

# THE POSITION OF MALI, BEAYE, AND BA'AJE IN LAND DAYAK (AUSTRONESIAN) AND WHAT IT MEANS FOR PROTO-LAND-DAYAK VOWEL RECONSTRUCTION

Alexander Smith  
*National University of Singapore*  
[smith.alexander.david@gmail.com](mailto:smith.alexander.david@gmail.com)

Carly J. Sommerlot  
*National University of Singapore*  
[csommerlot@gmail.com](mailto:csommerlot@gmail.com)

## Abstract

The internal subgrouping of Land Dayak languages (Sarawak, Malaysia and West Kalimantan, Indonesia) as well as the reconstructability of length distinctions in the vowel system of Proto-Land Dayak, have received recent attention that has resulted in changes to the internal subgrouping and a restriction of the full/reduced distinction in Proto-Land Dayak vowels. Earlier work on the reconstructability of the full/reduced distinction relied on assumptions about the subgrouping of certain languages that needed to be confirmed pending future research. In this paper, documentation of some previously unstudied Land Dayak languages is presented. The data from these languages, which was originally collected for syntactic analysis and description, supports a restriction in full/reduced to Proto-Land Dayak \*a but not other vowels, lending support to earlier proposals based on more limited datasets. The documentation of these languages also allows for a more accurate internal subgrouping of Land Dayak and shows the utility of language data beyond initial collection as well as the importance of archiving.

**Keywords:** reconstruction, subgrouping, Land Dayak, Austronesian, historical  
**ISO 639-3 codes:** day, byd, lra, sre, bth, sdo

## 1 Introduction

The documentation of languages and the utilization of digital archives for permanent storage of vast quantities of language data may provide opportunities for linguistic analysis beyond the scope of the original project. Indeed, language documentation projects are often carried out with the primary goal of making data available and accessible so that future audiences may utilize the corpus for yet-to-be-determined linguistic (or non-linguistic) analysis (Holton 2012, Woodbury 2014).

For the non-documentarian, the utility of an archive may depend crucially on a number of factors. A syntactician may find certain archives more useful due to the richness of syntactic data, while other archives may provide less utility. For comparative analysis, the principal concern of the linguist is the availability and accessibility of archived data from a range of related languages such that adequate genetic diversity is represented in the sample. Comparative analysis depends crucially on casting a “wide net” and working with data samples from as many relevant languages as possible. This becomes more crucial when the absence of a language or groups of languages from a data set may impact the linguist’s ability to accurately reconstruct a proto-language via comparative analysis.

The present research utilizes published, unpublished, and newly collected documentary data gathered as part of survey of languages in West Kalimantan, Indonesia, for the purpose of reconstruction. Specifically, we demonstrate how the documentation of these languages and the availability of linguistic data fill a data gap in the comparative study of Land Dayak, a subgroup within the larger Austronesian family. This project further exemplifies the importance of researchers making data on under- and undocumented languages widely

available, as projects that may have been conducted with certain specific research questions in mind can provide much needed data for projects conducted by others on the same languages.

The linguistic goal of the study is to show that a distinction between full and reduced vowels, /ā/ and /a/, was present in two primary branches of the Land Dayak subgroup, Benyadu-Bekati' and Bidayuh-Southern Land Dayak, warranting their reconstruction to Proto-Land Dayak. Such a reconstruction was proposed in Smith (2019) but relied on an assumption that sparse evidence for the contrast in Benyadu-Bekati' might be strengthened by subsequent research and new data sets. This research provides such new data sets in support of Smith's original hypothesis.

## 2 Language data used in this study

Borneo is a large island situated in central Insular Southeast Asia (ISEA) with high linguistic diversity but whose languages are often less well-documented than languages spoken elsewhere in ISEA despite the island's important position in comparative Austronesian Linguistics (Adelaar 1995, Blust 2010). This is particularly true for languages of the Land Dayak subgroup, with the most notable prior works including these languages being Connell (2013), Rensch et al. (2012), Smith (2017), and Sommerlot (2020). Rensch et al. (2012) focuses on Bidayuh languages spoken in Sarawak, Malaysia and contains an 83-page appendix of cognate sets from up to 27 Land Dayak languages. Smith (2017) adds data sets from Land Dayak languages in Kalimantan and proposes an updated subgrouping model based on the expanded dataset.

