elpublishing

Language Documentation and Description

ISSN 1740-6234

This article appears in: *Language Documentation and Description, vol 11*. Editors: Stuart McGill & Peter K. Austin

The Aslian languages of Malaysia and Thailand: an assessment

GEOFFREY BENJAMIN

Cite this article: Geoffrey Benjamin (2012). The Aslian languages of Malaysia and Thailand: an assessment. In Stuart McGill & Peter K. Austin (eds) *Language Documentation and Description, vol 11*. London: SOAS. pp. 136-230

Link to this article: http://www.elpublishing.org/PID/131

This electronic version first published: July 2014

This article is published under a Creative Commons License CC-BY-NC (Attribution-NonCommercial). The licence permits users to use, reproduce, disseminate or display the article provided that the author is attributed as the original creator and that the reuse is restricted to non-commercial purposes i.e. research or educational use. See http://creativecommons.org/licenses/by-nc/4.0/

EL Publishing

For more EL Publishing articles and services:

Website:	http://www.elpublishing.org
Terms of use:	http://www.elpublishing.org/terms
Submissions:	http://www.elpublishing.org/submissions

The Aslian languages of Malaysia and Thailand: an assessment

Geoffrey Benjamin

Nanyang Technological University and Institute of Southeast Asian Studies, Singapore

1. Introduction[1]

The term 'Aslian' refers to a distinctive group of approximately 20 Mon-Khmer languages spoken in Peninsular Malaysia and the isthmian parts of southern Thailand.[2] All the Aslian-speakers belong to the tribal or formerly-

[1] This paper has undergone several transformations. The earliest version was presented at the *Workshop on Endangered Languages and Literatures of Southeast Asia*, Royal Institute of Linguistics and Anthropology, Leiden, in December 1996. Expanded versions were presented at the *Fourth ASEAN Inter-University Seminar on Social Development*, Pattani, Thailand, in June 1999 and at the *Symposium on Language and Diversity*, Division of Linguistics and Multilingual Studies, Nanyang Technological University, Singapore, in March 2009. The present version has been completely expanded and updated. Earlier drafts received valuable comments and input from Nicole Kruspe, Niclas Burenhult and two anonymous referees, for which I am grateful; I alone am responsible for any defects that remain. My initial ethnographic and linguistic fieldwork on Temiar in the mid-1960s was supported by an Emslie Horniman Scholarship from the Royal Anthropological Instititute (London), and benefited greatly from the close cooperation of the Malaysian Department of Aboriginal Affairs (as it was then known) under the direction of Iskandar Carey and Baharon Azhar Bin Raffie'i. Some of the later fieldwork in Malaysia was supported by grants from the (former) University of Singapore and the Wenner-Gren Foundation. This paper was completed while holding a Visiting Senior Research Fellowship at the Institute of Southeast Asian Studies (ISEAS), Singapore, and a Senior Associateship at the Centre for Liberal Arts and Social Sciences (CLASS) at Nanyang Technological University.

[2] 'Mon-Khmer' is the conventional label for the Southeast Asian representatives of the larger Austroasiatic family, the other branch of which is represented by the Munda Languages of South Asia. It has been widely assumed that the two branches constitute genetically well-founded sub-families, but the view currently being proposed is that all the Austroasiatic subgroups are of more equal status and that the high-level division between 'Mon-Khmer' and 'Munda' should therefore be regarded as arbitrary (Sidwell 2010, Sidwell & Blench 2011). While this 'flat' view may turn out to be true, linguistic research continues in practice to treat the Munda and Mon-Khmer branches separately. For example, a planned handbook of the Mon-Khmer languages (Jenny & Sidwell, in preparation) is intended to parallel an already published handbook of the Munda languages (Anderson 2006). Accordingly, in this paper I shall continue to refer to Aslian as a subdivision of Mon-Khmer – but without prejudice as to any future dissolution of Mon-Khmer as a valid sub-family within Austrosiatic.

Geoffrey Benjamin 2012. The Aslian languages of Malaysia and Thailand: an assessment. In Peter K. Austin & Stuart McGill (eds.) *Language Documentation and Description,* Vol 11, 136-230. London: SOAS.

tribal populations known officially in Peninsular Malaysia as *Orang Asli* ('original people') and in southern Thailand as *Sakai* (an old and nowadays rather 'improper' Malay word for 'non-Malay tribespeople').

The term 'Aslian', which derives from 'Orang Asli', was first proposed privately by Gérard Diffloth in the 1970s, and later introduced into the literature in Diffloth (1974a) and Benjamin (1973: x, 1976a: 43). The label is admittedly less than satisfactory, but it is now well established in the linguistic literature and no usable alternative has yet been proposed. One drawback is that not all Orang Asli speak Aslian languages since at least 55,000 in Johor, Pahang, Negeri Sembilan and Selangor states of Malaysia are speakers of Austronesian languages, which with just one exception are really dialects of Malay.[3] This paper is concerned only with 'Aslian languages' in the narrow sense (genetically Mon-Khmer) and not 'Orang Asli languages' in general (a sociolinguistic term covering both Aslian and Aboriginal-Malay languages).

Relatively little was known about the Aslian languages until the 1990s, when intensive field-research on them was taken up again after some thirty years of relative inattention. However, misinformation still attends much of the public discussion of the status, history, relationships and characteristics of these languages. This paper is aimed at redressing the situation by surveying what is currently known about their circumstances, coupled with an assessment of the scholarly research on which that information is based. The question of endangerment will also be discussed: some Aslian languages are in imminent danger of extinction while the survival of others is by no means assured, even if they may seem to be currently viable.

What then is special about this survey? Why are the Aslian languages deserving of such attention?

First, practical concerns loom large. In terms of economic position and status, Aslian-speakers (along with other Orang Asli) currently remain at the bottom of the social ladder in Malaysia. Even their status as the original indigenous people of the Peninsula is currently challenged, on the almost incomprehensible grounds that they were not the bearers of 'civilisation', i.e., centralised state-formations – a role normatively ascribed by several political authorities in Malaysia solely to the Malays. This is not simply a cultural nicety: it has profound consequences for the rights of the Orang Asli to the lands they have occupied for millennia and which are now greatly threatened

[3] The exception is Duano (also known as Desin Dolaq or Orang Kuala) with at least 3,220 speakers in Malaysia and yet more on the island of Rangsang and its neighbours off the east coast of Sumatra in Indonesia (Kähler 1946-49; Pelras 1972 143-144, Sandbukt 1983).

by the spread of a plantation economy over which they have no control. Protracted legal cases relating to this are currently being pursued in several Peninsular states. The linguistic evidence, pointing as it does to a millennia-long and variegated Aslian settlement, is increasingly being called upon in the courts in support of the people's claims to a prior attachment to their land.[4]

Second, although the survival of the Aslian languages is undoubtedly under threat, the people themselves are not necessarily thereby linguistically incapacitated. Almost all Orang Asli are bilingual in Malay, the national language of Malaysia, and are increasingly able to read and write it.[5] But the vast majority still speak their own languages at home, employing them in almost all domains except those that require formal literacy. However, with the spread of mobile-phone text-messaging, *informal* literacy in Aslian languages has begun to emerge; and at least two of the languages, Temiar and Semai, are employed in daily radio broadcasts. Clearly then, Aslian speakers do take pride in speaking their own languages, and would regret their disappearance. For social scientists and others who share their concerns, the Aslian languages therefore provide a key *entrée* into the living culture of the Orang Asli. Moreover, the demonstration that these are neither mere jargons nor 'dialects' but lexically and grammatically rich languages should help in the task of gaining respect for these too-often denigrated populations.

Third, attempts are being made in some Malaysian states to introduce Aslian languages into the primary school curriculum as the medium of initial instruction. (The irregular school-attendance record of Orang Asli children continues to be a problem.) So far, these efforts have not proved very successful; but they are even less likely to succeed if the languages themselves are not adequately analysed and provided with a responsible orthography. Currently, there exist accurate technical accounts of only a few Aslian languages (readable only by those with a linguistics background), as well as some less accurate impressionistic accounts, but hardly anything has

---

[4] This may explain the failure or refusal of some sections of Malaysian officialdom to understand the linguistic situation. For example, in a press release (Ikram 1997) issued shortly before he left the Department of Orang Asli Affairs (JHEOA), a departing Director-General asserted erroneously that *all* the indigenous peoples of Malaysia spoke Austronesian languages and that this made them all 'Malays' in some sense or other. The distinctive linguistic, and hence historical, situation of the Aslian speakers was thereby silenced by an officer well placed to have done something positive about it. (The document also illustrates the assimilation of 'Austronesian' to 'Malay', and vice versa, that sometimes befuddles scholarly writing in contemporary Malaysia.)

[5] Wnuk (2010: 4) reports that almost all the Aslian-speakers of southern Thailand are bilingual in their own language (Maniq) and southern Thai, but it is not known whether any of them are also literate in Thai.

yet been produced that meets the criteria of both accuracy and practicality. The numbers of Aslian speakers may seem to be too small to warrant such an effort, but the evidence of projects aimed at maintaining bilingualism in 'small' languages in other parts of the world (e.g. Watahomigie & Yamamoto 1992, Craig 1992) suggests that it might aid in reducing the social problems that sometimes arise from minority status.

Fourth, the Aslian languages constitute a largely untapped source for Malaysian and Southeast Asian culture-history study. This has been recognised since Blagden (1906) and earlier, and was urged again by Diffloth (1979a) in an accessible article aimed at non-linguist students of Malaysian culture. Since then, and with the growth in knowledge of the Mon-Khmer languages, the same point has been made by several other authors. The fact that the primary relationships of the Aslian languages lie to the north, on the Southeast Asian mainland, presents a significantly different perspective on the Peninsula's remoter history than commonly adopted. Aslian studies have contributed to a revised understanding of the Peninsula's complicated linguistic stratification, involving the interplay of at least two inputs each of Mon-Khmer and Austronesian. (The latter is evidenced by the presence of non-Malay or pre-Malay loanwords in the Aslian languages.) The reverse is also true: the phonological and lexical conservatism of the Aslian languages provides a significant source of information on the linguistic history of the mainland that might otherwise be beyond recovery. Indeed, it could well be that 15 to 20 percent of the linguistic sources for Mon-Khmer history – and hence the history of Mainland Southeast Asia as a whole – is carried by the Aslian division of Mon-Khmer.

Fifth, the study of Aslian languages is contributing to further advances in general linguistics. As just noted, Mon-Khmer historical linguistics has increasingly benefited from the data on Aslian. A recently developed feature of Aslian studies has been to place clearer emphasis on the patterns of differentiation peculiar to language varieties spoken by nomadic foraging (hunter-gatherer) populations, as opposed to those spoken in the more sedentary farming-based communities. This question is important because, while such nomadic communities have accounted for about 99 percent of human existence, most linguistic theory has been based on languages spoken – or more usually *written* – in the sedentary circumstances of the one percent of human existence that has followed the emergence of farming. Data on some of the Aslian languages will therefore have an important bearing on this emerging concern. Another contribution comes from recent studies of Aslian grammar, in sedentary as well as nomadic populations. These have paid close attention to semantic processes that are often new to linguistics, and which are contributing to the shift away from the context-ignoring, variation-blind, universalising trend – labelled 'Simple Nativism' by Levinson (2003: 25-46) – that has dominated mainstream professional linguistics until recently. As

noted by Matisoff (2003: 48-51), Aslian lexica are unusually rich, and display unusual or unique phenomena of considerable theoretical interest from a cross-linguistic point of view.

