



PROJECT MUSE®

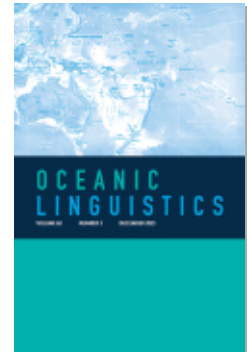
Lexical Evidence in Austronesian for an Austroasiatic
presence in Borneo

Juliette Blevins, Daniel Kaufman

Oceanic Linguistics, Volume 62, Number 2, December 2023, pp. 366-413
(Article)

Published by University of Hawai'i Press

DOI: <https://doi.org/10.1353/ol.2023.a913565>



➔ *For additional information about this article*

<https://muse.jhu.edu/article/913565>

Lexical Evidence in Austronesian for an Austroasiatic presence in Borneo

Juliette Blevins* and Daniel Kaufman†

*THE GRADUATE CENTER, CUNY AND †QUEENS COLLEGE, CUNY AND
‡ENDANGERED LANGUAGE ALLIANCE (ELA)

Divergence and diversity at the level of phonology and lexicon in many of the Austronesian languages of Borneo are widely recognized and well studied. However, the source of this divergence is debated. In this paper, lexical items in the languages of Borneo which lack secure Austronesian etymologies are the object of study. Some of these words show potential semantic and phonological matches with Austroasiatic forms, suggesting a possible early period of *in situ* contact between Austronesian speakers and speakers of Mon-Khmer languages on the island of Borneo.

Keywords: Borneo; Mon-Khmer; Austroasiatic; Contact; Loanwords; Prehistory

1. WHY ARE AUSTRONESIAN LANGUAGES OF BORNEO DIVERGENT?¹

Within the Austronesian language family, many languages of Borneo are widely recognized as being phonologically and lexically divergent,

1. The authors are grateful to four reviewers and to the editors for comments, criticism, and corrections.

For the purposes of clear exposition, we use “Mon Khmer” to refer to the Austroasiatic languages that are non-Munda, and that are included in the *Mon Khmer Etymological Dictionary* (MKED), our primary source for this study. It should be stressed from the start that the non-Munda Austroasiatic (“Mon Khmer”) languages appear to be a paraphyletic group, and do not form a subgroup within Austroasiatic. For this reason, “Proto-Mon-Khmer” (PMK) is a methodologically problematic term. Nevertheless, since this is the label given to the oldest reconstructions in the *Mon Khmer Etymological Dictionary*, we use the same label, which we ask the reader to interpret as something close to the Proto-Austroasiatic (PAA) form. For example, where we cite PMK *gma? ‘rain’ from MKED, this can be interpreted as PAA *gma? ‘rain’ (Sidwell and Rau 2015). Another problematic aspect of MKED is that reconstructions are based on several works, and may represent inconsistent views of the historical phonology. For example, Shorto (2006) believed that PMK had root ablaut, giving rise to alternations between short vowels, long vowels, and monophthongs vs. diphthongs, with distinct ablaut forms marked PMKA, PMKB, PMKC. The problem of inconsistency should be kept in mind, especially where vocalism is concerned. When specific reconstructions from the MKED are not cited, we use “Proto-Austroasiatic” (PAA) instead, with specific PAA forms taken from Sidwell and Rau (2015), or marked ** if they are our own. We use “Austroasiatic” to refer to the larger language family that includes Mon Khmer languages and Munda languages. When we speak of the languages of Borneo, we are referring to the Austronesian languages of Borneo. The term “Pan-Borneo” is used as a cover term for the language or languages that diversified into the Greater North Borneo languages and the Basap-Barito languages (Smith 2017).

Abbreviations used in this paper are: ACD = The Austronesian Comparative Dictionary (Blust and Trussell, ongoing); An = Austronesian; BD = Borneo Dictionary (online);

with a range of properties that are uncommon elsewhere within the family. In the realm of segmental contrasts, contrastive palatal consonants are found, final nasals may be pre-ploded, medial nasals may be post-ploded, and rare laryngeal series, like the voiced aspirates described for several dialects of Kelabit (Blust 2006), the voiceless sonorants described for the Sa’ban dialect of Kelabit, and the implosives of Bintulu, are attested (Blust 2013:67, 182, 184–85). Even more remarkable are the word-level properties, which make some languages of Borneo almost unrecognizable as Austronesian stock. These include variability and neutralization of vowel quality in nonfinal syllables and bulking of final syllables, with a shift toward iambic or even monosyllabic words. These word-level properties can be illustrated with some of the Borneo continuations of PAN *asu ‘dog’ in table 1. While the final syllable remains light in the majority of Austronesian languages (cf. Bunun *asu*; Ilokano *áso*; Malaweg *asú*; Toba Batak *asu*; Kambera *ahu*), many languages of Borneo are clearly different, with a range of distinct changes yielding a light-heavy or monosyllabic heavy syllable pattern. In this way, the languages of Borneo are similar to

TABLE 1. SOME BORNEO REFLEXES OF PAN *asu ‘DOG’ ILLUSTRATING SHIFT TOWARD IAMBIC WORD.

Subgroup	Language	*asu ‘dog’	Bulking of second syllable	Reduction/loss of first syllable
Kayanic	Busang	aso?	C-epenthesis	–
	Kelai	asaw	V ₂ -breaking	–
	Data Dian	aso:?	C-epenthesis, V ₂ -lengthening	–
	Mpraa	haw?	C-epenthesis, V ₂ -breaking	V ₁ -loss
	Modang	saø	V ₂ -breaking	V ₁ -loss
	Long Gelat	sa:	V ₂ -lengthening	V ₁ -loss
	Bahau	ho:?	C-epenthesis, V ₂ -lengthening	V ₁ -loss
Land Dayak		*kasu		
	Benyadu	kasu?	C-epenthesis	–
	Sungkung	kasokŋ	C-epenthesis	–
	Hliboi Bidayuh	kisúəkŋ	C-epenthesis, V ₂ -breaking	V ₁ -reduction
	Sanggau	kiu?	C-epenthesis, C-lenition	V ₁ -reduction
Kenyah	E. Penan	asəw?	C-epenthesis, V ₂ -breaking	–

C = central; CLICS³ = The Database of Cross-Linguistic Colexifications (List et al. 2019); MET = metathesis; MK Mon Khmer; MKED Mon Khmer Etymological Dictionary (SEAlang); Müller-Sch. = Müller-Schwaner (Borneo subgroup); N = north; NSC = Scott (1956); PAA (Proto-Austroasiatic); PAN, Proto-Austronesian; PMK Proto-Mon Khmer; PMP Proto-Malayo-Polynesian; PWMP Proto-Western Malayo-Polynesian; RKP = Puri (2001); RRNR = Rensch et al. (2006); Sm = Smith (2017).

Chamic languages, where parallel shifts toward iambic and monosyllabic word types have been documented in detail by Thurgood (1999) and attributed, in large part, to intense contact with Mon-Khmer languages in mainland Southeast Asia. For PAN *asu ‘dog’, compare the Borneo continuations in table 1 with Jarai *asəu*, W. Cham *saw*, and Wr. Cham *asuj*, *suw*, all from Proto-Chamic *ʔasow (Thurgood 1999:281).

Lexical divergence among the languages of Borneo is also well known (Blust 2010) and has recently been documented more fully by Smith (2017). In his comprehensive classification of the languages of Borneo, Smith makes regular use of innovative sound changes as well as lexical replacements to define subgroups. In some cases, lexical divergence from the mother tongue is extreme. Consider, for example, the remarks of Smith (2017:293) on PMP *hikan ‘fish’:

PMP *hikan ‘fish’ was replaced in nearly every subgroup in Borneo. It is retained only in Malayic and Land Dayak.² There does not appear to be any single word which can be reconstructed to a higher order subgroup as a replacement for *hikan, although *ajən has a wider distribution (it appears in Kayanic and Central Sarawak languages). Segai-Modang, Kenyah, and Tunjung do appear to share near-cognates for fish of basic shape *atuʔ/*atuk. Segai-Modang languages have no evidence of prolonged contact with Kenyah, as the majority of Kayanic loanwords in Kenyah are from the Kayan subgroup. Tunjung is sufficiently removed from both Segai-Modang and Kenyah to eliminate borrowing as an explanation.

Of particular interest are the many basic Austronesian vocabulary items replaced by forms with no known Austronesian etymology, like *ajən ‘fish’ and *atuʔ/*atuk ‘fish’ mentioned above. If this were an issue limited to the word for ‘fish’, or even an identifiable subpart of the lexicon, it might be possible to account for lexical replacements by taboos, antonymic language, or physical aspects of life in Borneo that gave rise to these changes. However, the lexical replacements without Austronesian etymologies put forward by Smith cover a wide range of semantic domains, from basic body parts, to color terms, to aspects of material culture.

Why are the languages of Borneo so phonologically and lexically divergent from the ancient Austronesian languages which gave rise to them? Why do they show areal phonological features typically associated with mainland Southeast Asia? Why do they show lexical replacement of common terms with words of apparent un-Austronesian origin? Blust (2001a, 2002, 2006, 2010) argues that Borneo is a “hot spot” of linguistic change, without attributing this to

2. A reviewer notes that even in Malayic and Land Dayak there is often lexical replacement, and if a language has a reflex of *hikan, it is unclear whether it is a retention or a borrowing from standard Malay.

any particular internal or external factor. Adelaar (1995), in contrast, tentatively attributes these properties to a possible Mon-Khmer substrate in Borneo. Based on phonological and lexical similarities between Land Dayak languages and the Aslian subgroup of Mon-Khmer languages spoken on the mainland, Adelaar suggests that these similarities could be due to Land Dayak people once speaking an Aslian or related Mon-Khmer language in Borneo, and later shifting to Austronesian.³ For a quarter of a century, the hypothesis of a possible Mon-Khmer presence in Borneo at the time of Austronesian settlement has not been seriously investigated, despite highly suggestive shared aspects of material culture including the traditional longhouse (Thurgood 1999:18; Blench 2010; Blust 2015). With the publication of the work of Smith (2017), and the increasingly comprehensive coverage of Mon-Khmer phonology and lexicon in the updated MKED as well as the comprehensive overview of Austroasiatic comparative-historical reconstruction (Sidwell and Rau 2015), a serious investigation of this question is now possible.

In this study, we focus on the lexicon, and ask: Are there words in the languages of Borneo without Austronesian etymologies that appear to be borrowings from (early) Mon-Khmer languages and, which, due to their distribution and/or meaning, are highly suggestive of *in situ* borrowing in Borneo? We do this by comparing Borneo lexical innovations compiled by Smith (2017) (with additional lexical material from Puri 2001; Rensch et al. 2006; and other sources) to Mon-Khmer reconstructions as gathered in the MKED, supplemented by recent work that has not yet been entered in the database (e.g., Kruspe 2010). The MKED is a digitized, IPA-standardized dictionary containing lexical data from individual MK languages, as well as reconstructions at every level of the non-Munda Austroasiatic family tree.⁴ To the extent that our mass comparison bears fruit, we believe it constitutes evidence for a Mon-Khmer presence in Borneo at the time of settlement by Austronesian speakers. The bulk of this paper is devoted to presenting hypothesized Mon-Khmer loan words that we have identified. As with any study of loan words, evidence for directionality includes phonological, morphological, semantic, geographical, and/or environmental considerations, and, in all cases, cognate-based factors. Since all Borneo lexemes examined here lack Austronesian etymologies, the existence of a potential Austroasiatic-internal etymology strongly suggests an MK source.

However, in order to proceed, we must establish a range of ground rules, as there are at least four factors that complicate this type of mass comparison. First, linguists as early on as Schmidt (1906) have suggested an ancient genetic

3. Adelaar (1995:51) also notes the possibility of there being “a third (unknown and now extinct) language spoken in Borneo and on the Malay Peninsula, and that its speakers in Borneo shifted to Land Dayak, while its speakers on the Malay Peninsula shifted to Aslian.” Since most of the Austroasiatic lexical comparisons in this paper are widespread within Mon-Khmer, and not limited to Aslian, we pursue the hypothesis that they were borrowed from a non-Munda Austroasiatic language into the languages of Borneo.

4. In addition to classical PMK reconstructions, MKED contains MK “clusters,”—lexemes that are similar in form and meaning, but which are difficult to derive from a single PMK form, indicated here by curly brackets {...}.

relationship between Austronesian and Austroasiatic languages based on both morphological (cf. Reid 1994) and lexical comparisons, as summarized in Blust (2013:696–704). Despite numerous problems with “the Austric hypothesis,” finding sound-meaning matches between Malayo-Polynesian words limited to Borneo and attested or reconstructed MK forms would be expected under the Austric hypothesis. In order to rule out a potential genetic signal, Borneo forms under consideration must be those with no plausible PAN or PMP/PWMP etymology, including derivations with PAN roots (in the sense of Blust 1988) that may be unattested elsewhere. A concrete example will illustrate the kind of data eliminated from this study. In arguing for a Central Sarawak group, Smith (2017:325) discusses the replacement of PMP *qali-matak ‘jungle leech’ with *tilaŋ ‘tiger leech’, a land leech marked by black, orange, and yellow stripes. Though, in this case, *tilaŋ might be compared to PMK *laŋ ‘with black markings’, or MK {clæŋ . . . } ‘leech (water)’ (cluster) these comparisons are not considered because Central Sarawak *tilaŋ ‘tiger leech’, may continue PAN *-laŋ ‘striped’. While it is very difficult to rule out all potential Austronesian etymologies for a particular lexeme, we have done our best to eliminate cases where an ancient relationship may be posited between PAN and PMK roots.

A second complicating factor in this mass comparison is the status of Malay. Though the origins of Malay are generally believed to be in the Malayic subgroup of western Borneo, speakers of Malay were established in southern Sumatra as early as the seventh century, and from there, the language spread to become a lingua franca of insular Southeast Asia for more than a thousand years, used by Malay traders, who had significant contact with MK languages of mainland Southeast Asia (Adelaar 1992, 2000).⁵ A Malayic word that appears to be borrowed from MK could have been borrowed any time after the migration of Malay speakers to Sumatra (or elsewhere), and hence, would not constitute evidence of an MK presence on the island of Borneo at the time of Austronesian settlement. In order to rule out MK loans into Malay that made their way back to Borneo languages via Malay contact, Borneo forms under consideration should either (i) have no Malay cognate; or (ii) on the basis of form or meaning be distinguishable from a direct Malay loan. In presenting suggested Mon-Khmer loans, Borneo lexemes are separated into two main classes: those without Malay cognates, which, we believe, are the strongest evidence for *in situ* contact (section 2.1) and loans with Malay cognates which satisfy condition (ii) above (section 2.2). A great deal of work on identifying Malay loans in the languages of Borneo has already been done by Smith (2017) and Rensch et al. (2006), and we follow their criteria where relevant.