Since Rensch et al.'s publication there has been significant syntactic description on Land Dayak languages, in the form of Sommerlot's (2020) dissertation. Data collection for this project was primarily focused on syntactic description and analysis of voice systems and A'-movement in both Malayic and Land Dayak languages. However, given the under- and undocumented status of all the languages included, significant lexical, sentence, and narrative data was additionally collected beyond the scope of the original project.

The present research utilizes word-level data from two languages, Beaye and Ba'aje, from Sommerlot's project to inform the internal subgrouping of the Land Dayak languages. A brief background of these two languages is provided below.

*Map 1. Approximate locations of where Beaye (Kumpang) and Ba'aje (Kase) are spoken*



### 2.1 Land Dayak languages: Beaye and Ba'aje

Beaye and Ba'aje are both spoken in the Landak Regency in West Kalimantan, close to the border of Malaysian Borneo. While data was collected in the university town of Ngabang, speakers originated from Kumpang (Beaye) and Kase (Ba'aje). The approximate locations of these two villages are shown in Map 1.

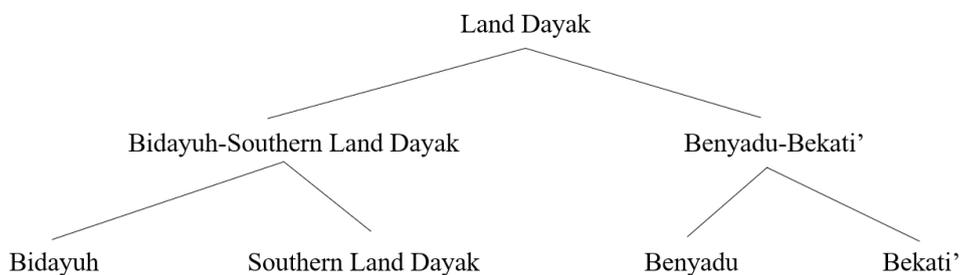
Prior to Sommerlot (2020), both Beaye and Ba'aje were undocumented in linguistic literature. The data collected consists of wordlists, phrases, sentences, and narratives. Sommerlot's dissertation provides in-depth description and analysis at the syntactic level, based primarily on the sentence-level data collected; the present study demonstrates how previously collected data can be utilized for reconstruction, based predominantly on the word-level data gathered.

### 3 Current subgrouping evidence

Land Dayak languages form a major linguistic subgroup in western Borneo, although there is some disagreement on precisely how Land Dayak fits into the larger Malayo-Polynesian subgroup (Rensch et al. 2012, Smith 2017, Smith and Rama 2022). Land Dayak has been traditionally divided into three main subgroups, Bidayuh, Benyadu-Bekati', and Southern Land Dayak. Bidayuh languages are mostly spoken in Sarawak, Malaysia, but some are also spoken across the border in Sanggau, Sintang, and Bengkayang regencies, in West Kalimantan, Indonesia. Benyadu-Bekati' and Southern Land Dayak are found primarily in the northern half of West Kalimantan, but may be found in southern areas as well.

There are two competing subgrouping proposals which both posit two primary subgroups. The first is from Rensch et al. 2012 and the second was first proposed in Smith 2017 and later corroborated in Smith and Rama 2022. Rensch et al.'s proposal groups Benyadu-Bekati' and Bidayuh together, to the exclusion of Southern Land Dayak. Smith's proposal groups Bidayuh and Southern Land Dayak together, to the exclusion of Benyadu-Bekati'. Smith and Rama (2022) use Bayesian Phylogenetic methods and find additional support for Smith's 2017 proposal. We therefore utilize Smith's subgrouping to inform the present study, shown in a simplified diagram in figure 1:

**Figure 1.** Current Land Dayak subgrouping (according to Smith 2017, Smith and Rama 2022)



The evidence for a primary division between Bidayuh-Southern Land Dayak and Benyadu-Bekati' is significant. Benyadu and Bekati' share a merger of PMP \*a and \*ə in penultimate syllables, the closing of word-final vowels with a glottal stop, and the coalescence of \*-ay and \*-aw to -e and -o. Bidayuh and Southern Land Dayak, on the other hand, do not merge \*a and \*ə in all penultimate syllables, but rather reflect a narrowly conditioned merger where \*ə merges with \*a in most environments but fails to do so in disyllabic words that begin with a labial consonant, final vowels are closed with -h, the diphthongs \*-ay and \*-au coalesce to -i and -u, and PMP \*-a is raised to -i. A table of these changes is presented below in 1. A more detailed discussion can be found in Smith (2017).