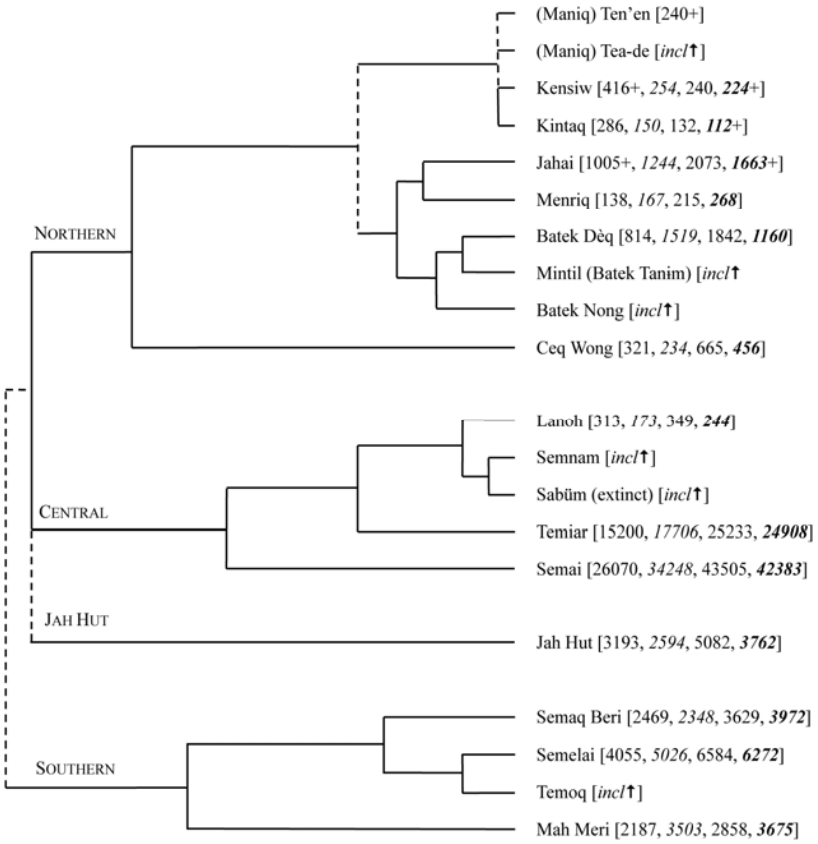
Finally, the Aslian languages as a group exhibit varying degrees of endangerment. Some are clearly moribund or endangered, while others appear for the moment to be reasonably healthy. But the long-term survival of all is less certain. As noted in more detail below, this is not necessarily a direct function of the number of speakers, as some of the 'smaller' Aslian languages may be more viable than those with more speakers. The relevant factors turn out to be not simply demographic, for questions of cultural identity and attitudes are also involved. This raises theoretical issues relating to the wider issues of language endangerment and extinction that are currently attracting close attention, and should be of special interest to readers of this publication.

1.1. Aslian demographics

In the earlier literature (Schmidt 1901, Blagden 1906, Wilkinson 1910, Pinnow 1959), the Aslian languages were known by a variety of labels that have now fallen out of use. The best known of these were 'Sakai' and 'Semang', roughly equivalent to what we would now regard as Central-plus-Southern Aslian and Northern Aslian, respectively. They have also been referred to, rather confusingly, as the 'Malaccan' languages, especially by scholars writing in German. A detailed survey of these changing usages is presented by Matisoff (2003: 3-9). The Aslian languages of southern Thailand are usually referred to collectively as 'Maniq'. In this paper I refer to the various languages by the names used by the primary linguistic author(s). These sometimes differ from the ethnonyms employed by the primary ethnographers, some secondary linguists, or Malaysian governmental agencies. However, I also provide the alternative names, if necessary, at appropriate places in the text.

The figure below is only approximately to scale. The genetic relationships indicated here are broadly correct according to the latest views. Dotted lines indicate undetermined relationships. Population figures in roman are from unpublished materials gathered by the Department of Orang Asli Affairs in 1996. Figures in *italic* are as recalculated by the Centre for Orang Asli Concerns (COAC) in 2000 from data gathered in 1999. The third set of figures, in roman, are for 2004, from unpublished materials supplied to the author by Juli Edo. The fourth set of figures, in ***bold italics***, are from 2008, as displayed in the Orang Asli Museum at Gombak. Figures for Maniq are from Wnuk 2010. (In 2008, non-Aslian, Malayic-speaking Orang Asli numbered around 53,000.)

Figure 1: Aslian languages: genetic relationships and population figures.



In Peninsular Malaysia and southern Thailand, Aslian-speakers currently amount to at least 100,000 out of some 160,000 Orang Asli. The latest obtainable population figures for the various Aslian-speaking groups are given in Figure 1. The figures are not strictly counts of speakers, but rather of members of ethnic groups, as administered by the Malaysian Department of

Orang Asli Development (Jabatan Kemajuan Orang Asli, JAKOA).[6] The numbers come from a variety of sources, as noted in the caption, and they are consequently not always consistent. The figures for the Maniq-speakers of south Thailand are not known for certain, but probably fall within the range 300-400.[7] According to Colin Nicholas (p.c.), the variation in the Malaysian figures is partly because the JAKOA's counts do not include Orang Asli living in urban areas, who therefore do not come under the JAKOA's primarily rural purview. The discrepancies have sometimes amounted to a 10-11% under-count in the case of Temiar and Semai. It is safe to assume, however, that all of these Temiar still actively speak their own language, and that the overwhelming majority of Semai can still speak Semai.

Some of the language labels in Figure 1 differ from the official listing of Orang Asli populations followed by the JAKOA. Semnam, Sabüm, Mintil and Temoq are categories not currently officially recognised (although the JHEOA did formerly list Temoq; see also Collings 1949), nor does the JAKOA distinguish between the various kinds of Batek. But these additional labels are necessary for linguistic and ethnological work. As indicated by the [*incl*↑] sign, population counts of these four groups are included in the immediately preceding figures.[8]

1.1.1. Aslian-speakers in southern Thailand

The data from southern Thailand adds very little to the total population of Aslian speakers, but does complicate the picture somewhat, for it suggests that there may be more linguistic, or at least dialectal, variety there than formerly suspected. One problem is that most of the Aslian-speakers of Thailand seem to have no names for their own languages – a situation not unknown elsewhere in Aslian, as for example with 'Chewong' (Howell 1984: 10-14). The label 'Maniq' (*Mani*ʔ, meaning 'own people') covers a range of three or more speech varieties to which other names, such as Tonga' (*Toŋa*ʔ), Mos (*Mɔs*) and Ten'en (Teanean, *Tɛn*ʔ*ɛn*) have also been applied in the literature – names that are no longer recognised by the people themselves (Wnuk 2010:

---

[6] Until 2010, the JAKOA was known as the Department of Aboriginal Affairs (Jabatan Hal Ehwal Orang Asli, JHEOA). Here JAKOA and JHEOA are used interchangeably, as appropriate.

[7] This estimate is based on figures in Phaiboon (1984: 1, 2006: 207), Hamilton (2002: 84-85), Bishop & Peterson (2003: 2), Burenhult (in press 2008: Table 1), and especially Wnuk (2010: 3).

[8] Mah Meri is also known as 'Besisi', 'Betisé'', or other such spellings; Jah Hut is sometimes written (more appropriately) as Jah Hĕt; Ceq Wong is sometimes written Chewong, Che' Wong or Siwang.

3). Phaiboon (2002, 2006) has presented comparative vocabularies for four 'Aslian dialects' in southern Thailand: 'Kansiw' (Kensiw), 'Yahay' (Jahai), 'Tean-ean' (Tɛnʔɛn) and 'Tea-de'. Wnuk (2010: 4) regards all varieties of Maniq as constituting a single language, but she does not otherwise comment on its relationship to the other Northern Aslian languages. Bauer (1993: 194) categorically regards these speech varieties (except Jahai, presumably) as varieties of Kensiw:

> 'Tonga' and 'Mos' do not exist; they are, in fact, Kensiw, located in Trang, Satul and Phattalung provinces in southern Thailand, some of whom refer to themselves as /maniʔ/ 'person' as do a number of other Aslian groups, including Kensiw of Yala. 'Kintaq' are linguistically Kensiw, and should therefore not be listed as a separate group.

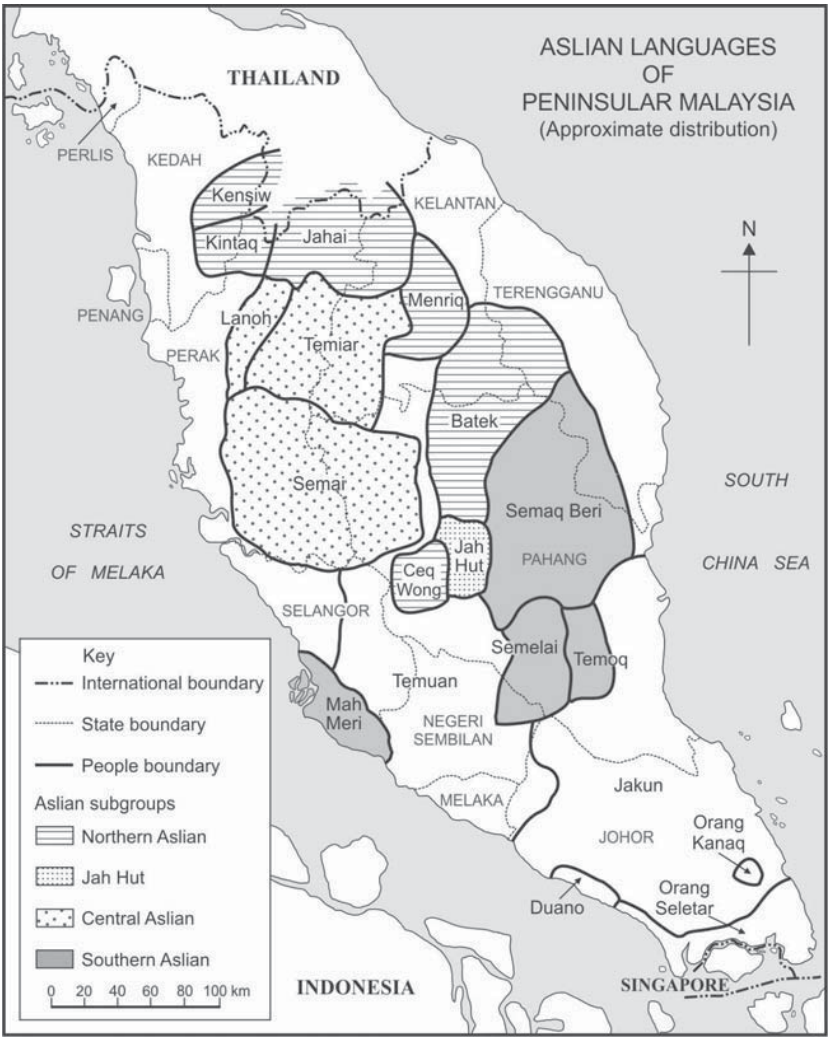
Other data suggests that the situation of Aslian in southern Thailand may not be quite so homogeneous. Using lexicostatistical methods – unfortunately not commensurable with the ones Benjamin (1976a) used for Malaysia, on which Figure 1 is mostly based – Bishop & Peterson isolated five speech-varieties that they thought might possibly represent up to three distinct (Northern) Aslian languages. Burenhult (p.c.) remarks of Tea-de that its position in his and his colleagues' work 'is unclear (it's not included in our database), but judging from Phaiboon's lists [2006] it's closer to Kensiw than Ten'en.'[9] Indeed, Diffloth (1975: 4-5) tentatively suggested that Ten'en (as 'Mos') shares significant innovations with the Jahai-Menriq group rather than with Kensiw. Dunn et al. (2011: 313-314) reaffirm the view that the degree of Aslian linguistic variety in southern Thailand has yet to be ascertained.

