language endangerment of Marori in this case was caused by the influx of migrants to

Kampung Wasur, the native territory of the Marori people, leading to an intensified contact situation which significantly weakened Marori's linguistic ecology. The accelerated change has occurred in the last two to three decades.

Intensified contact with outsiders is only one (external) factor with negative linguistic consequences impacting inter-generational transmission of minority languages (as seen in Marori). An equally critical factor is the motivation of the speech community (both collectively and individually) to maintain their languages. This is a topic in the cognitive domain of language learning/maintenance, to which we now turn.

4.2.4   *The cognitive motivational dimension*

Motivation is a complex notion concerning the *direction* and *magnitude* of human behaviour (e.g., why we do something, how long we are willing to sustain the activity and how hard we are pursuing it) (Dörnyei and Ushioda 2013:4). It plays a central role in first/second language (L1/L2) learning (Dörnyei and Ushioda 2013). For our discussion in this subsection, the relevant point is the motivation of members of minority speech communities to use their (L1) languages in domestic and public settings to allow a healthy L1 transmission from one generation of speakers to the next.

My recent research in the Indonesian context reveals the instrumental nature of motivation in language maintenance: the mother tongue is distinctively emblematic, and therefore instrumental, to augment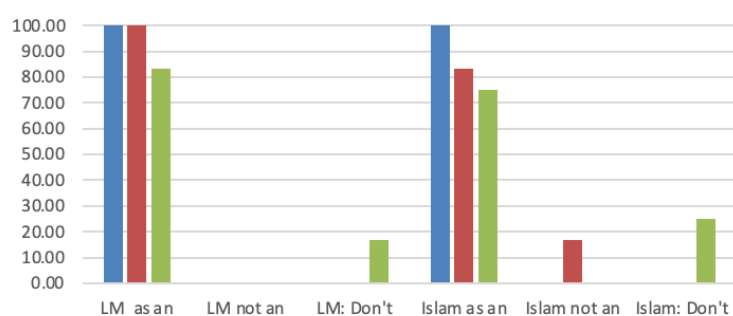 the speech community's ethnolinguistic identity in the relevant region. This is exemplified by Loloan Malay (LM) in western Bali. This language is distinctively emblematic as an identity of the

Figure 3: Linguistic and religious identities of the Loloan (Malay) people in western Bali.

Loloan (Malay) people in contrast to the Balinese language, which is distinctly symbolic/emblematic as the identity of the Balinese people. Importantly, however, the ethnolinguistic identity of the LM people is closely linked with their religious identity as Moslems. This is supported by the evidence shown in **Error! Reference source not found.**, which is based on survey data I collected in the region in 2019[14]. The survey asked the following questions to three generational groups (represented by different colours in the graph): (a) *Do you consider Loloan Malay is an important distinctive identity, distinguishing you/your Loloan community from others?* (b) *Do you consider Islam is an important distinctive identity, distinguishing you/your Loloan community from others?* As demonstrated by the left-most bars in **Error! Reference source not found.**, the LM respondents across all generations overwhelmingly consider both speaking their mother tongue (LM) and practicing Islam as part of their distinctive and collective identity. This finding provides evidence that multiple, distinctive identities are simultaneously functional in inter-group relations. That is, the LM people distinguish themselves from the Balinese people, who are Hindus and speak the Balinese language. Their Moslem identity feeds into, and also augments, the emblematic strength of the LM language as a distinctive identity of the LM people in Bali. It should be noted that this Moslem identity alone is not distinctive for the LM people in the larger context of Indonesia given that a large majority of Indonesians are Moslems.

---

[14] The survey undertaken in 2019 was part of a small collaborative research project with colleagues Prof Sosiowati and Ayu Widyastuti from Udayana University, and funded by an Udayana University grant in 2018-19. There were 45 local survey respondents comprising of different age groups: senior adults, middle-aged, and teens/young adults.