**Table 1.** Evidence for a primary division between *Benyadu-Bekati* / *Bidayuh-Southern Land Dayak*

PMP	Proto-Benyadu -Bekati'	Proto-Bidayuh-Southern Land Dayak
*-aCV(C)	*-aCV(C)	*-aCV(C)
*-əCV(C)		*-əCV(C)/-aCV(C)
-V	*-Vʔ	*-Vh
*-ay	*-e	*-i
*-aw	*-o	*-u
*-a	*-ah	*-ih

Although such differences may seem trivial, the internal subgrouping of Land Dayak has significant impacts on reconstruction. This is because any reconstruction must contain evidence from at least two primary branches in order to be considered valid. If, for example, evidence for some feature is present in Bidayuh and Southern Land Dayak, then that feature may be reconstructed to Proto-Land Dayak under Rensch et al.'s proposal, but is only reconstructable to the intermediate Proto-Bidayuh-Southern Land Dayak under Smith's proposal.

Regarding reconstruction, Smith (2019) reconstructs a typical consonant inventory for PLD as shown below in Table 1. Of note is the loss of PMP \*l, but an /l/ phoneme was likely present either in PLD itself or in the early history of its daughter languages as a product of borrowing from a Malayic source. The vowel inventory consisted of four main vowels, \*a, \*i, \*u, \*ə, with a fifth, \*ā, likely restricted to only a handful of words, from both native sources and through early borrowing. Smith notes that \*ə was pronounced [i], and that \*a was pronounced [ə] in the penultimate syllable but [a] in the final syllable.

The reconstructed consonant inventory is typical for most Austronesian languages of the area, and is shown below in table 2.

**Table 2.** Proto-Land Dayak consonant inventory (Smith 2019)

	labial	alveolar	palatal	velar	glottal
voiceless stop	*p	*t		*k	*ʔ
voiced stop	*b	*d		*g	
nasal	*m	*n	*ñ [ɲ]	*ŋ	
fricative		*s			*h
affricate			*j [dʒ]		
liquid		*r/(*l)			
glide			*y [j]	*w	

### 3.1 The Long-short problem in Land Dayak reconstruction

Disagreements in Proto-Land Dayak reconstruction centers on the vowels and whether a “full”/“reduced” distinction in the vowels is reconstructable to PLD itself. The argument for a full/reduced distinction began with Rensch et al. where a distinction was reconstructed for all vowels, \*i/\*ī, \*u/\*ū, \*ə/\*ē, and \*a/\*ā. Part of Rensch et al.'s justification for such a reconstruction is that long and short vowels are present in both Bidayuh and Southern Land Dayak subgroups, which, according to the subgrouping hypothesis utilized in that study, means that they may be reconstructed to PLD. As pointed out above, however, Smith (2017) and later Smith and Rama (2022), propose a fundamentally different subgrouping for LD, which places Bidayuh and Southern Land Dayak into a single subgroup, Bidayuh-Southern Land Dayak. This means that the critical distribution of full-reduced distinctions is found only in a single subgroup, and is therefore not reconstructable to PLD. To demonstrate, Table 3 lists words which Rensch et al. reconstructed with the full and reduced vowels \*ā and \*ā̄. In Bidayuh-Southern Land Dayak there are distinct reflexes from both vowels (reduced forms are highlighted), but in Benyadu and Bekati', both \*ā̄ and \*ā̄ are reflected with the same vowel, a.<sup>1</sup>

<sup>1</sup> PLD reconstructions in Tables 3 and 4 are from Rensch et al. 2012, all others are from Smith 2017: Appendix 5. PMP reconstructions are from Blust, Trussel & Smith 2023. LD reconstructions below the PLD level are all from Smith 2017: §2.6 with supplemental discussion in Smith 2019.