Jahai and Kensiw (there spelled Kensiu) are also spoken across the border in Malaysia. (The + sign in Figure 1 indicates that the given population figure does not include the Kensiws and Jahais of Thailand.) Under normal circumstances, such tiny speech communities would have to be judged highly endangered – as Phaiboon (2006: 209) does – but for Semang groups small

[9] Is the name 'Tea-de' related to the 'Tadeh' that, according to Tarapong Srisuchat (p.c.), may be in use for the Jahais of Narathiwat province? (He wasn't sure about the specifically Jahai identification.) Perhaps Tea-de is a variant of Tadoh, the Jahai name for a river just over the Malaysian border. That location corresponds to the position of Tea-de on the map in Burenhult et al. (2011: 3), but the proposed linguistic alignment with Kensiw would not. It may therefore turn out that Tea-de is a variety of Jahai, and that its position in Figure 1 is incorrect. Phaiboon's wordlist (2006: 210-223), on the other hand, does not appear to indicate that Tea-de is any closer to Jahai than to Kensiw.

population size does not necessarily mean imminent extinction, as I discuss further below.

Map 1: Aslian languages of Peninsular Malaysia



The map indicates the maximal known historical distributions of the languages rather than their present-day locations, which are more restricted. Temuan, Jakun, Orang Kanaq and Orang Seletar are Malayic dialects spoken by Orang Asli; Duano is an unclassified Austronesian language. (Map drawn

by Lee Li Kheng.)

1.2. Subgroupings and ethnology

The Aslian languages fall into four major subdivisions: Northern Aslian (Diffloth's 'Jahaic'), Central Aslian ('Senoic'), Jah Hut (a single language) and Southern Aslian ('Semelaic').[10] The approximate historically known locations of the languages are given in Map 1. A sketch map that also includes the Aslian languages of Thailand can be found in Dunn et al. (2011: 253); a more detailed map incorporating material on extinct languages is presented in Benjamin (1983).

The probable mutual genetic relationships of Aslian as currently understood are shown in Figure 1, which is based mainly on investigation of their historical phonology (Diffloth 1975). My own, less reliable, lexicostatistical calculations (Benjamin 1976a) produced the same result.[11] However, that study has since been superseded by completely fresh accounts (Dunn et al. 2011, Burenhult et al. 2011), based on new fieldwork on a wider range of language varieties and employing innovative statistical techniques. These new results are still in broad agreement with the two earlier studies, as the authors acknowledge. But there are nevertheless some significant differences of detail too, as illustrated in Figures 2 and 3. (Note, however, that Figure 2 primarily indicates degree of language interaction rather than genetic relationship; this is also true to a smaller extent of Figure 3.)

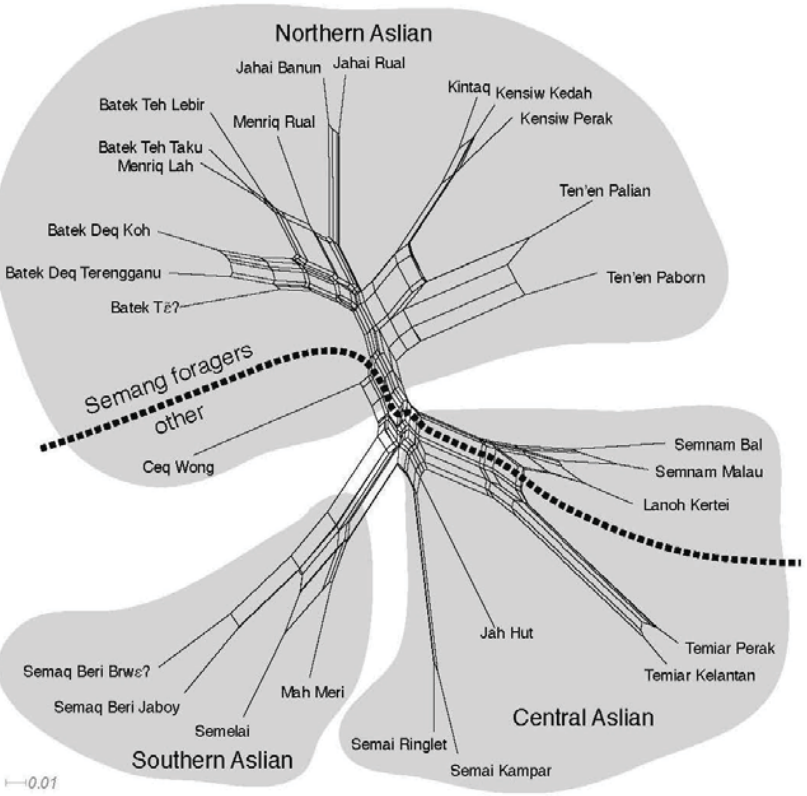
---

[10] Previously, Jah Hut was regarded as a Central Aslian language, but Diffloth & Zide (2005) present Jah Hut as a fourth primary branch of Aslian, a view supported by the more recent research incorporated into Figures 2 and 3 and by the current historical-phonology research of Timothy Phillips (p.c.).

[11] While that study has now been superseded by better formulated research, it contained hypotheses about the relative timings and trajectories of Aslian language movements in the Peninsula. Some of these suggestions still seem relevant, but others no longer fit with what has been discovered through later linguistic and archaeological investigations. The range of absolute dates I suggested on the basis of the glottochronological calculations must now be rejected. Even disregarding the questionable validity of glottochronology, the proposed dates have turned out to be far too early for any reasonable correlation with the Peninsula's archaeology. I now tentatively hold the view that the earliest advent of Aslian as a linguistic stock is associated in some way with the 'Tripod' culture, linked to the Ban Kao tradition of Central Thailand, that spread half-way down the western side of the Peninsula from approximately 3700 years ago (Benjamin 1997: 89-90).

Figure 2: Neighbour-Net clustering of lexicostatistical distances calculated from Aslian vocabulary lists

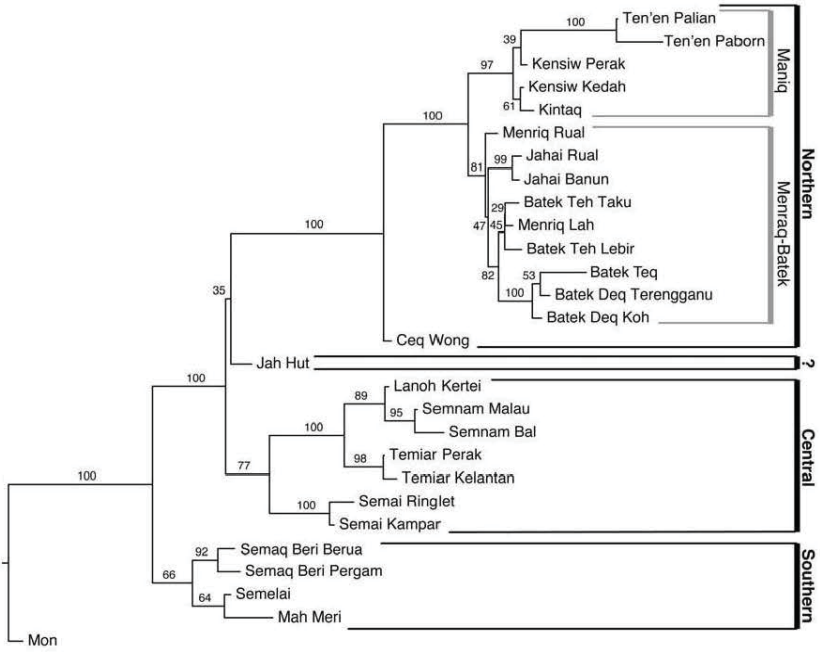
From Dunn et al. 2011: 305 (including the caption)
Reproduced by permission of John Benjamins, publishers of Diachronica.



The network is based on lexical data on 147 basic meanings in 27 Aslian language varieties. The method measures lexical similarity without a priori recognition of established sub-branches. The network should therefore not be regarded as a 'family tree'. The graph represents a matrix of distances without forcing the resolution of the major conflicts in the data.

Figure 3: Aslian family tree, rooted on Mon

From Dunn et al. 2011: 307 (including the caption)
Reproduced by permission of John Benjamins, publishers of Diachronica



This is a Maximum Clade Consistency tree, summarising the 750 post-burn in trees of the Bayesian phylogenetic tree sample with branch length equal to the median length of all congruent branches found in the sample. Branch length conflates rate of change and length of time to indicate the amount of diversification; numbers indicate the confidence we can have in each branch (expressed as a percentage).

Despite what is displayed in Figures 1-3, and despite the obvious congruences between them, it is not yet firmly established that the Aslian languages descend from a single proto-language. Given the emerging view that the relationships between the various branches of Mon-Khmer are of a generally 'flat' character (Sidwell 2010, Sidwell & Blench 2011), with relatively little nesting, the search for 'Proto-Aslian' may yet prove futile, and the earlier unitary models may need to be changed. Nevertheless, as hinted at in Figure 3, this does not preclude the view of Diffloth & Zide (2005) that the Aslian languages as a whole form part of a 'Southern Mon-Khmer' grouping along with Monic, and possibly also the Nicobar languages. Apart from Timothy Phillips's current research, there has been little work on Aslian historical phonology since Diffloth (1968, 1975, 1977) and the question therefore remains open.

The figures show that the Aslian-speakers consist of two rather large populations (Temiar, Semai), four middle-sized populations (Jah Hut, Semaq Beri, Semelai and Mah Meri), and several very small populations (the remainder). These different population sizes reflect the different modes of living that the people have followed over the centuries and (in some cases) millennia. In turn, the differences have had profound consequences for the sociolinguistic profiles of the Aslian languages, in particular that their degrees of endangerment vary considerably. The smallest groups have mostly been nomadic foragers living by hunting and gathering in the forest, supplemented by trading of forest products and casual farming or wage labour. The large Temiar and Semai populations have long been committed to sedentary swidden farming of grain and root crops supplemented by fishing, trapping and the trading of forest products. The middle-sized groups have followed a range of activities, but collecting-for-trade combined with farming has loomed larger among them than among the other Orang Asli. Each of these appropriative modes – foraging, horticulture and collecting-for-trade – *when dominant in its particular area* came to be institutionalised as a distinctive pattern of social communication, respectively 'Semang', 'Senoi' and 'Malayic'(Benjamin 1985a, 2011a).

This approach has sometimes been misunderstood. As used here, 'Semang', 'Senoi' and 'Malayic' identify kinship-based societal patterns rather than populations or 'ethnic groups'. I hypothesise that they label mutually dissimilatory lockings-in of the different societal traditions that were established long after the arrival of farming in the Peninsula. They are *not* meant to refer (as, for example, in the case of the Semang tradition) to some supposed continuation of unmodified Palaeolithic hunter-gathering practices. In other words, the emergence of the various sub-branches of the Aslian languages would probably have predated the emergence of these distinctive societal traditions, rather than resulting from them. Diffloth (2005: 79) dates the emergence of Aslian as a distinct subgroup of Mon-Khmer to around 3800

BP (before present), followed by the separation of Southern Aslian around 2800 BP and the still later separation between Northern and Central Aslian at around 2400 BP. As mentioned later, these dates accord well with a current view that would link the initial emergence of Aslian with the extension into the Peninsula of the Ban Kao ceramic tradition from central Thailand from around 3700 BP.[12]

The patterns of language differentiation in the Peninsula have mostly come to reflect the differences between the Semang, Senoi and Malayic societal traditions. In places, however, there have been 'mixed' societal patterns that did not undergo, or perhaps later rejected, these particular institutionalisations. (Hence, the qualification about areal dominance in the previous paragraph.) As noted below, these mixed patterns have proved important in interpreting the overall picture; they have particular relevance to understanding the circumstances of the Lanoh, Ceq Wong, Jah Hut, Mah Meri and Batek languages. With Figure 2 also in mind, let us now look at the various patterns in more detail.