A third complication of this lexical exercise is the sustained contact between Austronesian (pre-)Chamic speakers and speakers of MK languages, starting sometime around 600 BC and lasting over a millenia (Thurgood 1999:16). During this time, loans went from Chamic into MK languages, and from MK languages into Chamic. Contact languages include Bahnaric, Vietic,

5. For shared etyma between Malay, Khmer, and Thai, see Tadmor (2009:694).

Katuic, Mon, and Old Khmer. If an MK form could be interpreted as a loan from Chamic, we eliminate the comparison set from consideration.⁶ For example, Smith (2017:294) reconstructs pan-Bornean *aʔiŋ/*iʔiŋ and *akiŋ/*kaʔiŋ ‘waist’ based on Kayanic, Kenyah, Punan, Müller-Schwaner, and Berawan-Lower Baram forms. There is no Malay cognate here, but Thurgood (1999:338) reconstructs Proto-Chamic *kaʔiaŋ ‘loins; waist’ in his section on words with uncertain origins. Because it is possible that the Borneo forms originate from Chamic, they are not considered here (although such uninvestigated borrowings merit attention elsewhere).

A final complication of taking a random word or low-level reconstruction from languages of Borneo and looking for a possible loan source in the thousands of entries in the MKED is the issue of ensuring that the sound-meaning matches we identify are not chance occurrences. Language comparison for the purposes of historical reconstruction is reliant on the identification of regular sound correspondences. Loan phonology, in contrast, appears to be based for the most part on surface phonetic perceptual matching (Silverman 1992; Peperkamp and Dupoux 2003; Peperkamp 2005; Kang 2010; de Jong and Cho 2012), and in the world that was prehistoric Borneo, may have involved borrowing from one or more Mon-Khmer languages into one or more Austronesian languages. Without sound correspondences, we must rely on surface similarities in form, with all its attendant dangers. With small words and many dozens of diverse languages, the possibility of encountering chance resemblances is real. However, this source of false positives is constrained by restricting the search to only those Bornean words that have no Austronesian etymologies and by allowing semantic drift only where sound matches are exact and drift may be independently supported.

On the phonological side, we adhere to fairly strict phonetically based matching of hypothesized PMP and PMK sound systems and their near-descendants, as shown in table 2.⁷

As illustrated by these surface matches, voicing, place of articulation, manner of articulation, and nasality must match for consonants. For vowels, phonetic matching is required with tonic vowels, but vowels that are absent in MK may be epenthetic in Austronesian languages, and thus, are not required to show specific qualities (#23). A final phonotactically motivated correspondence is #32 in table 2, where an initial vowel in an Austronesian form may

6. See Thurgood (2007) on Acehnese as a Chamic language, earlier work on the Mon-Khmer influence on Acehnese, the migration of Acehnese speakers to Sumatra, and significant subsequent contact with Malay.

The possibility that the loans we identify are from Khmer is highly unlikely. First, clear cognate forms are absent. Second, most proposed loans are words for endemic Borneo species or local features.

7. Recall that “Proto-Mon Khmer” is to be interpreted as approximating PAA. The PMK inventory shown here includes reconstructed phones that occur in the MKED. While phonetic matching in table 2 is given for proto-phonemes of the respective languages, the same phonetic matches hold for sounds in daughter languages as well, and are used in all of our comparison sets. When referring to match numbers in table 2, we write #1, #2, etc.

TABLE 2. PHONETIC MATCHES PROPOSED FOR LOANS FROM MK INTO EARLY MALAYO-POLYNESIAN.

PMK	*p	*t	*c	*cr-	*k	*k/C	*ʔ	*m	*n	*ŋ	*ŋ
PMP	*p	*t	*c, *-t	*sr-	*k	*c	ʔ, ø	*m	*n	*ŋ, *j	*ŋ
match #	1	2	3	3a	4	4a	5	6	7	8	9
PMK	*b	*d	*j	*g	*g/C	*b	*d	*f	*l	*r	
PMP	*b	*d	*z, *-y, *c-	*g	*j	*b	*d	*d, *z, *y	*l	*r, *R	
match #	10	11	12	13	13a	14	15	16	17	18	
PMK	*s	*h	*w ⁸	*y ⁹	*#CC	*V ₁ V ₂					
PMP	*s	*h, ø	*w, ø	*j, ø	*#CVC	*V ₁ V ₂ , *V ₁ , *V ₂					
match #	19	20	21	22	23	23a					
PMK	*i	*u	*e	*o	*ɛ	*ɔ	*a	*ə	# σ #		
PMP	*i, *y	*u, *w	*e	*a, (o)	*ɛ	*a, (o)	*a	*a, *ə	#V σ #		
match #	24	25	26	27	28	29	30	31	32		

have no MK match, due to its epenthetic status in satisfying a disyllabic word template (cf. [Blust 2007](#)).

To exemplify the methodology, let us reconsider the lexical replacement noted earlier by which PMP *hikan ‘fish’ is replaced by a near-cognate *atu?/*atuk across Borneo ([Smith 2017:293](#)).¹⁰ First, we confirm that this form does not have any plausible cognates within the Austronesian language family by checking all databases available to us, most importantly the ACD. We then modify our understanding of *atu?/*atuk ‘fish’ based on data in [Puri \(2001\)](#) and [Rensch et al. \(2006\)](#). In this case, [Puri \(2001\)](#) lists seven cognate forms from N. Sarawak languages (Leppo’ Ké, Leppo’ Ma’ut *atok*; Uma’ Long *ata?*, *ato?* ‘carp’; Uma’ Lasan *atuk*; Uma’ Alim, Uma’ Badeng, Uma’ Bakung *atok*), while no cognates are found in [Rensch](#). While [Puri’s](#) forms are consistent with [Smith’s](#) reconstruction, Uma’ Long -ʔ < *-k, results in adjustment of the Borneo form to *atuk. The third step is to search the MKED for possible words meaning ‘fish’ or something closely related to ‘fish’, with the form [(*)ʔ(V)tuk]. In this case, a form fit is exact, but the meaning is not. As detailed in section 2.1.1 below, a PMK root *[ʔ]tuuk ‘to scoop up, root up’ is continued in Katuic and Monic with specific reference to dipping up fish, and the Katu *ʔatuuk* ‘dip up (fish)’ appears to be borrowed into other MK languages: Thavung (Vietic) *ʔatók* ‘to fish’; Chong (Pearic) *ʔaj tük* ‘fish trap’; Car (Nicobarese) *ha-tuək/ha-tüök* ‘to fish (w/out rod) with long lines’, suggestive of borrowed fishing technology. While the semantic match is not perfect, ‘fish’ (n.) vs. ‘to fish (by means of...)’, the phonetic match is, and the term is widespread in the Mon-Khmer speaking world. We decide that this comparison is worthy of discussion, and offer it in this paper.

Section 2, presents pan-Bornean ^x*atuk ‘fish’, along with other hypothesized loans from Mon-Khmer into languages of Borneo that meet the criteria

8. Syllable-final MK glides may color preceding vowels in suggested An loans.

9. Syllable-final MK glides may color preceding vowels in suggested An loans.

10. [Smith \(2017:293\)](#) reconstructs pan-Bornean ^x*atu?/*^xatuk ‘fish’ with ^x*atu? based on *Tujung mətu?* alone. See section 2.1.1 for further discussion.

outlined above, with sections in order of strongest to weakest. Section 2.1 includes pan-Borneo reconstructions that lack Malay cognates, while section 2.2 offers reconstructions limited to (Greater) North Borneo that also lack Malay cognates. Sections 2.3 and 2.4 provide pan-Borneo and (Greater) North Borneo reconstructions which have Malay cognates, but where there are good arguments that the Borneo forms are not Malay loans. Section 2.5 summarizes the loan data with respect to Borneo subgrouping and distributions, loan source within the Mon-Khmer language family, and semantic domain of borrowings. In section 3 we offer some concluding remarks on wider implications of this study for future work on Borneo as a potential contact zone. The appendix includes Borneo words that may appear to be *in situ* MK loans, but which fail one or more of the criteria outlined above.

Before turning to our own comparison sets, mention should be made of three words from Borneo languages that, in earlier work, have been argued to be MK loans with borrowing hypothesized *in situ* in Borneo: Proto-Malayic *pərut ‘stomach’ (Adelaar 1992:129), Land Dayak words for ‘die’ (Skeat and Blagden 1906:773; Adelaar 1995:90) and a range of words for ‘bathe’ (Adelaar 1995:90) with root *mu*. Malayic *pərut ‘stomach’ is discussed in the appendix, as it satisfies all but one criterion for possible MK loans laid out above. Central Sarawak and Land Dayak *kabas/*kəbəs ‘die’ (Smith 2017:338), in contrast, show matching forms only in Aslian languages, with no etymology in MK. Here, we agree with Smith (2017:338) that for *kəbəs, “without further lexical data and additional positive evidence it is difficult to evaluate the weight of this comparison.”¹¹ Terms for ‘bathe’ in Central Sarawak and Land Dayak appear to have a common root *-mu (Smith 2017:395), and also satisfy most of the criteria laid out above. However, this set is also relegated to the appendix, due to similar PAN and PMK roots, whose *m(V)- prefixed form yields *mu(h): compare PAN *buqbuq ‘pour, pour out’ and PMK *buh ‘to pour, sprinkle’.

In section 2, all data from the languages of Borneo is from Smith (2017) (Sm) unless noted otherwise. Other primary sources for Borneo languages are Rensch et al. (2006) (RRNR), Puri (2001) (RKP), Scott (1956) (NCS), and Smith (2018, 2021). On occasions, we have taken words from the Borneo Dictionary (BD). Data from Austronesian languages outside of Borneo as well as reconstructed Austronesian forms are from the ACD (Blust and Trussel 2020) unless noted otherwise. All data from Mon-Khmer languages including reconstructed Mon-Khmer forms are from the MKED unless noted otherwise. By restricting ourselves primarily to these sources, we have undoubtedly limited the amount of lexical material under evaluation, but have done so in the hope of including phonological and phonetic descriptions and reconstructions of a consistent

11. Though, perhaps consideration should be given to PMP *kabus ‘run out, come to an end’ as a potential source of the Borneo forms.

Another pair that is difficult to evaluate is Southwest Sabahan *dasam ‘rain’ (Sm:212), which has been suggested as a loan from Aslian based on Batek *lāsəm*, Semelai *lām* (Blench 2010:135).

nature, where questions regarding, for example, the sound systems of PAN, PMK, and all of their descendants are concerned.

2. MON-KHMER LOANS IN THE AUSTRONESIAN LANGUAGES OF BORNEO. The Mon-Khmer loans into Austronesian languages of Borneo suggested here are lexemes which do not appear to have any clear Austronesian etymologies, and, on the MK side, are lexemes which do have MK-internal etymologies. In presenting these, we use the following conventions: << ‘loaned from’; >> ‘loaned into’; ^{x*} ‘reconstruction of a non-inherited lexeme’. Loans are ordered alphabetically by gloss. Phonetic matches between the Borneo form and the suggested Mon-Khmer loan source adhere, for the most part, to those outlined in table 2. Comparison sets show Borneo data in the upper section, and MK data below. Note sections below each comparison set offer strengths and weaknesses of the comparison. MK languages/subgroups listed in the first line of the comparison after ‘<<’ (e.g., MK, Katuic . . .) are the languages with reconstructed or attested forms that could serve as potential loan sources for the particular item in question, satisfying the phonetic matchings in table 2, and with similar meanings. (If MK is not listed, this means that the PMK form, if reconstructed, is not a good phonetic and/or semantic match with the Borneo term.) In section 2.5, these loan sources are summarized and reviewed.

2.1. PAN-BORNEO LEXEMES WITH NO MALAY COGNATES. Comparisons in this class are the strongest of those proposed, since contact with MK languages outside of Borneo can, for the most part, be ruled out.

2.1.1. Fish₁.

Pan-Bornean ^{x*}atuk ‘fish’ << MK, Katuic, Monic
Cf. Pan-Bornean *atu?/*atuk (Sm:293)

North Borneo

Murutic

Bulusu	atuk kapon ¹² ‘hook’
--------	---------------------------------

Kenyah

Proto-Kenyah	*atuk (Sm:661)
--------------	----------------

Pawe, Gah, Laang, Sawa, Tau, Badeng	atok
--	------

Uma’ Long	ata? ‘fish’, ato? ‘carp’ (RKP:229, 250)
-----------	---

Kayanic

Segai-Modang

Long Gelat	tək, tək
------------	----------

Modang	təwk, tuk
--------	-----------

Gaai	təwk
------	------

Kelai	atok
-------	------

12. Bulusu *kapon* (Sm:639) is not translated.

Basap-Greater Barito

Barito

Tunjung	mə-tuʔ (cf. prefix in mə-rua ^k ŋ ‘twins’, Taboyan <i>ruaŋ</i> ‘two’)	
PMKA	*[ʔ]tuuk	‘to scoop up, root up’ ¹³
PMKB	*[ʔ]t[u]k	‘to scoop up, root up’

Katuic

Proto-Katuic	*tuuk	‘dip up (fish)’
Katu	(ʔa)tu:k	‘net fish by hand’
Katu	ʔatuuk	‘dip up (fish)’
Ngeq	tuuk	‘dip up (fish)’
Ta’Oi	tuuʔ	‘dip up (fish)’
Pacoh	tuəʔ	‘scoop fish with woven scoop’

Monic

Proto-Monic	*took	‘to scoop up; to dip up’
Old Mon	tuk	‘to draw (water)’
Mon	hətɔh	‘fish-trap constructed by damming section of ditch or small creek and baling out water, so that fish leaping dam are stranded’

Smith (2017:293) reconstructs the doublet *atuʔ/*atuk ‘fish’ with *atuʔ based on Tunjung *mətuʔ*, since, internal to Tunjung, *-k > -ʔ is unmotivated. We have added the RKP and BD forms to Smith’s original comparison set. The Bornean doublet could reflect: (i) two independent instances of borrowing; (ii) a single instance of borrowing, with an irregular development in Tunjung; or (iii) alternatively, the Tunjung form could be borrowed from a Kayanic or Land Dayak language where regular *-k > -ʔ occurred. Smith (2017:293) states that “Tunjung is sufficiently removed from both Segai-Modang and Kenyah to eliminate borrowing as an explanation.” However, earlier, it is shown that “Kayanic words have made their way into Tunjung” (Smith 2017:266). If Tunjung *-tuʔ* of *mətuʔ* ‘fish’ is treated as a borrowing from a North Borneo language, then this comparison set should be moved to section 2.2.

We suggest pan-Bornean (or North Bornean) ^{*}*atuk ‘fish’ borrowed from MK *tuuk with semantic extension of ‘fishing scoop’ to ‘fish (n.)’. The phonological match is regular, with an initial vowel added per #32.

13. A reviewer points out that Shorto’s initial bracketed glottal stop is based on a nonetymological ʔa- presyllable in Katuic, which is not relevant here.

Only MK reflexes with meanings related to fish and fishing are listed here.

An alternative is that the Bornean form is a Katuic loan. Katuic forms appear to be borrowed into other MK languages, strengthening this as a potential source for Bornean forms: cf. Thavung (Vietic) *ʔatɔk* ‘to fish’; Chong (Pearic) *ʔəj tɨk* ‘fish trap’; Car (Nicobarese) *hatu:k* ‘fish (generally)’ *ha-tuək* ‘to fish (without rod) with long lines’, suggestive of borrowed fishing technology.

A weakness of this comparison is the semantic distance between ‘fish (n.)’ and ‘dip up fish, use a dipnet to scoop fish up (v.)’. However, loan semantics in other MK languages noted above show similar semantic broadening, and the patient-oriented interpretation of Austronesian roots (Kaufman 2009) would make such a shift natural. A parallel shift appears to have occurred in Acehnese *suəŋ* and PC **saəŋ* ‘house’, which Sidwell (2005:235) identifies as borrowings from the Khmer root *saəŋ* ‘to build’. The MK root is attested as a verb, while its borrowings into Austronesian are resultative nouns. Outside of Austronesian, similar semantic shifts are in evidence: compare *catch* (v.) to *catch of the day* where *catch* refers to a fish.