The coupling of religious and ethnolinguistic identities means that a 'small' speech
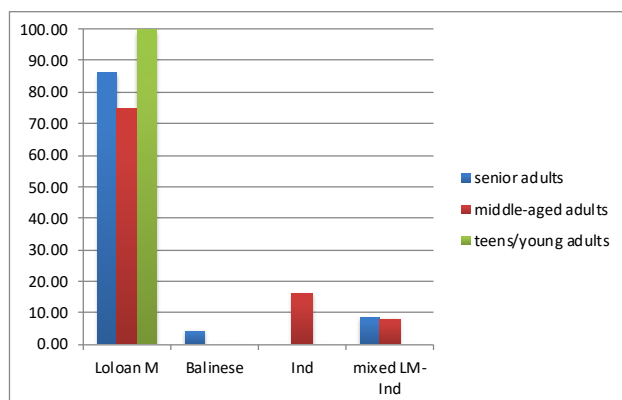

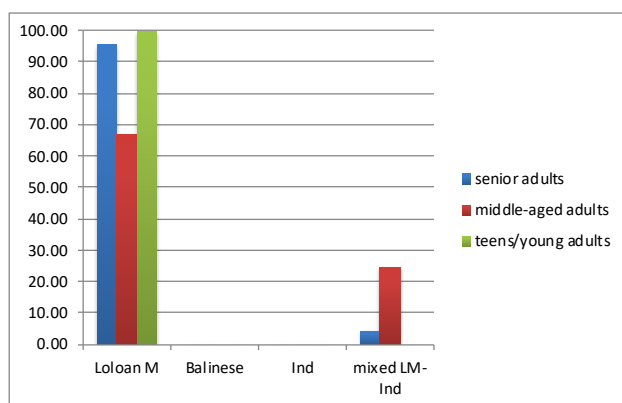*Figure 4: Languages used by Loloan Malay people when speaking with parents*

Figure 5: Languages used by LM community speaking with each other at social events in the village.

community such as Loloan Malay (approximately 2,500 speakers) does not necessarily have to be 'weak' in its linguistic vitality. This is because the awareness of one's religious-ethnolinguistic identity is constantly being re-enforced in domestic and public settings as a way to distinguish oneself from dominant groups. Evidence comes from the high frequency use of LM in domestic and public domains (Figure and Figure respectively) as a way to distinguish themselves from Balinese identities. Thus, while small in size, LM community members still demonstrate quite a healthy intergenerational transmission of language.

The strong instrumental motivation of the type seen in LM is, however, very rare in Indonesia, and perhaps elsewhere also.[15] For this reason, it could benefit from additional research in order to understand other potential factors motivating the high vitality of LM in spite of its minority status. As mentioned earlier from the survey data in **Error! Reference source not found.**, this motivation seems to arise from the tight link between LM's ethnolinguistic identity (i.e., a distinct language from Balinese) and emotionally charged religious identity (i.e., being Moslem). This analysis is supported by comparative data from

---

[15] This is perhaps because other areas of Indonesia are made up of mostly Moslem populations, so they do not have the distinction from Hindus as in Bali. Further research is needed to determine whether there are minority Moslem communities in certain parts of eastern Indonesia where Christians are the dominant group.

Marori and Enggano. While Marori and Enggano people equally consider language as an
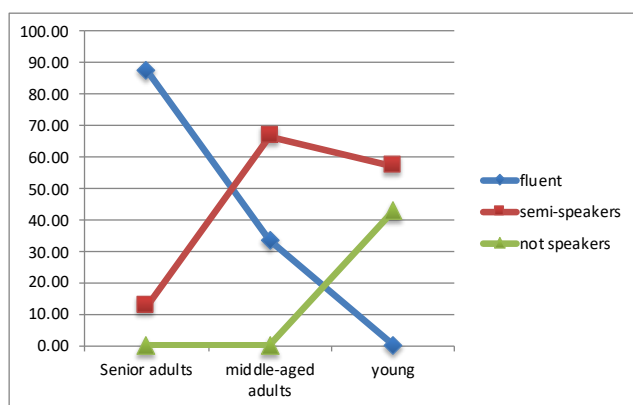

Figure 3: Self-reporting language proficiency across generations in Marori

important ethnolinguistic identity, this does not trigger the same strong collective motivation in language maintenance as in LM. For instance, when asked about future language prospects, all young Marori people said they would not impose or teach Marori to their children when/if

they had a family. This is not surprising given the fact that young Marori are not fluent in

their minority language (cf. Figure 3).