**Table 3.** Lack of full/reduced distinction in *Benyadu and Bekati'*

	PLD	Benyadu-Bekati'		Bidayuh-Southern Land Dayak			
		Benyadu	Bekati'	Singai	Biatah	Bukar	Golik
full	*bātu 'stone'	batu?	batu?	batuh	bātuh	bātuh	batuh
	*mānuk 'bird'	manuk	manu?	manuk	mānuk	mānuk	manuk
reduced	*rāŋit 'sky'	raŋit	raŋit	rəŋit	rāŋit	rāŋit	roŋit
	*dāun 'leaf'	dautn	daut	dəwən	dāwən	dāwət	doutn

There are two ways that these data may be analyzed: 1) a distinction between full and reduced vowels existed in PLD, and was eliminated in both *Benyadu* and *Bekati'*, or 2) there was no full/reduced distinction in PLD, and such a distinction is an innovation in *Bidayuh-Southern Land Dayak*. In order for the first scenario to be seriously considered, however, the proper evidence is necessary: languages that unambiguously subgroup within *Benyadu-Bekati'* which also have evidence of a past full/reduced distinction.

Smith (2019) proposed that such evidence indeed appears in *Mali*, a little-studied *Land Dayak* language of West Kalimantan. However, this proposal had only limited access to *Mali* linguistic data and the exact linguistic position of *Mali* was uncertain, forcing Smith to highlight the preliminary nature of his proposed evidence for an \*ā and \*ă distinction in PLD. In order for the *Mali* data to be seriously considered as supporting evidence for a distinction between ā and a in *Benyadu-Bekati'*, more robust subgrouping evidence is necessary. As will be discussed in the following sections, such evidence appears in Sommerlot (2020), where data from *Beaye*, and *Ba'aje*, in addition to *Mali*, supports the recognition of a full/reduced distinction as well as a subgrouping which places these three languages squarely within the B-B subgroup.

It was pointed out by a reviewer that unexplained distinctions can be reconstructed to a parent language even if they appear in only one branch. This is certainly true for distinctions that must otherwise be hand-waved away as irregular or unconditioned splits. This is not the case for the full/reduced vowel distinction, however. The source of the distinction in native vocabulary, as pointed out in Smith 2019, is from the merger of vowels after consonant deletion as in PMP \*tahəp-an > PLD \*tāpan. Other than a few unexplained full vowels that may be very early borrowings, all other reflexes of PMP penultimate vowels are reduced. The regular outcome is therefore a reduced penultimate vowel, no doubt brought about by the strong word-final stress system of LD languages. Because Smith (2019) was able to show that the full/reduced distinction arose through conditioned change, the fundamental question at hand here is whether the conditions that led to full vowels were present in PLD or not. These are not unexplained splits that may be remnants of an earlier distinction.

#### 4 New data: *Mali, Beaye, and Ba'aje*

After Sommerlot (2020), lexical data were made available from several newly recorded LD languages in addition to *Mali*. Relevant languages from Sommerlot for the present discussion are *Ba'aje* and *Beaye*. These languages appear to reflect a distinction between a full and reduced low vowel, \*ă and \*ā, in several words shown below in Table 4.

In these cases, inherited long \*ā is reflected with *a* in *Mali, Beaye, and Ba'aje*. Reduced vowels, on the other hand, harmonized with the following word-final vowel. For example, \*ă becomes *i* in reflexes of \*rāŋit 'sky' but *u* in reflexes of \*dāun 'leaf'. In cases where the final-syllable vowel reflects a word-final \*a, the reduced penultimate vowel is also reflected with *u*. With the full vowels, no harmonization is observed. There is limited data with *Ba'aje*, but even in the small dataset that we have available, differences between reflexes of full and reduced \*a are still observable.

**Table 4.** Full/reduced distinction in Mali, Beaye, and Ba'aje

PLD	Mali	Beaye	Ba'aje
*bātu 'stone'	batu	batu?	-
*mānuk 'bird'	manu?	manu?	manu?
*tāru 'three'	taru	tau	taru
*rāñit 'sky'	riñit	riñit	-
*dāun 'leaf'	duwutn	duwitn	-
*sāwa 'spouse'	suwi	suwi	-
*māta/*bāta 'eye'	buto	buti	bəti

There is clearly a distinction between full and reduced vowels reflected in the lexicon of these languages. Such a distinction may be reconstructed to PLD with evidence from Mali, Beaye, Ba'aje, and B-SLD, but only if it is shown that the relevant languages belong to the B-B subgroup. Evidence in support of this subgrouping is discussed next.

