# Austroasiatic loanwords in Austronesian languages

Waruno Mahdi
*Fritz Haber Institute of the Max Planck Society, Berlin, Germany*, mahdi@fhi-berlin.mpg.de

# Austroasiatic loanwords in Austronesian languages

WARUNO MAHDI

ABSTRACT

This paper investigates lexical borrowing from Austroasiatic into Austronesian languages. It does so for the following contact stages and interactions between these languages following the Austronesian overseas dispersal: (Stage 1) early contacts between Austroasiatic and Malayo-Polynesian particularly in the early Neolithic in the area encompassing mainland Southeast Asia, Northwest Kalimantan, and Sumatra, often resulting in the transmission of faunal terms; (Stage 2) interactions between speakers of Mon-Khmer and Malayo-Chamic languages during the early development of statehood; (Stage 3) exchange of terms in the period of early Khmer, Cham, and Malay kingdoms. Some of these transmissions can be shown to have taken place against the backdrop of the paramountcy of the kingdom of Funan. The latter stage also involves Sanskrit loanwords which were transmitted to Malayo-Polynesian via a Mon-Khmer language. The loanwords in this article are informative of Southeast Asia's language history as well as the region's cultural history.

KEYWORDS

Loanwords, Neolithic, statehood, Austronesian, Austroasiatic, Malayo-Polynesian, Mon-Khmer, Malay, Cham, Khmer, Aslian.

## 1. Preliminary remarks

The process of lexical borrowing between languages of different phyla is often rather complex. This makes it difficult to identify intermediate proto-forms, let alone to reconstruct their phonetic form. The present study includes material data from my presentation at the Eleventh International Conference on Austronesian Linguistics (Mahdi 2009), as well as from Uri Tadmor's presentation at the same conference (Tadmor 2009b).[1]

Austroasiatic borrowings in Malay and other Austronesian languages are of particular interest, but they are also fraught with problems. This is presumably the reason they were not included in the major Malay loanword project headed by Russell Jones (Jones 2007; Mahdi 2008: 322).

The relationship between Austroasiatic and Austronesian had already attracted much earlier attention when Wilhelm Schmidt grouped both into one hypothetical Austric language family (Schmidt 1906: 82). Although at first knowledge about Austroasiatic comparative linguistics remained somewhat incomplete, the Austric hypothesis continued to be entertained for some time, see George van Driem (1999), who nonetheless did not support it himself. Indeed, André-Georges Haudricourt (1966: 47) had already indicated that various alleged Austric cognate sets merely represented a Mon-Khmer substratum in Cham and Acehnese or occasional borrowings into Malay (sometimes without corresponding cognates in Cham).

I also no longer consider the Austric macrophylum a historical reality. As the title of this publication shows, I follow lexical borrowing from Austroasiatic into Austronesian languages. For reasons of space, I shall not address loanwords from Austronesian into Austroasiatic languages. I also omit early lexical transmission on the Asian mainland before the Malayo-Polynesian overseas dispersal, as it remains uncertain whether direct contacts between languages of the two phyla were involved, or if there was involvement of languages belonging to other phyla. The complexity of the problem is clearly illustrated by the hypothetical comparative analysis of Savaros Pou and Philip N. Jenner (1974). Furthermore, the early situation on the mainland was quite complex. Take *ag'* 'bow' in Table 1 as an example.[2] It is a frequently borrowed item in material-culture vocabulary. Its immediate precursor had undergone morphological derivation with limited distribution within Austroasiatic, the apparent donor phylum (see Table 1).

In the eastern part of the Austroasiatic speech area, 'crossbow' is expressed by reflexes of the same proto-form prefixed by *sn-*. The final glottal stop in the

[2] In this article, I use abbreviations for frequently used references as clarified in Appendix 1. The number sign <#> refers to the relevant lemma, whereas sub-entries are kept as in the original (A, b, c, et cetera). The names of Aslian languages quoted from SB06 have been re-edited following B76: 125-126.

proto-form is evident from the tone of the Vietnamese cognate *nā*$_{B1}$,[3] spelled as *ná*, which goes back to pre-Vietnamese *\*hnā?*. The prefixed Mon-Khmer form seems to have been borrowed into several neighbouring language phyla. To begin with, Hmong-Mien (formerly known as Miao-Yao) languages have *\*hnak* 'crossbow'. However, *\*hna[-i,-n,-ŋ]* 'crossbow' is more widespread. Then there is a cognate set in Tai-Kadai languages, reflecting *\*hnā?* 'crossbow'. Tibeto-Burman has Proto-Tibeto-Burman *\*sna* 'crossbow' (James A. Matisoff 2003: 172). The *\*a > ɔ* shift in the Early Middle Chinese reflex (see Table 2) apparently happened in various examples, as already noted by Paul K. Benedict (1972: 187). Bernhard Karlgren (1940: 150-151 #94z) cites it from the 周禮 *Zhōu lǐ* from the middle of the second century BCE.

| Austroasiatic *\*ag'* 'bow' |
| --- |
| **Munda** (P59 #6; B66 #27) |
| Kharia *k-a?* |
| Santali *ak'* |
| Mundari *a?* |
| Birhor *a?* |
| Korakur *a?* |
| |
| **Mon-Khmer** (Sh06 #266; S15 #3) |
| Palaung *a?* |
| Riang-Lang *ak*$_1$ |
| Praok *ak* |
| Danaw *ak* |
| Sre *a* |
| Semnam *āg* (SB06: 543) |

Table 1. Reflexes of Proto-Austroasiatic *\*ag'* 'bow' in Munda and Mon-Khmer languages.

| Austroasiatic |
| --- |
| **Proto-Mon-Khmer *\*sn-ā?*** (S00 #19; Sh06 #97) |
| Khmer *snā* |
| Stieng *sənā* |
| Loven *sənā* |
| Kuy *snā* |
| Sre *səna* |
| Biat *nā* |
| Vietnamese *nā*$_{B1}$ <*ná*> |
| **Hmong-Mien** |
| **Proto-Hmong-Mien *\*hnak*** (R10 #12) |
| Luoxiang Mien *ŋa*$_{D1}$ |
| Jinmen Mien *na*$_{D1}$ |

**Proto-Hmong-Mien** *\*hna[-i,-n,-ŋ]* (P70 #204; R10 #6)
Chengfeng Hmong *hnai*$_{A1}$
Qiandong Hmong *hnen*$_{A1}$
Chuangqiandian Hmong *hneŋ*$_{A1}$ ~ *nein*$_{A1}$
Petchabun Hmong *hneŋ*$_{A1}$
Suyung Hmong *hneŋ*$_{A1}$
Chiengrai Mien *hnā*$_{A1}$
Haininh Mien *na*

## Tai-Kadai

**Proto-Tai-Kadai** *hnāʔ* (L77 #4; H08 #593)
Thai *nā*$_{B1}$
Nong Khai *nā*$_{B1}$
Western Nung *nā*$_{B1}$
Po-ai *nā*$_{B1}$
Yai *nwa*$_{B1}$
Saek *nua*$_{B1}$

## Proto-Tibeto-Burman

**Proto-Tibeto-Burman** *\*s-na* (M03: 172)
Early Middle Chinese *nɔ'*
Chinese 弩 *nǔ*

Table 2. 'Crossbow' in Mainland Southeast Asian languages.

Austronesian has a proto-form prefixed with *\*pa-*. It underwent various semantic changes. Apart from the meaning 'bow', just as often 'arrow' and '[to] shoot' are noted.

**Proto-Austronesian *\*pa-naq* 'throw something at a target; shoot with bow and arrow'** (ACD under *\*panaq*)

**(Taiwan)**
Amis *panaʔ* 'bow'
Kavalan *pani* 'bow'
Tsou *pono* 'bow'
Pazeh *pa-pana* 'shoot with a bow'
Puyuma *panaʔ* 'arrow'
Paiwan *panaq* 'arrow'

**(Philippines)**
Ilokano *pána* 'arrow'
Isneg *pána* 'arrow'
Pangasinan *paná* 'arrow'
Tagalog *pánaʔ* 'arrow'
Cebuano *pánaʔ* 'arrow, spear projected; shoot an arrow, to spear'
Ifugaw *pána* 'bow and arrow'
Kapampangan *pana* 'bow and arrow; shoot with bow and arrow'
Bikol *pánaʔ* 'bow and arrow; archery'
Aklanon *panáʔ* 'bow and arrow; shoot with bow and arrow'

**(Southwest Malayo-Polynesian)**
Kadazan Dusun *pana* 'bow'
Karo Batak *panah* 'bow'
Malay *panah* 'arrow'
Balinese *panah* 'arrow'
Bare'e *pana* 'bow'
Wolio *pana* 'bow'
Muna *pana* 'bow'
Sundanese *panah* 'bow and arrow'
Makassarese *pana* 'bow and arrow'

**(East Malayo-Polynesian)**
Manggarai *pana* 'bow'
Watubela *fana-fanak* 'arrow'
Rembong *pana* 'bow and arrow'
Kei *fan* 'shoot with bow and arrow'
Biga *fan* 'shoot'
Nggela *vana* 'shoot with bow and arrow'
Lau *fana* 'shoot with bow and arrow'
Rotuman *fana* 'shoot with bow or gun'
Fijian *vana* 'shoot with an arrow or a gun; to pierce'
Hawaiian *pana* 'to shoot, as marbles, arrows, bow, bow and arrows'

Table 3. Reflexes of Proto-Austronesian *\*pa-naq.*

Note that the original meaning 'bow' is largely restricted to languages of Taiwan and West Malayo-Polynesia. Farther east it shifts to 'bow and arrow' and simply 'arrow'. In the east of the Indo-Malayan Archipelago and in Oceania, the reflexes frequently designate 'to shoot'. This confirms that it is not an original Austronesian proto-form, but introduced into the Austronesian region, whereby it was taken up by local Austronesian-speaking communities at different phases of their culture development. In the remainder of this article, I divide the interactions between Austroasiatic and Austronesian languages into several periods of socio-political development of their most prominent communities.

## 2. Earliest contacts between Austroasiatic and Western Malayo-Polynesian

Malayo-Polynesian migrations and language contacts in the far west of Southeast Asia continue to present a complex picture. The Aceh-Chamic migration into Indochina and subsequent movement to the north of Sumatra presumably began with the Sa Huỳnh culture on the Indochinese east coast, as already noted by Graham Thurgood (1999: 15).

Contacts with Malayic are more difficult to trace. The Malayic homeland was presumably in the western part of Kalimantan, with subsequent migration to Sumatra, as indicated by K. Alexander Adelaar (1992: 207) citing a personal communication from Robert Blust. This cross-over apparently also left behind sea-nomadic communities in the Riau-Lingga Islands. Malayic-speaking

seafarers were also involved in the maritime trade between Indochina and India, crossing the Isthmus of Kra near <u>Khao</u> Sam Kaeo (first mentioned in Chinese sources as *Dùnxùn* 頓遜, to be considered below in Sub-section 4).

As they lived in each other's proximity, Chamic speakers were in extensive contact with Bahnaric and Katuic speakers (see Paul Sidwell 2005: 211), more so than is otherwise characteristic of contact between Mon-Khmer and Malayo-Polynesian language speakers. I shall not discuss the specific features of that locally limited contact. Rather, I focus on lexical exchange between Austroasiatic and Malayo-Polynesian languages during this early period, highlighting numerous terms for fauna, flora, and parts of the human body.

## 2.1 Large bird of prey

There is an Eastern Austroasiatic etymon for a large bird of prey, which is typically referred to in the respective sources as hawk, kite, buzzard, eagle, or vulture. Some of the glosses given in the citations below might not be quite accurate, suggesting a generalized gloss as 'large bird of prey'. This word for 'large bird of prey' came into Malayo-Polynesian languages by at least two routes. One led to Aceh-Chamic, the other to Malayo-Javanic and Moken-Moklen.

