

# AUSTROASIATIC SUBGROUPINGS AND PREHISTORY IN THE MALAY PENINSULA<sup>1</sup>

Geoffrey Benjamin  
University of Singapore

## *Introduction*

During the course of several ethnographic field trips to Orang Asli (Aboriginal) groups in West Malaysia since 1964 I have collected basic vocabularies of 17 languages and dialects belonging to the Austroasiatic phylum. Initially the collecting was done for no better reason than to satisfy my curiosity, as my formal research brief was to work intensely on the basic ethnography and linguistics of just one Orang Asli group, the Temiar. (Useful general introductions to the Orang Asli will be found in Williams-Munt 1952 and Lebar et al. 1964: 176-186; 262-3.) But as the work of analysis proceeded and it became necessary to make additional summary investigations of non-Temiar groups also, I decided to gather as many lexical samples (based on the Swadesh list) of the Malayan Austroasiatic languages as opportunity allowed. Recently, with the help of colleagues, three additional vocabularies have been made available from groups I have not personally succeeded in contacting. The resulting list of twenty languages and dialects provides sufficient coverage of the overall pattern of divergence of the Malayan Austroasiatic languages to warrant the present attempt to work out their mutual relationships. In the absence of richer data, it is hoped that the resulting classification will

help to sketch out the most fruitful directions for future research on these languages.

### *Earlier Studies*

The languages of the Malayan aborigines have attracted scholarly attention since the early nineteenth century, and overall internal classifications have been proposed by several workers, most of whom used secondary materials gathered in the field by untrained collectors.

Although earlier writers had suggested that some of the languages of the Malayan aborigines were related to the Mon-Khmer languages, the first comprehensive study was P.W. Schmidt's *Die Sprachen der Sakei und Semang auf Malacca und ihr Verhältniss zu den Mon-Khmer Sprachen* which appeared in 1901. Basing his classification on distinctive vocabulary tests, Schmidt distinguished two major subgroups in the Malayan Austroasiatic languages: a northern 'Semang' group and a southern 'Sakei' group, the latter further divided into two subgroups, thus (square brackets enclose the names employed later in this paper):

Semang	[Northern Aslian]
Sakei I	[Temiar, Lanoh]
Sakei II	[Semai, Mah Meri]

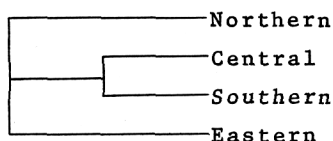
Schmidt excluded the 'Jakun' languages of the southern part of the Peninsula from his classification as not being Austroasiatic in affiliation.

In 1906 C. O. Blagden refined the classification he had proposed some decades earlier. Provided with more materials than Schmidt he produced the classification which with only minor modifications has served most scholars ever since. Again, a basic distinction was drawn between 'Semang' and 'Sakai' groups, but

finer distinctions were recognised in each of these subgroups than Schmidt had seen fit to draw. Blagden, however, was employing phonological criteria in addition to his very rich lexical material. His classification (Blagden 1906: 385f.) was as follows.

- A. Semang
  - 1) Semang-Pangan [Northern Aslian]
  - 2) Low Country Semang (extinct)
- B. Sakai
  - 1) Northern Sakai [Temiar, Lanoh]
  - 2) Central Sakai [Semail]
  - 3) Southern Sakai
    - a) South Western Sakai [Mah Meri]
    - b) South Eastern Sakai [Semelai, Temoq]
  - 4) Eastern Sakai
    - a) Inner Group [Jah Hut]
    - b) Outer Group [Semaq Beri]

Blagden further implies (p. 396) the following hierarchy of relationships between the various branches of Sakai:



What might well have become the next major advance unfortunately remained unpublished. I refer to the work of R.J. Wilkinson, who around 1909 arranged the collection of about sixty 'type vocabularies (of over two hundred words each) filled up by workers in all parts of the Peninsula' (Wilkinson 1910:7). Nothing now remains of these lists<sup>2</sup>, but a summary of the results was published. Wilkinson writes (1910:8)

The results of the enquiry have been the reverse of sensational. They do not bear out Mr. Blagden's theory that there are

(at least) three distinct linguistic groups in the Peninsula each with its own dialects and sub-dialects. The vocabularies tend to show that there are five dialects spoken and that these dialects have so many words in common that they may be regarded as belonging to one single language-group. Corresponding to these five main differences in speech there are important differences in race and culture, so that it may be taken for granted that there are five distinct tribes or races of aborigines; the Semang, the Northern Sakai [Temiar], the Central Sakai [Semai], the Besis [Mah Meril], and the Jakun [not Austroasiatic speakers].

In 1915 Wilkinson added the information (p. iv) that he had discovered a sixth 'Mon-Khmer dialect' in the Peninsula. This was spoken in 'the great mountain pass of Gunong Benom in Pahang' from where he had hitherto been unable to collect any word-lists. Though he gives no samples of this language, it represents in all probability the first report of the language that is nowadays referred to as Che' Wong.

With regard to the interrelationships of these languages Wilkinson goes no further than to remark

A word of caution has also to be uttered against the imprudence of relying too much upon word-lists. Word-lists of Central Sakai and of Besis (Southern Sakai) show very close affinities, while the grammars of the two languages are poles apart. Of the grammars of Semang, Benom and Jakun nothing whatever is known, and of Northern Sakai very little. We want more data.

Until recently the latest classification of the Malayan aboriginal languages to be based on new field data was that of Father Schebesta, the well-known



ethnographer of the Malayan Negritos. Though many of his guesses about the relationships of the more southerly languages can now be seen to be wrong, and although he did not in principle move beyond Blagden's classification, Schebesta did make two notable advances: he was the first to make a comprehensive report and use of the proper ethnic labels, and he greatly refined the classification of the Semang languages.

Schebesta's schema for the tribal names of Malaya is, except for minor changes, the one upon which standard administrative and scholarly usage is now based. (The original source of this schema in English is Schebesta 1926.) The final version of Schebesta's classification of the Semang languages (1952:85f.) reads as follows, with each major subgroup labelled with its distinctive term for 'fellow human being'.

- |                         |                                  |
|-------------------------|----------------------------------|
| 1) North Semang (Meni') | a) Tonga/Mos (S. Thailand)       |
|                         | b) Kensiu [Kensiu]               |
|                         | c) Kenta' (i) K. Nakil           |
|                         | (ii) K. Bogn [Kintaq Bong]       |
| 2) East Semang (Menra') | a) Jahay [Jehai]                 |
|                         | b) Menri' [Mendriq; Mintil]      |
| 3) South Semang (Batek) | a) Batek Nogn [Bateg Nong]       |
|                         | b) Batek Hapen                   |
|                         | c) Batek Kleb [properly, 'Kled'] |
|                         | d) Temo' [Temoq]                 |

Additionally, Schebesta distinguished a fourth division of Semang, West Semang (Sema') consisting of Sabub'n-Lanoh [Sabum; Lanoh]; but this he

known to belong with Temiar in Northern Sakai, all other authors have done.

Unfortunately, no lexical material of any value is available from the S. Thailand Semang, though recent information (Brandt 1961) suggests that they will survive as a group. The Temo' at the opposite end of Schebesta's list are quite obviously misplaced; I knew of them by hearsay only, and they belong linguistically and ethnologically with the other non-Semang groups that surround them in the south. The rest of Schebesta's classification is not worth discussing further, except to note his guess that 'the dialects styled by Blagden Southern and Eastern Sakai will...turn out to be Jakud'n dialects which approximate, through admixture, either to the Semang Men'ra' group or to one or both of the Sakai groups (Northern and Central).' In this assessment he was misled, as we shall see, by the high rate of borrowing between some of these languages.

A widely available classification is the one given by Williams-Hunt (1952:23). But this is no more than a variant of Blagden's scheme, made more complicated by the admission of non-linguistic criteria; it requires no further discussion here.

Finally, there are two classifications that linguists working outside the immediate area are likely to be using, Pinnow's (1959:4f.) and Voegelin's (1966). Neither of these classifications marks any advance, as they are clearly direct rehashes of Schebesta's and Blagden's respectively. Pinnow, however, was the first to propose an overall label for the Malayan Austroasiatic languages, which he calls the Malacca languages. Perhaps this is the

place to point out firmly that although older German sources refer to the whole of the Peninsula as 'Malacca', this term actually refers only to one of the smallest constituent states of the Malaysian Federation; a state, furthermore, in which only one of the languages (Mah Meri) covered by the term is spoken. To obviate any confusion I propose therefore that Diffloth's suggested term 'Aslian' be used henceforth for the Austroasiatic languages of the Malay Peninsula, including the immediately related languages of the Negritos of S. Thailand (but *excluding*, of course, the Austronesian languages spoken in the south of the Peninsula by Orang Asli of the so-called Jakun or Aboriginal Malay group).

In summary, all previous classifications have distinguished between a northerly Semang group of languages spoken by Negrito groups and a southerly Sakai group spoken mainly by non-Negrito groups. The Sakai languages, however, have not been consistently classified in relation to each other. My own investigation, based on entirely newly-gathered data, confirms the distinctiveness of the Semang group (my 'Northern Aslian'). But the Sakai group can now be seen to break down into two quite separate groups (my 'Central Aslian' and 'Southern Aslian'), each co-ordinate with the Semang group. In terms of Schmidt's and Blagden's classifications I have split; in terms of Wilkinson's classification I have clumped.

## 2. *The Present Survey*

### 2.1 *Selection of languages for study*

The first aim of this survey is to provide a basis for a rational classification of the various Malayan Orang Asli groups. A tradition has grown up of using an unwieldy classification in which

supposedly genetic, linguistic and cultural criteria are aggregated together as if they varied in a concordant manner. In fact, these three sets of criteria vary quite discordantly, and the only valid method of approach is to plot the pattern of variation of each set of characteristics separately: the present paper is concerned with the linguistic pattern (though an attempt is made later to relate this to the major culture-historical factors in Aslian ethnology). (Cf. Appendix I.)

The first task, then, was to obtain a sample of the language or dialect spoken by each of the Austroasiatic-speaking Orang Asli ethnic groups currently recognised in Malaysian administrative practice. This accounts for the following languages or dialects (cf. Min 1968:47): Kensiu, Kintaq, Jehai, Mendriq, Bateg, Che' Wong, Lanoh, Temiar, Semai, Jah Hut, Mah Meri, Semaq Beri, Semelai, Temoq.

However, closer investigation revealed that the official classification is rather too cavalier with ethnic distinctions that the Orang Asli themselves consider to be significant. In some cases this was corrected by enquiring about tribal names that had already appeared in the literature but which are no longer in current administrative usage, and in other cases by using hitherto unreported names obtained directly from Orang Asli informants. In this way the following categories were arrived at. Kintaq was split into K. Bong and K. Nakil (cf. Evans 1937:Cap. I; I have a sample only of the former dialect). Kelantan Mendriq was separated from the so-called 'Mendriq' of Pahang, which is here renamed Mintil in accordance with Negrito usage. Many dialects of Bateg were recognised, only two of which appear here--Bateg

Nong in Pahang which has been reported before, and Bateg Deq of Kelantan which is a newly-reported name. The Semnam and Sabum dialects were separated out from Lanoh proper (cf. Evans loc. cit.), while the latter was separated into two hitherto unreported sections, Lanoh Yir and Lanoh Jengjeng. The commonly recognised dialectal variability of Semai (cf. Diffloth 1968) was acknowledged by taking samples of two dialects, Semai I and Semai II.

## 2.2 *Present status of languages selected*

The present numbers and conditions of the speakers of the various Aslian dialects are given below; a map of their distribution is given in Fig. 1. The figures are taken from the 1969 Census, and the latest ethnographic and/or fullest bibliographic sources are given in parentheses. The two-letter abbreviations used here for the language names are also given.

*Kensiu* (Kensew)Ks: 98 Negritos, some still nomadic in north Kedah and over the Thai border, but most now settled since 1965 in a Government village near Baling in Kedah. (Schebesta 1954, 1957: Carey 1970.)

*Kintaq Bong* (Kənta? Bəŋ)KB: c.100 Negritos, still largely nomadic, but many semi-settled near Baling in Kedah and near Kelian Intan in Upper Perak. (Schebesta 1954, 1957; Evans 1937.)

*Jehai* (Jəhay)Je: 702 Negritos, living (unusually) in the mountains dividing Perak and Kelantan just south of the Thai border. Still largely nomadic, but many now settling down, especially where they meet up with Temiar. The 'classical' Negritos of Schebesta's writings. Probably contain several

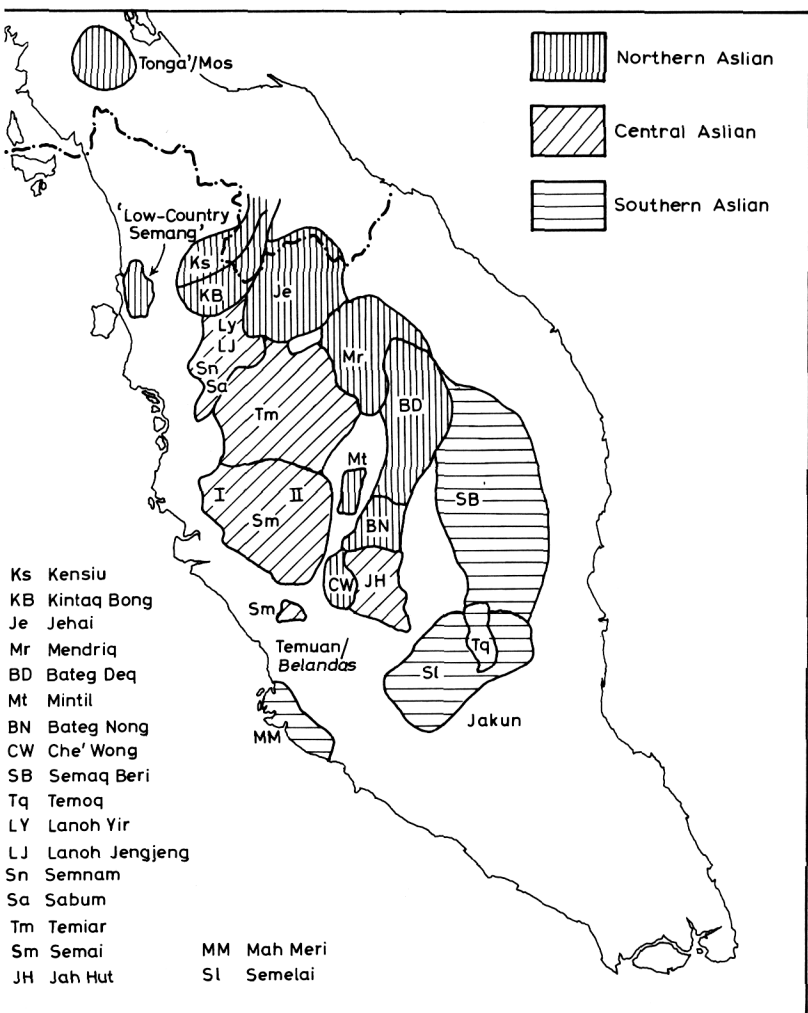


Figure 1. Malay Peninsula: Distribution of Aslian Languages (Approximate)

dialect groups not in contact with each other.  
(Schebesta 1927, 1954, 1957.)

*Mendriq* (Mənra?)Mr: 118 Negritos living in three or four dispersed semi-settled villages along the mid-reaches of the Kelantan River around Kuala Krai and Bertam; many of the so-called 'Bateg' on the lower Lebir River are also actually Mendriq.  
(Schebesta 1954, 1957.)

*Bateg Deq* (Bateg De?)BD: a dispersed group of about 300, nominally 'Negritos' but with a large proportion of non-Negrito elements in their physical and cultural make-up; some semi-settled on the Aring River in south Kelantan, but others wholly nomadic in that area and ranging over into Trengganu and Pahang.  
(Needham 1960; Benjamin MS; Endicott 1969:3f.)

*Mintil* (Mintj|)Mt: a small group of nomadic Negritos, numbering probably no more than 40 persons, ranging along the Tanum and coming out occasionally to Chegar Perah railway halt; no contacts with other Orang Asli groups, and only minimal contacts with Administration, by whom they are regarded as 'Mendriq'.  
(Benjamin MS.)

*Bateg Nong* (Bateg Nɔŋ)BN: a group of about 100 'Negritos' very similar to the Bateg Deq, mostly semi-settled in fairly accessible villages near Jerantut in Pahang. (Evans 1937; Endicott MS.)

*Che Wong* (Ce? Wɔŋ)CW: 272 semi-settled non-Negritos in three or four villages on the southern slopes of Gunong Benom between Raub and Kuala Krau in Pahang; their ethnological classification has always been problematic; Needham claimed that the proper tribal name is Siwang, but no other investigator has been able to confirm this. (Needham 1956.)

*Semnam* (Səmnām)Sn: a 'Lanoh' subgroup; Negritos semi-settled on the Ayer Bal river near Kuala Kenering Upper Perak.

*Sabum* (Sabūm)Sa: a 'Lanoh' subgroup; Negritos long to a nearly extinct group, now living semi-settled with other 'Lanoh' Negritos near Lenggong in per Perak.

*Lanoh Yir* (Lanoh YTr)LY: a Lanoh Negrito group living on the Sarah river, a tributary of the Perak adwaters; probably still mainly nomadic.

*Lanoh Jengjeng* (Lanoh Jəŋjəŋ)LJ: a semi-settled noh Negrito group living in association with Temiar the Ringat river above Grik in Upper Perak.

(N.B. Conventional published enumerations do not divide the Lanoh into subgroups; the total number of noh is put at 264, presumably including some other bgroups not listed here.) (For all Lanoh groups e Evans 1937; Schebesta 1954, 1957).

*Temiar* (Təmēr)Tm: the major Orang Asli group in rth Malaya, consisting of 9,929 swidden farmers tending over 2,000 square miles of jungle in Perak, lantan and northern Pahang; increasing contact with her ethnic groups at the periphery of their territory; their language has become somewhat of a lingua franca among Orang Asli groups, and is one of the two nguages in which special Orang Asli programmes are oadcast by Radio Malaysia. (Benjamin 1968, 1973.)

*Semai* (Səmay; Səmey)Sm: the largest single hno-linguistic group of Orang Asli, consisting of 6,506 swidden farmers, cash-crop farmers and wage-arners, living in many different environments, from ill jungle to urban fringes; one of the two major nguages of Orang Asli radio broadcasting, but



onsisting of several quite variable dialects, only two of which are treated here: *Semai I*, a lowland western dialect sample from Kg Redang Ponggol, near Bidor, Perak (this does not seem to correspond unambiguously with any of the dialects treated by Diffloth); *Semai II*, an eastern highland dialect from Fort Sin on the Betau river, Pahang (this probably corresponds to Diffloth's LIP dialect). (Dentan 1968; Diffloth 1968.)

*Jah Hut* (Jah Hēt)JH: 2,013 swidden farmers concentrated on the eastern slopes of Gunong Benom above Kuala Krau in Pahang. (Evans 1927:37-40; Polunin MS.)

*Mah Meri* (Hma? MərTh)MM: the 'Besisi' of earlier writers; farmers and fishermen living at various points along the coast of Selangor and Malacca, numbering 1,198 (Shahrum).