#### 4.1 *Where do these new languages fit in?*

Smith (2019) hypothesized that Mali and Beaye belonged to the Benyadu-Bekati' subgroup, and may therefore stand as evidence of a distinction between \*ă and \*ā in PLD.

The position of these languages, however, was not specifically defended in this earlier publication, and as stated in Smith 2019:118, "More data is needed to convincingly show both that Mali and Be' Aye' are in fact Benyadu-Bekati' languages and that the vowel restrictions [which reflect distinctions between full and reduced] will continue to play out." Smith goes on to state that, "If these languages [Mali and Be' Aye'] are shown to subgroup inside B-SLD, however, the current subgrouping model would restrict full vowel reconstruction to PB-SLD, and would not support reconstruction to PLD."

With additional evidence from Sommerlot (2020), it is now possible to show that Mali, Beaye, and Ba'aje are indeed Benyadu-Bekati' languages, and that the distinctions in vowels \*ă and \*ā may therefore be confidently reconstructed to PLD. The evidence is of several types: i) exclusively shared phonological innovations found only in Benyadu-Bekati' languages, including Mali, Beaye, and Ba'aje, ii) irregular phonological innovations exclusive to Benyadu-Bekati' which are also found in Mali, Beaye, and Ba'aje, and iii) regular phonological innovations, again, exclusive to Benyadu-Bekati', Mali, Beaye, and Ba'aje.

##### 4.1.1 *Lexical innovations*

Smith 2017:153-155 includes a list of several exclusively shared lexical replacement innovations which support the B-B subgroup. At least four of these innovations are present in Mali, and some are also present in Beaye and Beaje, although a smaller dataset from these two languages means that some of the evidence is exclusive to Mali. Table 5 provides a summary of the shared innovations. PMP forms are provided, followed by the PB-SLD and PB-B reconstructions with innovations bolded, and reflexes of the PB-B innovations in Mali, Beaye, and Ba'aje.

**Table 5.** *Benyadu-Bekati' lexical replacement innovations in Mali, Beaye, and Ba'aje*

PMP	PB-SLD	PB-B	Mali	Beaye	Ba'aje	Gloss
*adaduq/ *anaduq	*ambuh	*aŋhu	aŋo	aŋo	-	'long'
*zəlaq	*jaraʔ	*rataʔ	rata	-	rata	'tongue'
*(C)away	*jawi-n	*bahas	bas	bas	bas	'face'
*taŋis	*taŋis	*munseʔ	munse-a	-	-	'to cry'

#### 4.1.2 Regular sound change

There are a few sound changes that are exclusive among Benyadu-Bekati' languages as well (Smith 2017). One such sound change is the monophthongization of word-final diphthongs. In Benyadu-Bekati' these diphthongs monophthongized as mid vowels (\*ay > e, \*aw > o). In Bidayuh-Southern Land Dayak on the other hand they are monophthongized as high vowels (\*ay > i, \*aw > u). Table 6 organizes reflexes of \*atay 'liver' and \*suŋay 'river' in both Benyadu-Bekati' and Bidayuh-Southern Land Dayak languages. Mali, Beaye, and Ba'aje all show mid-vowel monophthongization whereas Bidayuh, represented below by Hliboi and Bukar, and Southern Land Dayak, represented by Jangkang, Kembayan, and Ribun, all have high-vowel monophthongization.