The Southern Aslian languages split apart and moved away from each other more cleanly than the Central and Northern Aslian languages, and there has been relatively little secondary lexical borrowing between them (Benjamin 1976a: 75). The ancestral populations must have become more interested in contacts with outsiders than with each other, as expected given the collecting-and-trading way of life they were following. This preference for living in autonomous communities, relatively closed off from each other but adapted primarily to relations with members of other societies, is still common to the Southern Aslian speakers, to the Malays (Melayu) and to those southern Orang Asli who have adopted different versions of the Malayic societal tradition (Benjamin 2011a: 181-190).

The Northern Aslian languages (except for Ceq Wong) required special statistical techniques to uncover their mutual relationships, or even how many languages there are (Benjamin 1976a: 60-66). This was because the majority

---

[12] Other ethnological syntheses that bring together linguistic findings with data from archaeology and human biology include Blagden (1906: 470), Bellwood (1993), Fix (1995, 2011), Baer (2000), Bulbeck (2004, 2011), Dunn et al.(2011), Burenhult et al. (2011). Archaeological studies relating to the Orang Asli are Solheim (1980), Adi (1985) and Nik Hassan Shuhaimi (1997). With the provisos just stated, these are mostly in agreement with the views presented here. Rather different ethnological views, based on old-fashioned migratory-wave theories, are still popularly current in Malaysia as well as in academic writings. But these are at odds with the more recent research-based findings, and are in need of revision.

of Northern Aslian speakers (the Semang populations)[13] have long maintained an almost continuous mesh of communication with each other from southern Thailand down to central Pahang, generated by their preference for small-group nomadism coupled with wide-ranging intermarriage (Benjamin 2011a: 177-180, Dunn et al. 2011: 308). Most Semang camps contain members of several different locality-based or ethnicity-based subgroups, and each individual appears to draw on a slightly different lexicon than everyone else. (See Endicott 1997 for a detailed discussion of this situation among the easterly Northern Aslian speakers.)

The Central Aslian languages display yet another pattern. These (like the Southern Aslian languages) broke away from each other at various times in the past but (as with the Northern Aslian languages) this did not lead to a cessation of intercommunication, for there remains a high rate of secondary lexical borrowing between them (Benjamin 1976a: 74-76). This resulted from continued but restricted communication, largely through intermarriage, between their sizeable sedentary populations, separated by natural barriers. The people have long been swidden-cultivators, moving every few years between village sites the locations of which are narrowly circumscribed by the valley walls and the locations of the neighbouring villages upstream and downstream (Benjamin 2011a: 180-181).

Each subdivision of the Aslian languages has thus encapsulated within its major pattern of differentiation the decision of earlier speakers to follow distinct ways of life: Northern Aslian foragers, Central Aslian horticulturalists, and Southern Aslian collectors-for-trade. But there are minor patterns within each division that constitute important exceptions. As already noted, and confirmed by Dunn et al. 2011 (see Figures 2 and 3, taken from that source), one of the Northern Aslian languages, Ceq Wong forms no part of the Semang social mesh; the Ceq Wong speakers, however, are not completely nomadic, and frequently lead a fairly sedentary life (Kruspe 2009a: section 1; Howell 1984: 18-23). Among the Central Aslian languages,

---

[13] As already noted, 'Semang' is used here as a label for a distinctive pattern of social organisation rather than a physical population. The term is roughly, though not completely, coterminous with 'Negritos' as used by JAKOA and some other writers, but biologically influenced terms are best avoided in linguistic or sociological contexts. The categories 'Semang' and 'Negritos' both exclude the Ceq Wong, who nevertheless speak a Northern Aslian language. On the other hand, 'Negritos' usually includes the Lanoh, who speak Central Aslian languages. Nicole Kruspe (p.c.) has commented on this as follows: 'could it be that the Northern Aslian speakers were originally Mongoloid like the Central and Southern Aslian speakers, and the Ceq Wong did not intermarry with the Semang? Until we have genetic samples I guess we can only speculate.' For further discussion, see Benjamin 2013.

the Lanoh dialects do form a close mesh (separate from the Northern Aslian mesh), but their speakers were traditionally mostly nomadic (though possibly secondarily so), and their lifeways have usually been classed as Semang-like (Dallos 2011: 29-66). The fact that the Lanoh are considered to be phenotypically 'Negritos', while Ceq Wong speakers are as non-'Negrito' as the remaining Orang Asli groups, is a further indication of the different directions taken by their patterns of intermarriage.

In contrast to the Aslian pattern, the still extant Austronesian languages of the Peninsula are relatively unvarying. Today, there is hardly any evidence left of a distinctive, non-Malay, Peninsular Austronesian. A few coastal populations, notably those who still maintain active ties with their relatives in Sumatra, speak an Austronesian language ('Duano' or 'Desin Dolaq', Map 1) distinct from Malay. On the basis of a significant Austronesian but non-Malay component in the vocabulary of all Orang Asli languages, Blagden (1906: 435f.) hypothesised that there were once several 'generically Malay' languages (equivalent, perhaps, to our current use of 'Malayic') falling outside of the range of the 'specifically Malay' dialects, spoken by tribal populations along the coasts and in the lowlands of the Peninsula. The last such to be recorded was the speech-variety known as 'Kenaboi' from Negeri Sembilan, collected in the late nineteenth century (Blagden 1906: 403-405, Hajek 1998, Benjamin 2012e). Kenaboi appears to have since become absorbed into the 'specifically Malay' dialect (Temuan) spoken by the Orang Asli of that area today.

The extant Austronesian languages of the Peninsula therefore result primarily from replacement, not differentiation, with Malay sweeping many of the Aslian languages away. But why should the speakers of those languages have been so ready to accept a new language in such a short time?[14] Some Southern-Aslian speakers can presently be observed to be shifting to Malay; the speakers of Northern and Central Aslian, however, still resist the loss of their languages, although they speak excellent Malay when communicating with outsiders. Institutional factors, such as the choice of Temiar and Semai as

[14] The Malay language as we know it today – the 'standard' varieties especially – may well be the product of creolisation processes dating back no more than about 500 or 600 years. The dialects of the southernmost Orang Asli, which are also spoken in the islands of the Straits and on the Sumatran mainland (Kähler 1960), probably provided the matrix out of which the 'Johor-Riau Malay' (written) standard was eventually created. The notion that Malay *replaced* other languages is therefore probably too simple an account for many areas (though it certainly applies to much of the progressive eating-away of the Southern Aslian languages). For amplifications of this view of the later external history of the Malay language see Benjamin (1993, 2009b). The most informative external history of Malay is Collins (1998).

the two main languages for use in Orang Asli radio programmes,[15] undoubtedly play a part in this, but the underlying difference is one of societal *consciousness*. The Northern-Aslian Semang populations, along with the Temiar and hill Semai, have until recently been orientated more towards their own immediate situations than to the outside world, while the southern Orang Asli and the Malays have been orientated more towards the outside world than to their fellows in other villages. The southern peoples, with their more 'transcendental' mode of consciousness (as I call it) and peasant-like socio-economic circumstances, have proved more susceptible to the spread of the exogenous but centralised and assimilatory culture of the traditional Malay states than the northerners, who have retained their variously 'immanent' or 'dialectical' modes of consciousness and tribal socio-economic circumstances up to the present day.[16]

Thus, any consideration of the degree of endangerment of Aslian languages – examined in more detail later – requires us to take account of the long-term interaction between those languages and Malay.[17] A simple

---

[15] These programmes commenced in 1959 and are currently broadcast as the *Asyik FM* channel of Radio Malaysia, between 8am and 10pm daily. The available frequencies are 6025 KHz (short wave, nationally), and on FM 89.3, 91.1 (Klang Valley only), 102.5 (western Pahang) and 105.1 MHz. The channel is also available over the Internet from the station's website, http://www.asyikfm.my, by clicking the *dengar sekarang* ('listen now') button. The Semai and Temiar announcers incorporate a substantial amount of Malay into their discourse, especially in the news bulletins, sometimes to the extent of submerging almost all Aslian elements. When taking live calls from listeners, however, Temiar and Semai are often used together. In the past, there have been occasional broadcasts in other Aslian languages, such as Mah Meri; but these no longer occur. Programmes in the Malayic dialects Temuan and Jakun are also broadcast daily, sometimes interspersed with phrases in Semai and Temiar.

[16] The contrast drawn here between 'immanent', 'dialectical' and 'transcendental' modes of consciousness has been discussed more fully elsewhere (Benjamin 1993: 348-349). For studies relating this approach, explicitly or implicitly, to the ethnology of the Malay Peninsula, see Benjamin (1985a, 2011a).

[17] See van Reijn (1975) for a discussion of older Malay loans in Aslian that have disappeared from modern Malay. The reverse process – loans from Aslian into Malay that have now disappeared – has also occurred in the past. This is evidenced by some striking entries in Bowrey (1701), a dictionary of Malay: for example, 'Nephew, by a brother, Cōmon', cf. Temiar *kəmɔn* 'nephew, niece', which has close cognates throughout Aslian and elsewhere in Mon-Khmer (Shorto 2006: 328-329, no. 1187). Other probable Aslian words found in Malay (but not more widely in Austronesian) include *ketam* 'crab' (Shorto 2006: 369, no. 1348), *helang* 'eagle' (Shorto 2006: 221, no. 714), *semut* 'ant' (Shorto 2006: 257-258, no. 873), *cucu* 'grandchild' (Shorto 2006: 79, no.43), *merak* 'peacock' (Shorto 2006: 160, no. 416), *pergam* 'imperial pigeon' (Shorto 2006: 362, no. 1319), *dekan* 'bamboo-rat' (Diffloth 2005: 78), and *jenut* 'salt lick' (cf. Temiar *jənɔɔd*, from *jɔɔd* 'to lick', possibly Shorto 2006: 313-314, no. 1106). As Wilkinson (1915: iv) noted: 'there are many instances in which the immigrant race

measure of this is provided by the differing percentages of Malay loans found in the basic vocabularies used in my 1976 lexicostatistical study (Table 1). The figures show that the highest rates of borrowing from Malay were found among the smaller and/or forest-collecting populations, especially where the lowland speakers of Northern Aslian languages have straddled old Malay routes through the forest. The lowest rates of Malay borrowings, on the other hand, are found among the large farming populations. This results both from their relatively remote situation and from their higher degree of economic self-sufficiency. Kruspe (2009a: section 3.3) has also commented on the limited degree of linguistic contact between Ceq Wong and Malay as resulting from the people's isolation. However, other factors are at play too. Among the Semai, the very high extent of dialect-fission (Diffloth 1968, 1977; Phillips 2005a) inhibits retention of the language as a whole, because each effective speech-community is actually rather small. The situation of Temiar, on the other hand, favours its retention. Historically, it was buffered from linguistic contact with Malay by the Aslian languages that surrounded it. This has had several other consequences: its lexicon has been enriched by multiple borrowings from other Aslian languages, and it has become something of a lingua franca among the northern Orang Asli.

Table 1

Aslian languages of Malaysia: Loan rates from Malay, percent (Benjamin 1976a: 73)

Northern Aslian

Kensiw	Kintaq	Jahai	Menriq	Batek Dèq	Mintil	Batek Nong	Ceq Wong
7	5	11	10	21	16	10	12

Central Aslian

Semnam	Sabüm	Lanoh	Temiar	Semai I	Semai II	Jah Hut
10	5	5	2	7	5	18

Southern Aslian

Mah Meri	Semaq Beri	Semelai	Temoq
25	10	23	14

---

has borrowed names of local animals and plants from the earlier inhabitants of the country'.