*Semaq Beri* (Səma? Bərēh)SB: an ill-defined group numbering 1,406, containing settled members (probably merging with Semelai in the south) and nomadic foragers ranging widely from around Tasek Bera in Pahang, through Trengganu, and into Kelantan; most other Orang Asli regard them as Semelai. (No useful literature available.)

*Semelai* (Səməlay)Sl: 2,391 wage-earners, swidden farmers and lake-fringe fishermen in the lowlands northwards from Segamat to the south bank of the Pahang river. (Collings 1949b; Hoe 1964.)

*Temoq* (Təmə?)Tq: a little-known group of c.100 nomads and casual cultivators ranging the Jeram river northeast of Tasek Bera in Pahang. (Collings 1949a.)

### 3 *Languages not discussed here*

In addition to the dialects discussed in this paper there are other Austroasiatic dialects in the Malay Peninsula. Of these some have totally died out, others will surely soon do so, while yet others should remain available for investigation for some time yet. A full listing of these additional dialects belongs more to an ethnological study so I will mention here only the linguistically more interesting ones.

It seems likely that most of the Negrito dialects of southern Thailand, including those spoken some distance north of the border (Evans 1927:2-14), will turn out to be closely related either to Kensiu or to Jehai. Likewise some of the hitherto unreported group names that Kirk Endicott and I have collected from Negritos in Kelantan, Trengganu and Pahang will most likely turn out to be closely related to the same group of dialects. Of great interest is the likelihood that a whole subdivision of the Negrito or northern Aslian languages has become extinct during the past century; this is the so-called 'Low-Country Semang' reported by various sources to have been spoken on the mainland opposite Penang island (Blagden 1906:390-1).

Among the groups classified ethnologically as Negritos but whose speech has long been recognised to belong to the Central Aslian group there are probably several dialects yet to be discovered. H.D. Noone (1936:52) suggested that this 'Lanoh complex' is the remnant of 'a series of different tribes whose ancestors were more numerous and ranged further afield both towards the sea and southwards among the foothills'. An inspection of Schebesta's 'Ple-Temer' material (1931) shows it to be not Temiar but some

dialect intermediate between Temiar and the Lanoh group. This suggests that Temiar might once have been joined to the latter group by a continuous dialect chain. My failure to find traces of the dialect described by Schebesta among the present-day inhabitants of the settlement where he collected it (Kuala Juheng, Perak) further suggests that many of these intermediate dialects might be rapidly disappearing.

Further south three languages stand apart from the other members of their respective subgroups: Che' Wong and Jah Hut in the Benom massif of central Pahang, and Mah Meri along the coast of Selangor. Such 'singles' are not unknown in the literature of language classification. But from time to time dialects have been reported from various parts of the Peninsula which may well turn out on closer examination to be immediately related to these three languages or to be intermediate between them and the other members of their groups. Unfortunately it has not been possible to gather fresh field data on these dialects sufficient to decide the issue here. For the record, however, it might prove useful if I mention some of the dialects reported in the older literature that might fit in with this discussion. The mysterious 'Beri Nyeg' or 'Jo-ben' of the upper Klau river on the west of Gunong Benom probably speak a language quite closely related to Che' Wong (Evans 1927:41-2; Ogilvie 1949:17-8). The 'Jah Chong' mentioned by Ogilvie (1949:18) are probably coterminous with the people called Jah Hut by other authors; but they may include speakers of dialects sufficiently different from Jah Hut to diminish somewhat its status as a 'single'. Blagden suggests (1906:397) that 'Besisi' (i.e. Mah Meri) is immediately related to a

ouple of dialects situated further inland, namely 'Orang Bukit of Ulu Langat' and 'Daley's Selangor Sakai' from the Kuala Lumpur area. To the best of my knowledge the Orang Asli of these two areas are now commonly regarded as Austronesian speakers, either of Malay proper or of some Jakun dialect; but careful language-hunting in that now rapidly urbanising area may still turn up clues as to the origins of the discontinuous distribution of the Austroasiatic languages along the west coast.

Another outlying dialect mentioned by Blagden (p. 396) is 'Orang Tanjong of Ulu Langat', which from this identification is quite clearly to be regarded as a form of Semai. Whether this isolated group of Semai speakers still exists I do not know, but a search might well provide additional information for the study of Semai dialectology already started by Diffloth (1968). They may well be the same people as the 72 Semais reported for Selangor state in the 1969 census.

Lastly, Blagden's mention of 'Southern Sakai' dialects in northern Johore suggests that forms intermediate between Semelai, Semaq Beri and Temoq might once have existed, lending greater credence to the possibility that these three languages are members of a dialect-chain. Perhaps the 'Semaq Palong' mentioned by Williams-Hunt (1952:44) belongs here.

#### .4 *Present state of Aslian linguistic studies*

Descriptive analysis of individual Aslian languages in terms of modern linguistic techniques has yet to make much headway. Only one study exists that attempts an overall analysis, Asmah Haji Omar's account of Kintaq Bong (1963); unfortunately, this has remained unpublished and unknown to many of the

linguists who could make best use of it. (But see Asmah's paper in this volume.) A slightly inaccurate account of the phonology and morphology of Temiar has been circulated in manuscript form for some years; a corrected and amplified version forms one of the papers in this volume (Benjamin 1973). Gérard Diffloth has been gathering material on Semai for some time and is already in a position to provide accurate data on that language (see his papers in this volume). Both Diffloth and myself have made phonemic analyses of several other Aslian languages in the course of our work, but these remain unpublished.

There is a small amount of less strictly organised literature on individual languages which should be used with great caution, even though some useful information can be gleaned from it. For Jehai there is Schebesta 1928; for Semai, Tauern 1914, Wilkinson 1915, and Dentan MS; for Temiar, Carey 1961; for an unidentified 'Lanoh' dialect, Schebesta 1931; for Semelai, Hoe 1964; for Mah Meri, Skeat 1906, vol. I:635.

### 3.0 *The lexicostatistical count*

#### 3.1 *Choice of test-vocabulary and criteria for cognacy*

Two main problems attend the setting up and use of a lexicostatistical test vocabulary for the Aslian languages: many of the items in the standard Swadesh 200-word list are inappropriate for various reasons; and in the absence of any comprehensive historical study of these languages it is often very difficult to decide whether two words are cognate or not.

The first of these objections is now widely recognised in the literature, and it is therefore not

cessary for me to justify one-by-one the omission of words from the list used here. Many groups of words, in particular the colour terms and deictics, have been severely curtailed as they do not fit at all well with the semantic structure of the Aslian languages. Many other words have been omitted because of the environments and cultures of the Orang Asli lead to the absence of the relevant concepts from their languages. Other words have been omitted so as to avoid errors arising from the vagaries of data collection in the field. A few words have been added to the list to give it a slightly more Southeast Asian cast, but it remains basically a 146-word selection from Swadesh's 200-word list. It is unfortunate that I did not come across Thomas's word-list (1960) especially tailored to the Mon-Khmer languages until after most of my Aslian lists were collected; it would be very useful, however, if any re-study of the Aslian languages were to be based on Thomas's very carefully tested list rather than on Swadesh's.

The lack of any Proto-Aslian reconstructions has made it necessary in some cases to employ somewhat arbitrary criteria of cognacy in calculating degrees of lexical relationship. In cases where cognacy is not immediately apparent, or where certain categories of loanwords are involved, decisions have been based on the following criteria.

The root morpheme is given primacy; this is almost always the final syllable of the word, so that the prefixes (usually k-, l-, and r-) and infixes (usually -n-, -r-, and -l-) are ignored. Thus, the following are regarded as cognates: cōp, kəcup (dig); kəpəpə?, kələpə? (wing); hup, hənum (breathe); ləd?et, lənət (know); mənī?, səmə? (person).

2) The initial consonants of the root morphemes are regarded as unlikely to vary between cognates, except in the case of palatals and liquids such as w-, y-, j-, ʔ-, r-, l-, etc. Thus dow, tiw (water); gəd, kēt (cut) are regarded as non-cognates.

3) Final consonants of cognates should retain a constant place of articulation, except in the case of certain regular sound-shifts (such as Central Aslian -ʔ = Southern Aslian -h). Thus, tɪʔ, ʔatēh (earth) are regarded as cognates; but bəw, təbəʔ (big), and ʔəneh, hənjut (heavy) are regarded as non-cognates.

4) Where several apparently related words differ along two or more different dimensions so that clear-cut clumping into cognacy-groups is not possible, they are regarded as non-cognates, as for example in the following cases: haj, wəp, nəm, haw (knife); jəroʔ, cərek, jələŋ (long).

5) No attempt is made to exclude inter-Aslian loanwords, even though this contravenes the canons of proper lexicostatistical technique. This does not affect the final calculation since most such cases stand out clearly from the cognacy rates of the immediately related languages. Besides, one of the aims of this paper is to calculate and discuss the significance of inter-Aslian loan rates. Thus, though SB hateʔ (tail) is clearly a loan from Northern Aslian it is treated here as a cognate.

6) Following from criterion 5, and notwithstanding the other criteria, the possibility of 'sampling errors' between speakers of different dialects is allowed for in cases where alteration of one or two phonetically similar phonemes would change the words into obvious cognates. Thus the following are all

garded as cognates: cō?, caw, cōh (dog); kawaw, wōd (bird); tleek, jeteek (sleep--a case of metathesis); ncem, ntɛp (near).

Certain acknowledgedly non-Austroasiatic words are treated as potential cognates in cases when there are reasons for assuming that they were incorporated into the Aslian languages before they split up into the present-day subgroups; but clear-cut loans from non-Austroasiatic languages (usually Malay) into individual Aslian languages are excluded from the count. (The same argument applies to the few suspected cases where genuine Austroasiatic forms have entered the Aslian languages from a source other than Proto-Aslian: in particular, some Khmeric loans seem to have entered Southern Aslian.) Thus the Austro-Aslian words kəbis, khəbəs, kəbos (die); ləbeh, ləbeh (any); siya?, səya? (salt) are regarded as cognates. t kəlko?, təkō?, etc. (claw) are regarded as non-cognate with MM kokont, as the latter appears to be a direct individual loan from Jakun.

## 2 *The calculations and subgroupings*

The test-vocabulary for all 20 languages and dialects is set out in Appendix II, which in addition to providing the relevant Aslian words, indicates the following decisions: cognates are grouped together; non-cognates are separated; loan-words are separated from 'valid' items; and proper indication is given in those cases where no item has been reported for any particular dialect.

These decisions were submitted to three different types of calculation and tabulation: a straightforward cognacy percentage matrix for all Aslian dialects, a counterindicative matrix for Aslian



dialects, and, for Northern Aslian alone, a matrix showing the so-called 'homotrophic' or 'characteristic vocabulary' indices.

### 3.21 *Cognacy rates*

The cognacy percentage matrix is displayed in Table 1. (The figures are rounded-off to the nearest whole number.) For each language-pair the tabulated percentage was calculated by scoring 1 for each case of cognacy and 0 for each case of non-cognacy, dividing the total number of cognate cases by the number of items compared, and multiplying the result by 100. However, in the case of two languages adjustments had to be made to correct inflated percentages resulting from the incompleteness of their test-lists: specifically, the figures for Kintaq Bong (KB) and Temoq (Tq) are carefully adjusted, and in most cases are slightly lower than the figures that would be arrived at by unadjusted calculation.

Examination of the cognacy-percentages leads to the conclusion that the Aslian languages fall into three major subgroups: *Northern Aslian* with a modal cognacy rate of about 47%, consisting of Kensiu, Kintaq Bong, Jehai, Mendriq, Bateg Deq, Mintil, Bateg Nong, and Che' Wong; *Central Aslian* with a modal cognacy rate of about 38%, consisting of Semnam, Sabum, Lanoh Jengjeng, Lanoh Yir, Temiar, Semai I and II, and Jah Hut; and *Southern Aslian* with a modal cognacy rate of about 38%, consisting of Mah Meri, Semaq Beri, Semelai, and Temoq. (In Diffloth's terminology these three subgroups are referred to as Jahaic, Senoic and Semelaic respectively.)

The cognacy rates alone, however, do not allow us to decide how these three subgroups are related to each other. Northern and Southern Aslian seem to be

Table 2

Aslian Languages Cognacy Matrix  
(percent, rounded-off)

Ks	81	KB	49	47	Je	57	56	54	Mr	41	42	42	55	BD	48	49	47	60	58	Mt	46	49	46	49	53	58	BN	38	40	34	34	34	32	45	CW	33	32	32	27	23	25	25	28	Sn	32	30	36	27	23	27	27	28	72	Sa	36	33	41	33	27	29	29	30	61	66	LJ	36	34	42	32	27	22	26	27	29	30	30	64	71	82	LY	31	28	32	27	22	19	22	23	29	34	37	38	36	47	Tm	25	27	27	24	19	22	23	29	35	37	40	39	52	84	Sml	26	28	29	26	20	22	23	29	35	37	40	39	52	84	Sml	24	27	25	26	26	23	30	33	29	31	28	28	29	40	39	JH	16	17	16	14	13	12	15	20	21	21	22	22	23	30	28	25	MM	23	25	21	25	21	22	26	30	31	31	33	35	35	33	33	28	35	SB	18	18	17	16	14	15	20	26	25	27	25	27	25	31	29	24	38	48	Sl	15	16	15	16	15	16	18	26	24	25	26	27	24	28	27	25	36	49	59	Tq
----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

equally distant from Central Aslian, with modal inter-group cognacy rates of 27% and 25% respectively. But the problem is that the relationship between Northern and Southern Aslian seems at first sight to be more remote than this, with a modal inter-group cognacy rate of only 16%. An attempt to resolve this difficulty will be made later in the paper; meanwhile, all three Aslian subgroups will be treated heuristically as coordinately related branches of Aslian.

Further subgrouping is nevertheless possible on the basis of these cognacy figures. Southern Aslian is a perfectly straightforward case, forming a dendrogram of the following shape:

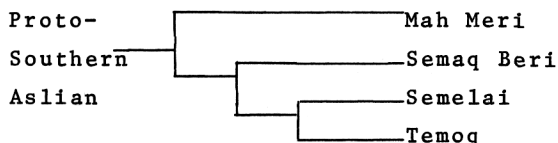


Figure 2

The Central Aslian figures likewise allow of the formation of a dendrogram, as follows:

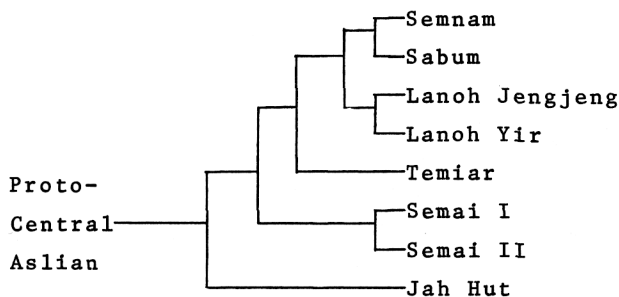


Figure 3

But whereas the Southern Aslian languages are quite clear-cut from each other, the Central group demonstrates two further features. Firstly, six of

the dialects can be merged into three (or even two) languages: a 'Semai' language containing (among others) dialects I and II; a 'Lanoh' language consisting of at least the Jengjeng and Yir dialects; and a 'Sabum' language consisting of the Semnam and Sabum dialects. Quite probably Lanoh and Sabum should be further merged into a larger 'Lanoh' language combining both dialect-clusters, but the cognacy figures alone do not allow us to decide the issue.

Secondly, there has been a quite high degree of borrowing between some of the Central Aslian branches, notably between Jah Hut and Semai, between Temiar and Semai, and between Temiar and Lanoh. The significance of such borrowings is discussed later.

It is the pattern of borrowings between Temiar and Lanoh that allows us to dispose of the possibility that Lanoh is more closely related to Temiar than to Sabum, which is the relationship implied by the pattern of shared phonological innovations between the three languages. Acceptance of this phonology-based classification would, however, entail a pattern of inter-language loans highly implausible from a geographical point of view.

The Northern Aslian languages, however, prove more recalcitrant to subgrouping on the basis of these cognacy figures. Only limited conclusions can be drawn by simple inspection: 1) Kensiu and Kintaq Bong are dialects of the same language ('Kensiu'); 2) Che' Wong is relatively distantly related to the other Northern languages, but has borrowed from Bateg Nong; 3) the remaining languages appear to fall into two major subgroups, a western one (Kensiu) and an eastern one (Mendriq, Bateg Deq, Mintil, and Bateg Nong); 4) Jahai is apparently more closely related to Mendriq

than to any other language, though the reverse is not true.

It might be thought that this resistance to clear-cut subgrouping results from the formation of dialect chains between populations who are, after all, nomadic in life style. But closer examination of the figures shows that though this has certainly occurred to a considerable degree, other factors are involved as well.

It is not always the case that the highest apparent cognacy rates are found between contiguous languages. Che' Wong, for example, appears to be more closely related to Kensiu some 200 miles away than to nearby Mintil; while Kensiu appears to be more closely related to Mendriq than to Jehai, which separates them by about 100 miles.

Clearly we are not dealing entirely with a chain-like situation. But, just as clearly, techniques other than straight cognacy counts are needed to separate Northern Aslian into its constituent branches. It is mainly for this reason that the counterindicative and homotrophic matrix methods also were applied to the lexical data.

### 3.22 *Counterindicative indices*

The method of counterindications attempts to measure for any pair of languages the likelihood that they are *not* immediately related. I do not propose to outline the calculation procedures here as they are easily accessible in the literature (Gleason 1959:27-8; Landar 1966:199f.). Suffice it to say that immediate relationship between any two languages is indicated on a counterindicative matrix when they are linked by a significantly low *minimum* figure. The

Ks		KB		Je		Mr		BD		Mt		BN		CW		Sn		Sa		LJ		LY		Tm		SmI		SmII		JH		MM		SB		Sl		Tq																																																																																																																																																																								
10	KB	23	24	17	13	28	Mr	20	18	31	16	BD	24	18	35	16	9	Mt	25	22	36	30	14	18	BN	28	26	45	40	29	38	31	CW	40	35	48	54	48	54	40	Sn	48	47	52	62	54	59	58	47	10	Sa	47	48	54	62	55	62	64	50	17	15	LJ	48	50	55	66	59	67	64	51	15	16	8	LY	51	47	53	61	59	63	60	52	26	26	21	20	Tm	55	45	55	58	51	58	58	43	34	40	41	45	27	SmI	54	42	50	53	48	55	55	44	31	36	37	40	27	3	SmII	42	34	48	47	33	46	41	29	36	40	44	49	42	21	24	JH	37	36	48	45	40	50	45	38	30	34	39	37	35	25	26	22	MM	49	48	61	53	47	54	54	48	38	46	48	50	43	39	37	41	19	SB	48	48	63	58	54	63	58	51	40	47	52	54	52	40	42	40	16	16	Sl	51	48	62	56	51	57	58	50	36	45	49	51	48	37	38	35	15	15	5	Tq

Table 2  
Aslian languages: Counterindicative indices

counterindicative matrix for the Aslian languages is displayed in Table 2.

Once again, the Southern and Central groups show unambiguous results, with the following pairs linked together as immediately related: Semelai and Temoq; Semai I and II; Lanoh Jengjeng and Lanoh Yir; and (less closely) Sabum and Semnam. This corresponds excellently with the subgroupings derived from the cognacy percentage matrix, with the further hint that Temoq and Semelai are more closely related than the cognacy method implies.