**Austroasiatic**

*k[]laeŋ* (Sh06 #714; S06 #122; S15 #431; SJ03 #102; SB06 #4)
Khasi *khlieŋ ~ ʔlieŋ* 'kite, eagle'
Khmer *khlaeŋ* 'kite'
Sre *klaŋ* 'kite'
Chrau *klaŋ* 'hawk, large raptor'
Kuy *k[ʌ]lāŋ* 'hawk'
Bru *klāŋ* 'hawk'
Sô *kəlāŋ* 'hawk'
Ta'Oi *kaʔlāŋ* 'hawk'
Kriang *kaʔlāŋ* 'hawk'
Biat *klāŋ* 'hawk, large raptor'
*Jeh klāŋ* 'hawk'
*Halang klāŋ* 'hawk'
*Lavi klāŋ* 'hawk'
Juk *kalāŋ* 'hawk'
Brao *klāŋ* 'hawk, i.a.'
Katu *kalāŋ* 'eagle'
Kammu Yuan *klāŋ* 'kite, hawk, eagle'
Kammu Yuan *klāŋ* 'kite, hawk, eagle'
Palaung *klaŋ* 'kite, hawk'
Praok *klaŋ* 'kite, hawk'
Riang-Lang *klaŋ₁* 'kite, hawk'
Lawa *klaŋ* 'kite, hawk'
Wa *klaŋ* 'hawk, eagle'
Central Nicobarese *kəlāŋ* 'white-bellied eagle'
Nancowry *kalāŋ* 'vulture'

Central Sakai *klaʔ* 'hawk'
Semai *kəlak* 'kite'
Kintaq-Bong *kəlaŋ* 'white eagle'

**Austronesian**

**Aceh-Chamic *kalāŋ** (T99: 322)
Acehnese *kluɘŋ* 'kite'
Phanrang Cham *kalaŋ* 'bird of prey'
Jarai *klaŋ* 'eagle, kite'
Rade *tlaŋ* 'bird of prey, hawk'
Chru *kalāŋ* 'kite, hawk, bird of prey'

**Moken-Moklen *qelaŋ**
Moken *kəlāŋ* 'eagle, hawk'

**Malayo-Javanic *qelaŋ 'eagle, hawk'** (N75: 170)
Malay *həlaŋ ~ ʔəlaŋ ~ laŋ*
Old Javanese *həlaŋ*
Sundanese *huɘlaŋ*
Madurese *laŋ*
(see also Minangkabau *alaŋ*)

Table 4. Austroasiatic and Austronesian words for 'bird of prey'.

The Malayo-Javanic set might ultimately derive from a Mon-Khmer origin (including an Aslian one), but the *\*k > \*q* shift remains unexplained. Perhaps there was mediation of Moken-Moklen, in which *\*k* is a regular reflex of Malayo-Polynesian *\*q*, whereas in Malay, Javanese, and Sundanese the regular reflex is *h ~ ø*. However, such a scenario still does not give a clear explanation of the presence of *\*q* in Malayo-Javanic languages.

Besides the above, there is Old Javanese *kalaŋkyaŋ ~ kālaŋkyaŋ* 'kind of bird of prey (hawk or buzzard)' (P.J. Zoetmulder 1982: 773), more precisely identified as 'white-bellied sea eagle (*Haliaeetus leucogaster*)' by Jiři Jákl (2022: 263-265). This is a compound word, and the first element is obviously a borrowing from Austroasiatic. The borrowing was direct, as seen from the initial *k-*, and not via Malayo-Javanic which had initial *\*q-> h-*. The second element looks like it could be *hyaŋ ~ yaŋ* < Proto-West Malayo-Polynesian *\*qiaŋ* 'deity', but the semantic consequences of such an assumption are not quite clear to me.

## 2.2 Peacock

In view of it also being attested in Munda languages (see Table 5), it is relatively well known that this bird name derives not just from Mon-Khmer but from a high-order Austroasiatic source. The form was borrowed into Chamic and the Bahnar cognates have been considered Chamic back-borrowings, particularly from Roglai (Sh06 #72). I do not place Acehnese *muraʔ* together with the Chamic reflexes, because it seems more likely to have been borrowed via Malay *məraʔ*. The word was also borrowed, apparently via Malay, into other

languages of Sumatra, Java, and Madura. In view of the exact position of its first syllable vowel, the form is likely to have been borrowed into Southeast Asia from a Munda source.

**Austroasiatic**

*m[]raək* 'peacock' (P59 #27; SJ03 #61; Sh06 #72; Sh06 #416)

**Munda**
Kharia *maráʔ*
Santali *maraʔ*
Mundari *maraʔ*
Ho *maraʔ*
Kurku *mara*
Sora *mārā-n ~ mār-ən*

**Mon-Khmer**
Old Mon *mrek*
Brao *brāk*
Juk *brāʔ*
Sre *bra*
Chrau *vrāʔ*
Biat *brāk*
Palaung *phraʔ ~ braʔ*
Riang-Lang *prak₂*
Kuy *maryaʔ*
Kantu *brāk*
Bru *rịạʔ*
Pacoh *rāk*
Bahnar *ʔəmraʔ*
Goʼlar Bahnar *həmraʔ*

**Austronesian**

Proto-Chamic *ˣʔəmrāk* (T99: 350)
written Cham *amrak ~ mrak*
Phanrang Cham *amraʔ ~ mraʔ*
Western Cham *mraʔ*
Rade *amrak*
Jarai *amrăʔ*
Chru *amrāʔ*
North Roglai *amrāʔ*

**Malayo-Javanic**
Malay *məraʔ*

> **via Malay**
> Old Javanese *mrak*
> Acehnese *mɯraʔ*
> Minangkabau *mɛrak*

Rejang *mərak*
Sundanese *mərak*
Javanese *mərak*
Madurese *mərak*

Table 5. Austroasiatic and Austronesian words for 'peacock'.

The Malay cognate was already indicated by Thomas Bowrey (1701), who spelled it *mārak* in his dictionary of this language. The ultimate Austroasiatic origin of the word seems obvious from the above.

2.3 Ant

The transmission of the word for 'ant' is somewhat more complex. Although the cognate set includes reflexes from both Western and Eastern Austroasiatic languages, reflexes from the original proto-form meaning 'to sting' seem to have only been retained in the eastern branch. A derivation with an *-m-* infix was retained in both East and West Austroasiatic languages. As already indicated by H.L. Shorto (2006: 258 #873B), this form exhibits cognates in Malay, Javanese, and Karo Batak. Shorto also cites the otherwise unattested and presumably erroneous Cham *hmōc* 'ant'. Meanwhile, I found further cognates in West Malayo-Polynesian languages:

**Austroasiatic**

*suəc* **'to sting'** (S06 #1217; Sh06 #873B)

**Mon-Khmer**
Kuy *sōc*
Sô *sūc*
Bru *sūjʔ*
Ta'Oi *sūc*
Sre *souc*
Biat *chōc*
Bahnar *sōc*
Kammu Yuan *hūc*
Riang-Lang *huc₁*
Praok *hɯc*
Proto-Semai *sɤɤc*

*s-m-uəc* **'ant'** (P59 #130; SJ03 #659; S06 #1218; S15 #517; Sh06 #873B-C)
**Munda**
Kharia *mu'j-ḍa*
Sora *mui-dā-n ~ muj-dā-n ~ mud-dā-n*
Santali *mu'j*
Mundari *mui'j*
Ho *mui*
Birhor *mui ~ mu'j*
Bhumij *mue*

**Mon-Khmer**
Juk *smōc*
Kuy *smōc*
Souei *smūjʔ*
Bru *səmūjʔ*
Sô *samūc*
Jru' *hmōc*
Bahnar *hmōc*
Brao *hmōc*
Jeh *mut*
Halang *mūt*
Bit *smūc*
Wa *muc*
Kammu Yuan *mūc*
Bo-Luang Lawa *ʔmauk*
Umphai Lawa *ʔmaut*
Mae-Sariang *ʔmaut*
Mon *həmot* 'k.o. ant'
Khmer *s-r-əmaoc*
Su' *rəmōt*
NK Lamet *rəmūc*

**Austronesian**

Malay *səmut*
Iban *səmut*
Javanese *səmut*
Balinese *səmut* ('small ant')
Rejang *səmut*
Karo Batak *səmut*
Dairi–Pakpak Batak *səmət ~ səmut* ('small ant')
(see also Toba Batak *semet-semet ~ sisemet* ('small ant'), with vowel assimilation)

Table 6. Austroasiatic and Austronesian words for 'sting' and 'ant'.

2.4 Crab

For another animal name, transmission from a Mon-Khmer source appears to have taken place in relative isolation. The Eastern-Austroasiatic cognate set is given below. There are only few borrowings in the Malay Peninsula, Sumatra, and Sarawak. I did not find any Chamic cognates.

**Austroasiatic**

***ktām* 'crab'** (H66: 47; Sh06 #1348)
Khasi *tham*
Khmer *kdām*
Kuy *ktām ~ tām*
Bahnar *kɤtām*
Halang *kɤtām*
Jeh *katām*

Sre *tām*
Semai *kəntem*

**Austronesian**

Moken *kətām ~ kətam*
Acehnese *gutwəm*
Malay *kətam*
Rejang *kətəm*
Iban *kətam*

Table 7. Austroasiatic and Austronesian words for 'crab'.

2.5 EEL

For this animal name, the original cognate set seems to be limited to some Mon-Khmer groups. Reflexes belonging to this cognate set appear to have undergone transmission into Western Malayo-Polynesian languages by several routes.

**Austroasiatic**

*\*nduŋ, \*dnduŋ, \*dnduuŋ, \*[l]nduŋ* **'eel'** (S06 #579; S00 #669; SJ3 #864; S06 #1379; P59 #124)
Kui *nthuŋ*
TaʾOi *handoŋ*
Bru *noŋ*
Souei *ʔanuŋ*
Chatong *ntṵŋ*
Kantu *ʔanduŋ*
Katu *ʔadoŋ*
East-Mnong *nduŋ*
Stieng *ndoŋ*
Chrau *nduŋ*
Köho *nduŋ*
Lavi *cuŋ*
Juk *duŋ*
Jruʾ *duŋ*
Nyaheun *duŋ*
Sapuan *duŋ*
Laveh *duŋ*
Kharia *ɖuŋɖuŋ*
Khmer *ɔndɔŋ, dɔndɔŋ, dɔnduŋ, ɔntɔŋ*
Boloven *duŋ*
Sre *nduŋ*
Biat *nduŋ*

**Austronesian**

*ᴿ[l,d][e,i]nduŋ* **'k.o. eel'**
Malay *lendoŋ, lindoŋ*
Acehnese *ndōŋ*
Karo Batak *duŋduŋ*

Gayo *dənduŋ*
Rejang *dənduŋ*
Iban *lindoŋ*
Balinese *linduŋ*
Sasak *linduŋ*
Kadazan *hinduŋ*
Tausug *induŋ*
Cebuano *induŋ*

Table 8. Austroasiatic and Austronesian cognate words for 'eel'.

A possibly related form without intervocalic -*d*- is reported for Aceh-Chamic, particularly in written Cham *lanūŋ* 'eel' (Étienne Aymonier 1891: 45) and Acehnese *linòŋ* 'eel' (J. Kreemer 1931: 164). Note that Acehnese also has *ndōŋ* 'eel', so that the form without intervocalic -*d*- apparently reflects a different proto-form.

2.6 ARMPIT

The exact reconstruction of the basic Mon-Khmer proto-form for 'armpit' is difficult to determine, because the vowel development remains unclear. There seem to be reduplicated cognates in some Philippinic languages. A probable cognate with the preposited element *kət*- is given below.

**Austroasiatic**

*\*ʔiək* **'armpit'** (Sh06 #269; S15 #39)
Riang-Lang *[ɔk₁]-yak₂*
Bahnar *[bɔk]-āk*
Danaw *kʰă-yɛ̀ˀk*
Palaung *[kəndəʔ]-iəʔ*

**Austronesian**

*\*kət-iək* **'armpit'**
Malay *kət-iak*
Minangkabau *kat-iak*

*\*yek-yek* **'armpit'** (ACD under PPh *\*yekyek₁*)
Bontok *yəkyək*
Bikol *yukyók*
Ifugao *yoyók*

Table 9. Austroasiatic and Austronesian words for 'armpit'.

Somewhat more widespread is a Mon-Khmer set of cognates with the prefix *\*kəl*-. There are cognates with the same meaning in several Western Malayo-Polynesian languages. Also note below an additional set of cognates with the prefix *\*il*- instead of *\*kil*-:

**Austroasiatic**

---

*\*kəl-ʔiək* '**armpit**' (Sh06 #269; S15 #39)
Yuan Khmu *kəlʔèk*
Cuang Khmu *klʔɛk*
Khmer *kliək*
Vietnamese *nách* (pre-Vietnamese *\*hn-āc* from an earlier *\*kn-āyk*)

**Austronesian**

---

*\*kilik* '**armpit**' (ACD under PWMP *\*kilik₂*)
Malay *kilek ~ kɛlɛk*
Javanese *kɛlɛk*
Malagasy *hélikʌ* ('armpit, carry under the armpit')
Tagalog *kilik* ('carry on the hip supported by the arm')
(see also Sundanese *kɛlɛk*)

*\*ilek* '**armpit**' (ACD under PWMP *\*ilek*)
Tausug *iluk*
Lun Dayeh *ilek*
Kelabit *ilək*
Lolak *iyok*

---

Table 10. Austroasiatic for 'armpit' with *\*kəl-* prefix, and Austronesian cognates.

