Dynamic multilingualism and language shift scenarios in Indonesia

MAYA RAVINDRANATH ABTAHIAN & ABIGAIL C. COHN

This article appears in:

Endangered languages and the land: Mapping landscapes of multilingualism

Proceedings of the 22nd Annual Conference of the Foundation for Endangered Languages (FEL XXII / 2018)
Vigdís World Language Centre, Reykjavík, 23–25 August 2018

Editors: Sebastian Drude, Nicholas Ostler, Marielle Moser


Cite this article:

First published: December 2018

Link to this article: http://www.elpublishing.org/PID/4016

This article is published under a Creative Commons License CC-BY-NC (Attribution-NonCommercial). The licence permits users to use, reproduce, disseminate or display the article provided that the author is attributed as the original creator and that the reuse is restricted to non-commercial purposes i.e. research or educational use. See http://creativecommons.org/licenses/by-nc/4.0/

Foundation for Endangered Languages: http://www.ogmios.org
EL Publishing: http://www.elpublishing.org
Dynamic multilingualism and language shift scenarios in Indonesia

Maya Ravindranath Abtahian
Department of Linguistics • University of Rochester •
514 Lattimore Hall • Rochester, NY, 14627-0096, USA
[maya.r.abtahian@rochester.edu]

Abigail C. Cohn
Department of Linguistics • Cornell University •
159 Central Ave • Ithaca, NY, 14853-4701, USA
[acc4@cornell.edu]

Abstract

Indonesia, an archipelago of roughly 14,000 islands, is home to approximately 700 languages. The instatement and development of Bahasa Indonesia as the national language is widely described as a highly successful example of language planning across the multilingual archipelago, and simultaneously as the catalyst for the endangerment of scores of local languages. In an ongoing project examining language shift in Indonesia and focusing on languages with a million speakers or more, we seek to explore this complex language ecology (Mufwene 2017) and the nature of dynamic multilingualism (Musgrave 2014) and language shift scenarios in Indonesia at multiple levels of inquiry, including sociolinguistic interviews, Indonesian census data, and a language use questionnaire. Kuesioner Penggunaan Bahasa Sehari-hari. Here we use our questionnaire data to look specifically at the relationship between identity (ethnolinguistic and regional) and language use in Indonesia with an analysis of how speakers classify and name language varieties that they report. We study the socio-geographic effect of “inner” vs. “outer” island in the Indonesian archipelago, comparing our census data results with further insight gained from our questionnaire data.

Keywords: language shift; Indonesia; Bahasa Indonesia; language endangerment; census

Introduction

Indonesia is home to approximately 700 languages (roughly 10% of the world’s languages) and is an archipelago of roughly 14,000 islands, ranging in size from among the largest in the world (New Guinea, Borneo, Sumatra) to dots of land as yet unnamed. The instatement and development of Bahasa Indonesia as a national language following the founding of the Republic of Indonesia in 1945 is widely described as a highly successful example of language planning across the multilingual archipelago (Heryanto 1995: 5). At the same time, increasing use of Indonesian as a lingua franca and language of daily use is arguably the catalyst for the endangerment of scores of local languages. To frame this shift only in terms of language loss, however, is to miss the dynamic complexity of the language ecology (Mufwene 2017) of Indonesia. In an ongoing project examining language shift in Indonesia, we seek to explore this complexity and the nature of dynamic multilingualism and language shift scenarios in Indonesia through the analysis of sociolinguistic interviews, Indonesian census data (Abtahian et al., 2016a), and a language use questionnaire, Kuesioner Penggunaan Bahasa Sehari-hari ‘Everyday Language Use Questionnaire’ (Cohn et al. 2013; 2014).

We start with a brief snapshot of the linguistic situation in Indonesia and an overview of our project, largely focusing on the “big languages” of Indonesia, those with over a million speakers. We then address the need to model factors contributing to language shift, building on Himmelman’s (2010: 46) observation that language shift is rarely the result of just one or two factors but rather “results from the specific and complex constellation of a variety of such factors… an endangerment scenario.” Then we turn to the socio-geographic effect of “inner” vs. “outer” island in the Indonesian archipelago, comparing our census data results with further insights gained from our questionnaire data and addressing the observation made by Musgrave (2014: 87) that the linguistic situation in Indonesia is better viewed as “changing patterns of multilingualism, rather than as shifts of large populations from one language to another.”