Replacement of PMP **hikan* ‘fish’ across much of Borneo (with the exception of Proto-Land Dayak ***iʔkan*, Bidayuh **iʔkan*, **iʔken*, **iʔke:ʔ* ‘fish’) by distinct lexemes may suggest a fishing taboo whereby the word for ‘fish’ was not spoken during active fishing.

2.1.2. Fish₂.

Pan-Bornean **ʔojən* ‘fish’ << MK, Bahnaric, Katuic, Palaungic
Cf. Greater North Borneo **ajən* ‘fish’ (Blust 2010); Proto-Müller-Sch. **ocen*
(Sm:282)

Central Sarawak

Melanau

Kanowit	jən
---------	-----

Kajang

Sekapan, Lahanan, Kejaman	jən
---------------------------	-----

Punan

Punan Aput	jen
Punan Lisum, Ukit, Buket	ajən

Müller-Schwaner

Kereho, Septuan, Aoheng	ocen (c<*z)
Hovongan	cien (c<*z)

Kayanic

Murik-Merap

Ngorek	sən (j,c<*z)
Merap	can (j,c<*z)

Basap-Greater Barito

Barito

Kadorih	ocin (j,c < *-z-)
---------	-------------------

<i>Basap</i>		
Segai Basap	ujən	
Tabalar Basap	aɣən	
PMK (cluster)	{chen, jaan . . . }	‘basket (general)’
<i>Monic</i>		
Mon	chen	‘basket (general)’
<i>Khmuic</i>		
Khmu	jaan	‘basket (general)’
<i>NBahnaric</i>		
Sedang	can	‘fish trap’
<i>Katuic</i>		
Bru	can	‘kind of fish trap which looks like a cage’
	cuan	‘kind of cone-shaped dipping net with criss-crossed handle’
<i>Palaungic</i>		
Danaw	tsən	‘fishing net’
<i>Nicobaric</i>		
Nancowry	ʔejun	‘large fish basket trap’

We suggest pan-Bornean $^{*}ojən$ as a borrowing from MK $^{*}ju\{a,\ə\}n$ ‘fish basket trap’. Note that no single PMK form is reconstructed for this word cluster. These MK forms may be irregular continuations of PMK $^{*}jaan$ ‘to fish with scoop-net’ (cf. Mon *càij* ‘to fish with scoop-net’, Lawa *ʔacuanj* ‘to net (fish)’, Khmer *chni:əj* ‘scoop net’ (infix), *cəj-n-coat* ‘to catch fish (by scooping them up) in a basket, gather fish into a basket’ (*coat* ‘to pour out, strain or drain off a liquid’)). Though PMK $^{*}-ŋ$ is typically stable, forms for ‘fish trap’ above would show irregular $^{*}-ŋ > -n$ influenced by words for ‘basket’ with final *-n*. The initial vowel in the borrowing could be interpreted either as the consequence of disyllabic bulking (Blust 2007), or, as the direct borrowing of a presyllable from a morphologically complex MK form (cf. Nancowry *ʔejun* ‘large fish basket trap’). Rounding of the initial vowel in Bornean forms is interpreted as anticipatory coarticulation of rounding in the following syllable. Voiced- and voiceless palatals in Borneo forms appear to reflect regular internal sound changes (as indicated by expected reflexes of PMP $^{*}z$), consistent with early borrowing, prior to language diversification in Borneo. However, it is also possible that an early form with $^{*}j-$ was borrowed, as well as a later form with $^{*}c-$.

A weakness of this comparison is the semantic distance between ‘fish (n.)’ and ‘fish basket trap’, or ‘basket’, although this is consistent with the patient orientation noted in section 2.1.1 above. Another weakness is the “cluster” status of the MK reconstruction itself.

As noted in section 2.1.1, replacement of PMP *hikan ‘fish’ across much of Borneo (with the exception of Proto-Land Dayak **ikan, Bidayuh *ikan, *iken, *ike:ʔ ‘fish’) by distinct lexemes may suggest a fishing taboo whereby the word for ‘fish’ was not spoken during active fishing.

2.1.3. Hornbill, helmeted.

Pan-Bornean ^x*tukuj ‘horny bulge (of hornbill); helmeted hornbill’

<< MK, Bahnaric

Cf. Proto-Kayanic *tukuj ‘helmeted hornbill’ (Sm:456)

Malayic

Ibanic

Iban	taŋkuj	‘horny excrescence on the beak of a hornbill; the bulge at the back of a person’s head’
	duŋkuj	‘the bulge of the forehead or back of the head’

Central Sarawak

Punan-Müller-Schwaner

Punan Bah	tukuəŋ	‘forehead’
Ukit	tukoŋ	‘forehead’

Kayanic

Proto-Kayanic	*tukuj	‘helmeted hornbill’
---------------	--------	---------------------

Basap-Greater Barito

Barito

Kapuas	tukuj	‘helmeted hornbill’
Bakumpai	tukuj	‘helmeted hornbill’
Maanyan	tukuk	‘helmeted hornbill’
Dusun Witu	tukuk	‘helmeted hornbill’

Basap

Basap	tukuj	‘helmeted hornbill’
-------	-------	---------------------

PMKA	*tg[uə]ŋ	‘bony ridge’
------	----------	--------------

Bahnaric

Bahnar	təkəŋ	‘profile, ridge (of nose), shin’
--------	-------	----------------------------------

Katuic

Bru	ŋkùəŋ	‘ridge of nose’
-----	-------	-----------------

The helmeted hornbill is endemic to Borneo, peninsular Malaysia and Sumatra. Austronesian speakers arriving in Borneo would not have seen this bird before. The helmeted hornbill is one of the largest hornbills in Borneo, and is culturally important there. According to Smith (2017:327), most speakers know the animal, if not by name, then by its distinctively long tail feathers which are collected for cultural purposes. Blust (2002:115) adds: “Perhaps nowhere else in the AN-speaking world does the hornbill attain such cultural

importance as in Borneo, but here, despite occasional reflexes of *kalaw, as in Kayan and Iban, names for the hornbill are often innovative.” Given these facts, that a word for hornbill might be borrowed from another language is not surprising.

We suggest that **tukuaŋ is borrowed from an MK form like *t(ŋ)ku(u)ŋ ‘bony ridge’, with reference to the horned bill of the bird, with devoicing of *tg > *tk (cf. Bahnar, Bru), but is otherwise identical to PMK.

It is possible that Proto-Kayanic *təkuəŋ ‘rhinoceros hornbill’ (Sm:465) has the same MK source.

2.1.4. Hornbill, rhinoceros.

Pan-Bornean **ti{n,ŋ}aŋ ‘rhinoceros hornbill’ << MK
 Cf. *tiŋaŋ ‘rhinoceros hornbill’ (Sm:465)

Kayanic		
Long Naah, Balui Liko	tiŋaŋ	
Central Sarawak		
<i>Kajang</i>		
Kejaman	tiŋan	
Lahanan	tiŋaŋ	
<i>Müller-Schwaner</i>		
Kereho, Septuan, Aoheng	tiŋaŋ	
Malayic		
Ketapang	tiŋaŋ	
Basap-Greater Barito		
<i>Barito</i>		
Kadorih	tiŋaŋ	
Ngaju	tiŋaŋ	
PMKA	*tniəŋ	‘forehead’
<i>Aslian</i>		
Semnam	cnəŋ	‘casque of a hornbill’
Kensiu	canōɣ	‘iliac crest’ (top border of the largest of the pelvic bones; bony surface you can feel when you press into your hips)
<i>Khasic</i>		
Mnar	təŋnaŋ	‘forehead’
Pnar	t ^h ŋa	‘forehead’
<i>Khmeric</i>		
Khmer	tnaŋ	‘joint; bulge’
Khmer	tŋah	‘forehead; Mount of Venus; pubis’
<i>Monic</i>		
Middle Mon	tnəŋ	‘forehead’

The rhinoceros hornbill is one of the largest of the hornbills in South East Asia and a symbol of Borneo. The bird is central to many traditional cultures, and its tail feathers are highly valued. Smith (2017:295) finds it “surprising that a word for hornbill cannot be reconstructed,” but the diversity of species in Borneo and their distinct behaviors, calls, and ecologies might lead one to expect diverse terms. That words for local hornbill species might be borrowed is unsurprising (see section 2.1.3).

We suggest that $^{x*ti\{n,\eta\}a\eta}$ is borrowed from an MK form like $^{*t\eta i\eta}$, $^{*t\eta a\eta}$ ‘forehead; bony protrusion; casque’. At least one language in Borneo suggests the borrowing of a cognate root with a meaning closer to the original: Ribun (Land Dayak) *tonuq\eta* ‘forehead’.

2.1.5. Leech.

Pan-Bornean $^{x*}j\grave{a}l\grave{a}w$ ‘leech’ << MK, Bahnaric, Palaungic

Central Sarawak

Punan-Müller-Schwaner

Proto-Punan	$^{*}jalaw$	‘earthworm’ (Sm:121,331)
Punan Lisum	jalow (jikit)	‘leech’ (jikit ?‘bite’)
Punan Derian	jelea’	‘worm’ (RKP:250)
Punan Tuvu’	j\grave{a}leh	‘intestinal worms’
Ukit	jalo	‘intestinal worms’
Punan Bah	jeluow	‘earthworm’
Beketan	jalow	‘earthworm’
Punan Aput	jalow	‘earthworm’
Punan Lisum	jalo	‘earthworm’
Ukit	jalo	‘earthworm’
Buket	jalo	‘earthworm’

Basap-Greater Barito

Barito

Kadorih	jorow	‘leech’
Dusun Witu	lelaw	‘leech’

Basap

Basap	j\grave{a}lo	‘leech’
-------	--------------	---------

PMKA	$^{*}gl\grave{a}w$	‘(kind of) leech’
------	--------------------	-------------------

Bahnaric

Nyaheun	gliw	‘leech’
Sapuan	gli:w	‘leech’ (water type)
Chrau	glu:	‘water leech’
Stieng	glu:	‘leech’

Palaungic

Proto-Palaungic	$^{*}gl\grave{a}w$	‘tadpole’
-----------------	--------------------	-----------

The initial /j-/ (cf. PMP *z) in Borneo forms is a match with MK *j (#12) or MK *g- in *gC clusters (#13a).¹⁴

The original meaning ‘leech’ is maintained in the Basap-Greater Barito languages, but shifted to ‘worm’ in Punan languages.¹⁵

A potential weakness of this comparison concerns an Austronesian look-alike. The An forms here must be distinguished from similar words ending in [ŋ] that may continue PWMP *gelaŋ ‘earthworm; intestinal worm’ (e.g., Southwest Sabah *liŋguaŋ*, *liŋkuaŋ* ‘earthworm’ [Sm:210].) The forms above are clearly distinct, as these languages continue PWMP *-ŋ as /ŋ/: compare pPunan *tolaŋ (Punan Tuvu’, Punan Lisum, Punan Aput, Ukit, Buket *tolaŋ*), Kadorih *turaŋ*, Dusun Witu *tulaŋ*, Basap *tulaŋ*, all from PWMP *tuqəlaŋ ‘bone’. The *g/j correspondences would also be irregular.

Compare also Central Sarawak *tilaŋ ‘tiger leech’ (Sm:325, 330), an An term that likely refers to the stinging land leech (aka jungle leech) which is striped black, orange, and yellow, possibly continuing PAN root *-laŋ ‘striped’.

Why would a word like ‘leech’ be borrowed? An Austronesian word for ‘leech’ is reconstructed with variants of the *qali- prefix, a prefix for words with a sensitive reference to the spirit world (Blust 2001b): PWMP *qali-matek ‘jungle leech’ (cf. PAN *qaNi-matek). This spiritual connection could result in higher rates of lexical replacement, including the replacement of native words with loanwords, or irregular developments of native terms. This might explain why, independently, Chamic borrowed a word for ‘leech’ from MK: Proto-Chamic ^x*plum ‘land leech’ from Bahnaric *pləəm (Thurgood 1999:328).

2.1.6. Steep, aslant, sloping.

Pan-Borneo ^x*siraŋ << MK, Monic, Khmeric, Bahnaric

Malayic

Ibanic

Iban	siraŋ	‘a cleft, fissure, depression’ (NCS:176)
	siruŋ	‘crooked’ (NCS:177)

Basap-Greater Barito

Barito

Ngaju Dayak	sidaŋ, siraŋ	‘oblique, slanting’ (as something that was cut at an oblique angle)
Malagasy	sírana	‘sloping’

PMK	*cra(a)ŋ	‘bank, embankment’ ¹⁶
-----	----------	----------------------------------

Monic

14. A reviewer critiques the original Shorto MK cognate set, saying that Tampuan *chliiw* and Khmu *cliə*, which were included, continue a distinct *cliiw, not *gləw. It is possible that the MK loan source had an initial palatal *j.
 15. A reviewer sees a central problem with a shift of meaning between ‘water leech’ and ‘worm’. Compare, however, Ilokano *alimá* ‘earthworm; leech’ < PWMP *qali-metaq ‘paddy leech’, or Kachok *klan* ‘leech, tapeworm’ (Olsen 2018:122), both clear instances of colexification of these two meanings in AN and MK, respectively.
 16. A reviewer suggests MK *jraŋ instead. This form is also compatible with our hypothesis.

Old Mon	craŋ	‘bank’
<i>Khmeric</i>		
Khmer	craŋ	‘(steep) bank’
<i>Bahnaric</i>		
Stieng	c ^ɛ r:aŋ	‘(steep) bank’

The two Barito forms are listed under “noise” in the ACD. Malay *curam* ‘steep, steep slope’ with final *-m* and initial *c-/si-* is distinct enough to rule this out as a Malay loan (table 1, 3a). The same holds for Malay *juraŋ* ‘cliff’.

Internal to MK, PMK *sraŋ ‘tooth, sharp projection’ appears to be unrelated.

The rugged terrain in Borneo and the difficulty of navigating steep embankments might make this a salient local term.

2.2. NORTH BORNEO LEXEMES WITH NO MALAY COGNATES.

Comparisons in this class are strong, since contact with MK languages outside of Borneo can, for the most part, be ruled out. If these loans predate the split of the North Borneo and Basap-Barito languages, reflexes in Basap-Barito have been lost. Alternatively, borrowing may have taken place after the split, suggesting a continued presence of MK speakers in Northern Borneo.

2.2.1. Ant, red.

*kəsá(?) ‘red ant’ << MK, Aslian, Katuic

Malayic

Ibanic

Mualang	kəsá?	‘ant’
Iban	kesa?	‘red ants’ (NCS:90)
PMKA	*ksəw?	‘red ant’ ¹⁷

Aslian

Semang	kaso?	
Kensiu	les kəsə?	‘red ant’ (les ‘ant’)

Katuic

Katu	saw, kasaw	‘red ant’ (sour)
Ngeq	kasa:w	‘red ants’

Monic

Proto-Monic	*(-)ksaw	‘red ant’
Mon	kechao	‘red ant’

Palaungic

Proto-Palaungic	*sə?	‘red ant’
-----------------	------	-----------

Pearic

Proto-Pearic	*ksu:	‘red ant’
--------------	-------	-----------

17. A reviewer suggests MK *ksu?. This does not significantly alter the comparison, as MK *ksə? would yield a loan-form with *a.