Returning to the basic Mon-Khmer proto-form mentioned above, a possible Chinese and even Tibeto-Burman cognate have been identified (B72 #448; Sh06 #269). Karlgren (1940: 334 #800M) notes "*Chuang*" as the earliest source. This would be the *Zhuāngzǐ* 莊子 manuscript from the late Warring States period, containing essays from the fourth to second centuries BCE (Martin Kern 2010: 74). Borrowing from Mon-Khmer is unlikely. Matisoff (2003: 326, 328 #a) notes a probable origin in Proto-Tibeto-Burman *\*g-yak* 'armpit', with the reflexes Lushai *zak* and Written Burmese *gyak-kəliʔ ~ chak-kəliʔ*. To these can be added Early-Middle Chinese *yiaiyk* and Chinese 腋*yè ~ yì* (Edwin G. Pulleyblank 1991: 364, 370).

It seems possible therefore that the Mon-Khmer forms were borrowed from Tibeto-Burman. However, there is also the possibility that Malayo-Polynesians sailing up the Irrawaddy or Brahmaputra (see Mahdi 1999: 166) transmitted the word to Tibeto-Burmans after having acquired it themselves from a Mon-Khmer source. Indeed, the above cited minor cognate sets in Malayo-Polynesian languages, reconstructed as *\*yekyek*, *\*kilik,* and *\*ilek*, seem most likely to be borrowings from Mon-Khmer.

### 2.7 Belly, stomach, intestine

The Mon-Khmer origin for the name of another body part is relatively transparent. There is a basic root form as well as various prefixed derivations in Mon-Khmer (Shorto 2006: 251 #844). It would appear that a derivation with the prefix *p-* was borrowed into Aceh-Chamic. A further cognate set exhibiting *\*p-* can be found in Malayic. The transmission into the Archipelago seems likely to have proceeded via Malay.

**Austroasiatic**

*\*rūc* **'intestine'** (S06 #904; S15 #658)
Kuy *ru̯atɛ*
Souei *rↄ̂j̃ʔ*
Ta'Oi *rↄ̄c*
Pacoh *rↄ̃jʔ*
Lameet *ɣóc*
Mương *roč*
Vietnamese *ruột*

**\*k-rūc** **'intestine'** (Sh06 #844)
Mon *krↄt*
Semai *krˠt* ('belly')

*\*p-rūc* **'intestine'** (S00 #404; Sh06 #844)
Mnong *pruəc*
Sre *proč*
Stieng *prↄ̃č*
Biat *prↄ̃č*
Köho *proc* ('belly')

**Austronesian**

**Proto-Aceh-Chamic \****pruac* **'stomach, intestine'** (T99: 360)
Acehnese *pruət*
written Cham *prwəč*
Rade *proč*
Jarai *proaĩʔ ~ pruaĩʔ*
North Roglai *puaiʔ*
Haroi *prŏaiʔ*
Phanrang Cham *proyʔ*

**Malayic \****pərut* **'stomach, belly, intestine'** (A92: 59)
Malay *pərut* 'stomach'
Minangkabau *paruɣʔ* 'stomach'
Hulu Banjarese *parut* 'stomach, belly'
Serawai *pəɣut* 'intestine'
Iban *pərut* 'stomach, belly, intestine'

Table 11. Austroasiatic and Austronesian related words for 'belly, stomach, intestine'.

2.8 Molar

There is a Mon-Khmer cognate set of words meaning 'molar'. This seems likely to be of Tibeto-Burman origin. Meanwhile, Tadmor (2009a: 694 Table 1) has noted the cognate Malay *gərəham ~ gəraham* 'molar', apparently borrowed from a Mon-Khmer source. There also are numerous further Malayo-Polynesian cognates:

---

**Tibeto-Burman**

---

**Proto-Tibeto-Burman \*_gam_ 'put into mouth / seize with mouth; jaw / molar'** (M03: 299, 300m)
written Tibetan *ħgam* 'put or throw in mouth'
written Burmese *ʔam-swâ* 'molar'
Trung *s-kam* 'molar' (consisting of *sa* 'tooth' + *kam*)
Old Chinese *\*gʼəm* 'hold in mouth'
Chinese 含 *hán* 'hold in mouth', 頷 *hàn* 'chin, jowl' (G12 #3818, #3824)

---

**Austroasiatic**

---

**Proto-Mon-Khmer \*[ʔ-,g-]ām 'molar'** (S00: 776; S06 #99; S15 #5, SJ03 #1024, Sh06 #1303, #1318)
Nyah Kur *niək kə-ʔam*
Danaw *am₁ pəiŋ₄*
Biat *gam*
Stieng *gəm*
Chrau *də-gam*
Köho *tər-gəm*
Katu *ta-ʔām*
Pacoh *taŋ-ʔām*
Bahnar *t[əŋ]-ām*
Jru' *kn-iəm*
Laveh *kan-ɨ:m*
Khasi *tyŋ-am*
Khmer *th-kam* ('jaw')

---

**Austronesian**

---

Malay *gərəham ~ gəraham*
Minangkabau *garam-an* (A92: 41)
Serawai *gəɣm-an* (A92: 41)
Iban *gaʔam* (A92: 41)
Rejang *gərŋəm*
Sundanese *carɛham*
Javanese *graham*
Madurese *ghərrəm*
Banjarese *garaham*

---

Table 12. Austroasiatic, Tibeto-Burman, and Austronesian cognates for 'molar'.

The above Austronesian cognates appear to have been borrowed from the same source, but probably transmitted via Malay rather than directly from a Mon-Khmer source. In contrast, Acehnese *ghuəm* and written Cham *tahoṃ*

(Étienne Aymonier and Antoine Cabaton 1906: 189) were perhaps borrowed directly from a Mon-Khmer source.

2.9 WART

In some Mon-Khmer languages there is a relatively restricted set of words meaning 'wart', for which there are a few borrowed cognates in Malayo-Chamic. I am including them here, although the time of the appearance in Mon-Khmer and borrowing into Malayo-Chamic is unclear and might have been relatively late.

---

**Austroasiatic**

> **Mon-Khmer** *\*ktuut* ~ *\*ktuət* **'wart'** (Sh06 #1009)
> Biat *tūt*
> Palaung *tot*
> Mon *kətot*
> Kuy *tāt*
> Stieng *tɔɬ*
> Chrau *cɔɬ*

**Austronesian**

> **Malayo-Chamic** *ˣktuət* **'wart'**
> Cham *katwaʔ*
> Acehnese *gɤtuət*
> Malay *kətuat*

Table 13. Mon-Khmer and Malayo-Chamic cognates for 'wart'.

---

Further research might identify more cognates in Malayo-Chamic or other related languages.

2.10 TWINS

Western Malayo-Polynesian languages display cognates for 'twins' which must have spread mainly via Malay (A92: 61 #1; ACD under 'twins'). This is confirmed by the fact that their proto-form features an *\*r* (and not an *\*R*) as diaphoneme. The ultimate origin of the form seems to be the Austroasiatic cardinal numeral 'two'. It seems noteworthy that Aslian languages have a prefixed *\*həm-*:

---

**Austroasiatic** *\*[b-]ʔār* **'two'**

> **Munda** (P59 #49)
> Kharia *u-bár*
> Santali *bar[-ea]*
> Birhor *bar[-ea]*
> Mundari *bar* ~ *bar[-ia]*
> Bhumij *bar-ia*

Korwa *bar-i*
Kurku *bar-ia*
Juang *am-bar*
Sora *bā-gu ~ bār*
Gutob *um-bār-o*
Remo *'m-bār*

**Mon-Khmer** (SJ03 #117; S06 #149; S15 #12, Sh06 #1562)
Old Mon *ɓār*
Bru *bār*
Ta'Oi *bār*
Chatong *bār*
Katu *ɓə̄r*
Pacoh *bār*
Sre *bar*
Chrau *vār*
Biat *bār*
Halang *bār*
Lavi *piɑr*
Juk *bār*
Jru' *bə̄r*
Nyaheun *bān*
Laveh *bār*
Bahnar *ɓār*
Jeh *bāl*
Kammu Yuan *pằr*
Riang-Lang *kə̂2-ar1*
Palaung *ar*
Khasi *ar*
Mương *hal*
Vietnamese *hai*

> **Proto-Aslian *həm-bār 'two'** (SB06 #271; Sh06 #1562)
> Sěpang & Air-Hitam Mah-Meri *h'mbār*
> Semaq-Beri *mar*
> (see also Bukit-Bangkong Mah-Meri *hmba*)

**Austronesian 'twins'** (A92 #1; ACD under 'twins')

> Toba Batak *hombar*
> Malay *kəmbar*
> Minangkabau *kambar*
> Serawai *gəmbaɣ* (with irregular *g-*, see A92: 61)
> Sundanese *kəmbar*
> Javanese *kəmbar*
> Balinese *kəmbar*
> Madurese *kəmbhar*
> Malagasy *kámbanʌ*
> Hulu Banjarese *kambar*
> Iban *gəmbar* (with irregular *g-*, see A92: 61)
> Makassarese *kambaraʔ*
> Bikol *kambál*
> Tagalog *kambál*

Kapampangan *kámbal*
Kasiguran Dumagat *kambál*
(see also Acehnese *kumbuə*)

Table 14. Austroasiatic cognates of *\*[b-]ʔār* 'two' and Austronesian loans in the meaning 'twins'.

The above Aslian set is most likely the direct source of the Malayo-Polynesian forms, although the shift of the initial *h > k,g* remains unexplained. Interestingly, the Austroasiatic cognates do not mean 'twins', but 'two', implying that the entire Malayo-Polynesian cognate set apparently originates from a single Austroasiatic source, more specifically, an Aslian, which was borrowed into Malay, involving the semantic shift 'two' > 'twin'. Further transmission into Western Malayo-Polynesian languages therefore proceeded via Malay.

## 3. Neolithic contacts

### 3.1 Thatching grass

The so-called cogon grass (*Imperata cylindrica*) is a species of long grass which occurs throughout Southeast Asia and is used in particular for thatching roofs, but also for weaving mats and bags. In some early sources, it is also glossed as 'jungle grass' or 'elephant grass'. A cognate set reflecting an apparently closest-to-original proto-form is represented in Munda languages. An affiliated cognate set in Mon-Khmer languages has an additional *\*p-* prefix. Cognates borrowed into Western Malayo-Polynesian languages do not include *p-*, hence must originate from Munda languages. They refer to the same *Imperata cylindrica* thatching grass. Another cognate set with initial *r-* is reported in Aceh-Chamic languages:

**Austroasiatic** *\*lāŋ* **'long grass'**

**Munda** *\*ə-lāŋ* (P59 #270)
Sora *alā́ŋ-ən*
Gutob *ʋlo*
Remo *lo*
Pareng *alāŋ-ən*
Kharia *ɔlɔ́ŋ*

**Mon-Khmer** *\*p-lāŋ* **'thatching grass'** (Sh06 #749; S15 #455; SB06 #34)
Old Khmer *plāŋ*
Kuy *plaŋ*
Palaung *pləŋ*
Riang-Lang *plaŋ*₁
Bo-Luang Lawa *plaŋ*
Umphai Lawa *plɔŋ*
Lameet *pláŋ*
Vietnamese *caɲ* <*tranh*>
Wa *ploŋ*

Khmu *plāŋ*
Khasi *phlaŋ ~ ʔlaŋ* ('grass')
Jah-Hut *pluoŋ*
Mah-Meri *pəloŋ* ('[thatch] roof')

**Austronesian**

**Malayo-Javanic \*alaŋ-alaŋ 'thatching grass'**
Malay *alaŋ-alaŋ ~ lalaŋ*
Iban *lalaŋ*
Minangkabau *alalaŋ ~ alaŋ-alaŋ ~ ilalaŋ*
Rejang *lalaŋ*
Madurese *lalaŋ*
Javanese *alaŋ*

**Aceh-Chamic \*[r]alāŋ 'thatching grass'** (Sh75: 88; T99: 301)
Acehnese *naluɯŋ*
Rade *hlaŋ*
Jarai *həlăŋ*
Chru *rəlāŋ*
Roglai *rɤlak*
North Roglai *ralāk*
Cham *ralāŋ*

Table 15. Austroasiatic and Austronesian cognate words for cogon grass (*Imperata cylindrica*).

The history of this form is somewhat complicated. There apparently was reduplication of the basic form with subsequent simplification of the reduplicated form *\*alaŋ-alaŋ > \*alalaŋ > \*lalaŋ*, ending in the loss of the reduplication in Javanese.