Dynamic multilingualism in Indonesia

The relationship between land and water sets the backdrop for the complex language ecology of Indonesia. Leow (2016: 4, citing Andaya 2006), argues that Southeast Asia’s ecological landscape is uniquely fragmentary and diverse, “made up of promontories and islands and surrounded by bodies of water that have seen thousands of years of movement,” and that this is mirrored in the human dimension, where the existing linguistic plurality in the region is then layered with centuries of colonialism that “produced new hybrids out of plurality.” The last 70 years of language planning for Bahasa Indonesia have added a layer of complexity to this linguistic plurality, as repertoires continue to narrow in some cases, and expand in others. In some cases local languages are being replaced by regional koiné, for example as Rural Jambi is replaced by City Jambi (Yanti 2010: 2-3) or the increasing importance of Kupang Malay (Errington 2014). Sometimes the new variety is a local colloquial variety of
Indonesian, as in the rapid adoption of Jakarta Indonesian, not only in Jakarta but also in other major urban areas (see Kurniawan in progress, for discussion). At the same time, in some speech communities, particularly those with complex speech registers like Javanese, the shift is better described as shifting patterns of multilingualism (Musgrave 2014) that offer a medium for changing performative functions of language (Goebel 2018).

In thinking about what has been lost and gained in this shifting language ecology, much of the attention to date has been on languages with small speaker populations. This focus follows from Krauss’s (1992) criteria that any language with fewer than 100,000 speakers be considered at risk, which includes 88.2% (623) of Indonesia’s languages (Lewis et al. 2014). Work on documenting endangered languages in Indonesia has largely focused on languages with small speaker populations, albeit recognizing that, as Florey writes, “[r]estricting the definition of ‘endangered language’ to those languages with small speaker populations disguises the extent of the problem” (Florey 2005: 59), and “[i]n spite of their large speech communities, the Javanese, Sundanese, and Madurese languages are actually endangered in that some of their domains of usage are being taken over by Indonesian, and, to a lesser extent, in that they are not always passed on to the next generation.” (Adelaar 2010: 25). In our project, we turn our focus to the “big language” communities of Indonesia with over a million speakers, generally considered protected from large-scale language shift. This focus grew out of the observation that even in language communities like Javanese (the 10th most widely spoken language in the world) language use patterns are changing (Errington 1998; Poejosoedarmo 2006; Setiawan 2012), and a lack of intergenerational transmission is also reported (Smith-Hefner 2009; Setiawan 2012; Kurniashih 2006). Methodologically, a focus on big languages allows us to consider more broadly the language ecology in each community and to compare these across different language communities in Indonesia in order to understand what factors are generalizable.

In Table 1, we list the 21 languages in Indonesia with over a million speakers (Lewis et al. 2014), by population. As we are making comparisons between Malayic [M] vs non-Malayic [non] ethnolinguistic groups and “inner” island (Java, Bali) [I] and “outer” island (Sumatra) [O], so we also indicate these distinctions.