With the Northern group, however, there are no strikingly low minima; but the following conclusions may be drawn. 1) The most obvious minima link Kensiu with Kintaq Bong (which is to be expected), and Mintil with Bateg Deq. The latter result is a useful one, as the straight cognacy figures suggest that Mintil is equally related to both Bateg Deq and Bateg Nong, while the counterindicative figures suggest that it is much less closely related to Bateg Nong than it is to Bateg Deq. (The other relatively low minimum, the 13 scored between Mendriq and Kintaq Bong, is discounted here because of the already mentioned inadequacies in the data on the latter language.) 2) On the other hand, the counterindicative figures strongly suggest that neither Jehai nor Che' Wong are immediately related to any of the other Northern languages. 3) Slightly higher minima link Bateg Nong with the pair formed by Bateg Deq and Mintil. 4) Mendriq cannot be unequivocally linked with any other of the Northern Aslian branches on the basis of its counterindicative indices any more than it can on the basis of cognacy rates; the indices do suggest, however, that Mendriq has exchanged vocabulary with Kensiu,

Bateg Deq and Mintil.

On the basis of the cognacy rates and counter-indicative indices alone, then, we can subgroup the northern Aslian languages only to the extent set out in the following dendrogram.

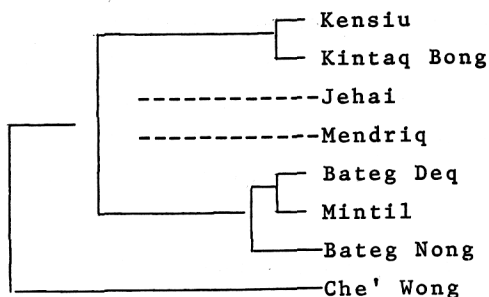


Figure 4a

The problem languages are Jehai and Mendriq, which appear to need yet further techniques before they can be satisfactorily positioned within the northern Aslian group.

### 2.23 Homotrophic indices

The 'homotrophic index' is Landar's name for what Gleason calls 'characteristic vocabulary index' (Gleason 1959:28-9; Landar 1966:202), and is aimed at tabulating a weighting of exclusively shared cognates between any pair of related languages. The method is well enough described in the sources to need no further exemplification here. A homotrophic matrix is examined for the maxima it contains: the higher the homotrophic index for any pair of languages, the more do they share items of vocabulary not shared by the other languages on the matrix. However, caution is necessary at this point since it is clear that if no special measures are taken to exclude inter-language loans from the tabulation (as in the present



case) then a high homotrophic index could indicate a large proportion of exclusively shared loans as well as of exclusively shared cognates. Loan-based homotrophic maxima can only be separated from cognacy-based homotrophic maxima by comparison with the results of the other two methods of calculation.

Table 3

Ks											
26	KB										
11	5	Je									
8	6	11	Mr								
3	3	4	6	BD							
4	4	4	12	10	Mt						
5	5	6	4	7	10	BN					
8	5	4	4	6	3	12	CW				

Northern Aslian languages:  
Homotrophic matrix  
(rounded-off)

The Northern Aslian homotrophic matrix (Table 3) enables us to make such a comparison and draw the following conclusions.

1) It confirms that Kensiu and Kintaq Bong are dialects of the same language. 2) It confirms the hypothesized close genetic relationship of Bateg Deq and Mintil. 3) It confirms the suspicion that the following language-pairs have exchanged vocabulary: Mendriq and Kensiu, Mendriq and Mintil, Bateg Nong and Che' Wong. 4) It suggests further that the following language-pairs have exchanged vocabulary: Mintil and Bateg Nong, Kensiu and Jehai. 5) It hints at a closer relationship between Kensiu and Che' Wong than their physical distance would seem to allow. 6) Most importantly, it confirms the suspicion that Mendriq and Jehai are immediately related.

Despite the obvious fact that the Northern Aslian languages have so borrowed from each other as

form a meshwork rather than a branching tree, we best study the pattern of borrowings if we first set up an ideal-type dendrogram to indicate the most probable hierarchy of genetic relationships. The combined weight of the three sets of calculations points to the following as the best-fitting representation of Northern Aslian relationships.

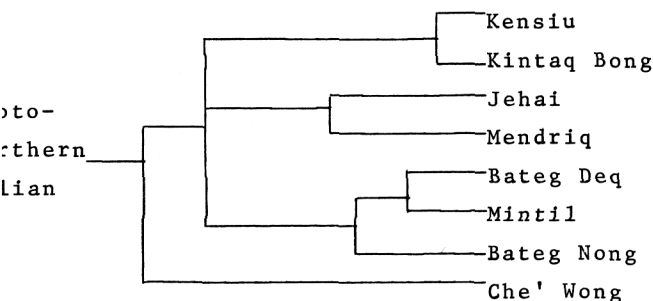


Figure 4b

### 3 *Non-lexical correlates of the proposed sub-grouping*

I have already quoted Wilkinson's warning against relying exclusively on lexical data in setting language classifications. The aim of the present paper, however, is to use the lexical data to set up heuristic classification of the Aslian languages in the hope that it will give direction to further research. But to do this job properly, I believe that the implications of the lexically-based classification should be fully worked out, even to the extent of setting up a detailed model of the prehistory of Aslian language differentiation, in the full understanding that future work may greatly alter the picture.

Nevertheless, the classification just proposed does receive some support from the sparse non-lexical

data at present available. Firstly, let us examine the extra evidence to support the three-part division into Northern, Central and Southern.

Phonologically, all Aslian languages share the same basic phonemic inventory (see Appendix II), with the following exceptions. 1) Northern Aslian has as rare additional phonemes /f/ and /z/. 2) While Central and Southern Aslian possess a phonemic contrast between long and short vowels in word-final syllables, this contrast is lacking from Northern Aslian. 3) Southern Aslian possesses a contrast between aspirated and non-aspirated syllable-initial stops which is lacking from Central and Northern Aslian.

Inspection of comparative vocabularies suggests that some regular sound changes correspond to the three-part grouping. I mention only two here: Southern and Central Aslian səCVC corresponds to Northern Aslian haCVC ('leaf', 'shoot', 'rotten'); Northern and Central Aslian word-final -ʔ corresponds to Southern Aslian word-final -h ('earth', 'fish', 'fruit', 'leaf', 'louse', 'rice', 'woods').

Morphologically, the Aslian numerals have long provided a key means of subgrouping, ever since Blagden's brilliant discussion (1906:454f.). A tabulation of the first three numerals (only Che' Wong and certain Southern languages go beyond three without borrowing from Malay) will demonstrate their closeness of fit with the three-part linguistic division.

	Northern	Central	Southern
e	nay, ney, noy	ney, nTh, ney- wey nanu?	moy, muy
o	biyeh, ber (duwa?)	nār	hmāh, mar, (duwa?)
ree	(tiga?)	ne?, ni?, (tiga?)	hmpe?, hmpe?

The pronominal system, however, sets Southern Aslian apart from the other two. Whereas the Northern and Central systems possess fully-blown pronominal paradigms containing singular, dual and plural number for all three persons, combined with an inclusive-exclusive contrast in the first-person terms, the Southern system lacks the dual number and the inclusive-exclusive contrast. In addition, the Southern Aslian pronominal paradigm possesses the peculiar feature that the second- and third-person terms are identical, which is not the case for Northern and Central Aslian.

Further study of these languages will surely reveal more contrasts of this kind.

#### 10 *Differentiation and interaction of Aslian languages*

The model of Aslian relationships developed above is intended to serve as a starting point for further investigation, and not as a self-justifying end-product. Let us see, then, how the dendrogram helps us to understand something of the past and present circumstances of the Aslian languages. There are two main problems here: inter-language loans, and Aslian linguistic prehistory. As we shall see, these two concerns are intimately connected.

It is of course now generally recognised that a dendrogram is a misleading way of representing

linguistic differentiation, for only in the rarest cases do languages split apart at a single moment in time never to contact each other again. The normal pattern involves a relatively long period of localised dialect formation during which most of the speakers retain contact and mutual comprehensibility. Some of the dialects may later move apart in such a way that the habit of intercommunication is lost, and a 'new' language comes into being. There is no reason however why the 'new' language should be as homogeneous and dialect-free as the usual naive model of linguistic differentiation would suggest. The chances are that so-called 'languages' are from their inception already marked by heterogeneity and incipient dialectal differentiation. This model, representing what is probably the most common pattern of linguistic differentiation, has been labelled the 'mesh principle' by Swadesh (1959:7f), who urged that it be made an explicit element in linguistic research. From here on in this paper we shall do just that.

If we assume that languages take a considerable time to split apart it should not be a matter for embarrassment that two or more coordinately related groups of languages usually show a considerable scatter in their constituent cognacy rates. On the contrary, this common observation should be turned to advantage and the variance in the figures deliberately used as a means of measuring the degree of mesh-formation that has occurred. To do this, however, we need to develop a simple set of procedural rules, as the published literature seems not to give any detailed guidance.

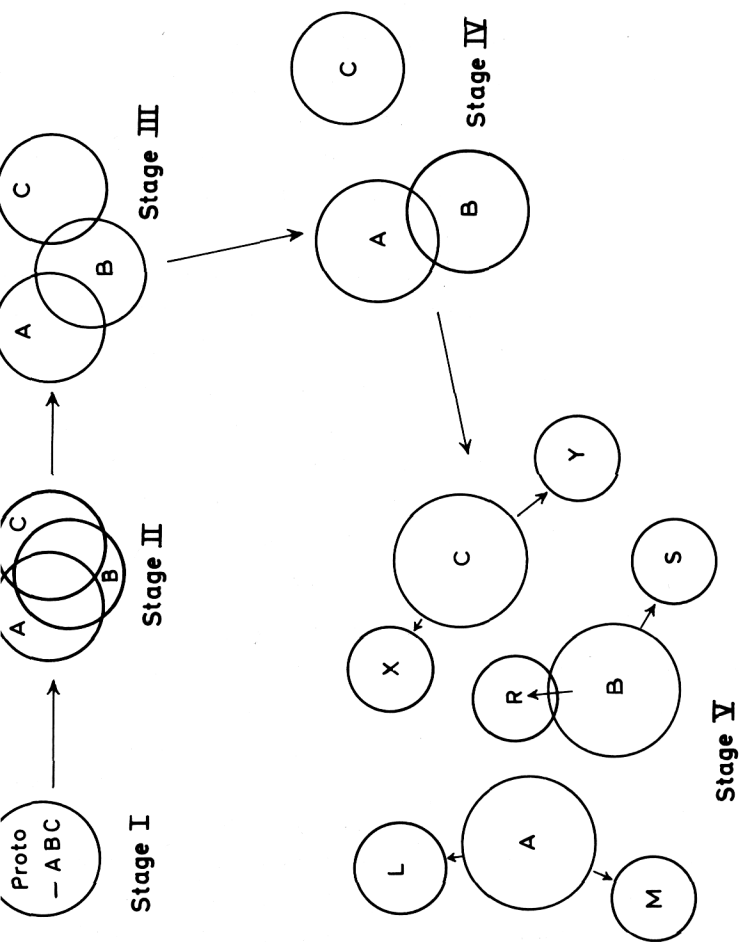
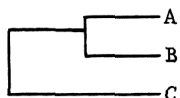


Figure 5

Consider the simple hypothetical case in which three languages are related genetically thus:



Use of the mesh principle allows the probable course of differentiation between them to be illustrated by the following series of Venn diagrams, where overlap of the circles indicates continuing mutual comprehensibility between dialects, and non-overlap indicates that the speech communities have finally split apart into mutually incomprehensible languages.

This situation is most likely to show itself on a cognacy percentage matrix in the form of a slightly lower cognacy rate for A and C than for B and C. The real-life situation, however, is likely to be more complex, as each of the languages A, B, and C will in turn be simultaneously undergoing further differentiation into dialects, as represented in stage V of Figure 5. The split between the A-B group and the C group would then be represented on a cognacy matrix by 18 distinct cognacy percentages. Under optimum conditions, then, for any two-way split there are four different kinds of cognacy figures available to play around with, and these provide the basis for the following set of procedural rules.

For any two coordinately related groups of languages:

- 1) The minimum reported cognacy percentage (e.g. A-C) will be taken as representing the earliest occurring split.

2) The highest reported cognacy rate which statistical calculation shows to fall still within the range of figures representative of a continuous process of mesh-formation will be taken as representing the latest-occurring split in that process (e.g. B-C).

3) The statistical mode of all the relevant cognacy rates will be taken as an indication of the degree of mesh-formation that has occurred: if the mode is higher than the latest-split cognacy rate, then the speakers of the various languages have probably remained in contact with each other; if the mode is equal to or less than the latest-split cognacy rate, then the speakers of the various languages probably lost contact with each other at an early stage.

4) Percentages more than 3% above the latest-split percentage will be taken to indicate that borrowing has occurred specifically between the two languages concerned: the difference between the apparent cognacy rate for the two languages and the latest-split cognacy rate for the groups to which they belong will be taken as the loan-rate (in %) between them.

The results of applying these calculations to the Aslian data are set out in Table 4, which tabulates the significant apparent loan rates between the various languages along with the proportion of the total vocabulary that each language has borrowed from Malay, in rounded-off percentages. This should be read in conjunction with the modal cognacy percentages given in Section 3.2 above. Let us now work out the implications of these figures.



[illegible]

### *Societal factors in language differentiation*

Firstly, the sociological aspects of Aslian language-differentiation and language-contact can be understood much more clearly with a knowledge of the modal cognacy rates for the three main groups and the intra-group loan rates.

The three main divisions of Aslian show quite distinct characteristics in relation to these figures. Northern Aslian has no intra-group loans and a low modal cognacy rate (38%); Central Aslian has many intra-group loans and a low modal cognacy rate (38%); Southern Aslian has many intra-group loans and a high modal cognacy rate (47%). Obviously this suggests that the speakers of these languages have generated different patterns of inter-group contact.

With the partial exception of the Che' Wong, the speakers of Northern Aslian languages have until very recently all been nomadic hunter-gatherers, living in small bands that move on every three or four days along traditional routes. There is a marked tendency towards exogamy and for individual families constantly to change their domicile from one band to another. Under such circumstances it is hardly surprising that mutual comprehensibility is possible over very large areas. Schebesta several times reported that the Malay speech he had first learnt in the Grik neighbourhood of north Perak was still understood by the Bateg nearly 200 route-miles away in central Pahang. Hence the markedly mesh-like pattern of Northern Aslian language distribution is to be expected.

With the exception of the Lanoh cluster, all the speakers of Central Aslian languages practise a long-established form of swidden farming, live in semi-

permanent villages, and usually move only within the confines of their home river valley. Although certain individuals travel a lot, the residential groups as a whole have a high degree of localisation. Under these circumstances the expected pattern of language differentiation would be one in which a few well delimited languages emerge as a result of earlier pioneer expansion, followed, as the population density rises, by consolidation of the various dialects, normally within the limits set by the major river basins. Mesh-formation would be restricted and it would be normal for the people to protest that they do not understand the speech even of their fellow tribesmen in the next valley. This is precisely the ethnographically reported situation for the Temiar, Semai and Jah Hut. Nevertheless, the figures show that these languages have exchanged vocabulary with other languages, though without breaking down the clear boundaries between them. Two factors probably are responsible for this: the widespread travels of individuals in search of spouses, employment or adventure; and the tendency of other smaller Orang Asli groups to attach themselves to the larger agricultural tribes in a sub-nuclear relationship for a season or two at a time.

The absence of mesh-formation and intra-group loans among the Southern Aslian languages points to a pattern of language differentiation in which language formation coincided with physical movements of the people away from each other so that contact was soon broken. Two factors would explain this pattern: the lower population density in the south suggests that pioneer, migratory swidden agriculture could have continued there until very recently; and the presence of ancient trade routes and settlements across the south

the Peninsula could have established very early the tendency of some southern Orang Asli groups to aggregate near centres of commerce and to engage as much in trading as in farming.

### *Interaction between contiguous languages*

Loans between individual languages, whether or not the same subgroup, demand more specific explanations. The simplest cases to deal with are those where exchanging languages are contiguous in distribution where the speakers of the different languages often come together. There are several examples of this in Table 4. In the north-west of the Peninsula all the Negrito groups are in physical and cultural communication with each other, such that it is virtually impossible to find a Negrito village in Perak with representatives of less than three different ethnic groups there; there is evidence that this is not merely a recent result of the formation of composite bands, but a long-established feature of Negrito social organisation. This explains the large block of loans linking Kensiu dialects with the Lanoh dialects, and the latter with Temiar and Jehai. (The apparent absence of loans between Jehai and Kensiu will be discussed later.) A similar explanation will account for the spectacular block of loans linking Bateg Nong, 'Wong and Jah Hut in the centre of the Peninsula, the loans between Mendriq and Bateg Deq in Kelantan. The Temiar and Semai do not normally live in joint villages, but along their mutual boundary it is usual for people to speak both languages well; nevertheless their high loan rates must be due to earlier close contacts which have since reduced in intensity.

#### 4.3 *Interaction between non-contiguous languages*

More difficult to explain are cases where loans appear to have taken place between languages that are not now contiguous. The assumption here must be that language movements have occurred: the apparent irregularities in Table 4 can then be used as a means of reconstructing the probable course of those movements. Furthermore, by comparing the loan rates of each of a group of immediately related languages with some other single language, it is possible to work out some form of relative chronology of language movements dated against language splits. Application of these techniques has produced the following suggestions as to the migrational history of individual Aslian languages.

1) *Jehai*: Whereas Mendriq and Kensiu, though contiguous in distribution have exchanged vocabulary, *Jehai* has no loans from either language. This suggests that Kensiu and Mendriq were once in contact and that *Jehai* has since intruded, relatively recently, between them. If so, it probably came from the south, as the *Jehai-Lanoh* loan rates are extremely high, suggesting an earlier even closer contact than they have at present (particularly as the *Jehai-Semnam* loan rate is some 10% lower). In general, *Jehai* shows no loans with the languages of the easterly Negritos, though it has very high loan rates with the languages (Lanoh, Sabum) of the westerly Negritos. This suggests that the move of *Jehai* into Kelantan and southern Thailand via the Belum and Pergau rivers must be relatively recent. It also suggests that Kensiu and Mendriq were previously linked up in southern Thailand, remaining in contact there until recently.

2) *Mendriq, Bateg Deq and Bateg Nong*: The movements of these three languages have left their traces

their loan relationships with Semaq Beri, which seems to have maintained a mesh-like connection with Northern Aslian languages through its direct contacts with Bateg Nong (10%) and Mendriq (9%). But its loan rate with Bateg Deq is rather lower (5%), suggesting either that Bateg Deq has only recently moved to its present location, or that the high proportion of borrowed loan words in Bateg Deq may have lowered its apparent loan rates with all other languages. The higher lower loan rates between Bateg Nong and the non-contiguous Semelai and Temoq suggest that these latter may be linked with Semaq Beri in a mesh relationship.

3) *Che' Wong*: This language shows a puzzling loan rate of 4% with Kensiu despite the distance which separates them. Assuming this figure to be valid, it would mean that during the early stages of Proto-Northern Aslian, Che' Wong did not immediately move away from the western group until after Jehai had split off and moved south. Che' Wong would then have retained contact with Kensiu for long enough to cause a rather but noticeable loan rate between them. Perhaps it was the return north of Jehai that spurred Che' Wong to move away eastwards. Che' Wong shows high loan rates with all the Southern Aslian languages except Meri, which suggests that its ancestor came into contact with Proto-(Semaq Beri-Semelai-Temoq) before the latter split apart but after Mah Meri had split off.