A basic Munda form apparently moved to the east in several transmissions: as *\*p*-derivation into Mon-Khmer; as *\*[r]*-derivation into Aceh-Chamic; and as reduplication into Malayo-Javanic (among which I include Rejang and Iban). The gradual simplification in Malayo-Javanic languages suggests an early date of the borrowing from Austroasiatic. This is also suggested by the initial *n* in Acehnese, resulting from a tendency to alternate between *r*, *l*, and *n* when the next consonant is *l*.

### 3.2 Leaf, sheet

In a number of Malayo-Javanic languages there is a count-word which reflects *\*helai* and is used for garments, cloth, or paper. The term seems to originate from a derivation of the Austroasiatic root for 'leaf'. The affiliated cognate set in Chamic retains the meaning 'leaf'.[4]

---

[4] For similar-looking forms in Dravidian and other languages, see Mahdi (1998: 398-399).

**Austroasiatic**

**Munda \*[u]-laʔ 'leaf'** (P56 #50)
Kharia *u'la*ʔ
Sora *'ōlā-n*
Gutob *ōlā*
Pareng *ōlā*
Remop *ōlā ~ ulā*

**Mon-Khmer \*s-laʔ 'leaf'** (SJ03 #10; S06 #22; S15 #421; Sh06 #230)
Old Mon *sla*
modern Mon *hla*ʔ
Kuy *slā ~ lhā*
Bru *sulā*
Ta'Oi *halā*
Katu *ʔalā*
Pacoh *ʔulā*
Stieng *lā*
Chrau *lā*
Jeh *lā*
Halang *lā*
Lavi *halā*
Jru' *lā*
Nyaheun *hlā*
Brao *chlā*
Bahnar *hlā*
Kammu-Yuan *láʔ*
Palaung *hla*
Riang-Lang *laʔ*ḷ
Lameet *láʔ*
Praok *la*
Lawa *hlaʔ*
Khasi *sla*
Mương *lá*
Vietnamese *la*ʙ₁ <lá> (from pre-Vietnamese *hlaʔ*)

   **Aslian \*sə-laʔ 'leaf'** (B76 #57)
   Kensiu *haliʔ*
   Kintaq-Bong *həliʔ*
   Jehai *haliʔ*
   Mendriq *haliʔ*
   Mintil *haliyʔ*
   Che'-Wong *haleʔ*
   Semnam *səlāʔ*
   Lanoh *səlaʔ*
   Temiar *səlāʔ*
   Semai *səlāʔ*
   Jah-Hut *hlaʔ*
   Semaq-Beri *salāh*

**Austronesian**

> **Chamic ˣsəla[ʔ] 'leaf'** (T99: 331)
> written Cham *halā*
> Rade *hla*
> Jarai *hla*
> Chru *səla*
> North Roglai *hlāʔ*
> Western Cham *hla*
> Phanrang Cham *hala*
>
> **Malayo-Javanic ˣhəlai 'sheet [of]'** (N75: 91)
> Malay *həlay*
> Sundanese *huulay*
> Old Javanese *həlay ~ həle*
> (see also Minangkabau *alay*, Hulu Banjarese *halay*)

Table 16. Austroasiatic and Austronesian cognate words for 'leaf'.

The Malayo-Javanic cognate set for 'sheet [of]' cited in Table 16 seems most likely to have been borrowed from Aslian. In Munda and Mon-Khmer languages, the ultimate vowel is *a* or *ā*. Only in some Aslian languages is there a shift to *e* or *i*, and in Mintil it is even diphthongized.

3.3 Chaff or husk (of rice)

After the husked rice grain is pounded, the unhusked grain (Malay *bəras*) is separated from the chaff (Malay *səkam*) by winnowing (Malay *mənampi*). It is therefore a feature of rice cultivation. As already noted by Tadmor (2009a: 693), the Malay word for 'chaff' seems to be of Mon-Khmer origin. This word was only borrowed into a few West Malayo-Polynesian languages, and only from an etymon with an *s*– prefix (as, for example, in Bru):

**Austroasiatic**

> ***[ə-, s-, ŋ-]kāmʔ 'chaff'*** (S06 #94; S15 #321; Sh06 #1313)
> Bru *sakām*
> Ta'Oi *ŋkām*
> Pacoh *ʔəkām*
> Kantu *ʔaŋkām ~ ŋkām*
> Mon *kam*
> Palaung *kham*
> Lameet *ŋkām*
> Lawa *kam*
> Riang-Lang *kham₁*
> Khmer *ʔɔŋkām*
> Kammu-Yuan *həŋkā̆m*
> Vietnamese *kām*B1 <cám> (from pre-Vietnamese *\*kāmʔ*; the final glottal is indicated by the B-tone)

**Austronesian**

> ˣ*səkam* **'chaff'**
> Malay *səkam*
> Acehnese *sukuəm*
> Gayo *səkam*
> Minangkabau *sakam*
> Madurese *səkəm*

Table 17. Austroasiatic and Austronesian cognate words for 'chaff'.

It is noteworthy that the borrowing into Malay (and thence into other Malayo-Polynesian languages) seems not to have been from Mon or Khmer, but apparently from a Katuic language such as Bru. The written Cham attestation *hakaṃ* (Aymonier and Cabaton 1906: 500) appears to reflect a different Mon-Khmer language. All of this would suggest early direct contacts between Malay and Mon-Khmer languages spoken in the interior of mainland Southeast Asia, circumventing Mon and Khmer. It is an indication of early Malay shipping up the Mekong River relatively far inland.

## 4. Early stages of statehood

### 4.1 Highland community

A polity referred to as *Dùnxùn* (頓遜) in an early Chinese source—the Liang Annals (梁書 *Liáng Shū*)—was considered by Paul Wheatley (1961: 19-21, 286) to be a Mon polity straddling the Isthmus of Kra. The Ming History (明史 *Míngshǐ*) indicates that Malacca (滿剌加 *Mǎnlàjiā*) had previously been reported to be the old country of *Dùnxùn* (W.P. Groeneveldt 1877: 129), effectively placing the latter on the Malay Peninsula. The crucial circumstance which made *Dùnxùn* important was that its territory straddled the Isthmus of Kra, simultaneously controlling a navigable river flowing eastwards to the Gulf of Thailand, and another flowing westwards to the Andaman Sea. As the Liang Annals note:

> More than 3,000 *lǐ* [*c.* 1,700 km] from the southern border [of 扶南 *Fúnán*] is the kingdom *Dùnxùn* situated on a sea mountain-path.

The quotation (retranslated here from the Chinese) is taken from Wheatley (1961: 16). He glossed *hǎiqí* (海 *hǎi* 'sea, maritime', 崎 *qí* 'mountain path, rugged') as 'ocean stepping stone'.[5] But to my mind, a literal translation probably best conveys what the Chinese writer had in mind: a mountain pass connecting the upper reaches of a river flowing to the sea in the east with one flowing to the sea in the west. Relatively small ships could be hauled over the pass, allowing them to cross the Isthmus of Kra without having to sail around the Malay Peninsula (which would have required waiting at the southern tip for the turn of the monsoon). It was therefore a mountain pass from sea to sea.

---

[5] For other interpretations, see Wheatley (1961: 20 fn. 3).

Turning to the name *Dùnxùn*, Shorto (1963: 583) interprets it as the Chinese rendering of a "proto-Mon" *\*Ḍūŋsun*, literally meaning 'five cities, or kingdoms'. The Liang Annals do indeed mention five "kings" in *Dùnxùn*. One might therefore conclude that *Dùnxùn* had originally been an Aslian tribal alliance of five clans subordinate to the Mons. This alliance must have profited from the presence in its territory of a mountain pass which connected two navigable rivers flowing in opposite directions. Although Old Mon *ḍuŋ* apparently meant 'polity, kingdom, country' in the Mon kingdom of Rāmaññadesa, Shorto (2006: 195 #581) is probably on the right track in assuming the corresponding proto-form to have meant 'clan territory'.

It seems likely that Malay-speaking sailors on the ships which sailed up and down these rivers pronounced the Mon name *\*Ḍūŋsun* as *\*Dusun*, subsequently applying it metaphorically to any hinterland community they reached for trading purposes by sailing up a river, who spoke a language they did not readily understand. Meanwhile, their contacts with Aslian communities on the mountain pass could explain the Aslian borrowings previously noted.

The tentative Malay rendering of this toponym as *dusun* seems to have disseminated across a remarkably wide area, confirming the early date of its origin. Hinterland regions and peoples in North Sumatra (J. Paulus 1917: 628), the Barito River Basin (Alfred B. Hudson 1967: 11, 14), and in Sabah (A.L. Gossens 1924) have been referred to as *Dusun*. The word also refers to a 'socio-political administrative unit amongst the Kerinci and the Rejang' (John N. Miksic 1989). Finally, it refers to 'small village, rural' and/or 'unsophisticated, boorish' in Acehnese *duson*, Gayo *dusun*, Minangkabau *dusun*, Sundanese *dusun*, Javanese *ḍusun* [kråmå style], Balinese *dusun*, Makassarese *rusuŋ*, and Buginese *dusuŋ*. The word *dusun* also occurs in the local Malay dialect of Roon (spoken in West Papua), where it means 'inland, forest' (David Gil p.c.).

In the above scenario of transmission, semantic development took place over a long period of time. The earliest borrowings, apparently transmitted by Malayic seafarers, referred to some upriver region one reached by boat.

## 4.2 Rice-fields

At the height of the power of the Mon kingdom of Ramaññadesa, diked and stepped submergible rice fields (paddy-fields) were developed on the Isthmus of Kra. In particular, at excavation sites in Satingpra they are dated to the period of 1900-1700 BP (Janice Stargardt 1983: 84 under "stage three"). This date must therefore have been much later than the beginnings of *Dùnxùn* alias *Dusun* discussed in the previous sub-section.

Tadmor (2009b) notes cognate pairs referring to such rice-fields, involving a Malay form on the one hand, and either a Mon or Khmer one on the other. Consider the attestations Old Mon *bnaŋ* 'unit of paddy-land' and Sre *brnāŋ* 'levelled land' in Mon-Khmer and Malay *bəndaŋ* 'irrigated rice-field, stretch of many paddy-fields'. H.C. Klinkert (1902: 169) glosses the latter as 'piece of arable land, either irrigated or not', but notes that others gloss it as 'paddy-fields' (*sawah*). In standard Indonesian Malay it means 'paddy-field' (Anton

M. Moeliono et al. 1988). I have not found further cognates in either Mon-Khmer or Malayo-Polynesian languages. They appear to reflect relatively late borrowing, and it is difficult to ascertain which of the forms is the precursor, but the paramount position of Mon-speakers at that time suggests that this language was the donor.

Tadmor (2009b) also cites another pair of terms: Old Khmer *raloŋ* 'a strip of rice-field or other land' and Malay *rəluŋ* 'unit of land area, 1.33 acres'. In pre-Angkorian Khmer, *raloŋ* meant 'channel, canal, waterway', while *sre raloŋ* meant 'a rice-field on the channel', with *sre* denoting 'wet rice-field' (Jenner 2009: 399, 549). A Malay cognate *rĕlong* is mentioned in R.J. Wilkinson (1901-1902), but it is not found in Klinkert (1902). It occurs as *rəluŋ* in standard Indonesian Malay where it stands for a unit of land area with two alternative measures, either ca 2,800 m$^2$ or ca 5,300 m$^2$ (W.J.S. Poerwadarminta 1976).[6] I have not found cognates in other languages.

## 4.3 WATER BUFFALO

The ancestral word for 'water buffalo, carabao' in Mon-Khmer languages, *\*grəbʌy*, is apparently a later development of a tentative Proto-Munda *\*Gəruayʔ* 'cattle/draught animal'.[7] The latter form was borrowed into several language phyla and sub-phyla, besides Mon-Khmer also Tibeto-Burman and Daic (Table 18). It is not to be confused with Sanskrit *gavaya* 'a species of ox, *Bos gavæus*', which appears to be a derivation of *go* 'cow, cattle' (Monier Monier-Williams 1899: 351, *gavayá₂*; 363, *gó*).

---

**Austroasiatic**

**Proto-Munda *\*Gəruayʔ* 'buffalo'** (the attestations below are from Z76: 1325)
Geta' *hrweʔ* 'cow'
Remo *gilaj* 'bullock'
Gutob *gula'j* 'bullock' (with *r/*u metathesis)
Kharia *orɛj* 'bullock'

**Mon-Khmer *\*grəbʌy* 'buffalo'** (H76: 467; S00 #580; S06 #763; SJ03 #1093; Sh06 #103)
Pearic *krəbaw*
Khmer *krəbʏy*
Stieng *krʏpɯ:*
Chrau *gəpū*
Köho *rəpu*
Bahnar *kəpō*
Jru' *kpʌw*
Su' *krəpi:*
Kantu *karpĩw*

---

[6] Moeliono et al. (1988: 739) erroneously cite the area in cubic metres.
[7] This reconstruction is mine. Compare *\*ɔreXj* 'draught animal' as reconstructed in Z76: 1319.