5 COVID-19 and community engagement

The discussion thus far has been focussed on language documentation research in a pre-COVID-19 context, noting that the pandemic hit Indonesia around February in 2020. This discussion is important as COVID-19 unexpectedly brought new challenges to doing (linguistic) fieldwork but resulted in a positive lesson learned about community engagement and capacity building, including leadership development and the decolonisation of field linguistics in Indonesia. At the time of writing this book chapter, the situation has gradually improved with researchers now allowed to travel to Indonesia for fieldwork, though people are urged to remain vigilant against the virus (e.g., continue to wear masks indoors and when in a crowd). In this final subsection, I briefly report and reflect on lessons learned from such COVID-19 restrictions, including the challenges and opportunities of (collaborative) research with local communities in Indonesia with a focus on the Enggano Project.

Community engagement is a central component of the Enggano Documentation Project since it involves locals in language documentation processes (i.e., data collection and

processing) and literacy material development for school children. Following the COVID-19 outbreak, doing close face-to-face fieldwork on Enggano Island was impossible between 2020 to mid-2022, and thus we resorted to remote fieldwork via our local and regional collaborators in Bengkulu. Bengkulu is the capital town of the Bengkulu province, the administrative province of Enggano. The town hosts the University of Bengkulu, which we collaborated with for the applied linguistics research strand of the project. Bengkulu has relatively good internet connection making remote fieldwork a possibility.

Our remote fieldwork was undertaken in two modes. The first mode was online via Zoom and included Enggano elders, who were invited to Bengkulu, and core research team members from Bengkulu University, Oxford, United Kingdom and Canberra, Australia. One key challenge related to accounting for time zone differences between the UK, Australia and Indonesia. Otherwise, the online mode worked well for transcription, grammar checking, elicitation tasks, interviews, and focussed group discussions (FGDs). Zoom functionalities such as screen sharing allowed all participants, in particular the Enggano elders, to watch video recording content and provide instant answers, corrections and feedback. In addition, Zoom supported high quality audio-video recording, allowing us to collect new data and to work on existing recordings (e.g. adding transcriptions). While the internet in Bengkulu is sometimes unstable, the weekly online remote fieldwork sessions via Zoom have otherwise been successful and productive in terms of remote data collection and processing. While we cannot fully replicate the experience of doing close on-ground fieldwork via Zoom, we (researchers) and the local language consultants mostly found remote fieldwork via Zoom convenient and had no problems in any aspect including calculating the payments for the local participants.

The second mode is 'indirect' or proxy fieldwork: our local RA and Research Collaborators (RCs) from Bengkulu went to the field on our behalf to do fieldwork. This was

only possible after the government relaxed the COVID-19 restrictions slightly allowing limited local mobility under certain conditions. This proxy fieldwork required the RA and RCs to attend remote training via Zoom first where clear instructions were given in terms of what to do in Enggano. Fortunately, the local RA had already received training (e.g., video recorder handling) prior to the COVID-19 outbreak. Given the impossibility for the project Chief Investigators (CIs) to go to the field in person, this proxy fieldwork mode was the only sensible option during COVID-19, in addition to the online mode mentioned earlier. The proxy fieldwork mode has been successfully carried out in line with modern language documentation practices to collect naturally occurring texts, record wordlists, facilitate FGDs with wider participants, and trial our teaching materials.

Interestingly, proxy fieldwork opens up new opportunities to decolonise linguistic research in Indonesia (cf. (Sawaki and Arka 2018, Arka 2018), with our local RA and RCs making greater contributions to the project's research activities. In a way, the Indonesian RCs, who are non-native Enggano speakers based in Bengkulu, represented the project's foreign CIs and were able to engage with the locals more. Thus, proxy fieldwork has caused a shift in the nature of local engagement and this shift has led to greater leadership responsibilities for local RCs.

While this shift supported the productivity of the project and created local capacity building opportunities, it did not come without its own set of challenges. For instance, not all linguistic research tasks, such as specific socio-cultural data collection methods via in-depth individual interviews or FGDs, can be easily and smoothly exercised by the local RA and RCs. These tasks require professional input as well as specific academic/linguistic skills and knowledge that can only be gained through formal education and experience. Capacity building to achieve the qualities comparable to those possessed by CIs cannot be realistically

achieved through project-based training of RAs/CIs as discussed above, but by bundling such training as part of more advanced formal tertiary education in the relevant disciplines.