<table>
<thead>
<tr>
<th>Language</th>
<th>Region</th>
<th>Pop in M</th>
<th>(non)</th>
<th>inner/outer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Java</td>
<td>Java &amp; Bali</td>
<td>84.3</td>
<td>non</td>
<td>I</td>
</tr>
<tr>
<td>Sunda</td>
<td>Java &amp; Bali</td>
<td>34.0</td>
<td>non</td>
<td>I</td>
</tr>
<tr>
<td>Bahasa Indonesia</td>
<td></td>
<td>22.8</td>
<td>M</td>
<td></td>
</tr>
<tr>
<td>Madura</td>
<td>Java &amp; Bali</td>
<td>6.7</td>
<td>non</td>
<td>I</td>
</tr>
<tr>
<td>Minangkabau</td>
<td>Sumatra</td>
<td>5.5</td>
<td>M</td>
<td>O</td>
</tr>
<tr>
<td>Batak¹</td>
<td>Sumatra</td>
<td>5.5</td>
<td>non</td>
<td>O</td>
</tr>
<tr>
<td>Betawi</td>
<td>Java &amp; Bali</td>
<td>5.0</td>
<td>M</td>
<td>I</td>
</tr>
<tr>
<td>Bugis</td>
<td>Sulawesi</td>
<td>5.0</td>
<td>non</td>
<td>O</td>
</tr>
<tr>
<td>Malay</td>
<td>Sumatra</td>
<td>4.9</td>
<td>M</td>
<td>O</td>
</tr>
<tr>
<td>Aceh</td>
<td>Sumatra</td>
<td>3.5</td>
<td>non</td>
<td>O</td>
</tr>
<tr>
<td>Banjar</td>
<td>Kalimantan</td>
<td>3.5</td>
<td>M</td>
<td>O</td>
</tr>
<tr>
<td>Bali</td>
<td>Java &amp; Bali</td>
<td>3.3</td>
<td>non</td>
<td>I</td>
</tr>
<tr>
<td>Musi</td>
<td>Sumatra</td>
<td>3.1</td>
<td>M</td>
<td>O</td>
</tr>
<tr>
<td>Makasar</td>
<td>Sulawesi</td>
<td>2.1</td>
<td>non</td>
<td>O</td>
</tr>
<tr>
<td>Sasak</td>
<td>Nusa Tenggara</td>
<td>2.1</td>
<td>non</td>
<td>I</td>
</tr>
<tr>
<td>Gorontalo</td>
<td>Sulawesi</td>
<td>1.0</td>
<td>non</td>
<td>O</td>
</tr>
<tr>
<td>Malay, Jambi</td>
<td>Sumatra</td>
<td>1.0</td>
<td>M</td>
<td>O</td>
</tr>
</tbody>
</table>

Table 1: Spoken languages of Indonesia with over 1 million speakers, based on Lewis et al., 2014

In Abtaihan et al. (2016a) we used a 1% sample of the 2010 Indonesian census, available through the Integrated Public Use Microdata Series (IPUMS; Minnesota Population Center 2014) to examine language shift toward Indonesian in the ten non-Malayic language communities with over a million speakers (those not closely related to Indonesian), according to a broad set of social categories: age, sex, religion, urbanization, education, and economic development. Our findings, summarized below, provided background for the questionnaire data analysis. We emphasize toward, as the census data available through IPUMS do not provide information as to what other languages the speakers are using, and of course do not provide longitudinal information for families in terms of shift away from other local languages. What the census data allow us to measure is probability of reporting to speak Indonesian for a combination of other social factors.

¹ We treat the closely related varieties of Batak, including Toba, Dairi, Simalungun, Mandailing, together.
Census data findings

Figure 1 demonstrates our baseline model. Our dependent variable is the answer to the question: What language does (RESPONDENT) use daily at home? ("Apakah bahasa sehari-hari yang digunakan (NAMA) di rumah?"); the Y-axis shows the predicted probability that a respondent speaks Indonesian for particular combinations of independent variables - in this figure age, sex, and urbanization.

These general findings demonstrate that younger speakers are more likely than older speakers to use Indonesian at home and urban dwellers are much more likely than rural dwellers to use Indonesian at home, with the effect of age being even greater in urban settings. In considering multiple factors at once, we are able to compare the relative strength of different factors. Thus while still highly significant, the effect of gender is much weaker than either age or urbanization, and in fact in this visualization the lines for male and female respondents are completely overlapping. (See Abtahian et al. 2016a for fuller discussion of the effects of sex, as well as the effects of economic development, education, and religion.)

This approach also enabled us to explore factors that are more specific to the Indonesian context, including ethnic group identity. As in many multi-ethnic and multi-linguistic societies, ethnic and linguistic affiliations are important aspects of group identity in Indonesia. With so many ethnic and linguistic groups varying in size, including some with populations in the millions, Indonesia offers an excellent test case where we can examine ethnic group as a social factor independent of our other social factors. These language communities are geographically spread around the Indonesian archipelago, so we could compare the effect of ethnicity alongside the locally meaningful distinction between the economically and politically central “inner” islands (Java, Bali, and Lombok) and the less central “outer” islands (Sulawesi and Sumatra).

