

The instatement and corresponding development of *Bahasa Indonesia* as Indonesia’s national language is widely described both as a highly successful example of language planning, and as the catalyst for the endangerment of scores of local languages in one of the most linguistically diverse regions of the world. And yet, the nature of this shift, perhaps more accurately characterized as a shift from one dynamic multilingual scenario to another (Musgrave 2014), is only partially understood. Building on Himmelmann’s (2010, p. 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*”, this paper is part of a larger project that seeks to understand language shift scenarios at multiple levels of analysis using census, survey, and interview data. We argue that a key to understanding linguistic practices, particularly in a complex multilingual landscape, is analyses of how speakers classify and name (label) language varieties that they report using. We correlate this here with analysis of speakers’ self-reported language proficiency and language use.

Our analysis is based on responses to *Kuesioner Penggunaan Bahasa Sehari-hari* (-- 2014), a questionnaire developed to collect detailed data from subjects from different language backgrounds in different parts of Indonesia. Providing an intermediary level of inquiry, it allows collection of extensive data in a short period of time, while also offering detail and insight into individual language choices. To date it has been conducted with 548 participants in 11 locations. Here we focus on results from four urban communities in Sumatra: Bengkulu, Jambi, Lampung, and Padang, analyzing data from 189 questionnaires (listed by location in Table 1).

In terms of number of languages listed (where the respondents have the flexibility to name the varieties they speak), we find that the clear majority list three or more varieties, including Indonesian, local languages and vernacular varieties of Malay, English, and other foreign languages. In our results increased heterogeneity 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%, as compared to a maximum of 12% in the other locations.) We also find that, despite expectations of ambiguity in the use of labels for Indonesian and closely related varieties of Malay, there is little ambiguity. Indonesian was consistently listed as a separate language alongside one or more local languages and/or vernacular Malay varieties. Further analysis of the relationship between labeling of varieties, self-reported language proficiency, and their relationship between language use reported in 34 different domains will be reported on, along with comparisons of these results to our analysis in three locations in Java and one in Bali where we find different labeling practices.

location	total	1		2		3		4 (+)	
Bengkulu	40	0	0	4	10%	17	42%	19	47%
Jambi	20	0	0	1	5%	10	50%	9	30%
Lampung	55	4	7%	19	35%	20	36%	12	22%
Padang	74	1	1%	8	11%	24	32%	41	55%

Table 1. Survey locations and number of languages listed

The comparison of batches of questionnaires from different locations allows us to consider intra-location, regional (here Sumatra), and cross-island comparisons, furthering our goal of building a more comprehensive picture of the interaction of language background, proficiency, use, as they impact language maintenance and language shift.