If these are MK loans, since MK *k-, *-s- and *-ʔ are all continued only in Aslian, then either borrowing preceded the break-up of PMK or is from Aslian. However, Mualang and Iban have sporadic addition of glottal stop word-finally (Smith 2017:192–93), so a borrowing without a final glottal stop should also be considered.

As these two Ibanic words have no known Borneo cognates, they might be considered suspect. On the other hand, the borrowing of at least one other MK term for ‘ant’ (see section 2.4.1), and the good sound/meaning match suggest loan status.

Note the form-meaning relation between PMKA *ksəwʔ ‘red ant’ and PMKA *ks[aw]ʔ ‘red’, PAA *-saw ‘red’.

At least one word for ‘ant’ was borrowed from Chamic into MK: Proto-Chamic *sidəm* (< PMP *sejem), Wr. Cham *hadəm*, Bahnar *hadam* (Thurgood 1999:303).

Seven different words for ant or termite are reconstructed to PMK, while only one such word (for termite) is firmly reconstructible to PAN.¹⁸ This may relate not only to the diversity of species in the original MK homeland, but also to the custom of eating many different ant varieties (and avoiding the eating of others). For an overview of edible insects in Borneo, see Chung (2010).

2.2.2. Barking deer₁ (*Muntiacus*).

North Borneo ^x*paus ‘barking deer’ << MK, Palaungic, Aslian

North Borneo

Murutic

Bulusu	paus	
Abai	faus (itom)	‘common barking deer’ (RKP:200)
Abai	faus (ria’)	‘Bornean barking deer’ (RKP:201)

Central Sarawak

Melanau

Kanowit	puyh	
---------	------	--

Land Dayak

Bidayuh

Bukar–Sadung	paəh	‘deer’ (RRNR:326)
--------------	------	-------------------

PMKA	*pus	
------	------	--

PMKB	*puəs	
------	-------	--

Aslian

Semai	pous	
-------	------	--

Bahnaric

Bahnar	poh	‘of deer, to bark’
--------	-----	--------------------

18. The ACD shows PAN *SayaN ‘termite’, with doublet *aNay ‘termite, white ant’, and *alujah ‘ant sp.’ which is based on Formosan (Saisiyat and Amis) data alone. At the PMP level, another term, *me(n)tik ‘ant sp. with venomous bite’ is found.

Katuic

Bru	pə:ɰh
-----	-------

Palaungic

Lawa	paus, pauh
Lamet	po:s

North Borneo ^x*paus appears to be a loan from an early MK form *paus ‘barking deer’. An *-s > -(y)h sound change is observed from PMP for some C. Sarawak languages, like Kanowit (Smith 2017:100, 394), so borrowing would have preceded that sound change.

The common barking deer (*Muntiacus muntjak*) is widely distributed in Southeast Asia, and the Bornean barking deer (*Muntiacus atherodes*) is endemic to Borneo. PAN *sakəC ‘barking deer’ is reconstructed on the basis of Formosan evidence alone, and this, as well as PAN *dekiŋ ‘to bark, of a deer’ (based on Puyuma and Malay) suggest that the barking deer was known to Austronesian speakers prior to their migration from Taiwan south. Since the barking deer is not endemic to the Philippines, the word may have been lost there. When Austronesian speakers moved yet again to the west, settling in Borneo, they encountered barking deer again, and, seemingly, innovated new terms for them. See section 2.4.2, and BARKING DEER₃ and BARKING DEER₄ in the appendix.

2.2.3. Cold.

Greater North Borneo ^x*səŋəm, ‘cold’ << ?MK, Bahnaric, Khasic
Cf. Near cognates *səŋəm, *siŋəm, and *səŋim (Sm:294)

Bulungan	səŋəm
----------	-------

North Borneo*Kenyah*

Uma Pawe, Lepo Gah	səŋim
Lepo Sawa, Lepo Tau, Badeng	səŋim

Central Sarawak*Kajang*

Lahanan	səŋim
---------	-------

Müller-Schwaner

Hovongan, Kereho	siŋəm
Seputan, Aoheng	siŋəm

Kayanic*Murik-Merap*

Ngorek	ŋəm
Merap	hŋəm

Kayan

Data Dian, Bahau	həŋəm
Busang	həŋəm

<i>Segai-Modang</i>		
Long Gelat, Modang	həŋam	
Kelai	sŋam	
PMK	*ŋam	‘immerse, soak’
<i>Bahnaric</i>		
Halang	həŋa:m	‘cold (water)’
Sedang	həŋiəm	‘cool’
Sapuan	baŋam, paŋam	‘cool’
<i>Khasic</i>		
Proto-Khasic	*-ŋam	‘cold’
War	fŋam	‘be cold’
Lynggam	təŋŋam	‘cold’
Khasi	saŋam	‘part of the jungle thick with trees and wet’
Khasi	sŋem	‘moist, wet’
<i>Monic</i>		
Nyah Kur	ləŋɔ̃m	‘cool’

Austronesian forms with both [i] and back non-high vowels [ə], [a], and [o] in the second syllable may suggest borrowing from an early MK language with a diphthong *iə in the second syllable. Alternatively, the vowel of the final syllable may show (irregular) variation due to flanking nasals.

Proto-Bidayuh *suŋoh, Proto-Bakati’ *saŋuoh ‘cold’ (Rensch et al. 2006) may be related to the Pan-Bornean forms suggested here, with *səŋom > saŋum > saŋu.

The classification of Bulungan is debated, but it does not appear to subgroup with Southwest Sabah or the North Borneo subgroup (Lobel 2016:5).

The PMK root *ŋam ‘immerse, soak’ has come to mean ‘cold, cool’ in derived forms: cf. Khasi *sŋem* ‘moist, wet’ and War *fŋam* ‘be cold’. This semantic shift is consistent with colexification in PMKA *lʔuət ‘wet, cold’ and supported crosslinguistically (see CLICS³). This is an important observation, since Iban *cəlap* ‘cold’ (borrowed into Land Dayak) may, ultimately have its source in MK *crləp, a derivative of PMK *ləp ‘to immerse’ (cf. Bahnar *həlyp* ‘to inundate’; Nyah Kur *crɔ̃p* ‘get wet all over the body’). However, this proposed loan is excluded from present consideration, since a potentially cognate root exists in PAN *-ləp ‘sink, submerge’.

It is unclear why *s-ŋam ‘be cold’ has not been reconstructed to PMK.¹⁹

19. A reviewer suggests that this is because the term was borrowed from Chamic into Bahnaric, and that the Khasian forms are chance look-alikes. We see no evidence of related lexemes in Chamic that would serve as the source for borrowing. Furthermore, the MK root appears in Khasic and Monic.

The association of meanings WET and COLD is unsurprising in thick jungle areas where sunlight is scarce.

2.2.4. Lung, lungs.

Greater North Borneo *sop ‘lungs; an inner cover’ << ?MK, ?Bahnaric, ?Katuic, Aslian

Kayanic

Segai-Modang

Modang	saup (Smith 2018)
--------	-------------------

Land Dayak

Proto-Land Dayak	*sop, *səp (RNNR:363)
------------------	-----------------------

Anah Rais, Biya, Sembaan	soop
--------------------------	------

Serambu, Bratak, Bau	əsop	‘animal lungs’
----------------------	------	----------------

Bidayuh Bau	ɔsuɔp	‘a lung’ (BD)
-------------	-------	---------------

Bau	əsəp
-----	------

Kembayan-A	səsə ^a p
------------	---------------------

Aslian

Jahai	sop
-------	-----

Kensiu	sɔp	‘lungs; texture like lungs’
--------	-----	-----------------------------

Semnam	sɔp
--------	-----

Temiar	sob	‘after-birth, placenta, the soft spongy substance expelled from the womb after birth’
--------	-----	---

Lack of cognate forms elsewhere in Borneo may be accidental: ‘lung’ is not an item on the extensive word list of Smith (2017), and forms here are taken from the description of Bidayuh from Rensch et al. (2006), as well as the Modang word lists of Smith (2018).²⁰

The only PMK-level reconstruction for ‘lung’ is *h-final: PMKB *tsɔh (cf. Proto-Bahnaric *-sɔh, Proto-Katuic *-sɔh). Whether Bornean *sop could be borrowed from a non-Aslian MK proto-language depends, first, on one’s reconstructed semantics (and morphology) of PMKB *sruup ‘to sheathe’ (cf. Proto-Bahnaric *sɔɔp ‘sheath, body hair’; Proto-Katuic *sɔp, *sɔɔp ‘covering’ (S&J:132), Sui sɔɔp ‘sheath-like, sock-like covering’; Khmer *stu:ap* ‘fibrous sheath; bract’).²¹ Question marks before MK, Bahnaric and Katuic express the questionable status of this semantic extension.

20. We treat these Borneo forms as distinct from Bulusu *opos*, Dumpas *opos*, Timugon *upos* ‘lung’ and other North Borneo words continuing *opos (Lobel 2016:159).

21. Munda may support PAA **sɔɔp ‘sheath, covering’: cf. Bodo-Gadaba *sopa* ‘shell (of egg), rind, skin (of vegetables)’; Gta’ *copa* ‘peel of fruits’(also used for egg shell, fish scales, skin of a mango seed, skin of brinjal) (SEAlang Munda Languages Project).

That some (non-Aslian) MK languages had *p-final words in this semantic range may find support in Chamic where an MK word for ‘lung’ was also borrowed: from (An *kulit ‘skin’+) MK ^x*sɔːʔ ‘lung’, one finds Chru *kəlsɔːʔ*, N. Roglai *kuliːʔ sɔːʔ*. But compare Acehnese *sūap* ‘lungs; placenta’, where final -p appears to be irregular (Thurgood 1999:325).

2.2.5. Pig-tailed macaque.²²

Bulungan area ^x*dok ‘pig-tailed macaque’ << MK, Bahnaric, Katuic

	‘pig tailed macaque’ (RKP:174)	‘slow loris’ (RKP:167)
--	-----------------------------------	---------------------------

North Borneo

Bintulu	beduuk (ACD)	
---------	--------------	--

Kenyah

Leppo’ Ké	dok	dok talun
Leppo’ Ma’ut	dok	dok kalun
Uma’ Long	doʔ	–
Uma’ Lasan	dok	dok malun, dok nalun
Uma’ Alim	dok	dok maliŋ, dok nalun
Uma’ Badeng	dok	dok talun
Uma’ Bakung	dok	dok talun

Dayic

Kelabit	bedhuk (ACD)	
---------	--------------	--

Central Sarawak

Punan

Punan Derian	(beruk)	duk alun
--------------	---------	----------

Kayanic

Murik-Merap

Merap	dauʔ, dao we	dauʔ laun
Pua’	(beruk atun)	dok alun

PMKA	*d̥uuk	‘(kind of) monkey’
------	--------	--------------------

Aslian

Semai	doʔ	‘monkey’
-------	-----	----------

Bahnaric

Proto-Bahnaric	*d̥ok	‘monkey’
Sre	dou	‘monkey’ (all macaques and semnopithecques)
Halang	d̥ok	‘monkey’

Katuic

Proto-Katuic	*d̥ok/*dook	‘k.o. monkey’
Katu	d̥ok, ʔad̥ok	‘red monkey with long tail’

22. Blench (2010:136) makes a similar comparison, though without the restrictive criteria detailed in section 1 that we employ here.

The pig-tailed macaque is found throughout Southeast Asia, including Borneo, and has a short, pig-like tail. Blust (2010) reconstructs Western Indonesian *bəduk and *bəRuk where both may be seen as irregular developments of PAN *beRek ‘domesticated pig’ (cf. Malay *beruk* ‘monkey’). Many languages across Borneo, including Punan Bah, Sekapan, Kejaman, Hoyongan, Kereho, Seputan, Aoheng, Lahanan, W Penan, and Basap, show reflexes of *bəduk. However, the forms shown above from the Bulungun area, all from Puri (2001:174, 167), are distinct in being both monosyllabic, and showing a non-high vowel /o/ in Kenyah languages. We suggest that they result from early borrowing of an MK form like [døk], [døk] ‘(kind of) monkey’. This borrowed form may have played a role in (analogical) reformation of Penan *mədok* and Punan Benalui *modok*. The same formative ^{*}dok appears in the terms for ‘slow loris’, as shown above. The fact that Punan Derian and Pua’ show *beruk* in the term for ‘pig-tailed macaque’ but *duk/dok*, respectively, in the word for ‘slow loris’ also supports these as distinct lexemes,—one an early borrowing, and the other, as suggested in the ACD loan under ‘monkey sp.’, an eventual loan into Malay with medial /r/ which spread via contact.

2.3. PAN-BORNEO FORMS WITH MALAY COGNATES. Comparisons in this class are complicated by spread outside of Borneo via Malay as a trade language. This has led, in some cases, to suggested PWMP reconstructions. MK loan sources explain both the extended distribution of these terms within Borneo, and their limited distribution in the Austronesian world via Malay contact. Wherever possible, evidence against the form as a Malay loan into other languages of Borneo is put forward.

2.3.1. Husk.

Pan-Bornean ^{*}səkam << MK, Khasic, Palaungic, Katuic
Cf. Malay *sekam* ‘husk’

Central Sarawak

Punan

Ukit	(ugo?)
------	--------

Land Dayak

Bidayuh- S. Land Dayak

Benyadu	saka ^m /saqa ^m
---------	--------------------------------------

Sanggau	sokam
---------	-------

Malayic

Ketapang	sokam
----------	-------

Kendayan-Salako

Kendayan	sakam
----------	-------

Ibanic

Keninjal	sokam
----------	-------

Iban	səkoʔ, sekuʔ	
Mualang	səkám	
Basap-Greater Barito		
Benuak	uya ^p m	
Tunjung	uya ^p m	
Basap	(səkəm)	
PMKA	*skaamʔ	‘chaff, husks of paddy’
<i>Khasic</i>		
Proto-Khasic	*skaam	‘husks’
Pnar	skam	‘husk’
<i>Palaungic</i>		
Proto-Palaungic	*skaam	‘chaff, bran’
<i>Katuic</i>		
Bru	sakaam	‘husk, chaff’

The sound-meaning match is good, and the regular sound changes that Land Dayak and Malayic forms (with the exception of Iban) have undergone, along with the less regular changes in Barito (initial *s- > h- occurred in Maanyan and Dusun Witu; intervocalic *-k- was continued, but *-q- > ø), argue against this as a Malay loan for all languages but Basap. (Ukit *ugoʔ* may be a loan from Malayic.)

This PMK form does not appear to be continued in Bahnaric. The MKED notes loans from MK into both Malay and Acehnese, and Tadmor (2009:693) notes this as a borrowing from MK into Malay. It is unclear if Semai (Aslian) *sekap* ‘chaff of grain’ is inherited or a Malay loan.

2.4. NORTH BORNEO FORMS WITH MALAY COGNATES. Comparisons in this class, like those in the previous section, are complicated by spread outside of Borneo via Malay as a trade language. This has led, in some cases, to suggested PWMP reconstructions. MK loan sources explain both the extended distribution of these terms within Borneo, and their limited distribution in the Austronesian world via Malay contact. Wherever possible, evidence against the form as a Malay loan into other languages of Borneo is put forward.