While I envisage the continuation of both online and proxy fieldwork modes in a post-COVID-19 era, given its obvious benefits, the direct engagement of CIs through face-to-face meetings onsite is still a necessary component of fieldwork. It can both complement and augment proxy and online fieldwork modes. Remote fieldwork cannot replace the physical presence of CIs, which is necessary for achieving the deepest level of engagement and insight in the field. My experience is that the same interpersonal exchanges between locals and CIs in the field is hard, if not impossible, to be replicated remotely.

6   Concluding remarks

This chapter discussed multidimensional aspects of language documentation and maintenance, including capacity building, in an Indonesian context. It highlighted the superdiversity of Indonesia's ethnolinguistic landscape, and demonstrated that minority languages were increasingly marginalised and endangered due to complex socio-historical-political factors. While Indonesia provides ample opportunities for linguistic research, continuing from the classic documentation which began in the late 19th century, its rich linguistic diversity remains otherwise largely under-documented, and modern documentation is a matter of urgency given the fragility of many minority languages. I have argued for the need to bundle modern language documentation projects with local capacity building efforts to help minority speech communities cope with language endangerment and maintenance challenges. Based on my own language documentation projects, I also argued that critical capacity issues, such as local engagement/participation, leadership and motivation, require more than just linguistic training.

The chapter also discussed how COVID-19 travel restrictions posed a challenge to language documentation and community engagement. I have discussed how the online methods and proxy modes opened up an opportunity for better capacity building and decolonisation of field linguistics in Indonesia. However, it remains to be seen to what extent different types/modes of documentation can be tailored productively for different research/documentation activities including community engagements to give the maximum benefits for all parties, particularly the local stakeholders.

References

Arka, I Wayan. 2005. "Challenges and prospect of maintaining Rongga: a preliminary ethnographic report." In *Proceedings of the 2004 Conference of the Australian Linguistics Society*, edited by Ilana Mushin. http://deneb.library.usyd.edu.au:8080/handle/2123/138.

Arka, I Wayan. 2010. "Ritual dance and song in language documentation: vera in Rongga and the struggle over culture and tradition in modern Manggarai-Indonesia." In *Language endangerment in the Austronesian world: challenges and responses*, edited by M. Florey, 90-109. Oxford: Oxford University Press.

Arka, I Wayan. 2011a. "Komplesitas pemertahanan dan revitalisasi bahasa minoritas di Indonesia: pengalaman proyek dokumentasi Rongga, Flores (Complexity of maintenance and revitalization of minority languages in Indonesia: Field experiences from Rongga Documentation Project, Flores)." *Masyarakat Indonesia, Majalah Ilmu-Ilmu Sosial Indonesia (Indonesian Journal of Social Sciences)* XXXVIII (1):35-56.

Arka, I Wayan. 2011b. *A Rongga-English dictionary*. Jakarta: Atma Jaya University Press.

Arka, I Wayan. 2012. *Kamus bahasa Rongga-Indonesia*. Jakarta: Atma Jaya University Press.

Arka, I Wayan. 2013. "Language management and minority language maintenance in (eastern) Indonesia: strategic issues." *Language Documentation and Conservation* 7:74-105.

Arka, I Wayan. 2016. *Bahasa Rongga: deskripsi, tipologi dan teori (Rongga: description, typology and theory)*. Jakarta: PKBB, Universitas Katolik Atma Jaya.

Arka, I Wayan. 2018. "Reflections on the diversity of participation in language documentation." *Language Documentation & Conservation* Language Documentation & Conservation Special Publication No. 15, edited by Bradley McDonnell, Andrea L. Berez-Kroeker, and Gary Holton:132-139.

Bradley, David. 1989. "The disappearance of the Ugong in Thailand." In *Investigating obselescence*, edited by Nancy C. Dorian, 33-40. Cambridge: Cambridge University Press.

Dörnyei, Z, and E Ushioda, eds. 2013. *Teaching and researching: Motivation. (2nd ed.)*. Harlow: Routledge.

Drabbe, Peter. 1950. *Lijst met woorden uit Papua-talen* Leiden: .