    Dakkang *karpu:w*
    Sedang *kopaw*
    Jeh *kapìaw*

**Tibeto-Burman**

    Old Burmese *klway ~ klwai* 'buffalo' (S74 #10)

**Daic**

    Thai *khwāi* 'buffalo' (L77: 242 #1, 287 #7)

Table 18. Words for 'buffalo' in Austroasiatic, Tibeto-Burman, and Daic languages.

Through various Mon-Khmer donor languages, the word was apparently borrowed into Western Malayo-Polynesian languages in several parallel trajectories of borrowing, particularly into Aceh-Chamic and, separately, into Malayo-Javanic.

**Austronesian**

    **Aceh-Chamic *kabau* 'water buffalo'** (T99: 322)
    Phanrang Cham *kapaw*
    Roglai *kabau*
    Jarai *kəbau*
    Rade *kəbau*
    Chru *kəbāu*
    Haroi *kəphiau*
    (see also Acehnese *kɯbuɯə*)

        **Malayo-Javanic *keRbau* 'water buffalo'**
        Malay *kərbaw*
        Teluk-Betung Lampung *kibaw*
        Javanese *kəbo*
        Madurese *kərbhuy*

Table 19. Words for 'water buffalo' in Aceh-Chamic and Malayo-Javanic languages.

The Malayo-Javanic forms regularly reflect *-eR-* (Malay and Madurese -ər-, Lampung-dialects -i-, Javanese -ə-) and therefore indicate an early date for the borrowing. Meanwhile, the Javanese reflex was also borrowed into Sundanese and Balinese (Table 20). The Malay form *kərbaw* was either borrowed in this form into some languages of Sumatra, for example, Toba Batak, Karo Batak, and Minangkabau, or as its vernacular Malay doublet *kərəbaw* which spread throughout the Archipelago where a Malay vernacular served as *lingua franca*. The following table presents a far from complete list. As indicated below, there also was transmission of the word farther north and east into Kavalan and Chamorro. This took place via uncertain carriers, probably Cebuano or other Philippine personnel in Spanish garrisons in North Taiwan from  circa 1629

till 1642 (see Raleigh Ferrell 1969: 19) and Guam since around 1740 (Donald M. Topping 1973: 6-7). The word was also transmitted to Fiji, presumably by Malay personnel of the colonial British.

**Austronesian**

**via Javanese *kəbo***
Sundanese *kəbo*
Balinese *kəbo*

**via Malay *kərbaw***
Toba Batak *horbo*
Karo Batak *kərbo*
Minangkabau *kabaw*

**via vernacular Malay *kərəbaw***
Ma'anyan *karɛβau*
Iban *kərəbau ~ kərəbo*
Murik *kələbaw*
Haria Saparua *kàrbō*
Kambera *karămboa*
Muna *karambau*
Gorontalo *olobu*
Tondano *kərəwou*
West-Bukidnon Manobo *kəravəw*
Maranao *karabao*
Cebuano *kalabaw ~ karabaw*
Tagalog *kalabáw*

**via secondary transmitters of the latter**
Kavalan *kaváu ~ kravau*
Chamorro *karabáo*
Waya Fiji *karavau* ('humpback ox') (ACD under 'water buffalo, carabao')
Bau Fiji (obsolete) *karavau* ('ox, bull or cow')

Table 20. Words for 'water buffalo' in Austronesian languages.

All these forms represent distinct stages of borrowing from vernacular Malay, which in turn reflected a Malayo-Javanic proto-form borrowed from some Mon-Khmer language.

Seen as a whole, the transmission of loanwords originating from a Proto-Munda \**Gəruayʔ* referring to 'cattle' or 'draught animal' underwent a profuse development, spreading over mainland and insular Southeast Asia, even reaching the western Pacific.

## 5. EXHANGES BETWEEN EARLY KHMER, CHAM, AND MALAY KINGDOMS

### 5.1 SILVER, MONEY

Malay-speaking traders had spread the word *salaka* 'silver, money' across the Archipelago since early times (Mahdi 1988: 359, 1996: 142). However, when the Khmer kingdom of Funan rose to paramountcy, it apparently also gained

suzerainty over a number of Malayic communities around the Gulf of Thailand (Mahdi 1994: 186-188). This led to the replacement of the word for 'silver, money' in Old Malay with a borrowing from Old Khmer. The same happened in Old Cham. The Old Khmer neologism had already spread through Mon Khmer languages which were under Funan influence.

The earliest inscriptions with one of the Old Khmer variant forms are dated between 578 and 726 CE (Jenner 2009: 327). However, the earliest use of the term in Old Khmer is probably older, from the time when Funan attained paramountcy in the second or third century CE. The Old Khmer word was borrowed into Old Cham as *pirak* in inscriptions XII-C and XVII at Mỹsơn (L. Finot 1904: 935, 951-952). The latter dates from the twelfth century CE. However, the Aceh-Chamic reflexes show various developments of the vowel location (Table 21). Hence, the written Cham reflex as well as the reflex in Phanrang suggest borrowing from a form like that in Köho. The Old Cham form, apparently retained in Acehnese, seems to have been the result of *r/i* metathesis.

There are no early notations of the word in Old Malay inscriptions, but Old Javanese *pirak* 'silver, money, wealth' already occurs in the Kawi version of the *Bhiṣmaparwa* (Zoetmulder 1982) dated ca 1000 CE (I Gusti Putu Phalgunadi 1995: 1). In Malay, the reflex is *pɛrak* 'silver' (Wilkinson 1901-1902). However, the Old Javanese and Old Cham cognates suggest that the original pronunciation in Old Malay must also have been *pirak*. Indeed, this seems to be confirmed by numerous borrowings transmitted via Malay. This includes instances in which there is an early borrowing with *i* and a later one with *e* ~ *ɛ* as a doublet:

**Austroasiatic**

**Mon-Khmer (H76 #3.3; SJ03 #197; S06 #187)**
Old Khmer *prak ~ prāk ~ prakk ~ prag* 'silver' (Jenner 2009)
Stieng *prāk*
Köho *priaʔ*
Bru *praʔ*
Kuy *prak ~ praʔ*
Souei *pəraʔ*
Sô *pəraʔ*
Ta'Oi *praʔ*
Chatong *praʔ*
Dakkang *prak*
Kantu *praʔ*
Pacoh *praʔ*
Lavi *prak*
Jru' *prak*
Houeikong Laven *prak*
Su' *prak*
Nyaheun *prak*
Laveh *prak*
Brao *prak*

**Austronesian**

**Aceh-Chamic** (T99: 360)
Acehnese *pira?*
written Cham *paryak*
Phanrang Cham *parya?*
Rade *prak* ('silver, money')

**Malayo-Javanic**
Malay *pɛrak*
Old Javanese *pirak*

### via Malay (A92: 86; ACD under 'silver')

**(Southwest)**
Karo Batak *pirak*
Toba Batak *pirak*
Balinese *pirak ~ perak*
Iban *pirak*
Kayan *pirək*
Kiput *pirək*
Kelabit *pirək*
Muna *pera*
Makassarese *pera?*
Mongondow *pera*
Serawai *pira?*
(see also Banjarese *pɛrak*, Hulu Banjarese *pirak*, Minangkabau *pirak*, Rejang *pirok*, Abung Lampung *pirak*, Sundanese *pɛrak*, Kadazan *piok*, Muna *pera*, Tidung *pīlàk*, Tarakan *péràk*)

**(Taiwan)**
Pazeh *pila*
C'uli' Atayal *pila?*
Saisiyat *pa-pila?*
Yami *pila* ('lead')

**(Philippines)**
Tausug *pilak*
Tboli *filak*
Tiruray *filak*
Maranao *pirak*
Cebuano *pílak*
Agutaynen *pilak*
Aklanon *pílak*
Hanunóo *pílak*
Bikol *pírak*
Tagalog *pílak*
Kapampangan *pílak*
Pangasinan *pilák*
Casiguran Dumagat *pílak*
Ifugaw *pilák*

Kankanaey *pilák*
Isneg *piráʔ*
Itbayaten *pilak*

**(Madagascar)**
Malagasy *firakʌ* ('lead')

Table 21. Austroasiatic and Austronesian words for 'silver' (and associated notions).

The *r* > *l* shift in the Taiwanese cognates were apparently acquired from a Philippine language at a time when Spanish military garrisons also included Philippine soldiers (see above, and Ferrell 1969: 19). Note that direct loans from Spanish retained the *r*, for example, Kavalan *broa* 'boat' and *kravau* 'buffalo' (Spanish *proa*, *carabao*).

In the Malayic regions of Southeast Asia, coins were also made of lead, and these remained in circulation until well into the sixteenth century. This explains the semantic shift from 'silver' to 'lead' in Malagasy and Yami.

## 5.2 GOLD, MACE

The Old Khmer word for 'gold' apparently derived from an earlier form meaning 'to shine' through the infixation of *-m-* (Table 22). Inscriptions with the Old Khmer form (and its doublets *māss* and *mas*) are dated between 578 and 777 CE (Jenner 2009: 371). Here, as with the word for 'silver' just discussed, Mon-Khmer cognates are restricted to languages spoken in areas under Funan influence.

A borrowing of this form is reflected in Old Cham *māh* 'gold', as in inscription XVII at Mỹsơn (Finot 1904: 951-952). In modern Aceh-Chamic languages, the borrowing is reflected as shown below. Unlike Old Malay cognates of the word for 'silver', the borrowed word for 'gold' is already found in early inscriptions. The Old Malay *mas* 'gold' occurs in lines 9 and 11 of the Naga inscription at Sabokingking, formerly known as Telaga Batu (Johannes Gijsbertus de Casparis 1956: 33), dated ca 840 CE.

Basic Malay words are typically di- or trisyllabic, and monosyllabic loanwords tend to be pronounced as disyllables. In the example of the word for 'gold', borrowed *mas* is pronounced either with syllabic nasal, *ṃas* (or *mmas*), or with pre-posed schwa, *əmas*. Therefore, Wilkinson (1901-1902: 37 and 646) notes the disyllabic doublet, spelled *ĕmas*, as main form and monosyllabic *mas* as secondary variant.

Subsequent loans into languages of the Archipelago from Malay are not as widely distributed as the word for 'silver'. Apart from the fact that gold was apparently not as widely used as money, there were pre-existing cognate sets for 'gold' in the Philippines – Sulawesi – Maluku area (Mahdi 1994: 182). Nonetheless, there are numerous loanwords derived from Malay *əmas* ~ *mas*:

**Austroasiatic**

---

**Early Mon-Khmer** ***yās* 'to shine'** (S06 #1873)
Old Mon *yās*
Mon *yèḥ*

**Mon-Khmer** ***y-m-ās* 'gold'** (Sh06 #1873)
Old Mon *yimās* ('shining [gold]')
Old Khmer *mās ~ māss ~ mas*
Khmer *mīəh*
Stieng *mā͡ih*
Sre *mah*
Biat *mā͡ih*
Bahnar *mayh*

**Austronesian**

---

**Aceh-Chamic 'gold'** (T99: 347)
Acehnese *muh ~ muih*
Rade *mah*
Jarai *mah*
Chru *mīh*
North Roglai *māh*
Haroi *mah*
Western Cham *mɨh*
Phanrang Cham *mɯ̆h*

**Malayo-Javanic 'gold'**
Old Malay *mas*

**via Malay** (A92: 56; ACD under 'gold')

**(Sumatra)**
Toba Batak *omas*
Serawai *əmas ~ mas*
Minangkabau *ameh*
(see also Karo Batak *mas*, Rejang *əmos*, Krui Lampung *amas ~ mas*)

**(Java, Bali, Lombok)**
Sundanese *əmas*
Old Javanese *əmas*
Javanese *mas*
Balinese *əmas*
Sasak *əmas*

**(Borneo)**
Iban *mas*
Bintulu *mas*
Melanau *mas*
Kenyah *mat*
Ngaju *amas*
Embaloh *amas*
Hulu Banjarese *amas*
(see also Tidung *àmàs*, Kadazan *amas*)

**(Sulawesi)**
Makassarese *ámmasaʔ*

**(East)**
Komodo *mas*
Manggarai *əmas*
Rembong *əmas*
Wetan *mas*
Selaru *mas*
Kai *mas*
Dobel *əmas*

Table 22. Reflexes of early Mon-Khmer *\*yās* 'to shine' and *\*y<m>ās* 'gold' in Austroasiatic and Austronesian languages.