2.4.1. Ant₂.

Malayic Bornean *səmut << MK, Bahnaric, Katuic, Khmeric
 Cf. Malay *səmut*

Bulungan	səmut
----------	-------

North Borneo

Murutic

Timugun Murut *incamut* ‘a very small, red house ant’ (Keith 1936)

Malayic

Ketapang	somot	
<i>Kendayan-Salako</i>		
Kendayan	samut	
<i>Ibanic</i>		
Keninjal	somot	
Seberuang	səmot	
Iban	səmuət	
	semut	‘an ant’ (NCS:169) (cf. <i>semeda?</i> ‘a fire-ant’ (NCS:168); <i>semukaw</i> ‘a kind of wasp’)
PMKB	*suuc	‘to sting’
PMKC	*suəc	‘to sting’
<i>Bahnaric</i>		
Proto-Bahnaric	*smo:c	
Sapuan	samo:c	
Laven	samo:c	
<i>Katuic</i>		
Proto-Katuic	*smuuc	
Bru	səmuuc	‘ant (small)’
Kui	səmooc	‘ant (small)’
<i>Khmeric</i>		
Khmer (Surin)	smo:c	
<i>Khmuic</i>		
Proto-Khmuic	*smu:c	
Khmu	mú:c	
<i>Monic</i>		
Mon	həmot	‘ant (small)’

Cognates of Malay *səmut* ‘ant’ are found in all Malayic languages of Borneo, and it has been suggested that the Malay form is a Khmer loan in Indonesian ([Tadmor 2009](#)), a hypothesis with no explanation for the words above.

All MK forms in this set appear to derive from an *-m- infix form of PMK *suuc, *suəc ‘to sting’ suggesting original reference to a stinging ant within MK.

If valid, this set is important in illustrating a surface match between final *[c] in MK and final [t] in Borneo languages, at least after back vowels [u] or [o] (table 2, #3).

As mentioned earlier, seven different words for ant or termite are reconstructed to PMK, while only one such word is firmly reconstructible to PAN (see footnote 13). This may relate to the diversity of species in the original MK homeland, and to customs of eating different ant varieties. For an overview of edible insects in Borneo, see [Chung \(2010\)](#).

2.4.2. Barking deer₂.

North Borneo ^x*kijaŋ ‘barking deer’ << MK

Cf. Malay *kijang* ‘deer, antelope’

Land Dayak

Benyadu-Bekati

Benyadu kija:kŋ

Bekati kijak

Bidayuh-S. Land Dayak

Jangkang kijakŋ

Golik kijakʔŋ

Malayic

Proto-Malayic *kijaŋ ‘deer, antelope’

PMKA *[j]əŋ ‘deer, venison’

Aslian

Temiar kasiŋ ‘deer’

Bahnaric

Bahnar kətəŋ ‘one of three types of local deer’

Khmeric

Khmer kdan ‘deer’

Monic

Nyah Kur kəcəŋ ‘mouse deer’

Palaungic

Lamet kəja:k ‘deer’

The suggestion is that North Borneo forms are loans from MK, with the MK source word having a form like *k(V)-j{a,ə,ɔ}ŋ). The MK *k(V)- formative may be *ki[i]h, *ki ə[h] ‘kind of deer or goat’.

As noted above, the barking deer is widely distributed in Southeast Asia, including Borneo, but it is not found in the Philippines. When Austronesian speakers settled in Borneo, they, seemingly, innovated new terms for them. See BARKING DEER₁ in section 2.2.2 and BARKING DEER₃ and BARKING DEER₄ in the appendix.

2.4.3. Stranger; guest, visitor.

North Borneo ^x*təmuay ‘stranger; guest, visitor’ << MK, Bahnaric, Katuic

Cf. Malay *tamu* ‘guest’

Central Sarawak

Punan

Punan Lisum tamuy ‘stranger’

Malayic

Ibanic

Mualang təmuay ‘stranger’ (vs. *tamu* ‘visitor’; Malay loan)

Iban temuay ‘a guest, visitor’ (NCS:193)

PMKA	*t ₁ mu[ɲ]	‘guest, visitor’
PMKB	*t ₁ muəj	‘guest, visitor’
<i>Bahnaric</i>		
Proto-Bahnaric	*tmɔːj	‘visitor, stranger’
Bahnar	təmɔːj	‘guest, stranger, foreigner, outsider’
Cua	təmɔːj	‘mountain people’
<i>Katuic</i>		
Katu	tamaaj	‘visit, visitor, stranger’
Kui	tma:i	‘stranger, visitor’
Bru	tampɔːj	‘visitor, guest’
Pacoh	təm.mɔːj	‘guest, visitor’
<i>Vietic</i>		
Proto-Vietic	*t-mɔːj	‘guest’

Note the presence of the final glide in Punan Lisum, Mualang, and Iban, and the meaning ‘stranger’. We take the final glide and divergent meaning together to argue against these forms as Malay loans. In contrast, Mualang *tamu* ‘guest’ is a Malay loan, with no final glide and distinct meaning. The Malay loan *tamu* ‘guest’ is also found in Beketan, Kendayan, Keninjal, Seberuang, and Bulusu.

Given the MK data above, the PMK meaning might include ‘stranger’. Note that co-lexification of ‘guest, stranger’ is not uncommon cross-linguistically. For some examples, see CLICS.³

Malay *tamu* ‘guest’ and *temu* ‘coming together at the same spot; meeting exactly’ may have a different history from the Borneo forms cited above. The ACD reconstructs PWMP *temu ‘to meet, receive; greet a guest’ where, with the exception of Isneg (Northern Philippines), cognates are all within a greater Malay contact zone.

2.5. SUMMARY OF SUGGESTED MK LOANS IN THE LANGUAGES OF BORNEO. Table 3 illustrates the distribution of MK loans with respect to the classification of Borneo languages from Smith (2017).

In the first column, borrowed meanings are listed alphabetically. Proposed loanwords from MK are in the second column. A tick in N-S means a reflex of the loan is attested in both the Greater North Borneo languages of the north and in the languages of the Basap-Greater Barito group in the south. A tick in a subsequent column means a reflex of the loan is attested in that subgroup.

Of the 15 loans identified, 7 are pan-Bornean, represented in both the Greater North Borneo group and in the Basap-Greater Barito group. Of those 7, two (*səkam ‘husk’ and *təmuay ‘stranger, guest’) have cognates in Malay. However, as we argue in section 2.2, Borneo forms do not appear to reflect direct borrowing from Malay in these sets.

TABLE 3. DISTRIBUTION OF SUGGESTED LOANS WITH RESPECT TO CLASSIFICATION OF BORNEO LANGUAGES.

Meaning	Loanword	N-S	NB	CS	Kay	LD	Mal	Bul	GB	Bas
ANT ₁ (RED)	^x *kəsá(?)						x			
ANT ₂	^x *səmut		x				x	x		
BARKING DEER ₁	^x *paus		x	x		x				
BARKING DEER ₂	^x *kijaŋ					x	x			
COLD	^x *səŋim		x	x	x			x		
FISH ₁	^x *atuk	x	x		x				x	
FISH ₂	^x *oʝən	x		x	x				x	x
HELMETED HORNBILL	^x *tukuj	x		x	x		x		x	
RHINOCEROS HORNBILL	^x *ti{n.ŋ}aŋ	x		x	x		x		x	
HUSK	^x *səkam	x		?		x	x			x
LEECH	^x *jələw	x		x					x	x
LUNGS	^x *sop				x	x	?			
PIG-TAILED MACAQUE	^x *dok		x	x	x					
STEEP	^x *siraŋ	x					x		x	
STRANGER, GUEST	^x *təmuay			x			x			
TOTAL LOANS	15	7	5	8	7	4	8	2	6	3

Bas, Basap; Bul, Bulungan; CS, Central Sarawak; GB, Greater Barito; Kay, Kayanic; LD, Land Dayak; Mal, Malayic; N-S, Pan-Borneo; NB, North Borneo.

NB: tabulations in Table 3 from the published version have been corrected here

One interesting generalization over the data is that there are no suggested loans that are entirely limited to the Basap-Greater Barito group. Another generalization is that the number of hypothesized loans decreases internal to the Greater North Borneo languages as one goes from North to South. Both of these generalizations support the view that these hypothesized loans originated somewhere in the northern part of the island, where North Borneo and Central Sarawak languages are currently spoken. Contact through trading across the South China Sea might explain these high concentrations of borrowings in northern Borneo, since these areas face the mainland, while south and east Borneo face away from the mainland. Under a loans-via-trade scenario, those borrowings reaching the Basap-Barito languages would have been very early loans, predating the north-south split, or later loans, arriving in the south not directly, but via northern languages. It is conceivable that the pan-Borneo cognates for NECKLACE (see appendix), a trade item, could be explained in this way. However, under the loans-via-trade scenario, we are at a loss to explain why 12/15 words identified as likely MK loans in table 3 have meanings pertaining to the local Borneo environment (ANT₁, ANT₂, BARKING DEER₁, BARKING DEER₂, COLD, FISH₁, FISH₂, HELMETED HORNBILL, RHINOCEROS HORNBILL, LEECH, PIG-TAILED MACAQUE, STEEP). Indeed, the inclusion of endemic species on this list makes it possible that, on arrival, Austronesian speakers, without words for local fauna, made use of the local words used by others. But this is not the loans-via-trade scenario. This is the loans-via-in-situ-contact scenario. This alternative hypothesis is consistent with the classification and linguistic prehistory of Smith (2017): early lexical borrowings

into the Austronesian languages of Borneo before the population split that gave rise to the Basap-Greater Barito group can, perhaps, still be identified and might include the 7 pan-Borneo sets in table 2. Later loans, entering Greater North Borneo languages after the hypothesized split of Western Indonesian in Borneo, would be limited to this subgroup, and within it, clustered in the north-central region (cf. [Smith 2017:420](#), map 12).

If it is the case that speakers of MK languages were inhabiting parts of northern Borneo at the time of Austronesian settlement and dispersal, what language or languages were they speaking?²³ And does the evidence of loans offered here allow for any specific proposals regarding the chronology of contact? Table 4 shows the distribution of loans with respect to MK internal subgrouping. The ordering of loans is as in table 3, alphabetical by gloss. Shaded words are pan-Bornean, and MK (sub)groups are arranged from most-to-least frequent in terms of observed possible source of borrowing. It should be remembered that the subgroups with ticks are possible loan sources based on the phonetic matchings laid out in table 2. If, as could be the case, a phonologically conservative MK was spoken in Borneo at the time of Austronesian settlement, that language could have served as the source of the loans with ticks in the MK column, eliminating

TABLE 4. DISTRIBUTION OF POTENTIAL SOURCE OF LOANS WITH RESPECT TO MK LANGUAGE FAMILY.

Meaning	Loanword	PMK	Bahn	Kat	Asl	Pal	Khas	Mon	Khmer	Khmu
ANT ₁ (RED)	*kəsá(?)	x		x	x	x		x		
ANT ₂	*səmut	x	x	x				x	x	x
BARKING DEER ₁	*paus	x	x	x	x	x				
BARKING DEER ₂	*kijaj	x	x		x	x		x	x	
COLD	*səjəm	?	x				x	x		
FISH ₁	*atuk	x		x				x		
FISH ₂	*ojən	x	x	x		x		x		x
HELMETED HORNBILL	*tukuj	x	x	x						
RHINOCEROS HORNBILL	*ti{n,ŋ}aj	x			x	x	x	x		
HUSK	*səkam	x		x	x	x				
LEECH	*jəlaw	x	x			x				
LUNGS	*sop	?	?	?	x					
PIG-TAILED MACAQUE	*dok	x	x	x	x					
STEEP	*siraj	x	x					x	x	
STRANGER, GUEST	*təmuay	x	x	x						
TOTAL LOANS	15	13	10	9	6	6	3	8	4	2

PMK = Proto-Mon Khmer; Bahn = Bahnaric; Kat = Katuic; Khmer = Khmeric; Asl = Aslian; Pal = Palaungic; Khas = Khasic; Mon = Monic; Khmu = Khmuic.

NB: tabulations in Table 4 from the published version have been corrected here

23. A reviewer rightly observes that another hypothesis is possible: vocabulary that appears to be borrowed from non-Munda Austroasiatic could reflect a distinct substrate residue from an earlier now lost language or language group that was spoken both in mainland Southeast Asia and in Borneo prior to the arrival of Austroasiatic and Austronesian speakers to these areas. For one proposal regarding a pre-Austronesian Australoid population in Borneo, see [Sellato \(1993\)](#). However, the Austroasiatic pedigree of all borrowings discussed in the body of this paper seems quite secure.

the other subgroups that are listed. For this reason, the sum totals in the final row must be taken as potentialities, and not more, or, perhaps, as some measure of how conservative a subgroup is with respect to certain phonological features (e.g., maintenance of initial clusters, maintenance of original voicing contrasts, etc.).

One generalization that can be drawn from table 4 is that most of the proposed loans identified here (13/15) are supported phonetically by the PMK reconstruction itself. Perhaps more interestingly, the one possible loan from Aslian, LUNGS, does not have a well supported MK-internal etymology (though see discussion in section 2.2.4) and does not extend outside of a small region of northern Borneo. If these words for ‘lung’ do continue an MK borrowing, their distribution may suggest late contact in Borneo, after the north-south split.

Three distinct phonological features are consistent with the final sets of words as MK borrowings. First, a voiced palatal stop [j] is in evidence (^x*kijaŋ; ^x*oŋə; ^x*jəlaw). Second, there are bare monosyllables (^x*sop; ^x*dok). Third, as in the majority of MK languages, all words are iambic, consisting of a single heavy syllable, or a light syllable (or sequisyllable) followed by a heavy syllable.

At least one of the suggested loans can be analyzed as morphologically complex in MK: ^x*kijaŋ ‘barking deer’ << PMKA *ki[i]h ‘kind of deer or goat’ + *[j]əŋ ‘deer, venison’. Morphological complexity supports the direction of borrowing, from MK into Austronesian.

Two potential loans from Aslian are Ibanic ^x*kəsá(?) ‘(red) ant’ and Proto-Land Dayak *sop ‘lungs’. This distribution might suggest that the Land Dayak region of prehistoric times, in the north western corner of Borneo, was a possible location for a local Aslian adstrate. However, there are alternative hypotheses: recall our comparison of ^x*kəsá(?) with PMKA *ksəw? ‘red ant’, PMKA *ks[aw]? ‘red’; and Land Dayak *sop ‘lungs; ?sheathe, sock-like cover’ with PMKB *sruup ‘to sheathe’ (cf. Proto-Bahnaric *səɔp ‘sheath, body hair’; Proto-Katuic *səp, *səɔp ‘covering’ (S&J:132), Sui sɔɔp ‘sheath-like, sock-like covering’; Khmer *stu:ap* ‘fibrous sheath; bract’) (see section 2.2.4).