4) *Temiar*: No striking movements seem to be involved in this case. The only puzzle is the high loan rate that Temiar, along with Lanoh, has with Semaq Beri, a language that is extremely unlikely even in its widest wanderings to have come into direct contact with Temiar. If these figures are to be trusted at all,

they suggest that a fairly dense mesh links Lanoh-Semai and Temiar with those Negrito groups that have direct contact with Semaq Beri in the east. The high loan rates between the Lanoh cluster and Semaq Beri would then be purely adventitious, resulting from the close relationship between Lanoh and Temiar and from the relatively high loan rates between Lanoh on the one hand and Kensiu and Mendriq on the other. An alternative explanation could be that Proto-Semaq Beri came into close direct contact with Proto-Lanoh-Temiar before the latter split up into its present dialects: the cognacy rates allow of such a possibility even though the geography is a little doubtful.

5) *Mah Meri*: As a group the Southern Aslian languages have exchanged vocabulary only with Semai among the Central Aslian languages, which indicates that contact between them came about after Semai had split away from Proto-Temiar-Lanoh. Mah Meri, however, has generally lower loan rates with Bateg Nong and Che' Wong than do the other Southern Aslian languages, which suggests that it moved off to the west before Bateg Nong and Che' Wong arrived in the south to make contact with Proto-Semelai-Semaq Beri. Hence, Proto-Southern Aslian must have been in contact with Proto-Semai, and Proto-Semelai-Semaq Beri must have remained in such contact even after Mah Meri moved away, until Bateg Nong and Che' Wong eventually intruded between the Southern Aslian languages and Semai. The most likely route along which Mah Meri would have travelled lies in territory now occupied by the Malay-speaking Temu and Belandas groups, which would confirm the suspicion expressed by both Blagden (1906:396) and Dentan (1961:176) that these latter are in origin Malayised Semai rather than true Jakun. Strikingly, a small Semai

### Modal time-dep

[illegible]



outlier is still to be found, south of the main body, in Selangor. Since Mah Meri has no loans with the other Southern Aslian languages it probably moved away quite clean, perhaps as a result of intrusion by a foreign group or as a result of a complete alteration in the life-style of the speakers of one or other of the two branches of Southern Aslian.

If we could now feed some absolute dates into the interpretation of the figures we would be able to construct an outline of the linguistic prehistory of Aslian. Under the circumstances a glottochronological experiment is fully warranted, despite the well-known pitfalls of this method.

#### 5.0 *Glottochronological counts and prehistory*

Table 5 displays the results of glottochronological calculations performed on the cognacy rates reported in Table 1. Time-depths in years-before-present (BP) have been calculated for those language pairs which correspond to the earliest- and latest-occurring separations for each of the splits implied by the dendrograms of language-relationships worked out in sections 3.21 and 3.22 above. A conservative retention rate of 86% has been employed in accordance with the tendency in recent glottochronological studies. Those who prefer to use lower rates in their calculations may easily convert the results presented here by means of a simple calculation.

The dates in Table 5 are given as maxima and minima at the 7/10 level of confidence, calculated according to the methods set out by Gudschinsky in her classic paper on lexicostatistical technique (1956). It is now possible to attempt a total reconstruction of Aslian linguistic history; the results of my attempt

do so are set out in the next section. But I must first explain why I believe this to be more than merely a mechanical exercise in fitting a flow-chart to a set of figures.

As Blagdon hinted long ago (1906:470f.), and as I have illustrated above, the statistical patterns of language differentiation bear a close relationship to the sociological and ecological circumstances of the people who speak those languages. A further possibility is that glottochronological datings, despite any doubts about their efficacy as absolute measures, may show a number of language splits to have been associated together in such a way as to suggest either that they all result from the same stimulus (or 'event', in an archaeological sense) or that they have directly stimulated each other in a serial manner. Possible cases of serial interactions of languages on each other have just been discussed, while the possible effects of trade and agriculture upon language distribution were discussed in section 4.1. The recent growth of interest in the archaeology and ethnology of the Malay Peninsula means that a glottochronology-based reconstruction would not stand for long without being tested against external data derived from quite different kinds of research. The following reconstruction might in its turn prove helpful to workers in related fields.

#### 1 *Reconstruction of Aslian linguistic prehistory*

Five phases may be distinguished in the linguistic prehistory of Aslian.

##### 1.1 *Emergence of Proto-Northern, -Central and -Southern Aslian from common Proto-Aslian*

1) Originally all Proto-Aslian speakers were nomadic hunter-gatherers. Their language was almost

certainly heterogeneous even before it entered the Peninsula, where quite possibly there were Negritos already present speaking a non-Austroasiatic language of their own (related perhaps to Andamanese).

2) Some form of sedenterisation, probably associated with rudimentary agriculture, arrived on the scene some time between 6610 and 6410 BP, somewhere in the South-Central area. Judging from the mainly *Austroasiatic character of the basic Aslian 'domestication' vocabulary*, agriculture was probably received from Austroasiatic-speaking groups further north with whom they were probably still in linguistic continuity.

3) The northern people remained nomads, retaining direct contact only with the central people, so that already before 5970 BP northerners and southerners were speaking distinct dialects. The central people, however, still retained linguistic continuity with both the northerners and the southerners. By this time Proto-Aslian would have separated off finally from other Austroasiatic languages.

4) Some form of 'trade' (i.e. localised, mutual advantageous contacts with newly arrived foreigners at points outside the original Proto-Aslian area) arrived between 5280 and 5060 BP in the south. This caused some of the southerners to move away from the others, so that by 4900 BP linguistic continuity had been lost between the central people and those southerners who had moved away to 'trade'; all the southerners still retained linguistic continuity between themselves, however.

5) Some form of 'trade', or perhaps colonial settlement, arriving in the north between 5120 and 4900 BP led to further divergence between the northerners

and the central people--the northerners possibly entering into occasional direct foraging or trading contact with the foreigners, much as their Negrito descendants do today. By 4750 BP some of these northerners had lost linguistic continuity with the central people, even though the residual northerners still formed a language mesh with the latter.

6) Proto-Southern Aslian finally separated off some time before 4470 BP, as pioneer agriculture led the southern farmers further south (possibly down the Tang Padang, Jelai and Tembeling routes).

7) The north-central mesh finally broke up into Proto-Northern Aslian and Proto-Central Aslian some time before 4210 BP.

## 12 *Separation of 'outlier' languages*

1) As a result of the preceding long period of mesh-relationship the Proto-Northern, -Central and Southern Aslian languages must each have been fairly heterogeneous from the beginning. Some of their dialects now separated out to become outlier languages at the following stages.

2) Proto-Jah Hut began to separate off from the other Proto-Central Aslian dialects some time after 4470 BP, finally emerging to the east as a distinct language some time after 3970 BP.

3) Proto-Northern Aslian divided into two blocks such that Proto-Che' Wong began to separate off to the west some time after 3900 BP.

4) A split in the remaining Proto-Central Aslian dialects beginning after 3700 BP resulted in Proto-Miar-Lanoh moving northwards away from Proto-Semai

and intruding into the Northern Aslian area.

5) Before 3440 BP Proto-Che' Wong had become distinct from the westerly Northern Aslian dialects, and by 3250 BP Proto-Semai had become finally separate from Proto-Temiar-Lanoh.

6) Proto-Mah Meri began to separate from the other Proto-Southern Aslian dialects some time after 3610 BP, becoming finally distinct before 3250 BP.

### 5.13 *Emergence of the major languages*

1) The western branch of Proto-Northern Aslian began to divide some time after 2710 BP, Proto-Kensiu moving westwards, Proto-Jehai-Mendriq southwards, and Proto-Bateg eastwards (probably to Ulu Pergau). By 2290 BP these had already established themselves as distinct languages, even though their speakers probably remained familiar with each other's speech and maintained actual contact. (It is possible that the stimulus for these moves was the development of more complex iron-working urbanising societies further north between 3000 and 2000 BP.)

2) Proto-Che' Wong moved south between Proto-Semai and Proto-Jah Hut until it abutted on the easterly Southern Aslian languages at about the same time (i.e. after 2570 BP) that the latter started dividing into Proto-Semelai to the west and Proto-Semaq Beri to the east. By 2190 BP Proto-Semelai and Proto-Semaq Beri were distinct dialects.

3) Some time after 2300 BP Proto-Lanoh-Sabum began to split off westwards from Proto-Temiar. This might well have been the result of a population explosion following increasingly efficient food-production in the central area, causing the peripheral people--who probably looked like the northerners anyway--to

some foragers in less densely settled territory. Some time before 2030 BP Proto-Lanoh-Sabum formed a distinct language, just south of Proto-Jehai-Mendriq and just west of Proto-Temiar.

4) Possibly as a result of the further outward expansion of the Temiar speakers, two intrusive splits ensued in the neighbouring Northern Aslian languages. Some time after 2180 BP Temiar intruded into Proto-Jehai-Mendriq causing it to divide into Mendriq north-eastwards (towards Bateg) and Jehai proper southwards. This process was complete before 1900 BP. The simultaneous eastwards expansion of Temiar coincided with a split in Proto-Bateg some time after 2240 BP, with Bateg Nong moving off southwards, probably up the Lebir valley, and Bateg Deq remaining in the mid-Menggiri area.

#### 14 *Emergence of the modern dialects*

1) Continuing population increase and, very probably, the spread of Malay and Indian settlers upriver from the coast now set in motion a further series of dialect splits and language relocations. It was probably at this time also that secondary Mon and Khmer influences made themselves felt, especially in the south, as a result of trade, mining and settlement.

2) Proto-Bateg Deq began to split apart some time after 1970 BP, with Bateg Deq proper moving south of the Lebir, and Mintil remaining behind in the Menggiri area; by 1640 BP they had completely separated. Mintil at this time remained in close contact with Mendriq, probably in the lower Galas valley, the two languages exchanging much vocabulary.

3) Some time after 1880 BP Proto-Semelai began to split up to form eventually a larger population of

farmers speaking Semelai proper and a smaller population of foragers speaking Temoq. Possibly at this time the Semaq Beri speakers also became secondarily nomadic. These changes, which were complete by 1620 BP, may have been started off by a chain reaction--the southward movement of Bateg Deq pushing Jah Hut further south until the latter in turn abutted on Proto-Semelai leading that language to split up.

4) Some time after 1780 BP Proto-Lanoh began to move northwards, probably up the Perak river as a result of Malay and Temiar population pressure, hiving off a group of Proto-Sabum speakers who remained behind. The Lanoh speakers began to abut on the Jehai speakers, the two languages exchanging much vocabulary. But continuing pressures pushed the Jehai speakers further upstream into the hills (not the normal habitat of Negritos) from where many Jehai speakers spread out into the headwaters of the Patani and Pergau rivers where they are still to be found. Lanoh and Sabum became distinct dialects by 1340 BP.

5) As the Jehai-speakers reached the Mendriq area on the lower Pergau, the latter language in turn spread south and west along the Nenggiri around the Lah area where the Mendriq still live. The Mintil meanwhile moved south to the headwaters of the Galas and over the Pahang border into the Tanum headwaters where they remain today in relative isolation (developing unique complex secondary changes in pronunciation).

#### 5.15 *The recent phase*

1) Several languages broke up into dialects, as follows. Since in most cases these dialects still fit into linguistic meshes, only nominal dates of separation can be given.

2) Proto-Sabum split apart into Semnam and Sabum  
oper between 1220 and 965 BP. Sabum later entered  
to close contact with Lanoh Yir.

3) Proto-Kensiu split into Kintaq Bong and Kensiu  
oper between 810 and 590 BP.

4) Proto-Lanoh split into Lanoh Yir and Lanoh  
ngjeng between 760 and 550 BP. Both dialects entered  
to close contact with Temiar and Kensiu.

5) Proto-Semai produced dialects I and II between  
0 and 470 BP, Semai II later entering into close  
contact with Temiar. (There are many Semai dialects,  
and the true picture must be more complex than this.)

6) Since the Pahang Rebellion of the 1890's (AD!)  
Bateg Deq has come into very close contact with Malay,  
hence their very high loan rate.

7) At the present time several ongoing language  
locations can be observed:

a) The retraction of Semai in the southwest  
in favour of Malay, the people probably becom-  
ing members of the Temuan-Belandas tribal group-  
ing.

b) The expansion of Temiar in the northwest  
around Grik as more and more speakers of Lanoh,  
Semnam, Sabum and Jehai drop their original  
languages.

c) The constant shuttling back and forth of  
Mendriq, Bateg Deq, Bateg Nong and Semaq Beri  
over the Kelantan, Trengganu and Pahang borders  
in the east. This is supported both by current  
ethnographic investigation and by careful exam-  
ination of Blagden's comparative vocabulary  
(cf. Appendix III).



d) The probable rapid loss of the Southern Aslian languages in favour of Malay as communications open up and schooling becomes widespread. Young Semelais (like some young Semais) frequently now speak only Malay to each other.

## 5.2 *Cultural correlates of the proposed reconstruction*

There are at least two specific ways whereby the above reconstruction may be cross-checked. From the archaeological point of view certain key 'events' are in principle likely to leave their traces in the soil and hence might one day be open to research. In particular, archaeological dating and placing of the following in the Peninsula would have great bearing on the present investigation: the date of the earliest appearance of a new human physical type that might reasonably coincide with the first arrival there of Austroasiatic speech; the date of the earliest beginnings of agriculture and/or sedentarisation; the date and place of earliest trading activities; the date of earliest colonisation by relatively more civilised foreigners. Even if the absolute dates provided by the archaeologists do not agree with those posited above, it will be possible to regard the language-based reconstruction as receiving support if the two sets of data bear proportionately the same relations to each other. It would then be a simple matter to change the lexical retention rate so as to fit the new evidence; indeed that would provide an excellent means of calculating the true retention rate.

A second line of approach would be through a similar lexicostatistical study on the non-Austroasiatic languages of the Peninsula. Specifically, if the reconstruction of Jakun linguistic prehistory were to show a similar pattern to that of the Aslian language

would further support the ideas presented here.

Indirect confirmation of the main features of this language-based reconstruction is provided by the studies of the cultural evidence for Malayan ethnogenesis that the author is currently undertaking (Benjamin, Forthcoming a,b). On present evidence it seems possible to reconstruct indigenous Malayan culture-history on the assumption that no more than two original distinct cultures (the Negrito, and another as yet unnamed) have been differentially acted on by the agencies of foraging, agriculture, trade and civilisation, in a specifiable sequence and in specifiable locations. The evidence so far amassed for this study indicates a striking degree of parallelism between the patterns of differentiation of the languages, the religious systems and the kinship systems that have evolved within the Peninsula.

Since the original Conference version of this paper was written, further (and more direct) confirmation of its findings has been provided by Solheim's recent archaeology-based reconstruction of Southeast Asian culture-history (Solheim 1970, 1972). The major points of convergence between our two reconstructions are as follows:

a) The Late Hoabinhian culture began to fade out between 10,000 and 8,000 BP in N.E. Thailand, but lasted for a while longer in some other areas. Hence my suggested date of 6500 BP for the earliest sedentarisation in Malaya would correspond well with the emergence of 'distinct Middle Neolithic cultures' by 8000 BP and with the hypothesis that grain cultivation had begun in N.E. Thailand by 7000 BP (Solheim 1970:151-2).

b) Solheim suggests (1970:152) that sea-faring was well developed in Southeast Asia by 6000 BP in response to the post-Pleistocene rise in sea-level, and that the by-now sophisticated metallurgy of copper and bronze had led to the development of trade in some areas by 5000 BP. This corresponds closely to my suggestion that trade had already affected the Proto-Aslians by 5000 BP. These sea-faring 'foreigners' would very likely have been of *Austronesian speech and may well thereby have been the source of the long-recognised Austronesian but non-Malay elements in the Aslian lexicon* (cf. Blagden 1906:435-8).

c) Solheim states (1970:152-3) that the major innovation of the fourth and third millennia B.C. was increased movement of people throughout Southeast Asia, especially all along the interior waterways. This is precisely what my reconstruction shows to have occurred in Malaya from 4340 BP up until the present time.

d) There is a strong possibility that indigenous pre-Khmer urbanisation (? centralisation) in Central Thailand during the first millennium B.C. (Solheim 1970:154) served as a stimulus for the radiation and differentiation of Northern Aslian between 2710 and 2290 BP.

e) Thereafter, Southeast Asia enters the protohistorical and historical eras. The reconstruction presented in this paper suggests that Indianisation, Malay settlement (probably from Sumatra and Borneo) and Mon-Khmer trading-mining contacts began to affect the Aslian speakers from a nominal date of 1970 BP -- an

exact correspondence with Solheim's (1972:41) nominal date of 1 A.D. for the beginning of the 'Conflicting Empires' period in Southeast Asian culture history.

The remarkably close correspondence between the independently pursued archaeology- and linguistics-based reconstructions is noteworthy. Assuming that further research in the field does not upset these conclusions, a purely linguistic result of some significance will be the confirmation that the suggested lexicostatistical retention rate of 86% per millenium is very much closer to the mark for work in Southeast Asia than the more widely used rate of 81%.

3

Obviously, a broad theory has been erected here on the basis of evidence that may not be sufficient to support it. Each language or dialect is represented by only one wordlist--a rather poor showing when compared to Wilkinson's insistence that each dialect be identified by at least three lists. The likelihood of error is fairly high, therefore, and this study must be done over again as soon as more comprehensive data become available. Nevertheless the overall results fit together in a way that suggests they will not be greatly altered after re-study; only in the case of the Maq Beri data is there some reason to suspect significant inaccuracies in the original word-list.

A second objection to the present analysis concerns the premiss upon which it is based, namely, the assumption that the three branches of Aslian recognised here go back coordinately to a single Proto-Aslian language. There is no doubt that there are at least two strata of Austroasiatic in the Peninsula (I have

data that amply confirm Blagden's suspicions on this score). Furthermore, some of the cognacy percentages reported here are rather lower than those quoted by Thomas and Headley (1970:404-5) as being typical of coordinately related subgroups within the Mon-Khmer family. While the Northern-Central Aslian and Central Southern Aslian rates are of the same order as those within Mon-Khmer, the lowest Aslian rates, viz., Northern-Southern Aslian, are of the same order as the cognacy rates between Temiar and both Nicobarese and Mon-Khmer. However, since the intra-Aslian lexical evidence does not support the idea of two separate proto-languages, it must be assumed that differentiation within the Aslian languages began earlier than within Mon-Khmer proper, but to be otherwise of the same order. This would also support the view, made plausible by both geography and simple inspection, that Aslian is more closely related to Mon-Khmer than to any other division of Austroasiatic. Indeed, it may well turn out to belong to Mon-Khmer proper, being that first subgroup to separate out from the common proto-language. (If so, Aslian may provide the best evidence yet for Benedict's 'Linking Austro-Tai' (see his paper in this volume).)