**Table 6.** *Mid and high vowel monophthongization in Land Dayak*

Benyadu-Bekati'					
PLD	Benyadu	Bekati'	Mali	Beaye	Ba'aje
*atay 'liver'	ate	ate	ate`	ate	ate
*suŋay 'river'	suŋe	suŋe	suŋe	-	-
Bidayuh-Southern Land Dayak					
PLD	Hliboi	Bukar	Jangkang	Kembayan	Ribun
*atay 'liver'	iti-tn	əti	oti	oti	oti
*suŋay 'river'	-	suŋi	-	suŋi	suŋi

#### 4.1.3 Lexeme-specific sound change

Some lexeme-specific sound changes also appear in Benyadu-Bekati' languages. For example, in reflexes of PMP \*daya 'upriver', the final vowel and preceding glide irregularly merged, producing Mali *dai*, Benyadu *dae*, and Bekati' (Rara dialect) *dae*. In another sound change, schwa in the penultimate syllable became *i* in reflexes of PMP \*ənəm (schwa usually merged with \*a in this environment). Examples include Mali *inəm*, Beaye *inum*, Ba'aje *inum*, Benyadu *inum*, Bekati' *inum*.

These changes are absent in Bidayuh-Southern Land Dayak. For example, vowel merger is absent in Singhi (Bidayuh) *doyx* as well as in Kembayan (SLD) *dəyəh*. Reflexes of \*ənəm 'six' also have the expected reflexes in Bidayuh-Southern Land Dayak, including in Singhi *nom*, Bukar *ənəm*, and Kembayan *nəm*. Lexeme-specific sound changes can provide powerful subgrouping evidence, but only when they target the same lexemes in all member languages. Since two specific changes target two specific words in Benyadu, Bekati', Mali, Beaye, and Ba'aje but the same words are regular in Bidayuh-Southern Land Dayak, the best interpretation of these facts is that the change took place in a proto-language (in this case, Proto-Benyadu-Bekati').

In summary, the evidence supports a hypothesis whereby Mali, Beaye, and Ba'aje subgroup together within the Benyadu-Bekati' subgroup of Land Dayak, which, according to Smith 2017 and Smith and Rama 2022, is one of two primary divisions. The evidence includes four exclusively shared lexical innovations \*aŋhu

'long', \*rata? 'tongue', \*bahas 'face', and \*munse? 'to cry', the merger of final diphthongs into mid-vowels, the irregular merger of -ya as -i in reflexes of PMP \*daya 'upriver', and the irregular raising and fronting of \*ə to i in reflexes of PMP \*ənəm 'six'.

#### 4.2 A new subgrouping of Proto-Land Dayak

It has been established that, 1) Mali, Beaye, and Ba'aje all reflect a distinction between a full and reduced \*a in penultimate position, and 2) these three languages likely belong to the Benyadu-Bekati' subgroup. The merger of \*ă and \*ā in Benyadu and Bekati' has implications for the internal subgrouping of the Benyadu-Bekati' subgroup as well. It follows from the observation that Mali, Beaye, and Ba'aje belong within the Benyadu-Bekati' subgroup that they must also be separated from the Benyadu-Bekati' languages which reflect the innovative merger \*ă, \*ā > a. The internal subgroup may therefore be further subdivided into one group with Benyadu and Bekati', and a second with Mali, Beaye, and Ba'aje. Furthermore, although the merger of \*ă and \*ā in Benyadu and Bekati' may be used to create a distinct subgroup, the interrelatedness of Mali, Beaye, and Ba'aje is not immediately clear. They may form a single subgroup, or they may be part of multiple, as yet undescribed, separate subgroups. Additionally, the name "Benyadu-Bekati'" does not seem appropriate as a descriptor of the larger subgroup which includes Mali, Beaye, and Ba'aje. Therefore, some renaming is in order. The larger subgroup may be labelled "Western Land Dayak", following the convention of cardinal direction naming for other LD subgroups (Specifically, Southern Land Dayak). The term "Benyadu-Bekati'" is now restricted to only those languages within the Western Land Dayak subgroup which have merged \*ă and \*ā as \*a. Mali, Beaye, and Ba'aje are grouped in a separate, for now unnamed subgroup. They are written in italics to indicate a lack of phonological innovations which define the subgroup, and they may be shown to belong to multiple additional subgroups if further research unveils such distinctions.