Figure 2 displays our baseline analysis divided by ethnic group, where we find dramatic differences in the overall probability of speaking Indonesian by ethnic group. This is partly influenced by size, where the languages with larger speaking populations are more likely to be maintained. More strikingly, we find that the ethnic group differences that we see may largely be attributed to inner vs outer island communities. All of the inner island language communities (Javanese, Sundanese, Madurese, Balinese, and Sasak) have both more maintenance of the local language and less difference between urban and rural speakers. In contrast, the languages spoken on the outer islands (Batak, Bugis, Acehnese, Makassarese, and Gorontalo) are uniformly further progressed in shift toward Indonesian.

This macro level approach offers insight into the contributions of a complex array of factors to language shift, but it doesn’t offer insight into how or why. To get at these questions we need qualitative data, but gathered in such a way that inter and intragroup comparisons are possible. For this we turn to our questionnaire data.

Language use questionnaire

Our language use questionnaire was developed for use throughout Indonesia (Cohn et al. 2013; 2014, also available in English). It provides an intermediate level of inquiry, allowing for the collection of more data in a shorter period of time from more locations than would be possible with detailed interviews in individual communities, but also offering far more detail and insight into individual language choices than can be gained from census data.

It includes questions about respondent’s background (geographic, ethnic, religious, educational and linguistic), language proficiency, and language use in 34 different domains, as well as the language background, proficiency, and use of their parents, grandparents, spouse and children (if relevant). It also includes several questions about language attitudes and use of technology. To date it has been conducted with 548 participants in 11 locations. Through the analysis of batches of questionnaires from different locations we can consider both intra-location and cross-location comparisons, furthering our goal of building more predictive models of the interaction of Bali here as we find that Sasak, spoken in Lombok, patterns together with the languages of Java and Bali.
language background, language mastery, use, and attitude, with potential language shift. The results demonstrate the expansion of Indonesian from school and official settings to extensive informal daily use; while at the same time highlighting the complexity of the linguistic landscape – respondents in the 11 locations reported using over 94 language varieties in addition to Indonesian (Abtahian et al. 2016b). They also exemplify the benefits of providing respondents freedom in naming language the language varieties they use, as they offer key insights into speaker’s perceptions about language varieties and language use (Abtahian et al. 2017).

Figure 4 demonstrates this flexible format for naming language varieties. In the leftmost column speakers are asked to list the languages and dialects that they use, without any predetermined categories or suggestions (as would generally be provided in a census). Participants were also asked to evaluate their skills in each of these varieties for “Understanding”, “Speaking”, “Reading” and “Writing” on a five-point scale. In this survey, collected in Bengkulu, Sumatra, the respondent listed four languages in the (highlighted) leftmost column: (Bahasa) Serawai, Indonesian, English, and Arabic. In the other columns they indicate the age when they first started using the language, the place where they learned the language, and their self-assessed competence in understanding, speaking, reading, and writing.

The analysis we present here is based on a subset of surveys in eight locations on three islands: Sumatra (190 from four locations), Java (194 from three locations), and Bali (28 from one location). In each location we have results from 28-80 participants, and we examine the number of labels, the consistency of labelling, and the types of labels used, as well as the relationship between labelling practices and ethnolinguistic and regional identity.

This particular subsample (shown in the map in Figure 3) allows us to compare within and between inner/non-Malayic (Java, Bali) speech communities and outer/Malayic (Sumatra)3 speech communities. We also consider a further distinction between the two groups, that the inner/non-Malayic languages have rich speech level systems (Balinese and Javanese) that the outer/Malayic languages do not.