As noted by Tadmor (2009b), in addition to the word for 'gold', Old Khmer exhibits a homonym which had been borrowed from Sanskrit *māṣa* 'a bean; a weight used for gold'. It is noted as English *mace* by Henry Yule and A.C. Burnell (1986, first published 1886: 530, *mace*ᵦ), glossed amongst other meanings, as 'a weight used in Sumatra, one-sixteenth of a Malay tael'. This is also the meaning given for the Old Khmer form by Jenner (2009), perhaps confusing it with the Old Malay form. Indeed, the Old Khmer word is glossed by Pou (2004: 373, *mās*ᵦ) as 'a measure of weight (for milk, honey, oil, ... or seeds)'.

I have not found *mas* in the meaning of a weight unit in Old Malay inscriptions. It is also not noted in Wilkinson (1901-1902). However, it is still noted in Klinkert (1902: 50) as a gold weight equal to '$1/16$ thaïl', while Bowrey (1701, *MA*) already glossed it among other meanings as 'the name of a gold weight sixteen whereof is accounted one Tial [sic]'. It is also cited as 'name of a weight in Sumatra equal to nearly 40 grains, being the 1-16th of a tail' by John Crawfurd (1852: 97-98).

Meanwhile, Zoetmulder (1982) glosses Old Javanese *mās ~ mas ~ əmas ~ həmas* among other meanings as 'a measure of gold, equivalent to 400 smaller units'. This no longer seems to be the meaning in modern Javanese. Meanwhile, the meaning of 'a unit of measure for gold' is reported for Bikol *amás*, Tae' *ammaʔ*, and Buginese *emmeʔ* (ACD under 'gold'), as well as for Iban, which has *emas ~ mas*. Just as with the word meaning 'gold', this is further evidence of an early transmission of the word via Malay.

5.3 Foxtail millet

Words for foxtail millet (*Setaria italica*) had been taken up in Austronesian languages since very early times, as demonstrated by the high-order proto-forms *\*beCeŋ* and *\*jawa*, of which the latter was borrowed from Prakrit. However, there is also a term which must have been borrowed from Mon-Khmer at a relatively later time. In case of Austroasiatic, the original base word has been reconstructed for Munda as *\*(h)oi* 'foxtail millet' (Z76: 1310). The word was then apparently adopted into Old Khmer during the period

of Funan, from where it spread to related languages spoken in the territory of Funan. It was also borrowed from Old Khmer into Chamic and Malayic languages.

**Austroasiatic**

**Munda *(h)oi** (Z76: 1310)
Sora *bur-oy*
Remo *wi-dar*
Gtaʔ *ū-hwe*
Mundari *oe*

**Mon-Khmer *s-kuəy** (Sh06 #1447; S15 #332)
Khmer *s-kūəy* ('Job's tears')
Riang-Lang $kay_1$-$k^huay_1$
Danaw $k^hwé_4$ $k^hraʔ_1$
Kammu-Yuan *həŋ-kɔ̄y*
Bumang $kɔi_{A1}$ ('millet')
Lameet *kāy* ('yam-like plant')

**Austronesian**

**Chamic *hakuəy**
Cham *hakɔ̄y*
Jarai *həkūay* ('millet')

**Malayic *səkuy**
Malay *səkoi*
Minangkabau *sakuay ~ sikuay*
Rejang *səkoi* ('millet')

Table 23. Words for 'foxtail millet' in Mon-Khmer and Malayo-Chamic languages.

A wider spread of this borrowing from Old Khmer was apparently prevented by the presence of words for millet borrowed earlier.

5.4 Obeisance, worship

There is a cognate set of words denoting a traditional gesture of obeisance: Old Khmer *saṃvaḥ*, Cham *sambah*, and Malay *səmbah*. As Tadmor (2009b) demonstrates, the donor must have been Old Khmer, because the word was apparently derived from a monosyllabic base word: Old Khmer *vaḥ* means 'to meet, come upon, touch', whereas *saṃ-vaḥ* is 'pay obeisance (by putting the palms of the hand together)' (Pou 2004: 436, 487).[8] For Pre-Angkorian Khmer, Jenner (2009: 512) only notes the disyllabic derivation. For Chamic, I have not found reflexes of the proto-form in various dialects, but only in the written form *saṃbaḥ* 'worship, tribute, greet'.

---

[8] See, for example, inscription K. 41 at Vat Prei Sva (G. Cœdès 1937-1966: VI.32) for the former and inscription K. 245 at Prasat Ta Kam (Cœdès 1937-1966: III.91) for the latter.

In Western Malayo-Polynesian languages, the reflexes were typically borrowed via Malay *səmbah*. Of the Tagalog doublet *sambá ~ simbá*, the latter form was also shared by Cebuano and West Bukidnon Manobo (Table 24). As a secondary development, the early Malay *səmbah* 'pay obeisance' then formed a compound with Malay *yaŋ* 'deity', resulting in *səmbahyaŋ*, a term for 'pray, prayer' (namely, 'pay obeisance to the deity'), which spread throughout the western half of the Archipelago:

**Austronesian**

**Malay *səmbah* 'pay obeisance (by putting the palms together)'**
Moken *səmah ~ ɲəmah*
Acehnese *sumah ~ sumbah*
Toba Batak *somba*
Gayo *səmbah*
Minangkabau *sambah*
Rejang *səmbəaʔ*
Sundanese *səmbah*
Javanese *səmbah*
Madurese *səmba*
Balinese *səmbah*
Banjarese *sambah*
Tagalog *sambá ~ simbá* ('pay obeisance, worship, go to church')
Cebuano *simba*
West Bukidnon Manobo *simba* ('worship, go to church')

**Malay *səmbah* + *yaŋ* > *səmbahyaŋ* 'pray, prayer'**

Minangkabau *sumbayaŋ ~ sambayaŋ*
Iban *səmbiaŋ*
Serawai *səmba(h)yaŋ*
Acehnese *sumayaŋ*
Toba Batak *sombayaŋ*
Gayo *səmiaŋ*
Rejang *səmiyaŋ*
Sundanese *səmbahiaŋ ~ səmbahyaŋ*
Javanese *səmbahyaŋ ~ səmbayaŋ*
Banjarese *sambahyaŋ*
Makassarese *sambayaŋ*
Buginese *səmpayaŋ*
Muna *sambahea*

Table 24. Reflexes of Malay *səmbah* and *səmbahyaŋ* in other Western Malayo-Polynesian languages.

The Malay precursors of the reflexes of *səmbah* and *səmbahyaŋ* examined above involve borrowings from Old Khmer *saṃvaḥ* rather than a source representing an earlier stage of Mon-Khmer language development. Therefore, the transmission over the Archipelago could only have proceeded relatively late.

## 5.5 Cannon

Malay *məriam* 'cannon', another historical loanword from Khmer, dates from an even later period, namely: the turn of Late-Medieval to Early-Modern Khmer. Etymologically, the original expression was *kāmphlɤːŋ meriəm* 'cannon' (literally 'firearm + large'), consisting of the prefixed *kām-* + *phlɤːŋ* 'fire' and *me-* + *riəm* 'eminent, senior' (Gustav Schlegel 1901; Joseph Guesdon 1930: 166, 1380). Khmer *meriəm* also occurs alone in the meaning 'cannon' (Guesdon 1930: 1380).

Relatively archaic Malayo-Javanic cognates show that the vowel in the last syllable was originally a schwa, and the last-syllable *a* observed in other cognates apparently results from later influence from Malay. Note that in almost all Malayic dialects and languages, the last-syllable schwa of a word automatically changes to *a*. Exceptions are vernacular Jakartan and Java Bazar Malay, in which a schwa in the last syllable is retained. As for more recent Malay influence in this respect, note, for example, the earlier cognate with a last-syllable schwa in the examples below (Table 25).

The earliest use of cannons in the Malay Archipelago dates from long after the shift of power centres from Sumatra to Java, even after the recession of Central Javanese empires such as Majapahit, and the advancement of coastal mercantile polities. Indeed, apart from cannons in Aceh and Malacca, the earliest reports of cannon in Island Southeast Asia come from Banten and Demak on the west and north coasts of Java. This explains the observed retention of a schwa in the final syllable, which would indeed be expected if the donor language was Javanese.

---

**Austronesian**

Acehnese *muriam ~ muruɯyam*
Toba Batak *mariam*
Rejang *məriam*
Iban *meriam*
Banjarese *mariam*
Makassarese *mariaŋ*
Buginese *mariaŋ*

**Original form with /ə/**
Balinese *mariyəm* (R. van Eck 1876)
Madurese *mariəm ~ məreəm* (P. Penninga and H. Hendriks 1936) ~ *mariyəm* (Asis Safioedin 1977)
Sundanese *mariəm* (Sugiarto et al. 1999)
Javanese *mriəm* (Sugiarto et al. 1999), *mriyəm* (Elinor Clark Horne 1974)

**Later form with /a/**
Balinese *mriam* (Sugiarto 1999), *mariyam ~ mriyam* (Charles Clyde Barber 1979)
Madurese *məriam* (Sugiarto et al. 1999)

---

Table 25. Words for 'cannon' in Austronesian languages.

Sundanese, Javanese, Madurese, and Balinese therefore feature a last-syllable schwa, confirming that the earliest introduction of cannons from Cambodia was into coastal polities in the island of Java.

## 5.6 Dirty

Finally, there is an example of borrowing of a word which is widely represented in Austroasiatic, but within Austronesian occurs only in Malayic languages.

---

**Austroasiatic**

---

**Munda** *\*[k]umu* **'dirty'** (P59 #114)
Mundari *humu*
Birhor *humu*
Ho *homu*
Kurku *kumu*

**Mon-Khmer** *\*kmuuʔ, \*kmuəʔ, \*kməʔ* (Sh06 #140; SJ03 #722; SB06 #116)
Old Khmer *kanmau* 'black, dark'
Khmer *khmau* 'black'
Palaung *kəmu* 'grease, dirty'
Bahnar *kəmɔʔ* 'dirty'
Halang *ʔmɔʔ* 'dirty'
Jru' *kmoʔ* 'dirt, filth'

   **Aslian** *\*kəməh* **'dirty'**
   Jehai *kāmah*
   Kintaq Bong *kamah*
   Sabum *kamah*

---

**Austronesian**

---

**Proto Malayic** *\*kamah/\*kumuh* **'dirty'** (A92 #10)
Minangkabau *kumuah ~ kuma*
Indonesian Malay *kumuh* (Zain 1957)
Serawai *kama(h) ~ kumuᵉ(h)*
Iban *kamah*

---

Table 26. Related words for 'dirty' in Austroasiatic and Malayic languages.

Malay attestations are strikingly missing from both Wilkinson (1901-1902) and Klinkert (1902). Not only is Sutan Mohammad Zain (1957) the earliest Indonesian Malay dictionary I found which cites *kumuh* 'dirty', it also notes it as a borrowing from Minangkabau. Indeed, both *kumuah* and the synonymous *kuma* are already cited by J.L. van der Toorn (1891). The word was probably borrowed by early Minangkabau migrants to Negeri Sembilan on the Malayan Peninsula, where they had been in contact with local Aslian communities since the fourteenth century CE (J.T. Newbold 1835: 242-243; J.M. Gullick 2003: 3-4).

## 6. Epilogue

The previous pages testify to a long history of lexical borrowing from Austroasiatic languages into Malayo-Polynesian. In the earliest period,

ancestors of the Aceh-Chamic and Malayic peoples who arrived in Indochina and the insular world directly to the south acquired names of specific local fauna from local Mon-Khmer languages.

With time, seafaring led to a dispersal of lexical items throughout the Archipelago by Malay seafarers. This sometimes led to the borrowing of parallel cognate sets following different phonological rules. For example, the word for 'silver' was originally transmitted by Malays as *pirak*, and, later, likewise by Malays as *pɛrak*. The word for 'cannon' discussed in the close to last sub-section even show Balinese and Madurese borrowings which differ phonetically from earlier borrowings of the same word in these languages.

In all this, we have two main particularities: one is the occurrence of distinct cognate sets in Austroasiatic languages which generated a set of borrowings in the territory under paramountcy of the Old Khmer kingdom of Funan. The other is the widespread transmission through Insular Southeast Asia via Malay. As the data in this article make clear, this latter process is not always characterized by uniform phonetic regularity.

REFERENCES

Adelaar, K. Alexander. 1992. *Proto Malayic; The reconstruction of its phonology and parts of its lexicon and morphology*. Canberra: Research School of Pacific Studies, Australian National University. [Pacific Linguistics Series C–119.]