Most of the Southern Aslian languages display a very high rate of Malay borrowings, as is consonant with the life-patterns of their speakers, which have long involved a relative loosening of ties among themselves in favour of trade links with Malays and others downstream. In the case of Mah Meri, this has gone so far that Skeat (in Skeat & Bladgen 1906, Vol. 1: 688, on 'Besisi') had some trouble discerning the Aslian element in the copious song texts he recorded in the 1890s. In Semelai, the pattern of borrowing is not uniform across all domains of the language, as indicated by Nicole Kruspe (p.c., slightly modified):

> Examining an area such as songs would skew one's perception of the situation. *Jampi* (spells) and many songs are almost exclusively in Malay. Many cosmological terms are based on borrowed Malay lexemes, while in everyday life they use Semelai lexemes. For example, in the underworld where the Spectre Huntsmen dwell there is a 'sea of blood', referred to by the Malay phrase, *lawot darah*. But the everyday Semelai word for blood is *maham*, not *darah*. Secondly, the borrowing does not always indicate replacement, but may be a means of increasing the richness of the language. The Semelai word for one type of ghost is *səkɔʔ*, but a division within this group are known by the Malay word for 'ghost', *hantuʔ*.

The use of Malay expressions or lexemes in religious contexts is indeed widely met with in Aslian languages, indicating that in former times the animistic practices of the two peoples were sometimes performed jointly.

1.3. External relationships

An underlying assumption in the preceding comments is that the present-day Aslian-speakers are the descendants in whole or part of the people who produced the various early artefacts and sites uncovered by archaeological work in the Peninsula. This runs counter to the standard popular view, however, for even today the 'origins' of the Orang Asli are considered by many writers on Malaysia to be something of a mystery. (The 'origin' of the Malays too is a very frequent topic of discussion, even in academic circles.) It has been the fashion to assume either that the prehistoric remains were left by other peoples (who are presumed to have moved southwards long ago to settle Island Southeast Asia and Australasia), or that each of today's Orang Asli 'tribes' has a separate origin to the north outside the Peninsula. (I have actually been asked by an educated commentator to state just which part of Yunnan the Temiar came from, as opposed to their neighbours the Semai!) Archaeologists and prehistorians have dropped this view, however. In any

case, it was never very widely held among the appropriate scholars. As already noted, important new studies discussing the earlier views and providing up-to-date syntheses have considerably reduced the mystery that some commentators continue to attach to the origins of the Orang Asli and their languages. It is still sometimes asserted, for example, that the Aslian languages bear some undetermined relationship to the Mon-Khmer languages of Mainland Southeast Asia without actually *being* Mon-Khmer languages,[18] a view that can no longer be sustained.

In the 1970s, Gérard Diffloth showed that the Aslian languages were indeed a well-constituted sub-group of Mon-Khmer proper. He first did so in his study of Jah Hut (Diffloth 1976a: 75-81) and then in a non-technical paper (Diffloth 1979a), but he has given further demonstrations several times since in other publications (see especially Diffloth 1984, 1997; Diffloth & Zide 2005). In recent years he has been engaged in the reconstruction of Proto-Mon-Khmer phonology through a lexical study of all branches of the family. He has proposed that Mon-Khmer might fall into three main subdivisions: Northern Mon-Khmer (including the Palaungic languages and Khasi, among others), Eastern Mon-Khmer (including Khmer itself and Vietnamese, among others), and Southern Mon-Khmer.[19] He suggests in turn that Southern Mon-Khmer probably consists of three sub-divisions: Monic (Mon in Burma and Thailand, and Nyah Kur in Thailand), Aslian and Nicobarese (several languages spoken on islands lying just north of Sumatra but belonging to India). While it takes a historical phonologist's eye (or ear?) to discern any special closeness between Nicobarese and Aslian, the closeness of Monic and Aslian is altogether more readily apparent. Two publications in particular –

---

[18] A recent example is Grimes (2000), as cited by Anderson (2009: 724). Anderson nevertheless closes his article with the sentence 'Among present-day specialists in Mon-Khmer linguistics, Gérard Diffloth deserves special mention.' The Mon-Khmer membership of the Aslian languages is accepted by all current linguists (including Diffloth) with first-hand experience of these languages, as well as by other historical linguists (such as Shorto 2006, Sidwell 2007, 2008) who have worked on Mon-Khmer as a whole.

[19] For a diagrammatic representation of this proposal (with an alternate set of higher-level labels), along with a suggested time-scale, see Diffloth (2005: 79). However, this view is not yet fully established. As already noted, recent work on the Mon-Khmer stock as a whole has questioned the status of some of the proposed intermediate subgroupings, on the grounds that the dozen or so divisions of Mon-Khmer are possibly of equal status – 'flat' rather than nested – and that this results from a very early and rapid expansion of the Austroasiatic family (including the Munda languages of India) in all directions from its homeland. Paul Sidwell (2010) and Sidwell & Blench (2011) in particular have argued for this approach – and also that 'Mon-Khmer' may not be as distinctive a subgrouping within Austroasiatic as is conventionally assumed.

Shorto (1971) and Diffloth (1984) – have demonstrated the connections many times over. As yet, however, no conclusive study of the historical details of the connection between Monic and Aslian has been published. Another specialist on Mon, Christian Bauer, has also been looking at Aslian (Bauer 1991, 1992a, b).

However, the possible place of Aslian within a Southern Mon-Khmer group is not the only problem. There are grounds for thinking that northern Peninsular Malaysia and the Isthmian regions of Thailand have witnessed not one, but three, layers of Mon-Khmer-speaking presence.[20] And there is evidence that Mon was present as the language of lowland civilisation in the Isthmian tracts and as far south as Perak and Kelantan until around 1200 CE. Khmer too seems to have left some traces in the same region, and even further south.[21] In addition to their basic character as Mon-Khmer languages, the Aslian languages also contain lexical evidence of possible secondary contact with both the Mon and the Khmer languages specifically (Benjamin 1997: 105-112; Bauer 1992a). Moreover, the Aslian languages also contain many words in their lexicon that are clearly of Austronesian, but not Malay, provenance (Blagden 1906: 435-438; Benjamin 2009a). There are many Malay loanwords too, often from earlier forms of Malay and providing evidence thereby for the history of that language (Blagden 1906: 432-435; Diffloth 1976a: 112). As writers have been noting for over a century, these features betoken a rather more complicated linguistic history for the Malay Peninsula than the regular view suggests.

---

[20] The hundreds of place- and river-names in the Aslian languages that are well-formed phonologically as Mon-Khmer words but which have no meaning in the present-day languages may provide an important clue to a Mon-Khmer presence prior to the advent of Aslian. But, given the paucity so far of relevant Mon-Khmer etymologies, these place-names may alternatively indicate the presence of an unidentified pre-Mon-Khmer stratum in the Peninsula and/or the Aslian propensity for forming new words by employing existing phonological patterns. This is a topic that needs further research. Burenhult (2005b: 25-29) has made a start, in his study of Jahai place-names; these, however, turn out to be mostly semantically transparent and linked to specific origin-tales.

[21] This view is in possible conflict with the claim that 'mainland' Austronesian languages were anciently spoken along a continuous tract of coastal land stretching from central Vietnam all the way to Peninsular Malaysia. While there is no reason to doubt the importance of (Austronesian) Chamic along the eastern parts of that stretch (Thurgood 1999: Chapters 2 and 3; Blust 1994), there is good linguistic and archaeological evidence in favour of Mon (and sometimes Khmer) as the main language of the various states that stretched around the Gulf of Thailand before Thai and Malay were later (between, say, the 12th and 14th centuries) imposed on the region as the main languages of civilisation.

2. Research on Aslian: a survey

2.1. Earlier work

Much ethnographic work on the Orang Asli has been carried out since the 1960s, and it is generally of a very high scholarly quality. The same can also be said for the study of Orang Asli prehistory, population biology, sociology and history. A general appraisal of Orang Asli research, now out of date, can be found in Benjamin (1989). Parkin (1991: 41-56, with a bibliography on pages 152-159) provided an accessible summary of the 'external' facts about the Aslian speakers, though the book as a whole is linguistically unreliable. These were both superseded by the exhaustive bibliography of Orang Asli studies assembled by Lye Tuck-Po (2001). The linguistic materials in the latter volume can be found on pages 158-168, but these came too early to incorporate the important work that has been done since the book was published. (This is also true of Matisoff's important survey of Aslian, mentioned later, which was written in 1983 but not published until 2003.) A more recent synoptic study of Orang Asli studies is Lye (2011).

Regrettably, the generally high quality of ethnographic work on the Orang Asli had not, until recently, been so true of studies on Aslian languages. There was a flurry of comparative interest earlier in the twentieth century (especially Schmidt 1901 and Blagden 1906; see also Pinnow 1959), but these have to be read with care, as they were mostly based on the second-hand records of amateur vocabulary collectors. Moreover, the language-names they employed are usually quite different from those currently used. (For a partial concordance of the older names with the current ones, see Benjamin 1976a: 125-126.) The earliest phonetic transcription of an Aslian language (probably Jahai) into IPA symbols was done over a century ago, by R. J. Lloyd in 1902 (see footnote 24 below). For the transcription itself, see Skeat & Blagden (1906: 628-629), and also Blagden's comment on it in Lloyd (1921-23: 27). The colonial administrator and scholar, R. J. Wilkinson (author of an outstanding Malay-English dictionary) also interested himself in Aslian, discovering languages not recorded by Blagden. Based on his own systematic collection of word-lists, he proposed a new classification of the languages (Wilkinson 1910), but this has not stood the test of time. He also published a useful word-list of 'Central Sakai', i.e., Semai (Wilkinson 1915).

The list of language-names used today derives from the work of Paul Schebesta (1928a), the major ethnographer of the Orang Asli in the earlier years of the twentieth century. Apart from discovering several new groups and regularising the ethnological terminology, Schebesta published grammatical sketches (translated and reworked by Blagden) on two Aslian languages: Jahai (Schebesta 1928b) and a Central Aslian language he called 'Ple-Temer' (Schebesta 1931). The latter differs significantly from the variety of Temiar

described in my own publications, being closer to the north-western dialect of the language that I came across along the Perak river in the 1970s, but which has yet to be described in print. It is possible that Schebesta's paper is the sole record of a language or dialect intermediate between Temiar and Lanoh which has since disappeared (see Noone 1936: 52). Schebesta's two studies are rich and valuable, but they are less than rigorous: scholars who do not already know something of Aslian will use them at their peril. Burenhult (2005a: 9-10), however, has presented a welcoming assessment of Schebesta's account of Jahai.

Little more was done along these lines for the next forty years until Iskandar Carey, a professional anthropologist who was heading the JHEOA at the time, produced a basic textbook of Temiar (Carey 1961) intended primarily for use by officers of the JHEOA; he also published a word-list of Menriq (Carey 1970). (Carey once told me how surprised he was that Dewan Bahasa Dan Pustaka decided to publish his Temiar book, as it was meant for internal circulation only.) While the copious texts reported in the book suffer from most of the transcriptional faults mentioned below, the texts themselves seem to be authentic examples of ordinary spoken Temiar. They are all the more interesting for being much more complex than, and contravening, some of the generalisations Carey makes in his short grammatical introduction. Robert Dentan, fresh from his initial period of ethnographic fieldwork (1963), took Carey's book as a model in producing an account of the Semai language. This has remained unpublished, but a revised version (Dentan 2003) is available online.

A detailed survey of Aslian studies from the nineteenth century up to the early 1980s was provided by Matisoff (2003) in a long and insightful essay written in 1983, and originally intended for inclusion in his planned volume on the languages of Mainland Southeast Asia. This remains a key study. Relying largely on data from Diffloth, Asmah and myself, as well as surveying the older literature, Matisoff goes well beyond what appears on the surface of the reports he uses, and he presents a thorough and illuminating exploration of much that would otherwise have gone unrecognised. I make use of some of his discussion below. Unfortunately, Matisoff's essay predates the more recent studies, and therefore could not take account of them.