Drabbe, Peter. 1954. "Talen en Dialecten van Zuid-West Nieuw-Guinea." In *Micro-Bibliotheca Anthropos 11*. Posieux (Freiburg). Anthropos-Institut.

Drabbe, Peter. 1955. "Spraakkunst van het Marind." In *Studia Instituti Anthropos II*. Vienna and Mödling: Anthropos Institut.

Himmelmann, Nikolaus P. 2006. "Language documentation: What is it and what is good for?" In *Essentials of Language Documentation*, edited by Jost Gippert, Nikolaus P. Himmelmann and Ulrike Mosel, 1-30. Berlin: Mouton de Gruyter.

Ravindranath, M , and A. Cohn. 2014. "Can a language with millions of speakers be endangered?" *Journal of the Southeast Asian Linguistics Society* 7 (64-75).

Sawaki, Yusuf, and I Wayan Arka. 2018. "Reflections on linguistic fieldwork and language documentation in eastern Indonesia." *Language Documentation & Conservation* Special Publication No. 15, edited by Bradley McDonnell, Andrea L. Berez-Kroeker, and Gary Holton:256-266.

Sosiowati, I Gusti Ayu Gde, I Wayan Arka, I Nyoman Aryawibawa, and Ni Made Ayu Widiastuti. 2019. "Domain change and ethnolinguistic vitality: Evidence from the fishing lexicon of Loloan Malay." *Language Documentation & Conservation* 13:586–617.

Sosiowati, I Gusti Ayu Gde, Ni Made Ayu Widiastuti, and I Wayan Arka. 2017. "Constructing Identity and Language Vitality Of Minority Language: Evidence From Loloan Malay." In *Proceedings of the 8th International Conference on Austronesian and non-Austronesian languages and linterature* edited by I Nengah et al. Sudipa, 273-282. Denpasar: Udayana University Press.

Steffensen, S. , and A. Fill. 2014. "Ecolinguistics: the state of the art and future horizons." *Language Sciences* 41: 6-25.

Sumitri, Ni Wayan, and I Wayan Arka. 2019. "Power in/behind Language in Ethnolinguistic Perspective: The Dynamics of Ritual Tradition in Rongga, East Manggarai." *Mozaik*:669-674.

Suparsa, I Nyoman. 2009. "Fonologi Bahasa Rongga." PhD, Graduate Linguistics Programme, Universitas Udayana.

Suparsa, I Nyoman, and I Wayan Arka. 2006. "Segmentasi mb, nd, dan ngg sebagai Sebuah Segmen Pranasal Stop dalam Bahasa Rongga (Segmenting and analysis of mb, nd and ngg as pranasal stops in Rongga)." In *KOLITA 4*, edited by B Kaswanti Purwo, 18-21. Atma Jaya University: Pusat Kajian dan Bahasa Unika Atma Jaya, Jakarta.

Suparsa, I Nyoman, and I Wayan Arka. 2009. "Diglosia, Kontak Bahasa, dan Bahasa Minoritas: Sebuah Studi Awal Kata Serapan dalam Bahasa Rongga (Diglossia, language contact, and minority languages: a preliminary study of loan words in Rongga)." *Linguistik Indonesia* 27 (2):131–143.

Verheijen, J. A J. . 1977a. *Manggarai Texts. I, II, III. .* Ruteng: Provinsi SVD Ruteng.

Verheijen, J.A J. . 1990. *Manggarai Texts". XVI.* Ruteng: Provinsi SVD Ruteng.

Verheijen, J.A.J. 1977b. *Bahasa Rembong di Flores Barat.* Ruteng: SVD Ruteng.

Verheijen, J.A.J. 1982. *Dictionary of Manggarai plant names.* Canberra: Pacific Linguistics.

Verheijen, J.A.J. 1967. *Kamus Manggarai I (Manggarai-Indonesia).* 'S-Gravenhage: Koninklijk Instituut voor Taal- Land- en Volkenkunde.

Verheijen, J.A.J. 1977c. "The lack of formative in affixes in the Manggarai language." *NUSA* 5:35-37.

Verheijen, J.A.J. . 1991. *Manggarai dan Wujud Tertinggi.* Jakarta: LIPI-RUL.

Figures