If this view is correct, observation of the mesh principle would lead us to expect that cognacy figures for Aslian vis-a-vis the other Austroasiatic languages would show a clear-cut break with all except the Mon-Khmer group. This latter would show a scatter of figures similar in kind to the scatter observed within the Aslian languages, indicating that Aslian remained in communication with some of the southerly Mon-Khmer dialects long after both groups had begun to split up. Ethnological and linguistic evidence (e.g. Blagden

06:449-452) that a relatively settled agricultural mode of existence was typical of some of the earliest Aslian-speaking groups would support this expectation, especially when confronted with my own suggestion (section 5.11 above) that agriculture and sedentarisation arrived on the scene only after Proto-Aslian had merged in the Peninsula but before complete separation had occurred between Proto-Aslian and the residual Proto-Mon-Khmer speech-community further north.

#### 0 *Suggestions for future research*

Even if the more hypothetical parts of this paper eventually turn out to be unfounded it is to be hoped that it may at least have pointed out the most fruitful directions for future research on the Aslian languages. The most urgent tasks would appear to be the following.

- 1) A thorough survey of the Aslian languages based not on the standard taxonomic categories but on new categories derived from fresh field research. A thorough hunt must be made for hitherto unreported languages, especially those spoken by groups already well recognised ethnologically and administratively but whose speech may well turn out to be other than expected. Special attention should be paid to small rapidly disappearing speech communities, as these may serve to indicate chain-like relationships where none could otherwise be suspected. Areas where such a search would probably be most fruitful are: Southern Maliland, the hills west of Baling in Kedah, the mid-reaches of the Perak river and its tributaries between Perik and Kuala Kangsar, the Kelantan-Trengganu-Pahang border areas, the slopes and foothills of Gunong Benom in central Pahang, the Jakun-Semelai hinterland south-east of Tasek Bera extending from Pahang into north-east Johor, and those parts of Selangor separating

Mah Meri and the Semai outlier from the main body of Aslian languages. Coupled with this should be dialectal surveys of the kind already begun for Semai by Diffloth.

2) For the purposes of historical linguistics, it is clear that the 'outlier' languages Che' Wong, Jah Hut and Mah Meri are crucial for the reconstruction of Proto-Northern, -Central and -Southern Aslian respectively. At the other end of the scale, attention should be paid to apparently insignificant dialects in the reconstruction of more recent stages; for example, the Lanoh and Sabum dialects are essential for a study of the history of Temiar.

3) From a more practical point of view, straightforward descriptive studies are needed; these would provide excellent topics for research by students of Linguistics at the University of Malaya. With data on Kintaq Bong, Temiar and Semai already collected (if not published) the following additional languages would seem to offer the widest scope in terms of practical accessibility to informants and of maximum diversity in structure: Mendriq, Che' Wong, Jah Hut, Semelai, and Mah Meri.

4) Research on secondary materials should not be neglected, however. It is clear from this paper that further lexicostatistical studies (especially with other Austroasiatic languages) and further exploration of the ethnological implications of language movements would be very useful.

## REFERENCES

- mah Haji Omar. 1964. *Bahasa Semang: dialek Kentakbong*. Unpublished research exercise, Department of Malay Studies, University of Malaya, Kuala Lumpur.
- \_\_\_\_\_. 1973. *The verb in Kentakbong*. Paper presented at First International Conference on Austroasiatic Linguistics, Honolulu.
- Benedict, P.K. 1973. *Austro-Thai and Austroasiatic: the 'Austro linkage'*. Paper presented at the First International Conference on Austroasiatic Linguistics, Honolulu.
- Benjamin, G. 1968. Headmanship and leadership in Temiar Society. *Fedn Mus. J. (N.S.)* 13:1-43.
- \_\_\_\_\_. 1973. *An outline of Temiar grammar*. Paper presented at First International Conference on Austroasiatic Linguistics, Honolulu.
- \_\_\_\_\_. M.S. Unpublished field notes on Negritos of Kelantan (1970) and of Kedah and Perak (1972).
- \_\_\_\_\_. Forthcoming, a. *Indigenous Religious systems of the Malay Peninsula*. Working Paper, Dept. of Sociology, University of Singapore.
- \_\_\_\_\_. Forthcoming, b. *Indigenous kinship systems of the Malay Peninsula*. Working Paper, Dept. of Sociology, University of Singapore.
- Blagden, C.O. 1906. *Language and Comparative vocabulary of Aboriginal dialects*: In Skeat and Blagden 1906, vol. II, pp. 379-472; 481-775.
- Brandt, J.H. 1961. The Negrito of Peninsular Thailand. *J. Siam Soc.* 18:
- Carey, I. 1961. *Tengleq kui serok: a study of the Temiar language with an ethnographical summary*. Kuala Lumpur: Dewan Pustaka dan Bahasa.
- \_\_\_\_\_. 1970. The Kensiu Negritos of Baling, Kedah. *J. Malays. Brch R. Asiat. Soc.* 43:
- Collings, H.D. 1949a. A Temoq word list and notes. *Bull. Raffles Mus. Ser. B* 4:69-85.
- \_\_\_\_\_. 1949b. Aboriginal notes. *Bull. Raffles Mus. Ser. B* 4:86-94. (Nos. 2-5 on Semelai.)
- Dentan, R.K. 1964. *Senoi-Semang*. In Lebar et.al. 1964, pp. 176-186.



- Dentan, R.K. 1968. *The Semai: a non-violent people of Malaya*. New York: Holt, Rinehart and Winston.
- \_\_\_\_\_. M.S. *An outline of Semai grammar*. (N.d., duplicated.)
- Diffloth, G. 1968. Proto-Semai phonology. *Fedn Mus. J. (N.S.)* 13:65-74.
- \_\_\_\_\_. 1971. Ambiguïté morphologique en semai. In J. M. C. Thomas and L. Bernot (Eds), *Langues et techniques, nature et société*. Paris: Klincksieck, pp. 91-93.
- \_\_\_\_\_. 1973. Minor-syllable vocalism in Senoi languages. Paper presented at the First International Conference on Austroasiatic Linguistics, Honolulu.
- Evans I. H. N. 1927. *Ethnology and archaeology of the Malay Peninsula*. Cambridge: University Press.
- \_\_\_\_\_. 1937. *The Negritos of Malaya*. Cambridge: University Press.
- Gleason, H.A. 1959. Counting and calculating for historical reconstruction. *Anthropological Linguistics* 1:pt 2:22-32.
- Gudschinsky, S. 1956. The ABC's of lexicostatistics (glottochronology). *Word* 12:175-210.
- Hoe Ban Seng. 1964. *Report on Aboriginal communities at Tasek Bera, Southwest Pahang: a study on the structure of Semelai society*. Unpublished research report, Dept. of Social Work, University of Singapore.
- Jimin bin Idris. 1968. Distribution of Orang Asli in West Malaysia. *Fedn Mus. J. (N.S.)* 13:44-48.
- Landar, H. 1966. *Language and culture*. New York: Oxford University Press.
- LeBar, F.M. et al. 1964. *Ethnic groups of mainland Southeast Asia*. New Haven: HRAF Press.
- Needham, R. 1956. Ethnographic notes on the Siwangs of central Malaya. *J. Malay. Brch R. Asiat. Soc.* 29:49-69.
- \_\_\_\_\_. 1960. Research projects in Southeast Asia. *Bull. Int. Com. Urgent Anthropol. Ethnol. Res.* 3:63-71.
- Noone, H.D. 1936. Report on the settlements and welfare of the Ple-Temiar Senoi of the Perak-Kelantan watershed. *J. Fed. Malay St. Mus.* 19:1-85.

- Alvie, C.S. 1949. Che' Wong word list and notes. *Bull. Raffles Mus. Ser. B* 4:11-43.
- Arnold, H.-J. 1959. *Versuch einer historischen Lautlehre der Kharia-Sprache*. Wiesbaden: Otto Harrassowitz.
- Arundin, I. M.S. Unpublished field notes on Jah Hut ethnography.
- Bebesta, P. 1926. The jungle tribes of the Malay Peninsula. *Bull. Sch. Or. Afr. Stud.* 4:269-278.
- \_\_\_\_\_. 1927. *Among the forest dwarfs of Malaya*. London: Hutchinson.
- \_\_\_\_\_. 1928. Grammatical sketch of the Jahai dialect. *Bull. Sch. Or. Afr. Stud.* 4:803-826.
- \_\_\_\_\_. 1931. Grammatical sketch of the Ple-Temer language. *J.R. Asiat. Soc.* 1931:641-652.
- \_\_\_\_\_. 1952. *Die Negrito Asiens*. Vol. I. Modling: Anthropos Institute.
- \_\_\_\_\_. 1954. *Ibidem*. Vol. II part 1.
- \_\_\_\_\_. 1957. *Ibidem*. Vol. II part 2.
- Brundage, P.W. 1901. Die Sprachen der Sakei und Semang auf Malacca und ihr Verhältniss zu den Mon-Khm̄er-Sprachen. *Bijd. Taal-, Land-, Volkenk.* 8:401-583.
- Brundage bin Yub.
- Brundage, W.W. and C.O. Blagden. 1906. *Pagan races of the Malay Peninsula*. 2 vols. London: MacMillan.
- Brundage, W.G., II. 1970. Northern Thailand, Southeast Asia, and world prehistory. *Asian perspectives* 13:145-62.
- \_\_\_\_\_. 1972. An earlier agricultural revolution. *Scientific American* 226: no. 4:34-41.
- Brundage, M. 1959. The mesh principle in comparative linguistics. *Anthropological Linguistics* 1:pt 2:7-14.
- Brundage, O. D. 1914. Versuch einer Sakai Grammatik mit Vokabularium. *Anthropos* 9:529-538.
- Brundage, D. 1960. Basic vocabulary in some Mon-Khmer languages. *Anthropological Linguistics* 2:pt 3:7-11.
- \_\_\_\_\_. and R.K. Headley. 1970. More on Mon-Khmer subgroupings. *Lingua* 25:398-418.

- Voegelin, C.F. and F.M. Voegelin. 1966. Languages the world: Indo-Pacific fascicle eight. *Anthropological Linguistics* 8:part 4.
- Wilkinson, R.J. 1910. *The Aboriginal tribes*. Papers on Malay Subjects: Supplement. Kuala Lumpur: F.M.S. Government Press.
- \_\_\_\_\_. 1915. *A vocabulary of Central Sakai*. Papers on Malay Subjects, 2nd Series, no. 3. Kuala Lumpur: F.M.S. Government Press.
- Williams-Hunt, P.D.R. 1952. *An introduction to the Malayan Aborigines*. Kuala Lumpur: Government Press.

# Appendix I

## Comparative Orang Asli Taxonomy

Group	Language	Physical/ Administrative	Traditional mode of subsistence	Culture- type (Schebesta)
Asi	Northern	Negrito	Foragers	Semang
Ataq	"	"	"	"
Ag	"	"	"	"
Ai	"	"	"	"
Adriq	"	"	"	"
Aeg	"	"	"	"
Al	"	"	"	"
Atil	"	"	"	"
Aeg	"	"	"	"
Ag	"	"	"	"
A'	"	"	"	"
Ag	"	Senoi	Farmers	Jakun
Am/	"	"	"	"
Anoh	Central	Negrito	Foragers	Semang
Aniar	"	Senoi	Farmers	Senoi
Anai	"	"	"	"
An Hut	"	"	"	Jakun
An	"	"	"	"
Anri	Southern	"	Farmers, Fishers	"
Anaq	"	"	"	"
Anri	"	Mixed	Foragers	"
Anelai	"	Proto-Malay	Farmers, Fishers	"
Anoq	"	"	Foragers	"

## Appendix II

### Aslian Comparative Vocabulary

**Orthography:** This follows the system set out in the phonemic inventory below.

#### 1. Consonants

	labial	dental	palatal	velar	glot
voiced stops	b	d	j	g	
voiceless stops	p	t	c	k	
fricatives	f,v		s,z		
semivowels	w		y		
nasals	m	n	ɲ	ŋ	
lateral		l			
flap		r			
laryngeal					ʔ,

Prenasalised and aspirated consonants are written as clusters: nt-, ɲj-, kh-, hm-, hŋk-, etc.

2. *Vowels:* All Aslian languages possess a 3 X 3 vowel system:

i	u	u
e	ə	o
ɛ	a	ɔ

Phonemic (double-) length is indicated by a macron: ā, ū, ē.

Phonemic nasality is indicated by a hook: ɛ̃, ʔ̃, ɔ̃.

#### *The Vocabulary*

The following conventions are employed in the word-lists.

- Words adjudged to be cognates are enclosed with the same pair of oblique slashes / /.
- Loanwords are enclosed at the end of the entry in square brackets [ ].

Where no Aslian word is reported for any language is indicated within the final square brackets by hyphen following the relevant language abbreviation.

The following language abbreviations are used and in the following order: Ks Kensiu, KB Kintaq Bong, Je Jahai, Mr Mendriq, BD Bateg Deq, Mt Mintil, BN Bateg Deq, CW Che' Wong, Sn Semnam, Sa Sabum, LJ Lanoh, LY Lanoh Yir, Tm Temiar, SmI Semai I, SmII Semai II, JH Jah Hut, MM Mah Meri, SB Semaq Beri, S1 Semelai, Tq Temoq.

1 ANIMAL: Ks ?ay, KB ?ay, Je ?ay, Mr ?ay, BN ?ay, LJ ?ay, LY ?ay / CW bənəh, Tq mənə? / Sn juko?, Sa juko? / Tm nam / Sm II mənhar / S1 pa?la? / [BD mənatan, Mt mənatan, SmI margās, JH binatan, MM bənatan, SB mənatan]

2 BACK: Ks kiyo?, KB keyo?, Je kərə?, Mr kəro?, BD kəro?, Mt keyow?, BN kəro?, Sn keyo?, Sa keyo?, LJ keyo?, LY kərɔ?, Tm kərɔ?, JH kəro? / CW həkən / SmI cəlōt, SmII cəlōd, MM jilət, SB cələn, S1 cələn / Tq jə?ak

3 BAD: Ks bəlog / KB dəyəm / Je la?əs, LJ la?əs, LY la?əs, Tm la?əs / Mr jəbec, BD jəbec, Mt jəbej, JH ja?buc / BN jalan / Sn tɔ? bə?əh / Sa papa? / SmI pə? bōr / SmII nec / MM bajaw / [CW məhə? n ?elo?, SB beh ?elo?, S1 da? ?elo?, Tq ta? mələ?] /

4 BELLY: Ks ?ej, KB ?ej, Je ?ɛj, Mr ?ɛc, CW ?ac, Sn ?ej, LJ ?ej, LY ?ej, Tm ?ej, SmII ?ej, MM ?o?əc / BD cən / Mt kəwt, BN kət, Sa kət, SmI kət, JH kūt / SB ləpəc, S1 ləpəc, Tq ləpəc

5 BIG: Ks tabə?, KB təbə? / Je cəkey, Sn cəkey, LJ cəkey, LY cəkey / Mr bəw, BD bəw, Mt bəw, BN bəw / CW mənə?, Tm mənū?, JH mənā? / SmI ntōy, SmII ntōy, S1 they, Tq nthoy / MM kadam / SB pa?əh /

[Sa ?ibū?]

- 6 BIRD: Ks kawaw, KB kawaw, Je kawəd, Mr kawaw, BD kawaw, Mt kawaw, BN kawaw, CW kawaw, SB kawəh / Sn cim, Sa cim, LJ cəp, LY cəp, Tm cəp, SmI cēp, SmII cēp, JH cēp, MM cim, S1 cim, Tq cip
- 7 BITE: Ks kap, KB kap, Je kam, Mr kap, BD kap, Mt kap, BN kap, CW kap, Sn kap, Sa kab, LJ kap, LY kab, Tm kab, SmI kap, SmII kap / MM negək / SB gan, S1 gingən / [JH gigit, Tq - ]
- 8 BLOOD: Ks məhəm, KB məhəm, Mr bəhəm, Mt məhəm, BN məhəm, CW məhəm, SmI bəhTp, SmII bəhTp, JH nəhīm MM maham, SB maham, S1 maham, Tq maham / BD yap Tm lōt / [Je darah, Sn dayah, Sa dayah, LJ darah LY darah]
- 9 BLOW: Ks pət, KB pət, Mr pət, BD pət, Mt pəwt, BN pət, CW pət, JH pət / Je pəhos / Sn təhōl, Sa təhōl, LJ təhōl, LY təhōl, Tm təhōl, SmI təhōl, SmII təhōl / MM ?əh, SB ?əh, S1 ?əh?əh, Tq bər?
- 10 BONE: Ks ?i?iŋ, KB ?i?eŋ, Je jə?iŋ, Mr jə?ik, BN ji?iŋ, CW jə?eŋ, Sn jə?aŋ, Sa jə?aŋ, LJ jə?eŋ, I jə?aŋ, Tm jə?āk, SmI jə?āk, SmII jə?āk, JH jə?aŋ MM jə?ak, SB jə?aŋ, S1 jə?aŋ, Tq jə?aŋ / [BD tulaŋ, Mt toylan]
- 11 BREAST: Ks ?am, KB ?am, Je ?əm, Mr ?am, BD ?am, Mt ?am, BN ?am / CW bu?, JH bō? / Sn mēm, Sa mēr LJ mēm, LY mēm, SmI mēm, SmII mēm / Tm bōt / MM tuh, SB tuh, S1 tuh, Tq tuh
- 12 BREATHE: Ks jəgjug, Je nəgjug, BN nəgjək, Sn nəgjug / Mr nəphəp, Mt nəfhawf, LJ hənum, LY hu Tm hənum, SmI bər-ləhəm, SmII pərhəm, JH pəŋhəm SB ?amhəm, S1 ?amhəm, Tq ?ahəm / Sa həghəg / MM hanəy / [KB -, BD napas, CW bənapas]

CHILD: Ks wōŋ, KB wōŋ, Je wōŋ, BN wōŋ, CW wōŋ /  
Mr kəwəl / BD ʔawəʔ, Mt ʔawəʔ, JH ʔewaʔ / Sn kōn,  
Sa kōn, LJ kəwōn, LY kəwōn, Tm kəwət, SmI kənōn,  
SmII kənōn, MM kənōn, SB kənōn, Sl kənōn, Tq kənōn

CLAW (NAIL): Ks kalkəʔ, KB kəlkaʔ, Mr kəlkaʔ, BD  
kalkəʔ, Mt kəlkaʔ, BN kəlkaʔ, CW təlkaʔ, SB təkəʔ  
/ Je cənros, Sn cənycs, Sa cənycs, LJ cērōs, LY  
cənros, Tm cənōs, SmI cənōs, SmII cənōs, JH  
cərēs, Sl cērūs, Tq ceros / [MM kokənt]

CLOUD: Ks ʔal, KB ʔal, CW ʔal / Je sagub, Sa sagub,  
LJ sagub, LY sagub, Tm sagub, SmI sagub, SmII  
sagub, SB sagub / BD mənʔəʔ / [Mr kabud, Mt kabut,  
BN kabot, Sn kabud, JH kabud, MM kabut, Sl kabut,  
Tq -]

CUT: Ks gət, KB gid, Mr git, Mt gəwt, Sn gēd, Sa  
gəd, LJ gəd, LY gəd, Tm gəd / Je gey / BD kəc, BN  
ket, CW kac, JH kēc / SmI koh, SmII koh / MM katik  
/ SB kəŋ / Sl təkəl, Tq təkəl