Results

Our first observation has to do with the striking consistency we find in the labelling of Indonesian (Bahasa Indonesia, Indonesia) across the three locations. Particularly in the four communities in Sumatra where the local languages are Malayic we expected to find ambiguity in the use of labels for Indonesian, as colloquial Indonesian in these cases may be mutually intelligible with local vernacular varieties of Malay (Errington 2014). Yet in all of these surveys Indonesian was consistently listed as a separate language alongside one or more local languages and/or vernacular Malay varieties, suggesting that speakers see no ambiguity as to whether these are identifiably different varieties. The total consistency we see in labelling Indonesian stands in contrast to a great deal of variation in labelling all other language varieties.

First, in terms of number of varieties identified by speakers, our overall results support the conclusions of Musgrave (2014), Goebel (2018), and others, that the increased use of Indonesian in these communities is not

---

3 Additional locations such as Betawi in Java and Batak in North Sumatra or Aceh in Aceh Province would allow us to compare inner/Malayic and outer/nonMalayic.
necessarily occurring at the expense of other local languages, although the nature of the multilingualism may be changing. Very few respondents listed only 1 language, and most of these were people who didn’t list Indonesian. Out of the 23 total respondents who listed only one language, only four of these only listed Indonesian (3 in Lampung, 1 in Semarang); the others listed a single local language (and generally also included information elsewhere that demonstrated that they do use Indonesian in daily life). 77% listed three or more and 42% listed four or more varieties, including local languages and vernacular varieties of Malay, Indonesian, English, and other foreign languages. We conclude that this reflects maintenance of the multilingual landscape of Indonesia, or perhaps more accurately, maintenance of an ideology of multilingualism on the part of respondents.

We also find differences between inner and outer island respondents, largely related to whether the language has speech levels or not (remember that the inner islands in this case are languages with highly codified systems of speech registers). For example, in the Javanese-speaking, centrally located city of Yogyakarta, speakers were most likely to use a combination of geographically-indexed and speech level labels, while in outer, Malayic-speaking Bengkulu Sumatra labels were more likely to refer to ethnically and geographically-indexed varieties.

In Bengkulu (Sumatra, outer, Malayic), in addition to very clear labelling of Indonesian, speakers labelled language varieties using language names or geographic locations, including local languages (e.g. Bengkulu, Besemah, Kaur), neighboring languages (e.g. Batak, Padang, Palembang), and ones spoken in other parts of Indonesia (Jawa). There are descriptive labels like “dialek o” and “dialek a” referencing particular linguistic features of dialects. And there are also a number of more subjective and descriptive labels related to home and family. For instance, a number of people listed Bahasa Ibu ‘mother language’, Bahasa Ibu-Bapak ‘mother-father language’, and Bahasa sehari-hari ‘everyday language’. In other words, respondents took full advantage of the flexibility provided by the survey in every way, including listing colloquial and in-group terms for language varieties that index both local and ethnic identities.

**Bengkulu (40) [6 No Indonesian]**
(Bahasa) Indonesia (33), Indo local (1)
Vernacular: Bengkulu, Basma, Batak, Bengkuku, Besemah, Jawa, Jawa Tengah, Kaur, Lębong, Lembak, Lintang, Lubuk Linggau, Manna, Malay, Padang, Pak-Pak, Palembang, Rejang, Serawa

Vernacular Regional: Bengkulu dialect, Bengkulu dialek O, Curup dialek o, Padang Dialek O

---

4 Dialek/dialek O and A/E/O (“O dialect”, “A dialect”, etc) refer to specific phonological variables that characterize regional varieties.

**Yogyakarta (56) [16 No Indonesian]**
(Bahasa) Indonesia (40)
Vernacular: Jawa, Banjar, Banka, Batak, Bugis, Mada, Mandar, Melayu, Palembang, Sunda
Vernacular Regional: (Jawa) Ngakap, Tegal, Jawa Bantul, Jawa Yogyakarta, (Jawa) Jogja, Jawa Magelang, Jawa Mojokerto, Jawa Pati, Jawa Solo, Jawa Trenggalek, Ponorogo
Javanese High: (Jawa) Krama, (Jawa) Krama alus, Jawa Madya, Krama Inggil
Javanese Low: (Jawa) Ngoko
Javanese Regional High: Jawa Inggil Yogyakarta, Jawa Krama Alus Yogyakarta, Krama Jogja, Krama Java Timur.
Javanese Regional Low: (Jawa) Ngoko Yogayakarta, Ngoko Jawa Jogia, Ngoko Java Timur
Foreign: English (33), Arab (15), French (6), German (2), Japanese (1)

Thus, in both Sumatra and Java we see labelling used in very nuanced ways, highlighting ethnolinguistic and geographic identity, with the additional attention to speech levels or registers in Java. National linguistic identity is also made distinct with the clear labelling of Indonesian. We find that these labels are not unlike those produced by perceptual dialectology maps (Preston 1989), and we thus interpret these labels as a reflection of folk ideologies about language and language varieties.