In section 2, our discussion of each proposed loan includes notes on semantic domain, where relevant. The two words for ‘ant’ which are thought to be borrowed may be best understood, as highlighting a dichotomy between stinging ants, and others, true to the proposed MK sources which for ANT₁ (RED) is the (stinging) red ant, and ANT₂ reflects a root *sooc ‘to sting’. This does not seem like the type of environmental knowledge that would result from trade. Rather, it suggests on-the-ground knowledge of Borneo species. Even more suggestive of *in situ* borrowing are the suggested borrowing of words for several endemic species, including the barking deer, two species of hornbill, and the pig-tailed macaque. As discussed above, a different explanation is in order for general words for fish. The observation of Smith (2017) that PMP *hikan ‘fish’ is replaced across much of Borneo by distinct lexemes may suggest a fishing taboo whereby the word for ‘fish’ was not spoken during active fishing. Again, borrowing would be most likely to occur *in situ*, as needed, not as consequence of overseas trade or Malay influence (which in these cases, has already

been ruled out). Overall, the semantic range of the suggested *in situ* loans identified is not unusual, centered around the local physical environment, with only two terms ‘husk’ and ‘stranger; guest’ suggestive of social and/or cultural contact.

3. CONCLUDING REMARKS. This study suggests the possible presence of speakers of MK languages in Borneo at the time of Austronesian settlement and early dispersal. If this was the case, evidence of these ancient MK speakers might be found in the genetic or archaeological record. Unfortunately, archeological and genetic research on the prehistory of Borneo remains in a relatively primitive state, despite remarkable findings attesting to an ancient human presence on the island. While it is now clear that Borneo has been settled for at least 40,000 years (Barker et al. 2007; Aubert et al. 2018), little is known about its inhabitants prior to the arrival of Austronesian speakers.

With regard to the archeological record, Soares et al. (2006:319) draw attention to the presence of artifacts displaying distinct features of the mainland, such as basket-marked and carved paddle impressed pottery in Sarawak (Bellwood 1997; Bulbeck 2008). Gua Sireh, in the region of Kuching in western Sarawak, is an especially important archeological site, containing pottery dated to 4,960–3,565 BP (Ipoi and Bellwood 1991) and in a style that is highly atypical for sites associated with the Austronesian expansion (Bellwood 2006). Both the style and the early date of the settlement are suggestive of a non-Austronesian origin.²⁴

With regard to material culture in the ethnographic present, it has been widely remarked that Bornean-style long houses are both unique to this area within the Austronesian world but strongly resemble MK domiciles on the mainland. Other aspects of physical culture which appear to be uniquely shared between the mainland and north-western Borneo have been noted by Blench (2010:137–39).

Genetic studies of the region are still in their infancy, and there is little agreement on the interpretation of results beyond several broad conclusions. One emerging consensus is that western Indonesia, as a whole, is highly distinct from both east Indonesia and the Philippines in showing high levels of admixture with mainland populations of probable Austroasiatic origin, in addition to other admixtures. Karafet et al. (2010) find a strong divide between western and eastern Indonesia somewhere between Sumba and Flores that suggests a mainland component in the west and a Papuan component in the east. In a more recent genetic study, Lipson et al. (2014) conclude that:

24. Benjamin (1987) notes the likelihood of a pre-Malayan Austronesian presence on the Malay Peninsula based on Austronesian loans into Aslian languages, which could not have originated in Malay. If, as Blust (2006) suggests, large-scale language leveling took place on the Malay Peninsula, then contact between speakers of Austronesian and Austroasiatic languages may have been longer and more intense than can be gleaned from current data. Since words with Austronesian etymologies are not considered in this paper, Blust’s suggestion has little bearing on our conclusions.

our signal indeed reflects gene flow from the mainland into ISEA from an ancestral population that is nested within the radiation of AA[Austroasiatic]-speaking populations, and hence it is likely that this source population itself spoke an AA language. [...] While a major AA contribution to western speakers of AN languages has not been proposed in the genetic literature, results from previous genetic studies are in fact consistent with these findings.

To our knowledge, no genetic study has yet examined Bornean populations as distinct from other western populations of the Indo-Malaysian archipelago, and there is, as of yet, no genome-wide sampling of Borneo (Soares et al. 2006). Thus, while there appears to be a strong case for Austroasiatic admixture in the region of Borneo, differentiating it from eastern Indonesia and the Philippines, there is no evidence yet that the signal is distinct from other islands of the Sunda Shelf.

Though speakers of MK languages are not known for long-distance seafaring as the Austronesians are, a history of seafaring seems probable, given the prehistoric presence of speakers of Austroasiatic languages on the Nicobar Islands, some 150 kilometers north of Sumatra. In addition, a recent proposal by Rau and Sidwell (2019) argues that Proto-Munda, an independent subgroup of Austroasiatic, had a homeland on the Mahanadi Delta 4,000–3,500 years ago, and that speakers of this language reached India via a maritime route around or across the Bay of Bengal from Southeast Asia. So, while more recent history shows MK people attached to their Southeast Asian homeland, there is growing evidence of seafaring in the distant past. With this noted, let us return to the question we started with.

Why are the Austronesian languages of Borneo different? The linguistic evidence we present in section 2 suggests a possible presence of MK speakers in Borneo close to the time of Austronesian settlement, and is consistent with the little that is known about the genetic, archeological and hypothesized culture and population history of the area. In all, 15 potential Mon-Khmer loans have been identified. In contrast to earlier studies, we have not found significant evidence of contact with Aslian languages or with Khmer. For each one of the four loans that could be from Khmeric there is a potential more widespread MK source as well. Most suggested loans resemble PMK reconstructions, suggesting that they could predate the splitting off of recognized subgroups within Austroasiatic. While the current study represents the most thorough and scientifically rigorous search for MK borrowings in the Austronesian languages of Borneo, it has only just scratched the surface, as it has been largely limited to etyma identified by Smith (2017) as local innovations.

Even if only a fraction of the suggested loans can find further support, the findings raise many new questions. Where and when did first contact between Austronesian speakers and speakers of MK languages take place in Borneo? Were multiple MK languages involved? Is early contact responsible for

divergent sound patterns noted for the languages of Borneo? Parallel questions have been asked in the context of Chamic and Acehnese, where MK influence is uncontroversial, but clear answers, especially with regard to MK substrates and donor languages, have been elusive. Sidwell (2005, 2006) notes that Chamic and Acehnese appear to preserve a branch of MK that is “otherwise unattested and now extinct—presumably the result of language shift” and that there is no evidence from borrowing from a single source. As a point of departure, we have found that the MK influence in Borneo cannot simply be attributed to contact with Khmer, Chamic, or Aslian languages, respectively. Furthermore, there is little overlap between the loans we have found, and Mon-Khmer loans into Chamic as documented in Thurgood (1999). Finally, it should be noted that the MK-like phonological patterns mentioned at the outset of the paper are not easy to pinpoint geographically or phylogenetically within Borneo. The Central Bornean Linguistic Area of Smith (2017) contains several phonological changes highly typical of mainland languages but not seen elsewhere among Austronesian languages. Despite the glaring areal pattern which groups Borneo together with MK languages of the mainland, Smith (2021:155–56) states, “this convergence is not occurring in a manner that can easily be explained as a consequence of diffusion, contact, substrata, or areal influence. Rather, the Central Bornean Linguistic Area consists of a small number of languages found here and there, often isolated from one another by hundreds of miles of mountainous terrain. [...] The problem of identifying a satisfactory explanation for the emergence of the Central Bornean Linguistic Area . . . will likely persist for some time.”

By extending future studies to the many excellent descriptions of Borneo languages that we have not yet had the opportunity to consult and those that will hopefully appear in the future, answers to these important questions may emerge, deepening our understanding of linguistic divergence in the Austronesian languages of Borneo.

APPENDIX

This appendix includes words from the Austronesian languages of Borneo that may appear to be MK loans but fail one or more of the criteria outlined in section 1.

A1. BANANA.

North Borneo *balak/*balat (<*barat) (Smith 2017:304)

Smith (2017:304) suggests a Borneo near-cognate *balak/*balat (<*barat) restricted to Central Sarawak languages, Land Dayak and Kenyah. However, a form *baRat/*baRak ‘banana’ is reconstructed for the Northern Philippines (Denham and Donohue 2009) with a possible relationship suggested for MK there. Compare Proto-Monic *braat (Old Mon *brat*, *bra:t*).

A genetic study of banana domestication suggests that *baRat ‘banana’ traveled from the Northern Philippines to Borneo (Perrier et al. 2011), making the status of potential MK cognates problematic.

An additional problem with the comparison is that PMK *priət ‘banana’ has initial *p- and only Monic shows a voiced reflex. Furthermore, only Phong *pra:k* (Khmuic) shows the final irregular [k] instead of [t]. This set of words merits further investigation.

A2. BARKING DEER₃.

Pan-Borneo *təlaʔus (Smith 2017:396)

The common barking deer (*Muntiacus muntjak*) is widely distributed in Southeast Asia, and the Bornean barking deer (*Muntiacus atherodes*) is endemic to Borneo. For two proposed loans with this meaning, see sections 2.2.2 and 2.4.2.

Evidence from Borneo suggests that the most widespread term for this animal goes back to *təlaʔus. Smith (2017:396) notes that “similar words for barking deer also appear in Pearic and Khmeric languages,” including Surin Khmer *təhlu:h* ‘a barking deer’. In his opinion, however, “it is not likely that similarities of this type are more than simple chance, as there is otherwise no apparent connection between the languages of Borneo and these specific MK languages” (op cit.). Given other suggested loans with similar meaning, as well as an apparent MK cognate in Acehnese *juəh*, *gluəh* ‘deer (small kind)’ (Cowan 1948), we do not believe this similarity is due to chance.

A3. BARKING DEER₄.

Greater North Borneo ^x*muncak ‘noisy forest mammal’

Land Dayak	
<i>Benyadu-Bekati</i>	
Benyadu	munseʔ ‘to cry’
Bekati	munçéʔ ‘to cry’
<i>Bidayuh-S. Land Dayak</i>	
Golik	muncak ‘slow loris’
Malayic	
Malay	muncak ‘barking deer’

Unlike the majority of loans examined in this paper, Malay *muncak* ‘barking deer’ has no obvious cognates in the indigenous languages of Borneo: Golik *muncak* ‘slow loris’ refers to another forest mammal (with a range of distinct vocalizations; Daschbach et al. 1981), and the Benyadu and Bekati forms above, if cognate, have a verbal meaning ‘cry’, possibly related to the call of the deer. We suggest that Malay *muncak* also originates

as an MK borrowing. A compound ****muup+*tja[a]k** < PMK ***muup** ‘mouth’ (cf. ‘language’ in Steng; ‘mouth, word, language’ in Riang) plus PMK ***tja[a]k** ‘sambhur deer, large forest mammal’. Since, at present, there is no evidence of an MK compound of this type, the comparison remains speculative.

A4. BATHE.

Greater North Borneo ***mu** ‘bathe’

Central Sarawak

Melanau

Proto-Melanau ***tə-mu?** (Sm:395)

Punan

Proto-Punan ***nəmu** (Sm:395)

Müller-Schwaner

Proto-Müller-Sch. ***nomu** (Sm:395)

Kayanic

Proto-Kayanic ***mu** (Sm:395)

Land Dayak

Proto-Land Dayak ***mamu** (Sm:395)

PMKA ***ɬuh** ‘to pour, sprinkle’

Aslian

Proto-Aslian ***mamuh** < ***mabuh**

Chewong **mamuh**

Semai **mamuh** ‘bathing, a bathe’

Semnam **mamu:h**

Temiar **muh** ‘bathe, to wash by immersion’

Smith (2017:395) discusses the North Borneo and Aslian forms in relation to continuations of PMK ***huum** ‘to bathe’, as opposed to the Aslian etymology in MKED, which is shown above. Given the sound-meaning matches for both disyllabic Proto-Land Dayak ***mamu** and monosyllabic Proto-Kayanic ***mu**, the comparison is strong, and, as originally suggested by Adelaar (1995), points to an MK source for borrowing in northern Borneo. The problem with this set is that PMK and PAN may have cognate roots: compare PMK ***ɬuh** ‘to pour, sprinkle’ with PMP ***buqbuq** ‘pour, pour out, as water or grain from a container’, presumably from an earlier root ***buq**. A reflex of PMP ***buqbuq** in Borneo is Iban *bubuh*, *mubuh* ‘pour out, pour away’. Since it is not difficult to imagine a series of changes from *mubuh* > *mbuh* > *muh* > *mu(?)*, the North Borneo forms above may, indeed have An etymologies after all.

Kruspe (2004:7–10) discusses an avoidance speech style in Semelai which involves a system of word substitution to avoid the utterance of tabooed words. Lexical replacement items can be drawn from many sources, including loans, or

words exclusive to the register that sometimes have cognates in other Aslian languages. Semelai *hūm* ‘bathe’ (from PMK *huum ‘to bathe’) is replaced by *mhmeh* (cf. Semoq Beri *hmeh* ‘to bathe’). It is possible that taboo avoidance played a role in this purported instance of borrowing.

A5. BETEL NUT.

Bornean Malayic *pinaŋ

Vietic

Thavung panâ:ŋ

Katuic

Katu pana:ŋ

Bahnaric

Cua pani:ŋ

Within MK, the forms above have no etymologies. Given the widespread trade of betel nut, the MK words are best analyzed as loans from Malay (or from Chinese; cf. Minnan *pin neng*, ACD). Further support for Thavung *panâ:ŋ* as a loan is Proto-Vietic *b-lu: ‘betel’. Compare also the Cua synonym *kli:ʔ* ‘betel nut’.

A6. BRING.

Kenyah Vo *jen* ‘bring’

In earlier work, Kaufman (2018) compares this form to Bahnar *je:n* ‘to bring, carry’ < PMKA *jun ‘to hand over, bring’. Although this form seems divergent from other Kenyah forms which include E. Penan *mihin*, Sawa *ŋgin*, and Badeng *ŋkin*, Rensch et al. (2006:198) suggest Proto-Bidayuh *di-gen, *di-ken ‘hold in hand’, where the root could well be an irregular reflex of PWMP *agem ‘hold, grip’ (cf. Bukat *ŋ-agem* ‘to hold’, Madurese, Balinese *agem* ‘hold’).

Palatalization *gen > *jen* would account for the Vo form.

A7. CHICKEN.

Proto-Central Sarawak *siaw

Smith (2017:289) suggests *manuk >> *siaw ‘chicken’ as a Proto-Central Sarawak innovation. In earlier work, Kaufman (2018) compares *siaw to Katu *siem* and Proto-Khasic *sɨiar, with irregular phonological matching for the final consonant. Smith (2017:324) states that “PMP *siwsiw ‘to cheep; a baby chick’ and *siap ‘baby chick’ appear unrelated” to *siaw, however, we suggest that *siaw is a continuation of PAN *siap ‘baby chick’. First, consider the directly continued forms: Abai Sembuak and Tingalan (West) (Southwest Sabah/Murut) *siap* ‘chick’; consider also the shift of meaning in Central Sarawak Bekatan *siap* ‘chicken’ (Land Dayak), and Smith’s own Proto-Kayanic

*siap ‘chicken’; finally, in North Sarawak, consider Proto-Kenyan *yap ‘chicken’ < *syap < *siap. Given these facts, it seems more likely that Proto-Central Sarawak *siaw replaced earlier *siap ‘chicken’.

A8. COUNT.