DANCE: Ks siseʔ, Je siseʔ, LJ siseʔ, LY siseʔ, Tm  
siseʔ, SmII ʔaseʔ / KB keʔeh, Mr kərjer, Sn kirjer  
/ BD kənsəŋ, CW kənsəŋ / BN təgah / Sa tənʔoh /  
[Mt -, SmI bəŋŋin, JH mənariʔ, MM mənariʔ, SB  
mənariʔ, Sl bənariʔ, Tq -]

DAY: Ks cəhey, KB cəhey, Mr cəhoy, Sn cəhōs / Je  
kətoʔ, BD kətoʔ, BN kətoʔ, CW kətoʔ, JH kətōʔ /  
Sa yis, LJ yTs, LY yis, Tm ʔis, SmI jis / SB  
təŋiʔ, Tq təŋiʔ / [Mt haiy, SmII hariʔ, MM ʔaiʔ,  
Sl ʔariʔ ]

DIE: Ks kəbis, KB kəbis, Je kəbis, Mr kəbəs, Mt  
kəbəs, BN kəbəs, CW kəbəs, Sn kəbəs, Sa kəbəs,  
LJ kəbəs, LY kəbəs, Tm kəbəs, JH kəbəs, MM kəbəs,  
SB kəbəs, Sl khəbəs, Tq kəbos / BD halot / SmI  
ndat, SmII dat



- 20 DIG: Ks bay, KB bay, Je bay, Mr bay, BD bay, Mt bay, BN bay, CW bay, Sn bay, Sa bay, LJ bay, LY bay, SB bay, Sl bay, Tq bay / Tm pæg / SmI cōp, SmII cōp, JH cəwɔm, MM kacup
- 21 DIRTY: KB kamah, BD kamah, Mt kamah, BN kamah, Sa kamah, LY kamah / Je jəʔɔm / LJ gəcəh Tm laʔəs / SmII nec / JH gəʔaj / SB jaʔjiʔ, Sl jaʔjiʔ, Tq jaʔjiʔ / [Ks -, Mr -, Sn -, SmI koton MM kətɔh]
- 22 DOG: Ks ʔiʔ, KB ʔeʔ / Je ʔacɪʔ, Sn ʔacēh, Sa cōʔ LJ ʔacēʔ, LY ʔacəʔ, Tm cəwɔʔ, SmI cōʔ, SmII coʔ, JH cuwɔʔ, MM caw, SB cōh, Sl cōh, Tq cōw / Mr ʔasuʔ, Mt ʔasowʔ / BN wəʔ / CW ʔeŋ / [BD ʔapen]
- 23 DRINK: Ks buʔ, KB buʔ / Je ʔəm, Mr ʔam, BD ʔam, ʔam / Sn ʔɔŋ, Sa ʔɔŋ, LJ ʔɔŋ, Tm ʔɔk / CW ʔuh, hɔh, JH woh, SB jaʔoh, Sl jəʔoh, Tq jəʔoh / SmI ŋɔt, SmII ŋɔt / MM cadow
- 24 EAR: Ks ʔəntiŋ, KB ʔənten, Je ʔənten, Mr ʔənten BD ʔənten, Mt ʔənten, BN ʔənten, CW hnten, Sn kəntɔk, Sa gəntɔk, LJ gəntɔk, LY gəntɔk, Tm gən SmI ntak, SmII ntak, JH ntaŋ, MM tən, SB ntən, tən, Tq ntən
- 25 EARTH: Ks tiʔ, KB teʔ, Je teʔ, Mr teʔ, BD teʔ, teʔ, BN teʔ, CW tɛʔ, Sn tēʔ, Sa tɛʔ, LJ tēʔ, LY tɛʔ, Tm tɛʔ, SmI tɪʔ, SmII teʔ, JH tēʔ, MM teʔ, SB ʔatɛh, Sl ʔatēh, Tq ʔatey
- 26 EAT: Ks ciʔ, KB ciʔ, Mr ciʔ, BD ciʔ, Mt cəiʔ BN ciʔ, CW cəyɛʔ, Tm cāʔ, SmI cāʔ, SmII cāʔ, JH cā MM nacāh, SB ncāh, Sl cāh, Tq cah / Je gey, Sn gey, Sa gey, LJ gey, LY gey
- 27 EGG: Ks makɔʔ, KB makɔʔ, Mr makɔʔ, BD makɔʔ, Mt makɔwʔ, BN makɔʔ / Je kətɛd, CW kətwɔd, Sn kətɛ

Sa tab, LJ tab, LY tap, Tm tab / SmI pəŋl̄ək, SmII pəŋl̄ək, JH pəŋləŋ / MM kəpəh, SB kapəh, S1 kapəh, Tq kapəh

EYE: Ks məd, KB mət, Je mid, Mr mət, BD met, Mt meit, BN met, CW ment, Sn mad, Sa mad, LJ mad, LY mad, Tm mad, SmI mat, SmII mad, JH man, MM ment, SB mad, S1 mət, Tq mōnt

FALL: Ks kəpəh / KB suwet, Sn səwəd / Je gibal / Mr rəs, CW rəwas / BD pəp, Mt pəwp, BN pəp / Sa kələ?, LJ kələ?, LY kələ?, Tm kələ? / SmI jēr / SmII təgəh / JH cərəh / MM gəwəc / SB gərək, S1 gərək / Tq sabəh

FAR: Ks mənji?, KB mənji?, Je mənə? ji?, Mr mənji?, BD mənji?, Mt mənji?, SmI nā?, SmII na? / BN cən, CW cən / Sn ləyə?, Sa ləyə? / LJ jərə?, LY jərə?, Tm jəro? / MM ləp, S1 ləp, Tq ləp / SB ŋay / [JH lajū?]

FAT (GREASE): Ks təc, KB təc, Mr təc, BD təc, Mt təc. BN təc / Je lə?əs, Sn lə?əs, Sa lə?əs, LJ lə?əs, LY lə?əs, Tm lən?əs / CW bəcə?, SmI bəncə?, SmII məncək / MM ləp / [JH gənəmə?, SB gəmə?, S1 gəmək, Tq gəmu?]

FATHER: Ks ?ey. KB ?ey, Je ?ey, Mr ?ey, Mt ?ey, BN ?ey / Sn dō?, Sa do?, LJ dō?, LY do? / Tm bəh / SmI mənə?, SmII ?abə?, JH ?ibe? / MM wə? / [BD pa?, CW yah, SB bapa?, S1 bapa?, Tq ?ayah]

FEAR: Ks ?əntəŋ, KB ?əntəŋ, BD ?əntəŋ, Mt ?əntəŋ, BN ?əntəŋ, CW hntəŋ, Sn tūŋ, Sa tuŋ, LJ tūŋ, LY tuŋ, Tm tūk, SB bəthəŋ, S1 bəthəŋ, Tq bəthəŋ / Je həgig / Mr ?əpac / SmI səŋəh, SmII səŋəh / JH bəhec / [MM gəli?]

FIRE: Ks ?əs, KB ?əs, Je ?əs, Mr ?əs, BD ?əs,

Mt ?os, BN ?os, CW ?os, Sn ?os, Sa ?os, LJ ?os,  
LY ?os, Tm ?os, SmI ?os, SmII ?os, JH ?os, MM  
?us, SB ?os, S1 ?us / Tq ?a?uh

35 FISH: CW kəyə?, Sn k̄ā?, Sa ka?, LJ ka?, LY ka?,  
Tm k̄ā?, SmI k̄ā?, SmII k̄ā?, JH k̄ā?, MM k̄āh, SB k̄ā  
/ S1 cəreh, Tq cəreh / [Ks ?ikan, KB ?ikan, Je  
?ikan, Mr ?ikan, BD ?ikan, Mt ?ikan, BN ?ikan]

36 FLOW: Ks wid, KB wət, Je wid, Sn wid, Sa wid, LJ  
wid, LY wid, Tm wid / Mr ?ayyt, BD ?ayyt, Mt  
?eyowŋ / BN təlɔh / SmI gəl, SmII gəl, SB gəl /  
Tq kəjbɔj / [CW məŋilir, JH məleleh, MM ?arus,  
S1 jərəm]

37 FLOWER: Mr bəkaw, BN bəkaw, CW bəkaw, Sn bəkaw,  
Sa bəkaw, LJ bəkaw, LY bəkaw, JH bokaw, SB bəkaw  
S1 bəkaw, Tq bəkaw / [Ks buŋa?, KB buŋa?, Je bon  
BD bonə?, Mt bonə?, Tm bon̄ā?, SmI bon̄ā?, SmII  
bonə?, MM bonə?]

38 FOOT: Ks can, KB can, Je can, Mr can, BD can, M  
can, BN can, CW cən / Sn juk, Sa juk, LJ juk, LY  
juk, Tm juk, SmI juk, SmII juk, JH jɔŋ, MM jək,  
SB jɔŋ, S1 jɔŋ, Tq jɔŋ

39 FRUIT: Ks kəbə?, KB kəbə?, Je kəmɔ?, Mr kəbə?,  
kəbə?, Mt kəbə?, BN kəbɔ?, Sa kəmɔ?, LJ kəmɔ?,  
kəmɔ?, Tm kəbə?, SmII kəb̄ē?, SB kəbə? / CW pəlɔ  
SmI pəl̄ē?, JH pələ?, MM pəl̄ēh, S1 pəl̄ēh, Tq pəl̄  
/ [Sn bɔh]

40 FULL: Ks habun / Je səbim, Mr səbim / Mt cəwəh,  
BN cəwəh / CW bək, Sn təkik, Sa təkik, LJ təkik  
LY təkik, Tm təkik, SmI tək̄ēk, SmII tək̄ēk, SB  
tək̄ēŋ / MM hməc, S1 səb̄ēŋ, Tq səbɔc / [KB -, BD  
pənuh, JH pənɔh]

GIVE: Ks ?ek, KB ?ek, Je ?eg, Mr ?ek, BD ?ek, Mt  
?eik, BN ?ek, CW ?ak, Sn ?og, Sa ?og, LJ ?og, LY  
?og, Tm ?og, SmI ?ōg, SmII ?ōk, JH ?ok / MM kajət,  
SB jon / S1 ha?gi?, Tq hagi?

GOOD: Ks bəd?əd, KB bad?əd, Je bəd?əd, BD bəd?ət,  
Mt bət?əut, BN bə?ət, Sn bə?ət, Sa bə?ət, LJ  
bə?ət, LY bə?ət / Mr ?abən / Tm məj / SmI bōr,  
SmII bōr / JH ?agan / MM ləp / [CW ?ilo?, SB  
?ilo?, S1 ?ilok, Tq mōle?]

HAIR (of head): Ks sək, KB sək, Je sөг, Mr sək,  
BD sək, Mt sauk, BN sək, CW sək, LJ sog, LY sog,  
Tm sōg, SmI sōk, SmII sōk, JH sək, MM suk, SB sup,  
S1 suk, Tq suk / Sn sintəl / Sa jamөг

HAND: Ks cas, KB cas, Je ciyas, Mr cas, BD cas,  
Mt ciyas, BN ciyas, CW cas / Sn tTη, Sa tTη, LJ  
tTη, LY tTη, Tm tTk, SmI tēk, SmII tək, JH tiη,  
MM thT?, SB tēh, S1 thT?, Tq thih

HE: Ks bə?, KB bə? / Je ?o?, Mr ?u?, BD ?o?, Mt  
?ou?, BN ?o?, CW ?uh / Sa ?a-, LJ ?a-, LY ?a- /  
Tm na- / SmI ki-, SmII kē?, MM hηki?, S1 kēh / JH  
yēh / SB he? / [Sn -, Tq -]

HEAD: Ks kuy, KB kuy, Je kuy, Mr kuy, BD kuy, Mt  
kəuy, BN kuy, CW kuy, Sn kuy, Sa kuy, LJ kuy, LY  
kuy, Tm kuy, SmI kōy, SmII kuy, JH kōy, MM khoy,  
SB khoy, S1 khōy, Tq khoy

HEAR: Ks kəjən, KB kəjən, Je kəjiη, Mr kəjiη, BD  
kəjiη, Mt kəjən, BN kəjiη, CW kəjən, Sn yək, Sa  
yək, LJ yək, LY yək, Tm kəyək, MM kayək, SB pəyən,  
S1 nyən / SmI cərηāy, SmII cərηay, JH tiηoy /  
[Tq dənən]

HEAVY: Ks hənjut, KB hənut, Je hənjut, Mr hənjut,  
BD hənjut, Mt hənjəut, BN hənut, CW hnut / Sn

ʔəneh, Sa ʔəneh, LJ ʔəneh, LY ʔənih, Tm ʔənuh,  
SmI nəh, SmII nəh, MM hɲəh, SB kəjəh, S1 kəjəh,  
Tq kəjuh / JH həl̥ə̃k .

49 HERE: Ks dɔh, BN dəh, CW doʔ, Tm doh, SmI deh,  
SmII deh, JH dɔ̃h / KB ʔəh, Je ʔəh, Mr ʔəh / BD  
dɔy / Mt həʔ, MM hoʔ / Sn nəh, Sa nəh, LJ nəh,  
LY nəh / SB haʔnaʔ, S1 noʔ / [Tq -]

50 HOLD: Ks cəkam, Mr kəm, Sn kōm, LJ kəwōm, LY  
kəwɔm, Tm kəwɔp, SB kəm / KB cɛp, BN cɛp, SmI ca  
JH cēp / BD bɔt, Mt bɔut / SmII kɔ̃d / MM kagul /  
S1 yɔk / [Je ʔaŋked, CW pəgaŋ, Sa pəgak, Tq -]

51 HUSBAND: Ks gəsuy, KB kəsuy, Je kəsuy, Mr kəsuy,  
Mt kəsuy, SmI gəns̥Tr, SmII ɲəns̥Tr, JH kəs̥Tr /  
BD tɛmkal / BN tiʔ, CW teʔ, Sa teʔ / Sn ʔəŋkon /  
LJ tɔw, LY tɔw, Tm tɔw / MM ləmɔl / SB ʔipəʔ / S  
kənlek, Tq kəŋlɔŋ

52 I: Ks yeʔ, KB yeʔ, Je yeʔ, Mr yeʔ, BD yeʔ, Mt ye  
BN jɛʔ, Tm yēʔ, SB yēh / CW ʔiŋ, Sn ʔiɲ, Sa ʔiŋ,  
LJ ʔic, LY ʔic, SmI ʔiŋ, SmII ʔəp, MM ʔəʔəp, S1  
ʔəp, Tq ʔaʔɔc / JH ʔihah

53 KNEE: Ks kalton, KB kalton, Je kalton, Mr kaltur  
BD kalton, Mt kaltɔn, BN kalton, CW kəlton, LY  
kalton, SB kəlton, S1 kaltun, Tq kəlton / Sn  
kayɔl, Sa kayɔl, LJ karɔl, Tm karɔl, SmI korɔl,  
SmII kurul, JH kərowal / [MM lutut]

54 KNIFE: Ks haj, KB haj / Mr wɛp, BD wɛp, Mt weɲp,  
BN wɛp, CW wap, Tm ʔəwɔj, SmI yɔ̃j, SmII yɔj, SB  
way, S1 wɔy, Tq way / Sn pɛd, Sa pɛn / JH ɲɛm /  
MM haw / [Je bɛsiʔ, LJ bɛsiʔ, LY bɛsiʔ]

55 KNOW: Ks həyɛb, KB həyɛb / Je ʔed, Mr lɛdʔet, M  
sinet / BN jahay, CW jəhay / Sn l̥Tɔp, Sa lip, LJ  
l̥Tɔp, LY lip / Tm lɛk / JH sɛr̥ə̃k, MM sɛrəʔ / SB

ke?, S1 ŋkhe?, Tq hake? / [BD tahu?, SmI paney,  
SmII mpaney]

LAUGH: Ks luk, KB lug, Je lug, Mr luk, Mt læug,  
BN lok, CW gewilək, Sn pəglūg, Sa lug, LJ lug,  
LY lug, Tm lug, SmI luk, SmII luk, JH lok, MM  
gələk, SB gələk, S1 gələk, Tq gələk / BD pilŋal

LEAF: Ks hali?, KB həli?, Je hali?, Mr hali?,  
BD hali?, Mt haliy?, CW hale?, Sn səlā?, Sa səla?,  
LJ səla?, LY səla?, Tm səlā?, SmI səlā?, SmII  
səlā?, JH hla?, SB salāh / [BN dawən, MM dawun,  
S1 dawun, Tq dawun]

LEFTSIDE: Ks ?əyil, CW yal, Sn yel, Sa wel, LJ  
yel, LY yel, Tm yel, SmI kəŋwTl, SmII wTl, SB  
sawel, S1 sawil, Tq sawel / KB wɨ?, Je wɨ?, Mr  
we?, Mt wɨy?, BN wə? / [BD kiri?, JH kiri?, MM  
kiri?]

LIVE: Ks gəs, KB gəs, Je gəs, Mr gəs, BD gəs, Mt  
gous, BN gəs, CW gəs, Sn gəs, Sa gəs, LJ gəs, LY  
gəs, Tm gōs / SmI rēs, JH rēs, MM rih, S1 ris,  
Tq ?aris / SmII suy / SB loh

LIVER: Ks kəlaŋis, KB kəlaŋis, Je kəlaŋis, Mr  
kəlaŋəs, BD kəlaŋis, Mt kəlaŋəs, BN kəlaŋəs / CW  
ros, LJ rTs, SmI rTs, JH rēs, MM gərih, SB gərəh,  
S1 gəris, Tq gərəs / Sn huf, Sa hup, LY hup,  
Tm hup / SmII nōs

LONG: Ks məŋji?, KB məŋji? / Je bətej, BD bətec,  
Mt bətec, BN bətəc / CW nisəŋ / Sn jərəē?, Sa jərə?  
LJ jərəē?, LY jərə?, Tm jəro? / SmI cərək, SmII  
cərək, JH cərəŋ / MM jələŋ, SB jələŋ, S1 jəlūŋ,  
Tq jələŋ / [Mr -]

LOUSE: Ks ci?, Je ci?, Mr ci?, BD ce?, Mt cei?,  
BN ce?, CW cə?, Sn ce?, Sa ce?, LJ cē?, LY ce?,

Tm cε?, SmI cē?, SmII cē?, JH cē?, MM cTh, SB cT  
S1 cTh, Tq cih / [KB -]

63 MALE: Ks tēmkal, KB tēmkal, Je tēmkal, Mr tēmkal  
BD tēmkal, Mt tēmkal, BN tēmkal, CW tūŋkal, SmI  
kērāl, SmII kērāl / Sn ?əŋkoŋ, Sa ?əŋkoŋ, JH  
kəra?koŋ / LJ tē? / LY tōw / Tm babēh / MM lēmōl  
S1 rēmōl, Tq rēmōl / SB ?ipə?