Antonissen, A. 1958. *Kadazan–English and English–Kadazan dictionary*. Canberra: Government Printing Office.

Aymonier, Étienne. 1891. "Première étude sur les inscriptions tchames", *Journal asiatique*, (Janvier–Février): 5-86.

Aymonier, Étienne and Antoine Cabaton. 1906. *Dictionaire Čam–Français*. Paris: Ernest Leroux. [Publications de l'École Française d'Extrême-Orient vol. VII.]

Barber, C[harles] Clyde. 1979. *A Balinese–English dictionary*. Aberdeen: University of Aberdeen.

Benedict, Paul K. 1972. *Sino-Tibetan; A conspectus*. London: Cambridge University Press.

Benjamin, Geoffrey. 1976. "Austroasiatic subgroupings and prehistory in the Malay Peninsula", in: Philip N. Jenner, Laurence C. Thompson, and Stanley Starosta (eds), *Austroasiatic studies*, pp. 37-128. Honolulu, HI: University Press of Hawai'i.

Berg, René van den, in collab. La Ode Sidu. 1996. *Muna–English dictionary*. Leiden: KITLV Press.

Bhattacharya, S. 1966. "Some Munda etymologies", in: N.H. Zide (ed.), *Studies in comparative Austroasiatic linguistics*, pp. 28-40. The Hague: Mouton.

Blust, Robert A. 1974. "A Murik vocabulary with a note on the linguistic position of Murik", *Sarawak Museum Journal* 22: 153-189.

Blust, Robert, Stephen Trussel, and Alexander D. Smith. 2023. CLDF dataset derived from Blust's "Austronesian Comparative Dictionary" (v1.2) [Data set; https://zenodo.org/records/ 7737548.]

Bowrey, Thomas. 1701. *A dictionary English and Malayo, Malayo and English*. London: printed by Sam. [Bridge for the Author – no page numbers.]

Capell, A. 1968. *A New Fijian dictionary*. Suva: Government of Fiji.

Casparis, Johannes Gijsbertus de. 1956. *Prasasti Indonesia II; Selected inscriptions from the 7th to the 9th century A.D.* Bandung: Masa Baru.

Cense, A.A. 1979. *Makassaars–Nederlands woordenboek.* 's-Gravenhage: Martinus Nijhoff.

Coedès, G. 1937-1966. *Inscriptions du Cambodge*, vols. I-VII. Paris: E. de Boccard/ École française d'Extrême-Orient. [Collection de textes et documents sur l'Indochine 3.]

Crawfurd, John. 1852. *A grammar and dictionary of the Malay language, with a preliminary dissertation*, vol II. *Malay and English, and English and Malay dictionaries*. London: Smith, Elder, and Co.

Daud, Bukhari and Mark Durie. 1999. *Kamus basa Aceh / Kamus bahasa Aceh/ Acehnese–Indonesian–English Thesaurus*. Canberra: Research School of Pacific and Asian Studies. [Pacific Linguistics C-151.]

Driem, George van. 1999. "Four Austric theories", *Mother Tongue* 5: 23-27.

Eck, R. van. 1876. *Eerste proeve van een Balineesch–Hollandsch woordenboek.* Utrecht: Kemink & Zoon.

Elkins, Richard E. 1968. *Manobo–English dictionary*. [Honolulu, HI]: University of Hawai'i Press. [Oceanic Linguistics Special Publication No. 3.]

Eringa, F.S. 1984. *Soendaas–Nederlands woordenboek*. Dordrecht / Cinnaminson, NJ: Foris Publications.

Ferrell, Raleigh. 1969. *Taiwan aboriginal groups; Problems in culture and linguistic classification*. Taipei: Academia Sinica. [Institute of Ethnology Academia Sinica monograph 17.]

Finot, L. 1904. "Notes d'épigraphie XI. Les inscriptions de Mi-sơn", *Bulletin de l'École française d'Extrême-Orient* 4: 897-977.

Genderen Stort, P. van. 1916. *Nederlandsch–Tidoengsch–Tinggálàn Dájàksche woordenlijst*. Batavia: Albrecht & Co./ 's Hage: M. Nijhoff. [Verhandelingen van het Bataviaasch Genootschap van Kunsten en Wetenschappen LXI, 5.]

Giles, Herbert A. 1912. *A Chinese–English dictionary*. Second edition, revised, and enlarged. Shanghai / Hongkong / Singapore / Yokohama: Kelly & Walsh; London: Bernard Quaritch.

Gossens, A.L. 1924. "A grammar and vocabulary of the Dusun Language", *Journal of the Malayan Branch of the Royal Asiatic Society* 2(2): 87-220.

Groeneveldt, W.P. 1877. *Notes on the Archipelago and Malacca, compiled from Chinese sources*. 's Hage: Martinus Nijhoff. [Verhandelingen van het Koninklijk Bataviaasch Genootschap van Kunsten en Wetenschappen 39.]

Guesdon, Joseph. 1930. *Dictionnaire cambodgien-français*. Paris: Les Petits-Fils de Plon et Nourrit.

Gullick, J.M. 2003. *A history of Negri Sembilan*. [Monographs of the Malaysian Branch of the Royal Asiatic Society, Issue 33. S.l.]

Hamidy, Badrul Munir, Azis Wahab, M. Yusuf Yacub, Syafruddin, and Zainal Hakim. 1985. *Kamus Rejang–Indonesia*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa.

Hapip, Abdul Djebar. 1977. *Kamus Banjar–Indonesia*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa.

Harahap, A. Bazar. 2007. *Kamus moderen bahasa Angkola Mandailing*. Jakarta: Forum Komunikasi Masyarakat Tapanuli Selatan dan Mandailing Natal.

Haudricourt, André-G[eorges]. 1966. "The limits and connections of Austroasiatic in the Northeast", in: Norman H. Zide (ed.), *Studies in comparative Austroasiatic linguistics* (*Indo-Iranian monographs* 5), pp. 44–56. The Hague / London / Paris: Mouton.

Hazeu, G.A.J. 1907. *Gajōsch–Nederlandsch woordenboek*. Batavia: Landsdrukkerij.

Headley Jr., R.K. 1976. "Some sources of Chamic vocabulary", in: P.N. Jenner, L.C. Thompson, and S. Starosta (eds), *Austroasiatic studies*, part I, pp. 453-476. Honolulu, HI: University Press of Hawai'i. [Oceanic Linguistics Special Publications No. 13.]

Helfrich, O.L. 1891. *Proeve van een Lampongsch–Hollandsche woordenlijst, bepaaldelijk voor het dialekt van Kroë*. Batavia: Albrecht & Rusche. [Verhandelingen van het Koninklijk Bataviaasch Genootschap van Kunsten en Wetenschappen 45/3.]

Horne, Elinor Clark. 1974. *Javanese–English dictionary*. New Haven, CT / London: Yale University Press.

Hudak, Thomas John. 2008. *William J. Gedney's Comparative Tai Source Book*. Honolulu, HI: University of Hawai'i Press. [Oceanic Linguistics Special Publication 34.]

Hudson, Alfred B. 1967. *The Barito isolects of Borneo; A classification based on comparative reconstruction and lexicostatistics*. Ithaca, NY: Cornell University. [Southeast Asia Program Data Paper 68.]

Jákl, Jiři. 2022. "The *Kalaṅkyaṅ* in Old Javanese literature; From bird of prey to mirror image of the poet", *Bijdragen tot de Taal-, Land- en Volkenkunde* 178: 252-271.

Jenner, Philip N. 2009. *A dictionary of pre-Angkorian Khmer*. Canberra: Research School of Pacific and Asian Studies. [Pacific Linguistics 597.]

Jones, Russel (gen. ed.). 2007. *Loan-words in Indonesian and Malay*. Compiled by the Indonesian Etymological Project. Leiden: KITLV Press.

Karlgren, Bernhard. 1940. *Grammata Serica, script and phonetics of Chinese and Sino-Japanese*. Stockholm: Museum of Far Eastern Antiquities. [Bulletin 12.]

Kern, Martin. 2010. "Early Chinese literature, beginnings through Western Han", in: Kang-i Sun Chang and Stephen Owen (eds), *The Cambridge history of Chinese literature*, vol. I *to 1375*, pp. 1–115. Cambridge: Cambridge University Press.

Kersten, J. 1984. *Bahasa Bali*. Ende: Nusa Indah.

Klinkert, H.C. 1902. *Nieuw Maleisch–Nederlandsch woordenboek*. Met Arabisch karakter. 2e verbeterde en vermeerderde druck. Leiden: E.J. Brill.

Kreemer, J. 1931. *Atjèhsch handwoordenboek (Atjèhsch–Nederlandsch)*. Leiden: E.J. Brill.

Kruspe, Nicole. 2010. *A dictionary of Mah Meri as spoken at Bukit Bangkong*. Honolulu: University of Hawai'i Press. [Oceanic Linguistics Special Publication No. 36.]

Larish, Michael David. 1999. *The position of Moken and Moklen within the Austronesian language family*. PhD thesis, University of Hawai'I, Ann Arbor. [University Microfilms International #9925288.]

Li, Fang Kuei. 1977. *A handbook of comparative Tai*. Honolulu, HI: University Press of Hawai'i.

Mahdi, Waruno. 1988. *Morphophonologische Besonderheiten und historische Phonologie des Malagasy*. Berlin / Hamburg: Dietrich Reimer. [Veröffentlichungen des Seminars für Indonesische und Südseesprachen der Universität Hamburg 20.]

Mahdi, Waruno. 1994. "Some Austronesian Maverick Protoforms–I & II", *Oceanic Linguistics* 33: 168-229, 431-490.

Mahdi, Waruno. 1996. "Confirmation of an etymology: Salaka 'Money Silver'", *Oceanic Linguistics* 35: 142.

Mahdi, Waruno. 1998. "Linguistic data on transmission of Southeast Asian cultigens to India and Sri Lanka", in: Roger Blench and Matthew Spriggs (eds), *Archaeology and language II*; *Archaeological data and linguistic hypotheses*, pp. 390-415. London / New York, NY: Routledge. [One world Archaeology 29.]

Mahdi, Waruno. 1999. "The dispersal of Austronesian Boat Forms in the Indian Ocean", in: Roger Blench and Matthew Spriggs (eds), *Archaeology and language III*; *Artefacts, languages and texts*, pp. 144-179. London / New York, NY: Routledge. [One world Archaeology 34.]

Mahdi, Waruno. 2008. Compte rendu (Jones 2007), *Archipel* 76: 318-322.

Mahdi, Waruno. 2009. "Some obscure Austroasiatic borrowings in Indonesian and Old Malay". [Presentation at the *Eleventh International Conference on Austronesian Linguistics*, Aussois, France, 22-26 June.]

Manik, Tindi Radja. 2002. *Kamus Pakpak–Indonesia*. Medan: Bina Media.

Matisoff, James A. 2003. *Handbook of Proto–Tibeto–Burman; System and philosophy of Sino-Tibetan reconstruction*. Berkeley, LA / London: University of California Press. [University of California Publications in Linguistics 135.]

McKaughan, Howard P. and Batua A. Macaraya. 1967. *A Maranao dictionary*. Honolulu, HI: University of Hawai'i Press.

Miksic, John N. 1989. "Urbanization and social change; The case of Sumatra", *Archipel* 37: 3-29.

Moeliono, Anton M. et al. (eds). 1988. *Kamus Besar Bahasa Indonesia*. Jakarta: Balai Pustaka. [BP No. 3658.]

Monier-Williams, Monier. 1899. *A Sanskrit–English dictionary*. Oxford: The Clarendon Press.

Moussay, Gérard. 1995. *Dictionnaire Minangkabau Indonesien Français*. Paris: Editions l'Harmattan. [Cahier d'Archipel 27.]

Newbold, J.T. 1835. "Sketch of the four Minangkabowe States in the interior of the Malayan Peninsula", *Journal of the Asiatic Society of Bengal* 4: 241-252.

Nothofer, Bernd. 1975. *The reconstruction of Proto-Malayo-Javanic*. 's-Gravenhage: Martinus Nijhoff. [Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde 73.]

Paulus, J. (ed.). 1917. *Encyclopædie van Nederlandsch-Indië*. 1e deel, 2e druk. 's-Gravenhage: Martinus Nijhoff / Leiden: E.J. Brill.

Penninga, P. and H. Hendriks. 1936. *Practisch Nederlands–Madurees woordenboek*. ['s-Gravenhage]: Van Dorp.

Phalgunadi, I Gusti Putu. 1995. *The Indonesian Mahabharata: Bhiṣmaparva*. Delhi: Aditya Prakashan *via* Pradeep Kumar Goel.