PMP *ihap ‘count, calculate’

In earlier work, Kaufman (2018) compares Hliboi Bidayuh (Lun Dayeh) *n̄nap* (Smith 2017:537), Proto-Punan *-iap (Smith 2017:495), and Proto-Müller-Sch. *-iʔap to MK words for ‘count’, including: Jahai *j̄ep* (Aslian); Tampuan *jaap* (Bahnaric); Bru [TS] *n̄ap* (Katuic); Phong *nap* (Khmuic); Nyah Kur [Central] *n̄ap* (Monic); and Sre [Koho] *nap* (Bahnaric). However, on further inspection, the Borneo forms look like continuations of PMP *ihap ‘count, calculate’, evidenced in the Philippines, Borneo, and continued as Proto-Chamic *iaap, Proto-Minahasan *iap. In this case, the direction of borrowing seems to be reversed. In contrast to the Bornean terms, there is no etymology for any of the MK terms above. It appears, in this case, that Austronesian terms for ‘count’ were borrowed into MK languages in the course of early trade, and during the intensive period of Chamic-Mon-Khmer contact (Thurgood 1999).

Other Bornean terms that could be etymologically related are W. Penan *jajap*, Vo *ʃap*, and Laang *jap* < PWLKen *jap ‘ten’ (Sm:256), as well as Kayanic forms continuing *japitan ‘nine’ (Sm:309) (cf. Proto-Kayanic *unjab ‘count’). These words for ten appear to relate to the end of the count, though they superficially resemble some MK words for ‘ten’, like Khmer *dap* (though cf. PMK *jət ‘ten’.)

A9. CRAB.

Pan-Bornean *kətam << MK, Bahnaric, Palaungic, Katuic
Cf. Malay *ketam*

Central Sarawak

Melanau

Kanowit	təke: (MET?)
---------	--------------

Land Dayak

Bidayuh-S. Land Dayak

Sanggau	kotam
---------	-------

Malayic

Ketapang	kotam
----------	-------

Ibanic

Keninjal	kotam
----------	-------

Seberuang	kətam
-----------	-------

Basap-Greater Barito

Barito	
Kapuas	katam
Ngaju	katam
PMKA	*ktaam
<i>Bahnaric</i>	
Proto-Bahnaric	*kta:m
Halang	kətaam
<i>Palaungan</i>	
Proto-Palaungan	*ktaam
<i>Katiuc</i>	
Proto-Katuic	*ktaam

The MKED includes the notes of Shorto (2006) on the borrowing of this form into Iban (*ketam*), Malay (*ketam*), and Acehnese (*götöəm*). The Malay form is clearly borrowed from an MK language, but this match is not included in section 2 because there is no solid evidence that the forms above are not due to subsequent borrowing from Malay. A synonymous word in the same language groups, which may be older is *karama? ‘crab’: cf. Benyadu *karama?*; Kendayan *karama?*; Iban *gərama?*; Mualang *gəyama?*; Lun Dayeh *kəra?*.

A10. GIVE.

Proto-Kenyah *na?

In earlier work, Kaufman (2018) suggests that this Proto-Kenyah form is an MK loan, citing PMKD *ʔan, pSBah *ʔa:n, and Khmu ʔan (Khmuic), and assuming otherwise unmotivated metathesis. This appears to be a chance resemblance. In contrast, compare Proto-Kenyah *na? with Tujung *nare?* and Benuaq *ñe:?*, related terms for ‘give’ in two Barito languages, where the monosyllabic form results from intervocalic *r-loss.

A11. GREEN₂.

Proto-Kenyah *biləŋ (Smith 2017:663)

In earlier work, Kaufman (2018) compares Proto-Kenyah *biləŋ to MK words for ‘green’ with no MK-internal etymologies including: Jahai *blʔeŋ* (Aslian), Kensiu *bəlʔuŋ* (Aslian), and Mlabri *bnliŋ* (Khmuic). We suggest Proto-Kenyah *biləŋ ‘green, blue-green’ as a continuation of PMP *buluŋ ‘dark blue, bluish black’, based on Kenyah (Long Atun) *biləŋ* ‘blue-green’ and *laŋaw biləŋ* ‘bluebottle (fly)’ (cf. Ilongot *bilin* ‘black’, Javanese *wuluŋ* ‘blue-black’, Rembong *buluŋ* ‘green, marine blue’, POC *buluŋ ‘dark green, dark blue’).

A12. LONGHOUSE.

Pan-Bornean ^x*bətəŋ ‘longhouse’ (Sm:291–92) << MK**Land Dayak**

Ribun bətəkŋ

Tamanic

Taman bətəŋ

Central Sarawak*Müller-Schwane*

Hovongan (lovu) petəŋ

Basap-Greater Barito

Kereho bətəŋ

Maanyan bətəŋ

Kadorih behtəŋ

PMKA *bdəŋ ‘walling material’

*Khmeric*Khmer p^htəəŋ ‘panel, leaf, partition’*Monic*Mon hətəəŋ ‘(wall or partition of) twilled texture
made from bamboo strips’

Compare also:

Bahnaric

Halang bruaŋ ‘longhouse’

Sedang vɛəŋ ‘room’

Katuic

Katu buoŋ ‘room’

Palaungic

Palaung bluŋ ‘bamboo matting dividing rooms’

Vietic

Chút puòŋ ‘room’

There are several weaknesses in this comparison. First, the pan-Bornean term ^x*bətəŋ ‘longhouse’ has a potential Austronesian etymology in PMP ^{*}bataŋ ‘tree trunk, fallen tree; log; stem of a plant; body; corpse; self; bridge of the nose; main part of something; main course of a river’. The Austronesian etymology is supported by the fact that some Borneo languages, like Bidayuh (Bukar–Sadong), colexify these meanings: *bataŋ* ‘a stretch of longhouse; tree trunk; a block or row of houses’ (ACD). Another weakness in the comparison is the voicing mismatch in the second consonant: PMK ^{*}d should correspond with Bornean ^xd. A final weakness in the comparison is semantic: though it

is the case that longhouses are distinguished from other houses by the many partitions or rooms, separated by walling material, the semantic distance between ‘longhouse’ and ‘walling material’ is significant.

While this may be a borrowed culture item, there is good reason to believe that Pan-Bornean *bətaŋ ‘longhouse’ is from PMP *bataŋ.

A13. NECKLACE.

Pan-Bornean ^x*koŋ ‘ring-shaped artifact’ << MK, Bahnaric, Khmeric reduplicated as Pan-Bornean ^x*koŋ-koŋ ‘necklace’

Malayic		
<i>Ibanic</i>		
Mualang	kuŋkoŋ	
Iban	kuŋkoŋ	‘a collar, necklace’ (NCS:95)
Basap-Greater Barito		
Taboyan	koŋkoŋ	
Basap	koŋkoŋ	
PMK	*kəŋ, *kəŋəŋ	‘ring-shaped artifact’
<i>Khmeric</i>		
Old Khmer	kəŋəŋ	‘ring, bracelet, necklace’
Khmer	kəŋ	‘ring-shaped artifact’
Khmer	kəŋəŋ	‘ring, circle; bracelet, necklace, anklet’
<i>Bahnaric</i>		
Sre	koŋ, koŋəŋ	‘ring’
Bahnar	kəŋəŋ	‘bracelet, metal collar’
<i>Katuic</i>		
Proto-Katuic	*kəŋ	‘bracelet’

We suggest that ^x*koŋ ‘ring-shaped artifact’ was borrowed from MK, and reduplicated to derive pan-Bornean ^x*koŋ-koŋ ‘necklace’. A similar case of borrowing is found in Chamic: Phan Rang Cham kəŋ < Proto-Chamic ^x*kəŋəŋ ‘bracelet’ which Thurgood (1999:145) suggests is borrowed from Proto-North Bahnaric.²⁵

While the Bornean root seems to be a loan, the status of necklaces as valuable trade items might weaken this item as evidence for ancient *in situ* borrowing in Borneo. Indeed, the Khmer term seems to have been widely borrowed on the mainland. However, the fact that the root is reduplicated in Borneo, but not in potential MK source languages, or in Chamic, suggests an old internal development that occurred before the diversification of languages in Borneo.²⁶

25. A reviewer suggests that the source of the Chamic loan is Khmer, not Bahnaric.

26. Bahnar [Pleiku] shows kəŋəŋ kjəŋəŋ ‘metal necklace’, but this does not appear to be a reduplicated form.

The distribution of this term, limited to Malayic and Basap-Barito is also suggestive of coastal trading.

Another weakness of the comparison is the existence of a widely attested PAN root **-kuŋ* ‘bend, curve’, as Iban *beŋkuŋ* ‘ring-frame, usually of rattan or creeper, e.g., in a fish trap, etc.’ < PMP **beŋkuŋ* ‘curved, crooked, bent’.

A14. RAIN.

Northern Borneo **gəma?*

North Borneo

Kenyah

Vo	ima?	‘rain’
----	------	--------

Land Dayak

Bidayuh-S. Land Dayak

Proto-Bidayuh	*umo:?	‘water’ (RRNR:314)
---------------	--------	--------------------

Proto-Bidayuh	*gumo:?	‘bathe’ (RRNR:314)
---------------	---------	--------------------

PMKA

**gma?*

Pearic

Proto-Pearic	*gma?	
--------------	-------	--

Aslian

Mah Meri	gəma:h	
----------	--------	--

Semaq Beri	gəma:h	
------------	--------	--

Khmuic

Proto-Khmuic	*kma?	
--------------	-------	--

Vietic

Proto-Vietic	*kma:	
--------------	-------	--

In earlier work, Kaufman (2018) compares the isolated Vo form with MK etyma. The Proto-Bidayuh doublet with initial *g-* vs. \emptyset strengthens the comparison (where $\emptyset < *k$), though the semantic match is weaker for Bidayuh.

A15. STIFF, RIGID.

PAN	* <i>-kaŋ</i> ₃	‘stiff, rigid; cramps’
-----	----------------------------	------------------------

PMP	* <i>kaŋkaŋ</i> ₃	‘cramps, stiffening of the limbs’
-----	------------------------------	-----------------------------------

Malay	cekəŋ	‘taut, astretch’
-------	-------	------------------

Malay	jeŋkaŋ	‘start or stiff in death’
-------	--------	---------------------------

Malagasy	rôhana	‘syphilitic rheumatism’
----------	--------	-------------------------

Dairi-Pakpak Batak	terkaŋ	‘stiff, inflexible’
--------------------	--------	---------------------

Toba Batak	baŋkaŋ dagaŋ	‘stiff, rigid’
------------	--------------	----------------

PMP	*keŋ ₂	‘stiff, as a corpse’
WMP		
Karo Batak	keŋ	‘stiff, stiffened, as a corpse’
CMP		
Manggarai	keŋ	‘taught, of a rope’
PMKA	*[ŋgəŋ]	‘stiff, bristling’
<i>Bahnaric</i>		
Sedang	khəŋ	‘durable, hard, stiff, strong’
Bahnar	təgəŋ	‘sticking up, erected’
Sre	gəŋ	‘stiff, straight, rigid’
<i>Katuic</i>		
Proto-Katuic	*kəŋ	‘hard’
<i>Khmeric</i>		
Khmer	kəŋ	‘to be hard, stiff; to be harsh, severe, strict’

An Austronesian root doublet *-kaŋ₃ ‘stiff, rigid; cramps’ and *keŋ₂ ‘stiff, as a corpse’ is reconstructed in ACD on the basis of the comparanda above. However, it is suggested in the MKD that Chamic *kəŋ* ‘motionless, paralyzed’ is a loan from Mon-Khmer. Therefore, we do not view this set as evidence for a borrowing of an MK root into Malayic (or Chamo-Malayic) that took place in Borneo.

A16. STOMACH, INTESTINES.

Proto-Malayic *pərut ‘stomach’ (Adelaar 1992:129) << MK, Bahnaric
Cf. Malay ‘stomach, belly, abdomen’

Malayic

Proto-Malayic	*pərut
Ketapang	parot

Kendayan-Salako

Kendayan	parut
----------	-------

Ibanic

Keninjal	pəyot
Seberuang	pəyut
Iban	pəruət
Iban	guloəŋ pəroət ‘intestines’
Mualang	pəyot

PMKA	*ruuc	‘intestine’
PMKB	*ruəc	‘intestine’

Bahnaric

Proto S. Bahnaric	*prɔ:c	‘intestines’
Sre	prɔc	‘intestines; internal organs’
Stieng	prɔ:c	‘intestine’
Mnong	pruec	‘intestine’

Adelaar (1992) notes this Malay form as a probable MK loan. There is no mention of it as a potential loan in Smith (2017), though all of the Malayic forms above are from Smith (2017:621).

The sound matches with Proto-S. Bahnaric are exact (including #3, matching final MK *-c with Malayic *-t.)

Outside of Malayic, there are a few forms for ‘intestines’ that may be cognate: cf. Dalat *paʔit*; Punan Aput *pait*; Benyadu *paraya*.

The problem with this comparison is that Chamic has also appeared to borrow the same term. Thurgood (1999:360) proposes Chamic *pruac ‘stomach; intestine, large’, suggesting Proto-Katuic *ruajʔ as the source, noting Acehnese *pruat*, Wr. Cham *pruəc* among other forms.

A17. TWIN.

Pan-Bornean *kambar ‘twin(s)’ << MK, Bahnaric, Katuic, Khmeric
Cf. Malay *kambar* ‘forming a match or pair, not of two who supplement one another, e.g., bride and bridegroom; but of mere equality or similarity, e.g., of twin children, or a worthy foe’; *anak kamar* ‘twins’

Malayic*Ibanic*

Iban	gambar	‘a person, thing at the side of another’ (NCS:57)
------	--------	--

Central Sarawak*Müller-Schwaner*

Hovongan	kambar
----------	--------

Land Dayak*Benyadu-Bekati*

Benyadu	kemerə
---------	--------

Bidayuh-S. Land Dayak

Sungkung	kambar
----------	--------

Golik	kombar
-------	--------

Sanggau	komayə
---------	--------

Basap-Greater Barito

Kadorih	kambar
---------	--------

Ngaju	kambar
-------	--------

Kapuas	kambar		
Bakumpai	kəmbar		
Maanyan	kam ^b ar		
		‘child’	‘two’ ‘twins’
PMKB	*kuən	*ɓaar	
<i>Bahnaric</i>			
Proto-Bahnaric	*kɔɔn	*ɓaar	
Bahnar	kɔɔn	baar	kəmaar
<i>Katuic</i>			
Proto-Katuic	*kɔɔn	*ɓaar	
Ngeq	(ceʔ)	baar	kampɬa:r
<i>Khmeric</i>			
Proto-Khmeric	*kɔɔn	*ɓaar	

The ACD lists Malay *kəmbar* along with apparent cognates in Casiguran Dumagat, Kapampangan, Tagalog, Bikol, Malagasy, Toba Batak, Old Javanese, Javanese, Balinese, and Makassarese, and states “A loan from Malay, which itself acquired the word from a Mon-Khmer source.” What the ACD does not mention is that similar terms are widespread in Borneo, not only in North Borneo, but in the Barito languages as well.

The case for *kambar* as an MK loan is strengthened by its bimorphemic status in MK < *kuun ‘child’ + *ɓaar ‘two’, as we illustrate here. From Malay, this term has spread into other Austronesian languages in Indonesia and the Philippines.

It is difficult to rule out the possibility that this word came into Malay from an MK language outside of Borneo. For Land Dayak, there is no source of word-final *r other than borrowings from Malay (Smith 2017). However, an interesting property of the Benyadu and Sanggau forms is a final epenthetic vowel, suggesting Proto-Land Dayak *kambara, which would account for the regular intervocalic reflexes of *r. Even so, the wide distribution of these forms and their lack of regular sound correspondence subgroup-internally suggest their status as Malay loans, where the Malay term was likely acquired outside of Borneo. The Iban cognate, with distinct voicing of the initial consonant, is the only possible sign of an *in situ* loan; if this were a Malay loan, an initial voiceless consonant would be expected.