Since then, a dictionary of 'Sengoi', i.e., Semai has been published (Means & Means 1986), followed by a Temiar dictionary (Means 1998). Both dictionaries are lexically fairly rich but they suffer from many of the transcriptional faults mentioned later. Being based on insecure phonetic analysis, they can be used safely only by those who already know something of the languages. The 'Sengoi' dictionary appears to be aimed largely at institutionalising the non-phonemic orthography for Semai used by the (Lutheran) Sengoi Mission, rather than to serve as a primary scholarly

resource. The Temiar dictionary contains a grammatical introduction, the information in which seems to have been obtained by interviewing Temiar *about* their language rather than by analysis of actual Temiar utterances or texts. This has generated several incorrect statements, such as that the infix *-a-* forms 'plural' verbs (Means 1998: 12), when in reality it serves as the middle-voice inflection, useable with singular subjects too (see Table 4 below). The body of the dictionary is consequently peppered with spurious 'plural' verbs. Both dictionaries contain English-to-Aslian sections, intended for Semai and Temiar learners of English.

A wordlist of Semai has also been circulated by an agency of the Malaysian Ministry of Education (Pusat Perkembangan Kurikulum 2000); this too, despite its 2512 entries, suffers from most of the transcriptional problems discussed later. A much improved version (SIL International 2002), produced with input from professional linguists, is more reliable in supplying a near-phonemic transcription in brackets immediately after the Semai headwords. (This seems to be essentially the same as a limited-distribution document prepared by Pusat Perkembangan Kurikulum 2004.) But the expected 'practical' orthography (in bold) still overrides the phonemic distinctions, presumably in a desire to look more Malay-like**: 'tenghok** [tenghòòg]**'**, vs. **'tapog** [tapok]', with <o> standing for both /ɔɔ/ and /*o*/. A careful revision and amalgamation of this confusing series of documents has finally been published in a more professional format (Basrim et al. 2008). Employing diacritics and apostrophes rather than IPA symbols, this dictionary nevertheless records Semai phonology closely, with the exception of failing to notate word-initial glottal stops. Its aims are practical rather than strictly linguistic, but it goes a long way towards bridging that gap. It is notable also for having been compiled primarily by native speakers. As Basrim says in his Introduction, 'this dictionary was produced by Orang Asli who wish to see their language made known to other people'. The definitions are in Malay, but there is an English index at the end of the volume. This important new dictionary more properly counts as 'current' work, to which I now turn.

2.2. Current linguistic research and publications

Gérard Diffloth began his interest in comparative Mon-Khmer with fieldwork on Aslian in the 1960s, while teaching Dravidian linguistics at the University of Malaya. He has since published several papers on Semai, Jah Hut and comparative Aslian historical phonology (see References). Nicole Kruspe completed a comprehensive PhD thesis (Kruspe 1999) on Semelai grammar, based on a scrupulous semantically-orientated analysis of her extended corpus of Semelai texts. This prize-winning study, later published in revised form (Kruspe 2004a), ranks as perhaps the most detailed account of any Mon-Khmer language, alongside Thompson's Vietnamese grammar. (In my view,

there is also no description of Malay or Indonesian that can match it for depth of analysis.) Semelai is now the most fully described Aslian language (see also Kruspe 2004b). Since 2007, Kruspe has been researching several other Aslian languages: Ceq Wong (Kruspe 2009a, 2009b, 2011a, in preparation a), Mah Meri (Kruspe 2010, 2011, in preparation b; Kruspe & Hajek 2009), Semaq Beri (Kruspe, in preparation c), and Batek. Currently, she is coordinator of *The Repository and Workshop for Austroasiatic Intangible Heritage* (REWAII), hosted at Lund University in Sweden. It is planned that this should house a significant collection of Aslian linguistic and cultural data.

Kruspe's 400-page dictionary of Mah Meri (Kruspe 2010) is the first – and so far the only – professional dictionary of any Aslian language. She has managed to produce a readable and accessible volume without skimping on phonemic or semantic niceties. The headwords are presented in a semi-practical transcription, fully mappable onto the phonemic structure, as explained in the volume's Introduction (Kruspe 2010: 6-10). This also contains a brief grammatical sketch (Kruspe 2010: 15-18). Each entry gives the pronunciation (in a 'narrow' IPA phonetic transcription), the meaning in both English and Malay, and (where relevant) short articles on the word's usage and/or its cultural significance, along with references to its historical appearance in Blagden 1906. For example, the entry that reads '**cɔ̤ʔɔ̤k** [tɕɔˈʔɔ̤k] *v.t.* 1) roast. *bakar.* 2) burn incense. *bakar.*' (Kruspe 2010: 131) is followed by examples of usage and a reference to item B468 on p. 550 of Blagden's vocabulary. Appendices provide English–Mah Meri and Malay–Mah Meri finderlists. The dictionary can therefore be used guesswork-free by almost anyone prepared to make a small effort, including native speakers, other Malaysians, casual enquirers, ethnographers, or linguists. However, the order of presentation is not alphabetic but 'Sanskritic' (like Shorto's two Mon dictionaries and Kruspe's own Semelai vocabulary, 2004a: 448-465), with back consonants listed before front consonants and front vowels before back vowels (Kruspe 2010: 11). This takes some getting used to, even for linguists. (The finder-lists, of course, are in conventional alphabetical order, which makes it easier to discover the wanted Mah Meri word.)

Niclas Burenhult completed a detailed PhD thesis on Jahai grammar in 2002; this has since been published, slightly revised, as Burenhult (2005a). Like Kruspe (2010), he presents his glossary (Burenhult 2005a: 168-208) in both phonemic and narrowly phonetic transcription. Burenhult (1999), a comprehensive bibliography of Aslian linguistics, was later incorporated into Lye (2001). Burenhult continues to produce studies of typological and areal issues, with special attention to phonology, morphology and semantics. (See Burenhult in the References, along with Levinson & Burenhult 2008 and Terrill & Burenhult 2008.) More recently, with Claudia Wegener, he has worked on the moribund Semnam dialect of Lanoh (Burenhult & Wegener 2009). Importantly, Burenhult, both alone and in collaboration, has produced

sophisticated discussions of the sociolinguistic character of nomadic and semi-sedentary hunter-gatherer communities (Burenhult in press; Dunn et al. 2011).

Sylvia Tufvesson is currently writing a fieldwork-based thesis on Semai expressives (Tufvesson 2008, 2011). Also currently working on Semai is Phillips (2005a, 2005b, 2009), who has been concentrating on dictionary-building and on Aslian historical phonology, the subject of his PhD thesis (2012), and the first such study since Diffloth's. Nancy Bishop and Mary Peterson continue to do detailed work on the Maniq (Northern Aslian) dialects of southern Thailand (see References). Ewelina Wnuk has completed a Master's thesis on Maniq phonology (Wnuk 2010), which has clarified many issues about the people as well as the language; she is now preparing a PhD thesis on Maniq. (See also her website and Wnuk & Majid 2012. Also working on Northern Aslian at Lund University is Joanne Yager (in preparation)).

I continue to work part-time on Temiar (see References). My early account of Temiar grammar (Benjamin 1976b) spawned a secondary theoretically-orientated literature concerned mostly with explaining the Temiar verb's complex reduplicative morphology. Adams (1989) includes secondary discussion of Aslian material. Majid et al. (2011) deals with reciprocal ('each other') constructions.

Kruspe, Burenhult, Tufvesson and Wnuk have collaborated in developing a large Aslian multimedia digital archive containing their recently collected data. This REWAAI archive is still under construction and not yet fully accessible for downloading.[22] One of their aims is to digitise the recordings that document cultural heritage, in the hope that they may be accessed and utilised not only by researchers but also by the language communities themselves. Together, these archives represent the most comprehensive documentation so far of any branch of Austroasiatic (and possibly of any Mainland Southeast Asian language grouping) according to modern standards. The Division of Linguistics, Nanyang Technological University, Singapore plans to organise an online archive of annotated spoken samples of Southeast Asian languages which should eventually incorporate some Aslian material. An 'Aslian Etymology' website, incorporating a variety of secondary data, is run by a group of Russian linguists (Starostin et al. 1998–2011), but this

[22] The archive structures can be inspected at the link: http://corpus1.mpi.nl/ds/imdi_browser/, by first clicking on 'DoBeS archive' and then (currently) 'Semang' and 'Semoq Beri and Batek'.

contains no first-hand material. The important journal *Mon-Khmer Studies* has begun to publish articles on Aslian languages after a lapse of many years.[23]

2.2.1. Malaysian and Thai research on Aslian

In Malaysia, some full-time linguists with a primary specialisation in Malay studies produced shorter studies of Aslian languages earlier in their careers. These include Asmah Haji Omar (1964, 1976) on Kintaq and Nik Safiah Karim & Ton Binti Ibrahim (1979a, 1979b) on Semaq Beri. Sadly, no Malaysian linguist did any such research during the 1980s and 1990s, apart from Teoh Boon Siong (part-author of Holaday et al. 1985, 2003) on Jah Hut, and some undergraduate exercises. Few Masters' theses and only one PhD thesis on Aslian linguistics have been produced from any Malaysian university, but in the last few years there has been a renewal of interest. Asmah Haji Omar returned to Aslian by editing a volume of 20 short linguistic and ethnographic studies (in Malay) on Mah Meri by her colleagues (Asmah 2006), the only book-length study of Aslian published in Malaysia. Two Masters' theses in Aslian linguistics were formerly in preparation at Universiti Sains Malaysia (Penang): Alang Sabak (a native speaker) on Semai ethnolinguistics and Che Mat Radzi on Kensiw morphology. But I have no further information on these. Yap Ngee Thai did some significant secondary analysis of Temiar morphology (Yap 2009a), and has begun fieldwork on that language (Yap 2009b). She also set up an online Basic Temiar course for students at Universiti Putra Malaysia (not currently functioning). Several studies, mostly on Northern Aslian, have been presented at conferences within Malaysia, but they are not all easily available. Among those that have been distributed in print and/or on the Internet are papers, singly or jointly, by Alias Abd Ghani, Fazal Mohamed Mohamed Sultan, Tengku Intan Suzila Tengku Sharif, Harishon Radzi, Salasiah Che Lah and Fauziah Mat. None of these researchers has yet made Aslian studies the centrepiece of their research, although some have expressed interest in doing so, including Tengku Intan Suzila, who has commenced a doctoral programme on the Batek language(s) of Pahang at Universiti Kebangsaan Malaysia. However, they have undertaken fieldwork on the languages, which in most cases is continuing.

---

[23] These can be accessed from the online archives at http://www.mksjournal.org and http://sealang.net/archives/mks. A planned handbook, *The Mon-Khmer Austroasiatic Languages* (edited by Mathias Jenny & Paul Sidwell, in preparation) will probably be produced in association with this journal. The volume will include grammatical sketches by Nicole Kruspe (Semaq Beri) and myself (Temiar). Northern Aslian will be covered in a jointly written overview chapter by Kruspe, Burenhult & Wnuk which will bring together fresh materials on Ceq Wong, Jahai and Maniq.

So far, however, most of this fieldwork has been too brief for the investigators to learn to speak the language and/or to collect extensive phonemically transcribed texts. This has generated some phonological inadequacies and lexical errors. A complicating factor is that few of the Malaysian publications have made reference to the literature on Aslian that has been published outside Malaysia since the 1970s. (There are even instances where Blagden 1906 has been treated as the latest word on the subject.) But the reverse is also true: the Malaysian linguists' publications have been little circulated, and are consequently mostly unknown to linguists working outside Malaysia. Currently, however, plans are under way to rectify this divide.