**Figure 2.** Proposed new subgrouping of Proto-Land Dayak

#### PLD

1. Western Land Dayak
  - a. Benyadu-Bekati'
  - b. *Mali, Beaye, Ba'aje*
2. Bidayuh-Southern Land Dayak
  - a. Bidayuh
  - b. Southern Land Dayak

Finally, although Rensch et al. reconstruct a full/reduced distinction for all vowels, not only \*ă and \*ā, Smith (2019) showed that all instances of "full" vowels \*ī, \*ū, and \*ē were the product of recent borrowing from Malay and not reconstructable to PLD regardless of their presence or absence in Benyadu-Bekati'. Mali, Beaye, and Ba'aje data do not change this, but the present study does confirm with new evidence that a full/reduced distinction between \*ă and \*ā can be reconstructed to PLD with data from both the Bidayuh-Southern Land Dayak and Western Land Dayak branches.

## 5 Conclusion

This research has shown that three languages, Mali, Beaye, and Ba'aje, show lexical and phonological innovations that are exclusive to the Benyadu-Bekati' subgroup of Land Dayak, therefore justifying their inclusion in the subgroup. Further, the languages reflect a distinction between a full vowel, \*ā, and reduced vowel \*a. This distinction provides crucial evidence from the Benyadu-Bekati' subgroup that this distinction was present in Proto-Land Dayak, was retained in Western Land Dayak and Bidayuh-Southern Land Dayak, and was later lost in Benyadu-Bekati' via merger (\*ā, \*a > \*a).

The research demonstrates the importance of continued language documentation in Borneo. The critical evidence for this reconstruction was only made available through the documentation of languages in West Kalimantan (Sommerlot 2020). The data, which are available in the dissertation and which will also be made available in future archiving work, were originally gathered for the purposes of syntactic analysis and were not meant for comparative analysis. However, syntactic fieldwork which keeps documentation in mind and produces a database that includes data beyond that useful for only syntactic analysis may have an impact on linguistic scholarship beyond their original intended use.

## References

- Adelaar, K. Alexander. 1995. Borneo as a cross-roads for comparative Austronesian linguistics. In *The Austronesians: Historical and Comparative Perspectives*, ed. by Peter Bellwood, James J. Fox, and Darrell Tryon. 75-95. Canberra: Research School of Pacific and Asian Studies, Australian National University.
- Blust, Robert. 2010. The Greater North Borneo hypothesis. *Oceanic Linguistics* 49:44–118.
- Robert Blust, Stephen Trussel, & Alexander D. Smith. (2023). CLDF dataset derived from Blust's "Austronesian Comparative Dictionary" (v1.2) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.7741197>
- Connell, Timothy. 2013. A sketch grammar of Matéq: a Land Dayak language of West Kalimantan, Indonesia. Master's thesis, University of Canterbury, Christchurch, NZ.
- Holton, Gary. 2012. Language archives: They're not just for linguists anymore. In Frank Seifart, Geoffrey Haig, Nikolaus P. Himmelmann, Dagmar Jung, Anna Margetts and Paul Trilsbeek (eds.), *Potentials of Language Documentation: Methods, Analyses, and Utilization*, 105-110. Language Documentation and Conservation Special Publication No 3.
- Rensch, Calvin R., Carolyn M. Rensch, Jonas Noeb, and Robert Sulis Ridu. 2012. *The Bidayuh Language: Yesterday, Today, and Tomorrow (Revised and Expanded)*. SIL E-Books, 33. Kuching: SIL International.
- Smith, Alexander D. 2017. The languages of Borneo: a comprehensive classification. Doctoral Dissertation, Department of Linguistics, University of Hawaii.
- Smith, Alexander D. 2019. A second look at Proto-Land Dayak vowels. *Oceanic Linguistics* 58(1):110-142.
- Smith, Alexander D. and Taraka Rama. 2022. Environmental factors affect the evolution of linguistic subgroups in Borneo. *Diachronica* 39(2):193-225.
- Sommerlot, Carly. 2020. On the syntax of West Kalimantan: Asymmetries and A'-movement in Malayic and Land Dayak languages. Doctoral Dissertation, Department of Linguistics, University of Texas at Arlington.
- Woodbury, Anthony C. 2014. Archives and audiences: Toward making endangered language documentations people can read, use, understand, and admire. *Language Documentation and Description* 12:19-36.

---

**Reviewed:** Received 27 March 2023, revised text accepted 5 September 2023, published 15 October 2023

**Editors:** Editor-In-Chief Dr Mark Alves | Managing Eds. Dr Paul Sidwell, Dr Sigrid Lew, Dr Mathias Jenny, Dr Kenneth Van Bik