REFERENCES

- Adelaar, Alexander K. 1992. *Proto Malayic: The reconstruction of its phonology and parts of its lexicon and morphology*. 119. Dept. of Linguistics, Research School of Pacific Studies, The Australian National University.

- . 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 Darell T. Tryon, 75–95. Canberra: Dept. of Anthropology, RSPAS, ANU.
- . 2000. Malay: A short history. *Oriente Moderno. Nuova serie. Anno 19 (80), Nr.2. Alam Melayu il Mondo Malese: Lingua, Storia, Cultura*. 225–42.
- Aubert, M., P. Setiawan, A. A. Oktaviana, A. Brumm, P. H. Sulistyarto, E. W. Saptomo, B. Istiawan, T. A. Ma'rifat, V. N. Wahyuono, F. T. Atmoko, J.-X. Zhao, J. Huntley, P. S. C. Taçon, D. L. Howard, and H. E. A. Brand. 2018. Palaeolithic cave art in Borneo. *Nature* 564:254–57.
- Barker, G., H. Barton, M. Bird, P. Daly, I. Datan, A. Dykes, L. Farr, D. Gilbertson, B. Harrison, C. Hunt, T. Higham, L. Kealhofer, J. Krigbaum, H. Lewis, S. McLaren, V. Paz, A. Pike, P. Piper, B. Pyatt, R. Rabett, T. Reynolds, J. Rose, G. Rushworth, M. Stephens, C. Stringer, J. Thompson, and C. Turney. 2007. The 'human revolution' in lowland tropical Southeast Asia: The antiquity and behavior of anatomically modern humans at Niah Cave (Sarawak, Borneo). *Journal of Human Evolution* 52(3): 243–61.
- Bellwood, Peter S. 1997. *Prehistory of the Indo-Malaysian Archipelago*. Honolulu: University of Hawai'i Press.
- . 2006. Borneo as the homeland of Malay? The perspective from archeology. In *Borneo and the Homeland of the Malays: Four essays*, ed. by James T. Collins and Awang Sariyan. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- . 2013. *Prehistory of the Indo-Malaysian Archipelago*: revised edition. Canberra: ANU Press.
- Benedict, Paul K. 1975. *Austro-Thai: language and culture*. New Haven: HRAF Press.
- Benjamin, Geoffrey. 1987. Ethnohistorical perspectives on Kelantan's prehistory. In *Kelantan Zaman Awal*, ed. by Nik Hassan Shuhaimi Nik Abdul Rahman, 108–53. Kota Bharu: Muzium Negeri Kelantan.
- Blench, Roger. 2010. Was there an Austroasiatic presence in island Southeast Asia prior to the Austronesian expansion? *Bulletin of the Indo-Pacific Prehistory Association* 30:133–44.
- Blust, Robert. 1988. *Austronesian root theory: An essay on the limits of morphology*. Amsterdam: John Benjamins.
- . 1997. Nasals and nasalization in Borneo. *Oceanic Linguistics* 36:149–79.
- . 2001a. Language, dialect and riotous sound change: The case of Sa'ban. In *Papers from the Ninth Annual Meeting of the Southeast Asian Linguistics Society*, ed. by Graham W. Thurgood, 249–359. Tempe: Program for Southeast Asian Studies, Arizona State University.
- . 2001b. Historical morphology and the spirit world: *qali/kali-prefixes in Austronesian languages. In *Issues in Austronesian Morphology: A festschrift for Byron W. Bender*, ed. by John Bradshaw and Kenneth L. Rehg, 15–73. Canberra: Pacific Linguistics.
- . 2002. Kiput historical phonology. *Oceanic Linguistics* 41:384–438.
- . 2006. The origin of the Kelabit voiced aspirates: a historical hypothesis revisited. *Oceanic Linguistics* 45:311–38.
- . 2007. Disyllabic attractors and anti-antigeminization in Austronesian sound change. *Phonology* 24(01): 1–36.
- . 2010. The greater north Borneo hypothesis. *Oceanic Linguistics* 49:44–118.
- . 2013. *The Austronesian languages*, 2nd edition. Canberra: Pacific Linguistics.
- . 2015. Longhouses and Nomadism: is there a connection? *Borneo Research Bulletin* 46:194–221.

- Blust, Robert, and Stephen Trussel. 2020. The Austronesian Comparative Dictionary, web edition. <http://www.trussel2.com/ACD>.
- Blust, Robert, and Alexander D. Smith. 2014. *A bibliography of the languages of Borneo (and Madagascar)* Reference Series No. 2. Phillips, Maine: The Borneo Research Council.
- Bulbeck, David. 2008. An integrated perspective on the Austronesian diaspora: the switch from cereal agriculture to maritime foraging in the colonization of Southeast Asia. *Australian Archaeology* 67:31–52.
- Borneo Dictionary (online). <https://borneodictionary.com/>.
- Chau Ju-Kua c.1250. 1911/1970. His Work on the Chinese and Arab Trade in the Twelfth and Thirteenth Centuries, entitled Chu-fan-chi. St. Petersburg/Taipei.
- Chung, Arthur Y. C. 2010. An Overview of Edible Insects and Entomophagy in Borneo. Paper presented at the UN-FAO Workshop on Forest Insects as Food: Humans Bite Back. February 19–21, 2008, Chiang Mai, Thailand.
- Cowan, H. K. J. 1948. Aanteekeningen betreffende de verhouding van het Atjesch tot de Mon-Khmer talen. *Bijdragen tot de Taal-, Land-en Volkenkunde van Nederlandsch-Indie* 104:429–514.
- Daschbach, N. J., M. W. Schein, and D. E. Haines. 1981. Vocalizations of the Slow Loris, *Nycticebus coucang* (Primates, Lorisidae). *International Journal of Primatology* 2:71–80.
- De Jong, Kenneth, and Mi-Hui Cho. 2012. Loanword phonology and perceptual mapping: comparing two corpora of Korean contact with English. *Language* 88:341–68.
- Denham, Tim, and Mark Donohue. 2009. Pre-Austronesian dispersal of banana cultivars west from New Guinea: linguistic relics from eastern Indonesia. *Archaeology in Oceania* 44:18–28.
- Hudson, Alfred B. 1978. Linguistic relations among Bornean peoples with special reference to Sarawak: An interim report. *Studies in Third World Societies* 3:1–44.
- Ipoi, Datan, and Peter Bellwood, P. 1991. Recent research at Gua Sireh and Lubang Angin. *Bulletin of the Indo-Pacific Prehistory Association* 10:386–405.
- Jenner, Philip N. 2009. *A Dictionary of Angkorian Khmer*. Canberra: Pacific Linguistics, Research School of Pacific Asian Studies, ANU.
- Kaboy, Tuton. 1965. *Punan vocabularies*. *Sarawak Museum Journal* 12:188–200.
- Kang, Yoonjung. 2010. The emergence of phonological adaptation from phonetic adaptation: English loanwords in Korean. *Phonology* 27:225–53.
- Karafet, Tatiana M., Brian Hallmark, Murray P. Cox, Herawati Sudoyo, Sean Downey, J. Stephen Lansing, and Machael F. Hammer. 2010. Major East-West division underlies Y-chromosome stratification across Indonesia. *Molecular Biology and Evolution* 27:1833–44.
- Kaufman, Daniel. 2009. Austronesian nominalism and its consequences: A Tagalog case study. *Theoretical Linguistics* 35:1–49.
- . 2018. Lexical evidence for an Austroasiatic presence in Borneo. Talk presented at Cornell University. <https://bahasawan.com/wp-content/uploads/2020/01/Cornell-Borneo-handout.pdf>.
- Keith, Henry George. 1936. Some Ulun-no-Bokan (Murut) words from North Borneo. *Journal of the Malayan Branch of the Royal Asiatic Society* 14:314–22.
- King, Victor T. 1993. *The Peoples of Borneo*. Cambridge, MA: Blackwell Publishers.
- Kruspe, Nicole. 2004. *A grammar of Semelai*. Cambridge: Cambridge University Press.
- . 2010. *A dictionary of Mah Meri as spoken at Bukit Bangkok*. Oceanic Linguistics Special Publication in Linguistics. Honolulu: University of Hawai'i Press.
- Lipson, Mark, Po-Ru Loh, Nick Patterson, Priya Moorjani, Ying-Chin Ko, Mark Stoneking, Bonnie Berger, and David Reich. 2014. Reconstructing Austronesian

- population history in island Southeast Asia. *Nature Communications* 5, Article number: 4689.
- List, Johann-Mattis, Christoph Rzymski, Tiago Tresoldi, Simon Greenhill, and Robert Forkel. 2019. The Database of Cross-Linguistic Colexifications, reproducible analysis of cross-linguistic polysemies [CLICS³].
- Lobel, Jason William. 2016. *North Borneo Sourcebook: Vocabularies and Functors*. Honolulu: University of Hawai'i Press.
- Mahdi, Waruno. 2008. Review of Russell Jones (gen. ed.) *Loan-words in Indonesian and Malay*. *Archipel* 76:318–22.
- Mellars, Paul. 2006. Going east: New genetic and archaeological perspectives on the modern human colonization of Eurasia. *Science* 313:796–800.
- Olsen, Emily Long. 2018. The sound patterns of Kachok in the context of Bahnaric and North-Bahnaric studies. Doctoral diss., The Graduate Center, CUNY.
- Peperkamp, Sharon. 2005. A psycholinguistic theory of loanword adaptations. In *Proceedings of the 30th Annual Meeting of the Berkeley Linguistics Society*, ed. by M. Ettliger, N. Fleischer, and M. Park-Doob, 341–52. Berkeley: The Society.
- Peperkamp, Sharon, and Emmanuel Dupoux. 2003. Reinterpreting loanword adaptations: The role of perception. *Proceedings of the 15th International Congress of Phonetic Sciences*, 367–70. Barcelona: Causal Productions.
- Perrier, Xavier, Edmond De Langhe, Mark Donohue, Carol Lentfer, Luc Vrydaghs, Frédéric Bakry, Françoise Carreel, Isabelle Hippolyte, Jean-Pierre Horry, Christophe Jenny, Vincent Lebot, Ange-Marie Risterucci, Kodjo Tomekpe, Hugues Doutrelepon, Terry Ball, Jason Manwaring, Pierre de Maret, and Tim Denham. 2011. Multidisciplinary perspectives on banana (*Musa* spp.) domestication. *Proceedings of the National Academy of Sciences* 108(28): 11311–18.
- Pinchi, V., P. Barbieri, F. Pradella, M. Focardi, V. Bartolini, and G-A. Norelli. 2015. Dental ritual mutilations and forensic odontologist practice: a review of the literature. *Acta Somatologica Croatica* 49(1): 3–13.
- Pou, Saveros, and Philip N. Jenner. 1973. Some Chinese loanwords in Khmer. *Języki Obce w Szkole*. 1–90.
- Puri, Rajindra K. 2001. *The Bulungan ethnobiology handbook*. Center for International Forestry Research. Bogor, Indonesia.
- Rau, Felix, and Paul Sidwell. 2019. The Munda maritime hypothesis. *Journal of the Southeast Asian Linguistics Society* 12:35–57.
- Reid, Anthony. 1988. *Southeast Asia in the Age of Commerce: 1450-1680. Volume one: The lands below the winds*. New Haven and London: Yale University Press.
- Reid, Lawrence A. 1994. Unravelling the linguistic histories of Philippine negritos. *Language contact and change in the Austronesian world* 77:443.
- Rensch, Calvin R, Carolyn M Rensch, Jonas Noeb, and Robert Sulis Ridu. 2006. *The Bidayuh language: Yesterday, today and tomorrow*. Kuching, Malaysia: Dayak Bidayuh National Association.
- Schmidt, Wilhelm. 1906. Die Mon–Khmer-Völker, ein Bindeglied zwischen Völkern Zentralasiens und Australasiens. *Archiv für Anthropologie*, Braunschweig, new series, 5:59–109
- Scott, N. C. 1956. *A dictionary of Sea Dayak*. London: School of Oriental and African Studies. SEAlang Mon-Khmer Etymological Dictionary [MKED].
- Sellato, Bernard J. L. 1993. The Punan question and the reconstruction of Borneo's culture history. In *Change and development in Borneo*, ed. by Vinson H. Sutlive, 47–81. Borneo Research Council Inc. Williamsburg, VA: Department of Anthropology, The College of William and Mary in Virginia.
- Shorto, Harry L. 2006. *A Mon-Khmer comparative dictionary*, ed. by Paul Sidwell, Doug Cooper, and Christian Bauer, PL 579. Canberra: The Australian National University.

- Sidwell, Paul. 2005. Aceh-Chamic and the Aceh-Chamic language family. In *Chamic and Beyond: Studies in mainland Austronesian languages*, ed. by Anthony Grant and Paul Sidwell, PL-569, 211–46. Canberra: Pacific Linguistics, The Australian National University.
- . 2006. Dating the separation of Aceh-Chamic and Chamic by etymological analysis of the Aceh-Chamic lexicon. *Mon-Khmer Studies* 36:187–206.
- Sidwell, Paul, and Felix Rau. 2015. Austroasiatic comparative-historical reconstructions: An Overview. In *The Handbook of Austroasiatic Languages*, ed. by Mathias Jenny and Paul Sidwell, 221–363. Leiden: Brill.
- Silverman, Dan. 1992. Multiple scansions in loanword phonology: Evidence from Cantonese. *Phonology* 9:289–328.
- Skeat, Walter William, and Charles Otto Blagden. 1906. *Pagan races of the Malay peninsula*. Volume 2. London: Macmillan.
- Smith, Alexander D. 2017. The languages of Borneo: A comprehensive classification. Doctoral diss., University of Hawai'i at Mānoa.
- . 2018. Modang and Woq Helaq wordlist. Kaipuleohone archive. <http://hdl.handle.net/10125/70562>.
- . 2021. The historical phonology of Hliboi, a Bidayuh language of Borneo. *Oceanic Linguistics* 60(1): 133–59.
- Soares, Pedro A., Jean A. Trejaut, Teresa Rito, Bruno Cavadas, Catherine Hill, Ken Khong Eng, Maru Mormina, Andreia Brandão, Ross M. Fraser, TseYi Wang, JunHun Loo, Christopher Snell, TsangMing Ko1, António Amorim, Maria Pala, Vincent Macaulay, David Bulbeck, James F. Wilson, Leonor Gusmão, Luísa Pereira, Stephen Oppenheimer, Marie Lin, and Martin B. Richards. 2006. Resolving the ancestry of Austronesian-speaking populations. *Human Genetics* 135:309–26.
- Tadmor, Uri. 2009. Loanwords in Indonesian. In *Loanwords in the World's Languages: A Comparative Handbook*, ed. by Martin Haspelmath and Uri Tadmor, 686–716. Berlin: De Gruyter Mouton.
- Thurgood, Graham. 1999. From Ancient Cham to Modern Dialects: Two Thousand Years of Language Contact and Change. *Oceanic Linguistics Special Publication* 28. Honolulu: University of Hawai'i Press.
- . 2007. The historical place of Aceh-Chamic: The known and the unknown. Paper presented at the First International Conference of Aceh and Indian Ocean Studies. February 24–27, 2007, Banda Aceh, Indonesia.