64 MANY: Ks jēnuh, SB bəjəh / KB balə?, JH ba?lɛ? /  
Je bēnoled / Mr kōm, BD kōm, Mt kēum / BN bōŋ /  
Sn ləbēh, Sa ləbēh, LY ləbēh, Tm ləbēh / LJ  
bə?asTl / SmI jə?oy, SmII jə?oy / MM hnōm / S1  
rəy / [CW bəso?, Tq -]

65 MEAT: Ks sej, KB sej, Je sej, Mr sec, BD sec, Mt  
sēic, BN sec, CW sac, Sn sej, Sa sej, LJ sej, LY  
sej, Tm sej, SmI sec, SmII sec, JH sec, SB ŋsec,  
S1 sec, Tq sōc / [MM dagik]

66 MOON: Ks gēcī?, KB kēce?, CW kēcε?, LJ gēcē?, Tm  
gēcε?, SmI gēcē?, SmII gēcē? / Mr bəlɛp / [Je  
bolan, BD bolan, Mt bōulan, BN bolan, Sn bolan,  
Sa bolan, LY bolan, JH bulan, MM bulan, SB bolan  
S1 bulan, Tq bolan]

67 MOTHER: Ks na?, KB na?, Mr na?, BD na?, Mt na?,  
Sa na? / Je bə?, BN bō?, Tm bō? / Sn nō?, LJ nō?  
LY nō? / SmI ?amē?, SmII ?ame? / JH ?ide?, MM  
gade?, Tq gado? / SB mēy / [CW ma?, S1 ma?]

68 MOUNTAIN: Ks cəba?, CW cəba? / Je jəlmōl, LJ  
jəlmōl, LY jəlmōl, Tm jəlmōl / Mr tol, Mt tēul,  
BN langōn / Sn pō?, Sa pō? / SmI lōt / SmII cənə  
/ SB bənēm, S1 bənēm, Tq bənōm / [KB -, BD gunōl  
JH gunōŋ, MM gunōŋ]

69 MOUTH: Ks hap, KB hap, BD hap, CW hōŋ / Je tənə  
Mr tənət, Mt tənət, BN tənət, Sn tənTd, SB kənə

Tq kənūt / Sa nag, LJ nāg, LY nag, Tm nāg / SmI  
mpāk, SmII mpāk, MM pak / [JH mulut, S1 mōlet]

NAME: Ks cō?, KB cō?, CW cō? / Je kēmōh, Mr kēmōh,  
BD kēmōh, Mt kēmōh, BN kēmōh, Sn kēmōh, Sa  
kēmōh, LJ kēmōh, LY kēmōh, Tm kēmōh, SmI muh,  
SmII muh, JH ?amōh, SB jēmōh, Tq kēmōh / [MM  
gēlāh, S1 gēlar]

NEAR: Ks tēdeh, KB tēdeh, Je pēnahdeh, Mr pēdeh,  
BD pēdeh, Mt pēdeh, BN pēdeh, CW dēwah, LJ  
pēlēndōh, LY pēlēndōh / Sn tēday, Sa tēday / Tm  
?ēpōn / SmI rē?, SmII rē? / JH pēem, SB ntēp, Tq  
hntēp / MM mīn / S1 dēkhes

NECK: Ks tēngkōg, Je tēngkōg, BN tēngkōk, Sn tēngkōg,  
Sa tēngkōg, LY tēngkōg / KB cēnōd / Mr ?uŋut, BD  
nut, Mt ?ēnēut, LJ tanēn, Tm tanēn, SmI tanēn,  
SmII tanēn / CW lēnē?, JH lēnē?, S1 lēnē?, Tq  
lēnē? / SB sēmārōŋ / [MM lēhēh]

NEW: Ks kabil, KB kabil / Je pay, Tm pāy, SmI pāy,  
SmII pāy, JH pay, MM mpay, SB sēpay / BD ze?, Mt  
?ēi?, BN re?, CW rē?, S1 ?arēh, Tq ?arēh / [Mr  
baru?, Sn bayu?, Sa bayu?, LJ barū?, LY baru?]

NIGHT: Ks hekut, KB hekut, Je hērkəd, Mt hekout,  
BN hēkwet, Sn hēŋōd, Sa hēŋōd, LJ hēŋōd, LY hēŋōd,  
SmII mēŋōt, JH sēŋāt / Mr tēwəy / BD haŋəp / CW  
bētom, S1 pētom / Tm layəg / MM dēy / SB pēlet /  
Tq gayəw / [SmI kēlem]

NOSE: Ks mōh, KB mōh, Je mōh, Mr mōh, BD mōh, Mt  
mōh, BN mōh, CW mōh, Sn mūh, Sa muh, LJ muh, LY  
muh, Tm mēh, SmI mōh, SmII mōh, MM muh, SB muh,  
S1 muh, Tq muh / [JH hidūŋ]

NOT (with verbs): Ks bēya?, KB bēya?, Mr bēra?,  
Mt bēya?, BN bēra? / Je mēŋid / BD nēŋ / CW hē?,



JH hēt / Sn tō?, Sa tō?, LJ tō?, LY tō?, Tm tō?,  
SmI ti?to?, SmII ti?to? / MM ɲot / SB beh, S1 beh  
Tq biy

77 OLD (inanimate): Ks sēwah / Je manah, Mr manah,  
BD manah, Mt manah, BN manah, LJ manah, LY manah,  
Tm manah, SmI manah, SmII manah / CW nəŋ, JH nēŋ  
Sn līw, Sa liw / MM lē?, SB ?ale?, S1 lē?, Tq  
?ale? / [KB -]

78 ONE: Ks nay, KB nay, Je ney, Mr nay, BD ney, BN  
nay, CW nōy, Sn nīh, Sa niy, LJ niy, LY niy, Tm  
ney, JH neywey, MM muy, SB moy, S1 mūy, Tq muy /  
SmI nanə?, SmII nanə? / [Mt sa?]

79 PERSON: Ks mēni?, KB mēni?, Sn sēmā?, Sa sēma?,  
LJ sēmā?, LY sēma?, MM hma?, SB sēma?, S1 sēmā?,  
Tq sēma? / Je mēnra?, Mr mēnra?, CW bəri? / BD  
batek, Mt bateik / BN lē? / Tm sən?ōy, SmI sən?ōy  
SmII sən?ōy / JH jah

80 PLAY: Ks pēnguh, KB pēnguh, Mr pēnguh, BD pēnguh,  
Mt pēngouh, BN pēnguh / Je min, Tm man, SmI man,  
SmII man, SB ?anman / CW niha? / Sa hiŋka?, LJ hi  
LY hiŋka? / JH si?se? / [Sn ?usig, MM nakhal,  
S1 nakol, Tq mayin]

81 QUIVER (for darts): Ks lēg, CW lək, Sn lēg, Sa  
lēg, LJ lēg, LY lēg, Tm lēg, SmI lək, SmII lək,  
JH lək, MM hlək, SB lēg, S1 lək, Tq luk / Je  
banə?, Mr banə?, BD banə?, Mt banə?, BN banə? /  
[KB -]

82 RAIN: Ks hēj, Je hej, Mr həc, Mt hāj / BN lēsəm,  
CW sēm / Sn mī?, Sa mi?, LY mi?, SmI mani?, SmII  
mani?, JH mī?, MM gēmāh, SB gēmāh / LJ tēh, Tm  
tēh / Tq bə?ah / [KB -, BD ?oʝan, S1 ?ari?]

83 RED: Ks tēhɔŋ, Je tēhup, Sn tēhūp, Sa tēhon / KB

təhəŋ, Mr pəhəŋ / BD bərakəc / BN rəgih, LJ rəgih, LY rəgih / CW bərtay / Tm cələk / SmI rəŋān, SmII rəŋān / JH rəŋap / MM gəcaŋ / SB dəŋat / [Mt meyah, S1 mirah, Tq -]

RICE (unhusked): CW be?, Sn bā?, Sa ba?, LY ba?, Tm bā?, SmI bā?, SmII bā?, JH ba?, MM be?, SB babāh, S1 babāh, Tq babah / [Ks padiy, KB -, Je padey, Mr padi?, BD padi?, Mt padi?, BN padi?, LJ bəras]

RIGHTSIDE: Ks tem, KB tem, Je tim, Mt təum, BN tēm, CW tēm, Sn tēp, Sa tēp, LJ tēp, LY tēp, Tm tēp, SmI kəntəp, SB satam, S1 satam, Tq ?atəm / Mr həŋjan / [BD kanan, JH kanan, MM kanan]

ROAD, PATH: Ks gələŋ / Je har, Mr har, BD har, Mt hay / BN həh / CW dəŋey, Tq dəŋay / Sn nūŋ, Sa nōŋ, LJ nəwōŋ, LY nōŋ, Tm nōŋ, SmI nōŋ, SmII nōŋ, JH nōŋ / SB tərūŋ, S1 tərūŋ / [KB -, MM -]

ROOT: Ks dəpag, Mr dəpak / Je jə?is, BN jə?is, CW jə?es, Sa yēs, SmI rə?is, SmII rə?is, S1 res / BD ?awey, Mt ?awey / Sn jəlōg / LJ yōw / LY catək, Tm cantīŋ / SB jəmə?, Tq jəmə? / [KB -, JH jaŋkar, MM jaŋkāh]

ROTTEN: Ks sə?, KB ?ənsə?, Je ?ənsə?, Mr ?ənsə?, Mt sə?, Sn sanū?, Sa sanu?, LJ sasō?, LY sasə?, Tm sō?, SmI sō?, SmII sō? / BD ha?ət, BN ha?ət, CW ha?ət, JH si?in, MM su?ət, SB sə?it, S1 sə?it, Tq sə?ət

SALT: Ks siya?, KB səya?, Mr səya? / Je mpəj, Sn ?əmpōj, Sa mpəj, SmI mpōc, SmII mpōc / BN tēpəl, CW tupəl, JH pol, SB tēpəl / [BD gərəm, Mt gaem, LJ garam, LY garam, Tm garam, MM garam, S1 garam, Tq garam]

- 90 SAY: Ks penah / Je pēnsēŋ / Mr bērase? / BD  
kədəh, Mt kədəuh, BN kədəh / CW bēt / Sn tuh,  
Sa tuh / LJ caŋkāy / LY cal / Tm rō?, SmI ŋrō?,  
SmII ŋrō? / JH kēŋlōŋ, S1 khələŋ / SB panyēŋ /  
Tq ?abōŋ / [KB -, MM cakap]
- 91 SCRATCH: Ks kəwɔj, KB kay, Je kaj, Mr kac, BD ka  
BN kac, CW kikɔc, JH kāk, MM kakac / Mt kəus, /  
Sn gih, Sa gis, LY gis, Tm gTs, SmI gēh, SmII gi  
SB gahgeh, S1 gahgeh / LJ gərēŋ, Tq mərac
- 92 SEE: Ks dēŋ, CW daŋ, Sn dah, Sa dak, LJ dāk, LY  
dak / Je ?el, Mr ?el / BD tēt, BN tēt / Tm nēh,  
SmI nēŋ, SmII nēŋ / JH lēh / MM cələw / SB cēn /  
Tq ?elēŋ / [KB -, Mt caŋɔ?, S1 jəŋɔk]
- 93 SHARP: Ks hətēh / KB cəma?, Je cəme?, BN cuma?,  
JH cəma? / CW cibəd / LJ jələt, SB lot / LY pəhe  
Tm pəheŋ / S1 ?əluh, Tq luh / [Mr makan, BD maka  
Mt makan, Sn makan, Sa tajap, SmI tajap, SmII  
tajap, MM tajam]
- 94 SHOOT (blowgun): Ks həluh, KB həluh, Je haluh,  
Mr haluh, BD haluh, Mt haluh, BN haluh, Sn səluh  
Sa səluh, LJ səluh, LY səluh, Tm səluh, MM ləh /  
CW hapud, SmI pōt, SmII pōt, JH put / SB bah?ɔh,  
S1 ?ah?ɔh, Tq ha?ɔh
- 95 SHORT: Ks cəmhət, Je cəhad, Mr cənhət, BD cənhət  
BN cənhət, CW cohət / KB cəŋket, JH siŋket / Mt  
tuwɔip / Sn kəldēl, LJ kəldēl, LY kəldel / Sa  
pənpet, Tm ?apet / SmI patT?, SmII patē? / MM  
jələt, SB jələ?, S1 jələ?, Tq jələ?
- 96 SIBLING (younger): Ks beh, KB beh, Je ber, Mr  
ber, BD bel, Mt beeh, BN ber, CW bar / Sn pə?,  
Sa pə?, LJ pē?, LY pə?, Tm pə? / SmI mənəŋ, SmII  
mənəŋ / Tq yē? / [JH ?adT?, MM ?adi?, SB ?ade?,  
S1 ?adT?]

7 SIBLING (elder): Ks ʔoʔ / Je pɛʔ, Mr pəʔ / BD toʔ, Mt tɔwʔ, BN toʔ, CW toʔ / Sn kəloʔ, Sa kəlōʔ, LJ kəlōʔ, LY kəloʔ, Tm kəlōʔ / SmI tənɛʔ, SmII tənēʔ / JH ʔiʔɛm / MM yeʔ, Tq yeʔ / SB ʔeʔeʔ, S1 ʔiʔeʔ / [KB -]

8 SING: Ks pənlon, KB pənlon, Je pənlon, Mr pənlon, BD pɪnlon, Mt pənlon, Sn pɪnlōn, Sa pɪnlon, LJ pɪnlōn, LY pɪnlon / CW gəŋhaŋ / Tm gabag / [BN sewan, SmI dindan, SmII nanɪʔ, JH bəsoraʔ, MM pantut, SB nanɪʔ, S1 nanɪʔ, Tq nanɪʔ]

9 SIT: Ks ŋɔk, KB ŋɔk, Je ŋɔg, Mr ŋɔg, BD ŋɔk, Mt ŋɔuk, BN ŋɔk, CW ŋɔk, JH ŋəʔŋōk / Sn kət / Sa kɔc, LJ kōc, LY kɔc / Tm gəl, SmI gēy, SmII gēy / MM khōm, SB khom, S1 khōm, Tq khom

0 SKIN: Ks katəʔ, KB kətoʔ, Mr kətəʔ, BD kətəʔ, Mt kətəʔ, BN kətəʔ, SmI gətēʔ, SmII gətēʔ, JH kətūʔ, SB gətūh / Je səmpoʔ, LY səmpoʔ, Tm səmpōʔ / CW hɔʔ / Sn katɛk, Sa katɛk / LJ səŋkɔʔ / [MM kulit, S1 kulit, Tq kulit]

1 SKY: Ks kəpɪŋ, KB kəpɪŋ / Je kətoʔ, BD kətoʔ, Mt kətɔuʔ, BN kətoʔ / Mr barəŋal / Sn balɪŋ, Sa balɪŋ, LJ balɪŋ, LY balɪŋ, Tm balɪk, S1 malɪŋ / SmI gəsɪk, SmII səɪk / [JH laŋit, CW laŋit, MM laŋit, SB laŋit, Tq laŋit]

2 SLEEP: Ks tik, Je tɪg, Mr tek, Bd tek, Mt tɛk, BN tek, Sn teg, Sa teg, LJ teg, LY teg, JH ticēk, MM gətik, SB jətek, S1 jətek, Tq jətek / CW ʔɔm / KB ləbit, SmI bēt, SmII bēt / Tm səlog

3 SMALL: Ks ʔajɔʔ, Je ʔajɔʔ / KB kanot, Mr kanet, BD kaneʔ, Mt kanit, BN kanet, CW kanet, JH kapet, MM kənen, SB kəʔnet, Tq kanit / Sn ʔahūʔ / ʔisɛt / LJ ʔamēs, LY ʔames, Tm ʔames / SmI macon,

SmII macɔ? / S1 rakɛt

- 104 SMELL (sniff): Ks ?ɔɪ, Je ?ɔɪ, Mr ?ɔɪ, BD ?ɔɪ, F  
?ɔɪ, Sn ?ɔɪ, Sa ?ɔɪ; LJ ?up, LY ?up, Tm ?ūɪ, JH  
?ɔɪ, SB hup / SmI nōy, SmII nōy / MM loʔom / S1  
khələm / [KB -, Mt -, CW ciyɪm, Tq -]
- 105 SMOKE: Ks yey, KB ?iyey, Je ?iʔey, BN yey, CW h  
LJ ?iʔey / Sn jiʔēm, Sa jəʔim / LY ?agəy / Tm ca  
SmI cas, SmII cas / MM jələk, SB jələk, S1 jələk  
Tq jələk / [Mr ?asəp, BD ?asəp, Mt ?asəp, JH  
?asap]
- 106 SMOOTH: Ks gəhəh, KB gəhəh / Je bəjlaj, Mr bələc  
BD bələc, Mt bələc, CW bələc, Sn səlej, Sa səlej  
LJ bələj, LY səlej, Tm səlej, SmI sələc, SmII  
sələc, JH bələc / MM hliw / [BN -, SB liciɪ, S1  
liciɪ, Tq liciɪ]
- 107 SNAKE: Ks ?ikɔb, KB ?ikɔp, Mr jəkɔb, Mt jəkoup,  
BN jəkɔb / Je tajɔ?, Sn tajū?, Sa tajɔ?, LJ tajū  
LY tajɔ?, Tm tajū?, SmI tijT?, SmII tijT?, MM  
tejɔw, SB tijɔh, S1 tijɔh, Tq tejɔw / CW talun,  
JH talun / [BD ?olar]
- 108 SPEAR: Ks ?ad, Mt ?ad, BN ?at, CW ?ət / BD bulus  
Sn bulūs, Sa bulus, LJ bulūs, LY bulus, JH bulus  
/ Tm tarɔg, SmI tarɔk, SmII tarɔk / [KB -, Je  
mata?, Mr mata?, MM tohok, SB tohok, S1 ləmēɲ,  
Tq tohok]
- 109 SPIT: Ks bej, KB bej, Je kəbej, Mr kəbec, Mt  
kəbeic, CW kəbaj, Sn kəbej, Sa kəbej, LJ kəbēj,  
LY kəbej / BD gətɔf, BN tɛf, Tm gətɔh, SmI tɔh,  
SmII gətɔh, JH tuwɔh, MM bathɔy, SB təhtɔh, S1  
tahtɔh, Tq thɔh
- 110 SQUEEZE: KB lam / Mt ləmac / BN rit, Tm reyed,  
SmI rTɔd, SmII rēd, JH rɛt / CW wɛt, Sn wɛd, Sa

wəd, LY həlwən, SB wəd / [Ks pules, Je cəpid, Mr polas, BD rames, LJ cəpəd, MM pərah, S1 pəcət, Tq ramas]

1 STAB: Ks ceg, KB cek, Mr cik, BD cek, Mt ceik, BN cēk, SB cək / Sn cōg, Sa cōg, SmI cək, SmII cək, JH cək / LY cələg, Tm cələg / [Je lawan, CW tikam, LJ lawāt, MM tikam, S1 tikam, Tq tohok]

2 STAND: Ks hənjan, KB hənan, Je hənjan, Mr hənjan, BD hənan, Mt hənan, BN hənan, CW hɲɔɲ / Sn kajəh / Sa təd, LJ təd, LY təd, Tm təd / SmI jinjek, SmII jinjek, JH jinjəwɔɲ, MM jək / SB ʔoʔaw, S1 ʔuʔaw, Tq ʔoʔaw

3 STICK (rod): Ks ʔad, Je ʔad, Mr ʔad, Sa ʔad, LJ ʔad, LY ʔad, Tm ʔād / Mt cəɲəɲ, SmI kənəɲ, SmII kəniɲ, JH kərəɲ / Sn ʔamāɲ / [KB -, BD tɔŋkat, BN tɔŋket, CW tɔŋkat, MM tɔŋkat, SB tɔŋkat, S1 tɔŋkot, Tq tɔŋkat]