In Thailand, fieldwork on Mon-Khmer languages has been given some priority, and several important studies of the Northern Aslian languages of southernmost Thailand have been done, especially from Mahidol and Thammasat Universities. I made use of one of these sources earlier in this paper. For other work, published and unpublished, see the listings under Bauer, Bishop, Peterson, Phaiboon, Pothisarn and Theraphan in the References. (I am only familiar with Thai materials written in English, but there are yet other studies written in Thai, which I cannot read, some of which are listed in the References.)

2.3. Non-linguistic sources

Several ethnographic fieldworkers have gathered reliable lexical data on various Aslian languages, employing phonologically accurate transcriptions. The following have been especially careful in their transcriptions: Endicott (1972a, 1979) on Batek Dèq; Gianno on Semelai (1990, with an accessible account of the phonology and morphology on pp. 158-163); Lye (1997, 2004) on Batek Iga'; Vogt (so far unpublished) on Batek Nong; and Dallos (2011) on Lanoh. Other ethnographers, such as Roseman (1991, especially pp. 185-189) and Jennings (1995) on Temiar, have taken advice from Diffloth or myself on transcribing their field languages, and persuaded their publishers to employ an accurate IPA-derived orthography. But the general quality of the linguistic data contained within ethnographic accounts of the Orang Asli otherwise remains unsatisfactory, and much valuable information is being lost. This is as true of work done by Malaysian scholars (including researchers who are themselves Orang Asli), as of the work of non-Malysians. Some researchers are too much in thrall to the spelling conventions of romanised Malay to risk using extra symbols. The vowels in particular are recorded poorly.

As detailed in Section 4.2, all Aslian languages have at least nine places of vowel articulation. In many cases these are further cross-cut by phonemic oppositions between oral and nasal and/or between long and short. The

symbols <a,e,i,o,u> just will not capture these rich vowel systems, which amount to thirty or more phonemic distinctions in some cases. Consequently, all-important distinctions are lost in the records. The differences between: *o* and *ɔ*, *e* and *ɛ*, stressed *ə* and *ʉ*, *i*, *ɯ* or *ɤ*, long and short vowels, and word-final *ʔ* and *k*, regularly go unrecorded. Other problems that arise are: the omission of syllable-initial or word-initial *ʔ* (<*en*> for *ʔɛn*, <*en en*> for *ʔɛnʔɛn*); the use of syllable-final or word-final vowels instead of consonants (<*senoi*> for *sɛnʔɔɔy*); and the failure to recognise word-final palatals for what they are (<*poit*> for *pɔɔc*).[24] Some writers have represented the languages as completely monosyllabic, which they are not (<*cheb chib*> or <*cheb-chib*> for *cɛbciib*), and word-boundaries frequently go unrecognised. British and Australian writers are wont to add <*r*> and <*h*> to indicate vowel quality or length, such as <*ter*> for *təəh* or *tʉʉh*. And almost all such writers are confused about the differences between affixes, clitics and separate words, so that the hyphen either gets misused, or not used when it should be. These same problems, of course, are precisely those that would inhibit the development of literacy in the Aslian languages themselves. Many Aslian-speakers are now literate – but in Malay, and to a small extent in English, rather than in their own languages. Unfortunately, some of the recent studies by Malaysian linguists have continued to exhibit some of these transcriptional faults – especially by writing *e* for *ə*, *k* for *ʔ*, *u* for *w*, *i* for *y*, *-ⁱt* for *-c* word-finally, and ignoring vowel nasality. (When committed by linguists, these faults must be regarded as failures of phonemicisation rather than simply a matter of transcription.)

---

[24] This mistake also appears to have been made by the early phonetician R. J. Lloyd (in Skeat & Blagden 1906, volume 1: 629) in an otherwise careful transcription from a cylinder recording made by W. W. Skeat in 1899 of a Northern Aslian (probably Jahai) song performance. He wrote ['oʔ gɛr'sɔjd: tɛlɛ'mɔjn] for what might possibly have been *ʔoʔ gɛrsɔj tɛlɛmɔɲ*, or more likely *ʔoʔ gərsɔj tələmɔɲ*. (IPA [j] is pronounced *y*.) Skeat's original transcription (p. 627) displays the same feature: *O-gersóydd tĕ̆lĕmóyn*. He was probably working from spoken dictation and obviously heard [ə], while Lloyd heard [ɛ] in the sung version. The verse is translated as 'We rub them and they lose their stiffness.' Niclas Burenhult's comment (p.c) on this is

> The text is perfectly interpretable in Jahai as *ʔoʔ grsɨc t-lmɔy* [3SG feel REL-be.soft] 'he/she/it feels what is soft.' I'd transcribe the 'feel' verb with a final voiceless palatal stop /*c*/, not the voiced one /*j*/. The final /*y*/ in *lmɔy* is nasalised and thus phonetically almost indistinguishable from the palatal nasal /*ɲ*/.

Thus, whatever the other details, the final consonants were indeed palatals rather than the dentals assumed by Lloyd. He nevertheless transcribed the typical rising offglide in the vowels preceding those palatals.

In an attempt to improve matters, I wrote a detailed two-part article (Benjamin 1985b, 1986a) directed specifically at non-linguists working on Orang Asli studies to help them pronounce, write down and type the languages as accurately as they would wish when recording any other ethnographic details. (With the emergence of Unicode-based computer typography, many of the recommended typing conventions are no longer useful.) Unfortunately, my efforts seem to have had no significant results: those who would have done a good job anyway did so without that extra help, while those who were going to do a less than accurate job seem to have ignored my suggestions. The only response to it came from linguists (Bauer 1992b, for example) – for whom it was not intended, despite the wealth of detail presented.

3. Sociolinguistic issues

3.1. Extinctions

Several Aslian languages are known to have died out. Northern Aslian ('Bila', 'Wila' or 'Low-country Semang') word-lists were collected from the Province Wellesley coast opposite Penang island in the early 1800s (Blagden 1906: 390-391), but these languages have long since disappeared. The names Bila and Wila are probably variants of the Malay *bela* 'to nurture, look after', in reference to the purportedly dependent status of the people; they were probably related to the Kintaq group. In the 1920s, I. H. N. Evans (1936: 12-13, 23-26) found Northern Aslian speakers in Kedah, in valleys further to the west than the currently westernmost language, Kensiw. He also found such languages in lowland areas of Perak where none exists today.

At the other end of the Northern Aslian range, in Pahang, Batek Iga' was down to just one speaker by 2011, according to Lye Tuck-Po (p.c.). The Batek people of Kelantan gave both Endicott and myself information about groups over the state border in Terengganu, speaking related languages (such as Batek Tɛ̃ʔ), but who seem to have decreased in numbers. Decades later, Kruspe (p.c.) reported that there remained a community of approximately 26 speakers of Batek Tɛ̃ʔ on the Sungai Besut in northern Terengganu, but that the language was moribund. She met a couple of Batek Tɛ̃ʔ who had married into different Batek subgroups and who no longer remembered their first language. She has suggested that the loss of Batek Iga' and Batek Tɛ̃ʔ was initially precipitated by the communist Emergency (officially 1948-1960, but ending even later in some Orang Asli areas), and more recently by the loss of their original territory. Forced regroupment into smaller areas outside their original range and subjection to external influences has affected the language ecology of the area.

A similar pattern has emerged along the middle reaches of the Perak valley. I have already mentioned Schebesta's 'Ple-Temer' from near Gerik in Hulu Perak as relating to a variety that may no longer exist. Sadly, according to Burenhult (p.c.) and Burenhult & Wegener (2009), Sabüm and Lanoh Yir, Central Aslian languages and dialects spoken near Gerik, appear to have become extinct since I recorded their vocabularies in the 1960s. Burenhult thinks that Yir is probably the same variety as the one he labels Kertei, 'remembered (but not actively spoken) by maybe 3-4 people'. He has collected a 600-item wordlist and two short narratives from an elderly rememberer. The status of Lanoh Jengjeng is uncertain: it may be still be spoken or remembered by a few people now living in Temiar settlements east of the central reaches of the Perak river. Dallos (2011: 33-34) provides up-to-date information on the circumstances of the sub-populations to which these and other labels are supposed to have referred, while remarking that they were never applied consistently and that enforced resettlement has in any case led to much amalgamation.

Near Kuala Lumpur, a short vocabulary of a 'mixed' language (or jargon?) labelled 'Rasa' appears in Blagden's list and on his map (Blagden 1906: opp. p.386), but it too seems to have disappeared, and has so far resisted identification. In Pahang and Johor, Southern Aslian languages possibly related to Mah Meri were once reported further inland than the coastal distribution it has today. In both states, these communities included the ancestors of some of today's upriver Malay communities. In the nineteenth century, they were reported to be pagan Aboriginal communities, some of whom spoke not Malay but a Southern Aslian language related, presumably, to Semelai (see, for example, Miklucho-Maclay 1878). This may be the 'Semaq Palong' ('marsh people'?) mentioned by Williams-Hunt (1952: 44) as living to the south of the Semelai until a few decades previously. Alternatively, the label may refer to the place-name Palong in eastern Negeri Sembilan, and shown on Blagden's map as falling within his 'Southern Sakai' speech area. Or it could simply mean 'dugout-canoe people', a description that would fit with Semelai practice.[25] Rosemary Gianno, who has done detailed ethnographic work among the Semelais (Gianno 1990), told me that she came across several examples of Malay-speaking populations in Pahang which were formerly Aslian-speaking at (among other settlements) Kampung Pasir Kemudi (Sungai Kuantan), Sungai Mentiga and Tanjung Rengas (Kuala Pahang).

[25] The locations of these extinct languages are all indicated in Benjamin (1983), and fuller accounts are given in Benjamin (1976a, 1989).

It is reasonable to assume, then, that there may be some Aslian languages that have not yet been discovered. Several Aslian speech-varieties *have* in fact been discovered, as noted earlier, by Kruspe and Burenhult since 1996. It is unlikely, however, that there are any Orang Asli *populations* that have not yet been accounted for. I came across Mintil (which should perhaps more neutrally be called Batek Tanum or Tanɨm), a hitherto unrecorded Northern Aslian language from northern Pahang, among the patients of the Orang Asli hospital at Ulu Gombak, just outside Kuala Lumpur, in the late 1960s. The only published data on this language appears in my lexicostatistical vocabulary-lists (Benjamin 1976a: 102-123).[26]

These extinctions did not necessarily occur through the death of the speakers, although that must have happened in some cases. More usually, the reason was cultural and linguistic assimilation, either to the Malay community or to one of the larger Orang Asli groups, such as the Temiar, Temuans or Jakuns. Indeed, there is reason to believe that the currently Malay-speaking Temuans of southern Selangor were once Semai-speakers. (The Temuans of Negeri Sembilan, on the other hand, may possibly have had a different linguistic origin, perhaps as Semelai or Mah-Meri-speakers.) For example, at least some of the Temuans living on the southern side of the Sepang river in Selangor probably descend from the Mah Meri ('Besisi') speakers that Skeat (in Skeat & Blagden 1906: 510) identified as living there in the late nineteenth century.

3.2. Degrees of endangerment

How secure are the Aslian languages? Some young Aslian speakers take pride in speaking their own language, and would regret its disappearance.[27] For example, one of my Temiar friends occasionally sends me letters (and now SMS text messages) in his own Central Aslian language. These he types in a painfully achieved *ad hoc*, but inconsistent, orthography based on his knowledge of written Malay. (I present a sample in Section 3.3.) He could more easily do the job in Malay, but Temiar is his clear preference. At Tasik Bera (Pahang) in July 1998, I found that the villagers were speaking nothing

[26] The major syllables of this language frequently consist of diphthongal and (uniquely in Aslian) sometimes triphthongal, nuclei: [kəlkɔəʔ] 'claw', [lɔuyɔus] 'straight' (from Malay *lurus*), [kəui] 'head'.