Finally, an interesting possible effect was observed in the data in terms of the patterns of labelling as they relate to ethnic heterogeneity (based on 2010 census data analysis by Ananta et al. 2015). Looking by province, the groups range in their ethnic heterogeneity from highly homogeneous (Yogyakarta, Central Java which are both 97% Indonesian) to highly heterogeneous (Lampung, only 14% Lampung and 64% Javanese due to transmigration in the 1980s and 1990s). In our results increased heterogeneity of the speaker population is seen to be associated with
fewer labels, as Lampung, more ethnically and linguistically diverse than the other locations, shows more respondents listing only one or two languages (42%) and only 22% listed four or more.

Conclusions

Although predicting linguistic behaviour on the basis of speaker-assigned labels for language varieties is not straightforward, the classification and labelling of language varieties plays a crucial role in the process of enregisterment (Agha 2003, Johnstone et al. 2006), which in turn informs linguistic practices (Preston 2011). In Indonesia, where Heryanto (2007) argues that even the concept of language was imposed by European colonialism and did not reflect local understandings of language, the (emerging) classification, labelling, and enregisterment of language varieties that are not tied to the nation-state is closely linked with local and ethnic identities.

The results of our survey data demonstrate Musgrave’s view that the linguistic situation in Indonesia is better viewed as changing patterns of multilingualism, and not necessarily a loss of the fluid and rich linguistic repertoire of language varieties and styles available to individual speakers and communities (as discussed by Goebel 2018). At the same time, there is striking evidence of rapidly shifting linguistic practices, particularly in urban areas and especially in highly ethnically heterogeneous areas such as Lampung. This is a reminder that while being a big language offers certain protections, patterns of language shift and potentially loss are at play irrespective of size of a language community.

The relationship between land and water offers an interesting backdrop for the complex language ecology of Indonesia, and in our broader analysis we consider additional factors such as size of island, location of the community, and size of speaker population. The comparison of batches of questionnaires from different locations allows for intra-location, regional, and cross-island comparisons, furthering our goal of building a more comprehensive picture of the relationship between these locally meaningful social and geographical factors and language maintenance and shift.

Acknowledgements

We thank our collaborators on this project, particularly Yanti and Tom Pepinsky. We also thank the audience members at FEL XXII, NWAV 44, 45, and 46 and ISMIL 43 and 44 as well as other colleagues who have given valuable feedback.

References


Abtahian, Maya R., Abigail C. Cohn, Sonya Waddell and Yanti 2016. Tergantung (“It depends”): A quantitative analysis of language choice in Indonesia. NWAV 45, Vancouver, BC.

Abtahian, Maya, Abigail C. Cohn, and Yanti. 2017. Language labels as ethno-archaeological facts in Indonesia. NWAV 46, Madison, Wisconsin.


Cohn, Abigail, John Bowden, Timothy McKinnon, Maya Ravindranath, Rindu Simanjuntak, Bradley Taylor, and Yanti. 2014. Assessing the impact of Indonesian on the maintenance of local languages. Proceedings of KIMLI, Lampung, Indonesia.


Heryanto, Ariel. 2007. Then there were languages: Bahasa Indonesia was one among many. In S. Makoni and A. Pennycook (eds.) Disinventing and Reconstituting Languages, 42–61. Clevedon: Multilingual Matters.


Mufwene, Salikoko. 2017. Language vitality: The weak theoretical underpinnings of what can be an exciting research area. Language 93. e202–e220. Available at: muse.jhu.edu/article/680465