4 STONE: CW təmɔʔ, JH təmɔʔ, SB təmɔɲ, Tq təmun / [Ks batuʔ, KB batuʔ, Je batuʔ, Mr batuʔ, BD batuʔ, Mt bateuʔ, BN batuʔ, Sn batuʔ, Sa batuʔ, LJ batuʔ, LY batuʔ, Tm batūʔ, SmI mbatūʔ, SmII mbatūʔ, MM batuʔ, S1 batuʔ]

5 STRAIGHT: BN jəsɪɲ / Sn pəlūɲ, Sa pəlun / LJ sijek, LY sijek, Tm sijek / JH dəɲɔɲ / SB ləsoh / Tq yihyah / [Ks bətul, KB -, Je bətul, Mr bətul, BD bətul, Mt lɔuyɔus, CW lurus, SmI tɛgaʔ, SmII tɛgaʔ, MM lurus, S1 rulus]

6 SUCK: Ks jəhəd, Je jəhəd, Mr jəhət, Sa jəhəd, LJ jəhəd, LY jəhəd / KB jɔt, BD jɔt, Mt jaut, BN jɔt, Sn jɔd, Tm jɔd / CW buʔ / SmI ɲɔʔ, SmII ɲɔʔ, Tq ɲuʔ / SB sək / [JH sədɔt, MM sədut, S1 ʔisap]

- 117 SWELL: Mr kən, Mt kəun / CW səwah / LJ his, LY his, Tm hīs / SmI ʔas, SmII ʔas, S1 ʔos / [Ks kəmaŋ, KB -, Je kəmaŋ, BD kəmaŋ, BN kəmaŋ, Sn ʔibūʔ, Sa kəmak, JH bəŋkak, MM bəŋkak, SB bəŋkak Tq buntit]
- 118 TAIL: Ks hatiʔ, KB hətiʔ, Je hatiʔ, Mr hatɕʔ, BD hacɕʔ, BN hatiʔ, CW hateʔ, Sn səntaʔ, Sa səntaʔ, LJ səntaʔ, LY səntaʔ, Tm sentāʔ, SmI sentāh, SmII sentāh, JH səntaʔ, SB hateʔ / S1 pəs, Tq pas / [Mt ʔikəih, MM ʔikūh]
- 119 THIN: Ks hetəh, KB hetəh, Mr hərtəl, BD hərtəl, Mt hetəl / BN səpɿy / Sa pəhən, LJ pəhən, LY pəhən, Tm pəhən / SmI ŋsey, SmII ŋsey, MM sēh, S1 shey, Tq sey / [Je lipis, CW nipis, Sn lipis, JH mimpis, SB nipis]
- 120 THIS: Ks ʔmh, KB ʔəh / Je təh, Mr təh / BD dəʔ, CW doʔ / Mt həʔ, MM nahəʔ / BN ladəh, Tm doh, SmI deh, SmII dih, JH dəh / Sn nəh / LJ dec / S1 naʔ, S1 nəʔnəʔ, Tq ʔanəʔ / [Sa -, LY -]
- 121 THOU: Ks bəʔ, KB bəʔ, Mr beʔ / Je pay, Mt paŋ / BD məh, Sn mTh, Sa mih, LJ mTh, LY mih, JH ʔimāh / BN məʔ, CW mɐʔ / Tm həʔ, SmI hēʔ, SmII hēʔ, MM hiʔ, SB hiʔ / S1 kāh / [Tq -]
- 122 THREE: CW pet, Tm neʔ, SmI niʔ, SmII niʔ, MM hmpeʔ, SB mpeʔ, S1 hmpeʔ, Tq hmpeʔ / [Ks tigaʔ, KB -, Je tigaʔ, Mr tigaʔ, BD tigaʔ, Mt tigaʔ, BN tigaʔ, Sn tigaʔ, Sa tigaʔ, LJ tigāʔ, LY tigaʔ, JH tigāʔ]
- 123 THROW: Ks həwid, Mr həwit / Je həg, LJ həg / BD pəntəŋ / Mt paŋkaʔ, Sn pəkaʔ, Sa pəkaʔ / LY bədəl, Tm bədəl / SmI pɛc, SmII pɛc / MM piŋgok / SB hmel / S1 chək / Tq yoj / [KB -, BN baliŋ, CW baliŋ, JH limpar]

- TIE: Sn bæg, Sa bæg, LJ bæg, LY bæg, Tm bæg, SmI bæk, SmII bæk, MM kabæk, SB bæk, S1 bæk, Tq bæk / [Ks ?iket, KB -, Je ræbed, Mr ?iket, BD ?iket, Mt tamat, BN siyet, CW siyat, JH jækæt]
- TONGUE: Ks lætig, KB lætig, Je læntig, Mr læntik, BD læntik, Mt læntiyk, BN læktik, CW latek, Sa læntag, LJ læntag, LY læntag, Tm læntāg, SmI læntāh, SmII læntāh, JH læntak / Sn pæled / SB læpch, S1 læpes, Tq læpes / [MM lidah]
- TOOTH: Ks nus, KB yus / Je hæp, Mr hap, BD hap, Mt hap, BN hap / CW læmup, Sn læmap, Sa læmop, LJ læmop, LY læmop, Tm mop, SmI læmōp, SmII læmōp, JH læmōp, MM jæmoy, SB læmop, S1 læmop, Tq læmop
- TREE: Ks ?ihy?, KB ?ihy, Je jəhy?, CW jəhy?, Sn nihy?, Sa nehy?, LJ jəhū?, LY jəhu?, Tm jəhy?, SmI jəhu?, SmII jəhu? / BD tom, Mt toum, BN tom / JH təkoh, MM təkoh / SB dəlɔŋ, S1 dəlōŋ, Tq dəlɔŋ / [Mr kayu?]
- TURN: Ks wælwel, Mr wælwel, Sn ?ilwəl, Sa ciwel, LJ kəwēl, LY cərwel, Tm wel, SmII riylwal / Je bəri?le? / Tq yoc / [KB -, BD pusin, Mt pusin, BN kəseh, CW pusin, SmI posik, JH posin, MM kisāh, SB pusin, S1 pusin]
- TWO: KB biyeh, CW ber, Sn nay, Sa nay, LJ nār, LY nar, Tm nār, SmI nār, SmII nār, JH nar, MM hmāh, SB mar / [Ks duwa?, Je duwa?, Mr dəwa?, Mt duwa?, BN duwa?, S1 duwa?, Tq duwa?]
- VOMIT: Ks kə?, KB kə?, Je kə?, Mr kə?, BD kə?, Mt kə?, BN kɔ?, CW kə?, Sn ko?, Sa ko?, LJ kō?, LY kō?, Tm ko?, SmI kē?, SmII kē?, JH kū?, MM khu?, SB ko?, S1 khu?, Tq ku?



- 131 WALK: Ks cəb, Je cəb, Mr cəp, BD cəp, Mt cəup, BN cəp, CW cəcəb, Sa cib, LJ cib, LY cib, Tm cTb, SmI cTb, SmII cTb, JH cib / MM cə? / SB səgwag, S1 suwak, Tq səwak / [KB -, Sn -]
- 132 WASH (bathe): Ks ʔənlay, KB ʔənlay, Je ʔəlay, Mr ʔəlay, BD nay / Mt sɔuc, BN sɔc / CW mamuh, Sn mamuh, Sa mamuh, LJ mamuh, LY mamuh, Tm muh, SmI mamuh, SmII mamuh, JH maʔmūh, SB mahmēh, Tq mahmeh / MM hūm, S1 hūm
- 133 WATER: Ks bətew, KB bətew, SmI tēw, SmII tēw, JH tɔw / Je tɔm, Mr tɔm, BD tɔm, Mt tɔm, BN tɔm, CW tam / Sn ʔɔŋ, Sa ʔɔŋ, LJ ʔɔŋ, LY ʔɔŋ, Tm ʔɔk / MM dow / SB jəʔoh / S1 dak, Tq dak
- 134 WE (plural inclusive): Ks heʔ, KB heʔ, Je heʔ Mr hiʔ, BD heʔ, CW hēʔ, SmI hTʔ, SmII hēʔ, JH ʔihēʔ MM hēh, S1 hēʔ / Mt yeʔ, SB yēh / Sn ʔeʔ, LY ʔēʔ Tm ʔēʔ / Sa ʔin, LJ ʔic / [Tq -]
- 135 WET: Ks pəcuʔ, Je pəcɛʔ, Mr pəcəʔ BD pəcəʔ, Mt pəcəʔ / KB məʔaj, BN maʔac, CW məʔac, Sn məʔaj, Sa məʔac, LJ məʔaj, LY kəʔaj, Tm kəʔaj, SmI kəʔaj SmII kəʔaj, SB maʔac / JH geseʔ / MM təkɔŋ, S1 təkoh, Tq təkoh
- 136 WHAT?: Ks ləw, KB low, Mr low, BD ʔayləw, Mt ʔayləw, BN ʔayləw, Sn loʔ, LJ tah loʔ, LY mah loʔ Tm loʔ / Je tah bayoʔ, Sa yɔʔ / CW məpɔcɔʔ / SmI māh, SmII māh, SB hmɔh / JH mənəkɔʔ / MM namaʔ S1 madēh / [Tq -]
- 137 WHEN?: Ks jənħəh, KB jənħəh / Je mapoʔ, SmI mpa JH poʔ nɔn / Sn bel, Sa bel, LJ bel, LY bel, Tm bēl, SmII mbil, SB bel / [Mr masaʔ ʔaloʔ, BD masaʔ ʔayləw, Mt biləʔ, BN masaʔ ləw, CW biləh, MM belēh, S1 bilaʔ, Tq -]

- 8 WHITE: Ks bəltaw, KB pətaw, Sa pəltaw / Mr bəkən / BD bəyɛl / Mt bəyɛik, LJ bəyūg, LY bəyūg, Tm bəyūg, SmI biyēk, SmII biyēk / BN halək / [Je puteh, CW putih, Sn puteh, JH putih, MM potih, SB poteh, S1 putih, Tq poteh]
- 9 WHO?: Ks tom, KB tom / Je maken, Mr maken / BD ʔo-ləw, Mt ʔay-ləw, BN tək-ləw / CW bicəʔ, Tm cəʔ / Sn diyoh, LJ tah dəʔoh, LY jəʔoh / Sa dəyəʔ / SmI bəʔ, SmII bəʔ / JH nah / MM humaʔ / SB hŋkəʔ / S1 kadəh / [Tq -]
- 0 WIND: Ks bewaʔ, KB bewaʔ, Je bərwaʔ, MM buwaʔ / Mr hanən / Mt seiʔ / Sa siyag / LJ parək, LY parək / Tm həlnŋl / SmI pinəy / SmII pəs / JH mənhəŋ / SB təhəl / [BD ʔaŋin, BN ʔaŋin, CW ʔaŋin, Sn ʔaŋin, S1 ribut, Tq -]
- 1 WIFE: Ks kəndeh, KB kəndeh, Je kəneh, Mr kəneh, BD kəndeh, Mt kənɛih, SmI kənah, SmII kənah, JH kənah / BN jiʔ, CW jəʔ / Sn babəʔ / Sa kədoy, MM kədoh / LJ leh, LY leh, Tm leh / SB kəmpən, S1 kəmpən, Tq kəmpən
- 2 WING: Ks kələpəh, KB kələpəh, CW kələpaʔ, Sn kələpoh, JH kərəpaʔ / Je kəpəŋ, Mr kəpəŋ, LJ kəpəŋ, LY kəpəŋ, Tm kənyēk, SmI kəpēk, SmII kəpēk / BN pəwic, Sa pəyej, S1 pərəc / MM kəmphek, SB kəpak / Tq kəŋlək / [BD sayap, Mt sayap]
- 3 WIPE: Ks jəd, Je jəd / Mr məh / BN təmpəs, SmI pəs, SmII pəs / LJ sēʔ / Tm gTd / [KB -, BD gəsək, Mt ŋgəsət, CW sapuʔ, Sn gəsəʔ, LY gəsəʔ, JH sapūʔ, MM sapuʔ, SB saʔpoʔ, S1 sapuʔ, Tq gəsək]

- 144 WOMAN: Ks mabeh, KB mabeh / Je babo?, Sn babō?  
 Sa babo?, LJ babō?, LY babo?, Tm babō? /  
 Mr ya?lɛw, BD ya?lɛw, Mt ya?lɛw / BN koŋ, CW koŋ  
 JH kəra?kəŋ / SmI kərdōr, SmII kərdōr, MM kədōh  
 SB kədōr, S1 kərdōr, Tq kərdor
- 145 WOODS: Ks kaheb, KB kaheb, Je ba?heb, Mr kaheb,  
 BD hep, Mt hep, BN hep, LJ dənheb, LY dənheb /  
 CW berte? / Sn bəyɿ?, Sa bəyi?, JH bəri?,  
 MM mərɿh, SB bərɿh, S1 bərɿh, Tq bərih / Tm  
 sərək, SmII sərək / SmI jəres
- 146 YOU (plural): Ks bom, Je bom / KB yam / Mr  
 jəmpəh, BD jənpəh, CW jin / BN gin / Sn yuh,  
 Sa yuh, LJ yūh, LY yuh / Tm ɲob / SmI ŋkēh,  
 SmII ŋkē? / JH yon / MM hi?, SB hi? / S1 jē?,  
 Tq ?aji? / [Mt -]

---

<sup>1</sup>The vocabularies upon which this paper is based were collected during field trips financed at various times by the following foundations and scholarships: Henry Ling Roth Scholarship (University of Cambridge), Horniman Scholarship and Esperanza Trust (Royal Anthropological Institute), Ministry of Education and Science (London), Wenner-Gren Foundation (New York), University of Singapore. The author also wishes to thank Dr. Asmah Haji Omar for permission to use her Kintaq Bong vocabulary, Kirk Endicott for collecting the Bateg Nong vocabulary specially for this paper, and a traveller who wishes to remain anonymous for collecting the Temoq vocabulary on tape. The following people kindly made additional lexical material available: Ivan Polunin (Jah Hut), Barry Hooker (Semai), Gérard Diffloth (Che' Wong, Semai). Their material does not appear here, but it has proved useful as a check on the author's own vocabularies. The author alone is responsible for the views expressed in this paper, though he has greatly benefitted from discussions with Kirk Endicott and Gérard Diffloth.

<sup>2</sup>An unsigned notebook containing a comparative Aboriginal vocabulary in an orthography very like

at used by Wilkinson, and now in the possession  
the National Museum, Kuala Lumpur, may be  
lkinson's work.

### *Appendix III*

#### Concordance with Blagden's Vocabulary

Almost all of the linguists who have used Aslian data in comparative Austroasiatic studies have relied entirely on C. O. Blagden's 'Comparative Vocabulary of Aboriginal Dialects' printed as an appendix to vol. 1 of Skeat and Blagden 1906. However, few of the dialects listed in Blagden's Vocabulary are identified by anything other than the collector's name and the locality of collection; language or tribal names are hardly ever reported. Usage of Blagden's Vocabulary is therefore something of a hit-or-miss matter. It is to be hoped that the following table of those identifications that seem most probable will prove helpful to users of Blagden's otherwise extremely valuable material.

KENSIU: Pang. Jalor

KINTAQ BONG: Sem. B. Max., Sem. Stev.

JEHAI: Jehehr, Sem. Jarum, Sem. Plus (misplaced)

MENDRIQ: Pang. Gal., Pang. Sam

BATEG DEQ: Kerbat, Lebir, Pang. K. Aring, Pang. U. Aring, U. Kel.

SEMNAM-SABUM: Sak. Jer., Sem. K. Ken. (?Semnam), Sem. Martin, Sem. Per. (?Sabum)

LANOH: Po-Klo (?=Schebesta's 'Ple-Temer'), Sŭm. (?Lanoh Yir)

TEMIAR: Sak. Br. Low, Sak. Croix, Sak. Kerb., Sak. Korb. Gb., Sak. Korb. Lias, Sak. Morg., Sak. Plus. Sak. Tan. Ram., Tem. Cl., Tembi, Tumm. Stev.

SEMAI: Darat, Jelai (?Semai II), Or. Berumb., Sak. Bat. Pad., Sak. Blanj. Cl., Sak. Blanj. Cliff., Sak. Chen., Sak. Em., Sak. J. Low, Sak. Martin (?Semai

I), Sak. Ra., Sak. Sel. Da., Sak. Slim, Sak. Sung.,  
Sak. Tap., Sak. U. Bert., Sak. U. Kam., Sak. U.  
Tap., Sen. Cl., Sen. Cliff., Serau, Sin. Stev.,  
Tan. U. Lang.

H HUT: Kerdau, all Krau dialects, Sak. Guai

H MERI: Bers. Stev., all Bes. dialects

MAQ BERI: U. Cher., U. Tem.

MELAI: Bera, Semilai Coll. Nya., Serting

MOQ: Or. Hu. Jo., Pal., U. Ind.

ntil, Bateg Nong and Che' Wong appear not to be  
presented in the vocabularies available to Blagden.

## Appendix IV

### Sources of Aslian Word-lists

Except where otherwise stated, all the word-lists used in this study were collected by the author. For those languages contiguous to Temiar the latter language was used in questioning, supplemented by Malay. Other languages were investigated through the medium of Malay. Many of the word-lists were collected not *in situ* but at the Orang Asli Department's hospital at Ulu Gombak, Kuala Lumpur, with the kind permission and cooperation of Dr. Malcolm Bolton and Cik Ruslan bin Abdullah.

KENSIU: Sireh s/o Kundang, Kg Lalang, Siong, Baling, Kedah.

KINTAQ BONG: Taken with slight modification from Asri 1963, with the author's kind permission.

JEHAI: Ləkəwas s/o Kotey, Klian, Perak.

MENDRIQ: Penghulu Buloh s/o Mawə?, K. Lah, Bertam, Ulu Kelantan.

BATEG DEQ: ?ək s/o Səli? and ?odaŋ s/o Saɔ?, Ulu Aring, Ulu Kelantan.

MINTIL: Bərahim s/o Tale?, s. Tanum, Lipis, Pahang.

BATEG NONG: Collected by Kirk Endicott from near K. Tembeling, Pahang.

CHE' WONG: Yasih s/o Tukeh, s. Pasu, tributary of s. Gali, above Raub, Pahang.

SEMNAM: Təmohŋh s/o Bəwata?, place of residence unrecorded.

SABUM: Layaŋ d/o Buloh, Lenggong, Perak.

LANOH JENGJENG: Batin Kəladih s/o Gaŋcār, s. Ringat Temengor, Perak.

LANOH YIR: Təwā? Bahul s/o Kolim, s. Sarah, Ulu Perak.

MIAR: Based on the author's own speaking knowledge of the dialect spoken in the lower Perolak valley, Ulu Kelantan.

MAI I: Tentam s/o Husin, Kg Redang Ponggol, near Telok Anson, Perak.

MAI II: Unidentified informant from Fort Sin, Pahang.

H HUT: Kamarudin s/o Bujang, Kg Paya Pasu, K. Krau, Pahang.

H MERI: Ibrahim s/o Indun, s. Judah, Carey Island, off Selangor coast.

MAQ BERI: F/S Mat Yunus, s. Ganti, Maran, Trengganu.

MELAI: Nihit d/o Dahit, Bukit Serok, near Tasek Bera, Pahang.

MOQ: Gathered anonymously from a headman on s. Jeram, near Kg Aur, Pekan, Pahang.