Pinnow, H.-J. 1959. *Versuch einer historischen Lautlehre der Kharia-Sprache*. Wiesbaden: Otto Harrassowitz.

Pou, Saveros. 2004. *Dictionnaire vieux khmer–français–anglais / An Old Khmer–French–English dictionary*. Paris: L'Harmattan.

Pou, Saveros and Philip N. Jenner. 1974. "Proto-Indonesian and Mon-Khmer", *Asian Perspectives* 17: 112–124.

Poerwadarminta, W.J.S. 1976. *Kamus umum bahasa Indonesia*. Jakarta: Balai Pustaka. [BP No. 1866.]

Pulleyblank, Edwin G. 1991. *Lexicon of reconstructed pronunciation in Early Middle Chinese, Late Middle Chinese, and Early Mandarin*. Vancouver: University of British Columbia Press.

Purnell, Herbert C., Jr. 1970. *Towards a reconstruction of Proto-Miao-Yao*. PhD thesis, Cornell University, Ann Arbor. [University Microfilms International #70-17,096.]

Ramos, Teresita V. 1971. *Tagalog dictionary*. Honolulu, HI: University of Hawai'i Press.

Ratliff, Martha. 2010. *Hmong-Mien language history*. Canberra: The Australian National University. [Pacific Linguistics 613.]

Reid, Lawrence A. 1971. *Philippine minor languages; Word lists and phonology*. Honolulu, HI: University of Hawai'i Press. [Oceanic Linguistics Special Publications No. 8.]

Richards, Anthony. 1981. *An Iban–English dictionary*. Oxford: Clarendon Press.

Safioedin, Asis. 1977. *Kamus bahasa Madura–Indonesia*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa.

Said, M. Ide. 1976. *Kamus bahasa Bugis–Indonesia*. Ujung Pandang: Pusat Pembinaan dan Pengembangan Bahasa.

Schmidt, W[ilhelm]. 1906. "Die Mon–Khmer–Völker, ein Bindeglied zwischen Völkern Zentralasiens und Austronesiens", *Archiv für Anthropologie* 33: 59-106.

Schlegel, Gustav. 1901. "Mèrjam (etymology of the word)", *T'oung pao*, séries 2, 2: 324.

Shafer, Robert. 1974. *Introduction to Sino-Tibetan*. Wiesbaden: Otto Harrassowitz.

Shorto, H.L. 1963. "The 32 *Myos* in the Medieval Mon Kingdom", *Bulletin of the School for Oriental and African Studies* 26: 572-591.

Shorto, H.L. 1971. *A dictionary of the Mon inscriptions from the sixth to the sixteeenth centuries*. London / New York, NY / Toronto: Oxford University Press. [London Oriental Series 24.]

Shorto, H.L. 1975. "Achinese and Mainland Austronesian", *Bulletin of the School of Oriental and African Studies* 38: 81-102.

Shorto, H.L. 2006. *A Mon–Khmer comparative dictionary*. Canberra: Australian National Univerity. [Paul Sidwell (main ed.), Doug Cooper and Christian Bauer (assist. eds), Pacific Linguistics 579.]

Sidwell, Paul. 2000. *Proto South Bahnaric; A reconstruction of a Mon-Khmer language of Indo-China*. Canberra: Australian National University. [Pacific Linguistics 501.]

Sidwell, Paul. 2005. "Acehnese and the Aceh-Chamic language family", in: Anthony Grant and Paul Sidwell (eds), *Chamic and Beyond; Studies in mainland Austronesian languages*, pp. 211-216. Canberra: The Australian National University. [Pacific Linguistics 569.]

Sidwell, Paul. 2006. *The Katuic languages; Classification, reconstruction and comparative lexicon*. Munich: Lincom Europa.

Sidwell, Paul. 2015. *The Palaungic languages; Classification, reconstruction and comparative lexicon*. Munich: Lincom Europa.

Sidwell, Paul and Pascale Jacq. 2003. *A handbook of comparative Bahnaric*, Vol. 1: *West Bahnaric*. Canberra: Australian National University. [Pacific Linguistics 551.]

Skeat, Walter William and Charles Otto Blagden. 1906. *Pagan races of the Malay Peninsula*. London: Macmillan & Co.

Stargardt, Janice. 1983. *Satingpra I: The Environmental and Economic Archaeology of South Thailand*. Oxford: B.A.R. [BAR International Series 158.]

Stokhof, W.A.L. 1982. *Holle Lists; Vocabularies in languages of Indonesia* vol. 3/4. *Central Moluccas: Ambon (II): Buru, Nusa Laut, Saparua*. Canberra: Research School of Pacific Studies. [Pacific Linguistic Series D-50.]

Stokhof, W.A.L. 1983a. *Holle Lists: Vocabularies in languages of Indonesia* vol. 7/1. *North Sulawesi: Gorontalo group and Tontoli*. Canberra: Research School of Pacific Studies. [Pacific Linguistic Series D-61.]

Stokhof, W. A. L. 1983b. *Holle Lists: Vocabularies in languages of Indonesia* vol. 7/2. *North Sulawesi: Philippine languages*. Canberra: Research School of Pacific Studies. [Pacific Linguistic Series D-60.]

Sugiarto, dkk. 1999. *Kamus Indonesia – Daerah: Indonesia, Jawa, Bali, Sunda, Madura*. Edisi kedua, direvisi, dan diedit oleh Widya Kirana. Jakarta: Gramedia Pustaka Utama.

Tadmor, Uri. 2009a. "Loanwords in Indonesian", in: Martin Haspelmath and Uri Tadmor (eds), *Loanwords in the world's languages; A comparative handbook*, pp. 686-716. Berlin / New York, NY: De Gruyter Mouton.

Tadmor, Uri. 2009b. "Lexical borrowing between Mon-Khmer and western Austronesian". [Presentation at the *11th International Conference on Austronesian Linguistics*, Aussois, France, 22-26 June.]

Thurgood, Graham. 1999. *From ancient Cham to modern dialects; Two thousand years of language contact and change*. Honolulu: University of Hawai'i Press. [Oceanic Linguistics Special Publication 28.]

Toorn, J.L. van der. 1891. *Minangkabausch–Maleisch–Nederlandsch woordenboek*. 's-Gravenhage: Martinus Nijhoff.

Topping, Donald M. 1973. *Chamorro reference grammar*. Honolulu: University of Hawai'i Press.

Topping, Donald M., Pedro M. Ogo, and Bernadita C. Dungca. 1975. *Chamorro-English dictionary*. Honolulu, HI: University Press of Hawai'i.

Tuuk, H.N. van der. 1869. "Proeve van een vergelijkende woordenlijst van Lampongsche tongvallen", *Tijdschrift voor Indische Taal-, Land- en Volkenkunde* 17: 569-575.

Warneck, Joh. 1977. *Toba-Batak–Deutsches Wörterbuch*. Den Haag: Martinus Nijhoff.

Wheatley, Paul. 1961. *The golden Khersonese; Studies in the historical geography of the Malay Peninsula before A.D. 1500*. Kuala Lumpur: University of Malaya Press.

Wielenga, Douwe Klaas. 1909. *Schets van een soembaneesche spraakkunst: naar, t dialect van Kambera.* Batavia: Landsdrukkerij.

Wilkinson, R.J. 1901-1902. *A Malay–English dictionary*. Singapore: Kelly and Walsh.

Wolff, John U. 1972. *A dictionary of Cebuano Visayan*. Ithaca, N.Y: Department of Asian Studies, Cornell University. [Southeast Asia Program Data Paper No. 87.]

Woollams, Geoff. 1996. *A grammar of Karo Batak, Sumatra*. Canberra: Research School of Pacific and Asian Studies. [Pacific Linguistics Series C–130.]

Yule, Henry and A.C. Burnell. 1986. *Hobson-Jobson*. Edited by William Crooke. Calcutta / Allahabad / Bombay /Delhi: Rupa & Co. [First published 1886.]

Zain, Sutan Mohammad. 1957. *Kamus moderen bahasa Indonesia*. Djakarta: Penerbit Grafica.

Zide, Arlene R.K. and Norman H. Zide. 1976. "Proto-Munda cultural vocabulary; Evidence for early agriculture", in: P.N. Jenner, L.C. Thompson, and S. Starosta (eds), *Austroasiatic studies*, Part II, pp. 1295-1334. Honolulu, HI: University Press of Hawai'i.

Zoetmulder, P.J., in collaboration with S.O. Robson. 1982. *Old Javanese–English dictionary*. 's-Gravenhage: Martinus Nijhoff.

APPENDIX 1. ABBREVIATIONS USED

| | |
|---|---|
| A92 | K.A. Adelaar (1992) |
| B72 | P.K. Benedict (1972) |
| ACD | R. Blust, S. Trussel, and A.D. Smith (2023) |
| B66 | S. Bhattacharya (1966) |
| B76 | G. Benjamin (1976) |
| G12 | H.A. Giles (1912) |
| H08 | T.J. Hudak (2008) |
| H66 | A-G. Haudricourt (1966) |
| H76 | R.K. Headley Jr. (1976) |
| L77 | F.K. Li (1977) |
| M03 | J.A. Matisoff (2003) |
| N75 | B. Nothofer (1975) |
| P59 | H.-J. Pinnow (1959) |
| P70 | H.C. Purnell Jr. (1970) |
| R10 | M. Ratliff (2010) |
| S00 | P. Sidwell (2000) |
| Sh06 | H.L. Shorto (2006) |
| Sh75 | H.L. Shorto (1975) |
| S06 | P. Sidwell (2006) |
| S15 | P. Sidwell (2015) |
| S74 | R. Shafer (1974) |
| SB06 | W.W. Skeat and C.O. Blagden (1906) |
| SJ3 | P. Sidwell and P. Jacq (2003) |
| T99 | G. Thurgood (1999) |
| Z76 | A.R.K. Zide and N.H. Zide (1976) |

APPENDIX 2. LANGUAGES CITED

| | |
|---|---|
| Acehnese | J. Kreemer (1931); B. Daud and M. Durie (1999) |
| Angkola Batak | A.B. Harahap (2007) |
| Balinese | J. Kersten (1984) |
| Banjarese | A.D. Hapip (1977) |
| Banjarese, Hulu | A.D. Hapip (1977) |
| Buginese | M.I. Said (1976) |
| Cebuano | J.U. Wolff (1972) |
| Cham, written | É. Aymonier and A. Cabaton (1906) |
| Chamorro | D.M. Topping, P.M. Ogo, and B.C. Dungca (1975) |
| Chinese | H.A. Giles (1912) |
| Dairi–Pakpak Batak | T.R. Manik (2002) |

| | |
|---|---|
| Early Middle Chinese | E.G. Pulleyblank (1991) |
| Fiji, Bau | A. Capell (1968) |
| Gayo | G.A.J. Hazeu (1907) |
| Gorontalo | W.A.L. Stokhof (1983a) |
| Iban | A. Richards (1981) |
| Javanese | E.C. Horne (1974) |
| Kadazan | A. Antonissen (1958) |
| Kambera | D.K. Wielenga (1909) |
| Karo Batak | G. Woollams (1996) |
| Kavalan | R. Ferrell (1969) |
| Lampung, Abung | O.L. Helfrich (1891) |
| Lampung, Krui | O.L. Helfrich (1891) |
| Lampung, Teluk-Betung | H.N. van der Tuuk (1869) |
| Ma'anyan | A.B. Hudson (1967) |
| Madurese | P. Penninga and H. Hendriks (1936); A. Safioedin (1977) |
| Mah-Meri, Bukit-Bangkong | N. Kruspe (2010) |
| Makassarese | A.A. Cense (1979) |
| Malay | R.J. Wilkinson (1901–1902) |
| Manobo, West-Bukidnon | L.A. Reid (1971); R.E. Elkins (1968) |
| Maranao | H.P. McKaughan and B.A. Macaraya (1967) |
| Minangkabau | G. Moussay (1995) |
| Moken | M.D. Larish (1999) |
| Muna | R. van den Berg (1996) |
| Murik | R. Blust (1974) |
| Old Chinese | B. Karlgren (1940) |
| Old Javanese | P.J. Zoetmulder (1982) |
| Old Khmer | S. Pou (2004); P.N. Jenner (2009) |
| Old Mon | H.L. Shorto (1971) |
| Rejang | Hamidy et al. (1985) |
| Saparua, Haria | W.A.L. Stokhof (1982) |
| Sundanese | F.S. Eringa (1984) |
| Tagalog | T.V. Ramos (1971) |
| Tarakan | P. van Genderen Stort (1916) |
| Tidung | P. van Genderen Stort (1916) |
| Toba Batak | J. Warneck (1977) |
| Tondano | W.A.L. Stokhof (1983b) |