[27] Phaiboon (2006: 209), however, expressly states that the Northern Aslian Maniq languages of Thailand 'belong to a minority group whose descendants feel embarrassed to speak them.'

but Semelai among themselves, even though they were living on both sides of a paved road regularly traversed by non-Semelais.

The majority of Orang Asli, perhaps all, regularly speak more than one language without difficulty, e.g. a Kensiw (Northern Aslian) man I once met in Kedah who, though completely non-literate, spoke excellent Kensiw, Jahai (also Northern Aslian), Temiar (Central Aslian), Malay and Southern Thai (spanning three linguistic phyla), and tried hard to get me to teach him some English (from a fourth phylum).[28] Kruspe informs me that Semelais employed in the Senoi Praaq (the paramilitary unit manned by Orang Asli) also enjoy displaying their knowledge of Semai, the unit's *lingua franca*. Her impression is that Semelais adopt the majority Aslian language of their place of residence, never Malay. This was also true of mixed marriages, whether with Semai, Jakun or Malays. She met only one married pair who spoke Malay together, a Semelai–Jah Hut couple who had converted to Islam. More generally, Orang Asli field staff employed by the Department of Orang Asli Affairs and their families reported that they spoke the Aslian language of the place they were posted to, frequently changing language as they are moved around. The effects of such intra-Aslian multilingualism have so far been touched on only by Kruspe (2009a: section 3.2).[29]

Given this close interaction between the speakers of various Aslian languages, it is not surprising to find that they have sometimes undergone mutual replacements. Consequently, some of the languages with small numbers of speakers have become especially vulnerable. Temiar, for example, is currently expanding at the expense of the neighbouring Lanoh languages, some of which have recently become extinct, leaving the remainder highly endangered (Burenhult & Wegener 2009: 284). A further factor is the emergence of a pan-'Asli' identity within the *Bumiputera* (indigenous) section of the Peninsular population in contradistinction to the Malay majority. The emergence of Temiar and Semai as pan-Asli *lingue franche* – Temiar in the north and at the Gombak hospital, and Semai within the ranks of the paramilitary Senoi Praaq – is likely to put further pressure on the retention of the remaining Aslian languages. This might well in future hasten the establishment of just a few Aslian languages as the criteria of unitary 'Asli' distinctiveness, to the detriment of the other Aslian languages.

---

[28] I am grateful to Anthony Diller for confirming the good quality of this man's Southern Thai on the basis of a recording I made of him using the language in conversation.

[29] For a detailed discussion of earlier interactions between different Aslian languages based on statistical calculations of the inter-language loan rates, see Benjamin (1976a: 73-81).

In recent years there has been much infrastructural change in Orang Asli territories, brought about by population resettlement, the building of major highways and the development of a plantation economy. This has already led to changes in the linguistic profile, and this will undoubtedly continue. The increased range of communication will make it easier for speakers of different Aslian languages to get together, but it will also allow the inroad of Malay and other majority languages (English, Tamil, Chinese) into places that were formerly insulated from them. On the other hand, it has also led to a small-scale local adoption of an Aslian language by outsiders. I have noticed, for example, that Temiar is now spoken by more Malays and Chinese in the interior of Kelantan than in the recent past; but the numbers are still very small. Such cases often result from intermarriage. More generally, it is difficult to foresee the effects of these infrastructural changes, and very little formal research has been done on the topic. One example, however, is a study of the situation of Batek in the National Park of Pahang (Tengku Intan Suzila et al. 2012) which shows that tourism, administrative practice and the experience of being subjected to research by linguists all have an effect on the sociolinguistic situation. They point out that further research on the language will therefore need to incorporate an awareness of the concomitant ethical issues.

The factors just mentioned are affecting all Aslian languages. But other factors are involved as well, and the degree of endangerment of the languages is not uniform, nor is it necessarily directly proportional to the numbers of speakers. The discussion that follows must therefore be regarded as conditional, in the face of the broader processes of modernisation just mentioned.

Apart from the now mostly moribund Lanoh languages, I suspect that the language in greatest danger is Mah Meri. Kruspe (2010: 5) reports that although it is 'still the first language acquired by children, and most are monolingual when they enter preschool at the age of four … it is increasingly common among young men to speak Malay with their peers.' In a more recent comment (p.c.), she writes:

> I think that by far the most destructive influence on Mah Meri has been the total environmental devastation that has taken place over the last century. Consider that they were already employed as labourers in the logging of their forests at the turn of last century. When virtually every aspect of your culture has been erased, there isn't a lot left to talk about that requires your mother tongue, or that your mother tongue can express. The Mah Meri probably represent the future of a lot more Aslian groups.

Mah Meri is spoken in the vicinity of other Orang Asli communities who speak a Malay dialect as their sole home language – the Temuans (also known as Belandas), Jakuns (also known as Orang Hulu) and the Orang Seletar.[30] While some of these latter populations probably descend from speakers of Southern or Central Aslian languages, most are probably descendants of the same population from which some of the Malays – Orang Melayu in the narrow sense of the term – also derive. It is for this reason that the general label 'Aboriginal Malay' sometimes applied to these groups is an appropriate term.[31] The example presented by a still-pagan and tribal population that speaks Malay may well reassure Southern-Aslian speakers that their culture is not threatened merely by changing language. In the past, ethnographers who had worked among Aboriginal Malay populations told me that they encountered very strong feelings of ethnic identity there, coupled with a high degree of resistance to Melayu-isation. This may no longer be the case, however: Temuans are currently undergoing conversion to Islam at a rate that is probably similar to that of many other Orang Asli populations.

Further north there is no Aboriginal Malay population, so any thought of shifting to Malay as the customary language would mean simply *becoming* Malay (*masuk Melayu*), with its attendant conversion to Islam. While many Semai and Temiar have indeed become Muslims, they have done so while retaining their own languages – so far, at least.[32] Semai and Temiar, however, are not identical where endangerment is concerned, although both languages possess relatively large numbers of speakers.

Semai is split into more than forty quite variable dialects (Diffloth 1977), not all of which are mutually intelligible. This greatly reduces the likelihood

---

[30] The uxorilocal residence pattern of the Temuan communities living next to the Mah Meri has led the latter to lose young men whenever they marry a Temuan woman, with a consequent further dilution of the speech-community. Blagden (1906: 396) too noticed this pattern over a century ago, and ascribed certain language shifts to it.

[31] Aboriginal Malays also live in Indonesia, across the straits in Riau, Jambi and neighbouring provinces of Sumatra. Skeat (1906, Vol. 1: 87) was probably correct when he referred simply to 'Land Jakun' and 'Sea Jakun' when discussing the ethnology of these populations. (See also Kähler 1960 and Pelras 1972.) The term 'Proto-Malay' – with its unfortunate reference to outdated social-evolutionary and migratory-wave theories – is still used in place of 'Aboriginal Malay' by some writers and by the JAKOA in its English-language publications. In its Malay-language documents, the JAKOA sometimes uses the more acceptable, but rather ambiguous, term *Orang Melayu Asli* (which can denote both 'Aboriginal Malay' and 'pure Malay').

[32] Some lowland Semai communities in Perak are reported to be Malay-speaking. I do not know whether this is an autonomous matter or whether it is the consequence of Malayisation and conversion to Islam.

of the continued existence of Semai as a language: each dialect, with a current average of around 1100 speakers each, is therefore on its own to some extent. And, as Robert Dentan (the main ethnographer of the Semai) has pointed out (Dentan 1968: 15), the Semai did not have a strong sense of overall group identity at the time of his fieldwork. Highland Semai sometimes referred to the lowland ones as 'those Malays'; the latter returned the compliment by calling the highlanders 'those Temiars'. On the other hand, Semai is employed daily in radio broadcasts by the state broadcasting company. According to Anthony Williams-Hunt (p.c.), the dialects that get this all-important airing are mainly those of Parit, Kelubi and Tapah, all in Perak. And, as already noted, Semai is also spoken as a *lingua franca* in the Senoi Praaq paramilitary regiment.

Temiar culture is broadly similar to highland Semai culture, but there are important sociolinguistic differences. First, although Temiar has a few dialectal variations (mainly into a northern and southern dialect) all forms of Temiar are mutually intelligible. Temiar sometimes remark that the language is 'the same' everywhere, even though they recognise regional differences in accent and rhythm. The relatively uniform character of Temiar, coupled with its highly regularised pattern of verbal morphology (see Table 4), suggests that a process akin to language standardisation has occurred – however unlikely that might seem in a segmentary tribal society. There are also reasons to think that Temiar, unlike Semai, expanded eastwards from Perak into new territory (Kelantan) in a relatively short time: this too would have had the effect of making the language rather uniform.

A second feature of Temiar, as already remarked, is that it has been buffered on almost all sides from wider Malaysian society by Aslian-speaking populations: the Lanoh complex to the west, Jahai to the north, Menriq to the east, and Semai to the south. In the last century or so, Malay also impinged on Temiar in a few places: at Lasah, Jalong and Temengor in Perak, and around Kuala Betis and Bertam in Kelantan. This central location has had two main effects. Borrowing from the other languages has greatly increased the Temiar word-stock, generating a high degree of synonymy: often there are several words available for the same or closely similar meanings. In this respect, then, Temiar has moved in the opposite direction from that usually taken by endangered languages. Another important effect has been to make Temiar into a *lingua franca*: all Lanoh, most Jahai, all Menriq, and some Batek and Semai speak Temiar as a second or third language. This is true also of several Malays in the Bertam and Gua Musang areas of Kelantan, some of whom are related by kinship to Temiars. I estimate the number of such second-language Temiar-speakers at around 3000.

Perhaps because of this feature, Temiar has gained a further lease of life under the new circumstances generated by the JAKOA, and in particular at its

hospital complex at Ulu Gombak, just north of Kuala Lumpur. There, Orang Asli of all groups live and work together, communicating freely. Although they all speak Malay, many of them also speak Temiar as an expression of 'Asli' solidarity, as already mentioned. As a consequence, Orang Asli from groups not traditionally known to the Temiar, such as the Mah Meri or Jah Hut, have also learnt to speak Temiar, an achievement that some of them take particular pride in. Nor is this restricted to Aslian-speakers, for I have even come across a few Jakuns and Temuans who take pleasure in conversing with me in Temiar.

The situation of the Northern Aslian languages is different again. (I refer to those spoken by historically nomadic populations, who mostly followed the Semang societal tradition; Ceq Wong has followed a different trajectory.) Although these populations are small, but mostly constant over the decades, I would not judge the languages to be endangered simply because of that. There are several reasons why some of them at least may yet survive for some time. First, the Semang populations have long been held in some disdain by their Malay neighbours, who (in earlier times, at least) regarded them as not fully human. As Endicott has shown (1972b: 47ff.), Malay animism ascribed three 'souls' – *semangat*, *roh* and *nyawa* – to normal human beings, including the majority of Orang Asli. But the Semang were thought to have only two, lacking a *roh*, the soul that marks the boundary between humanity and animality. As a consequence, Semang people were not normally invited into Malay houses, although other Orang Asli were welcome, and social relations were restricted to trade or temporary labour-hiring. It may well be that these attitudes no longer persist, but I personally witnessed Malay villagers behaving in that way north of Baling, Kedah, when I was living and travelling with some Kensiws there in 1972. Semang people are quite aware that they are often seen as members of a pariah population. Even now, Semang nomadism (real or imagined) seems to be the feature that most disturbs Malay commentators. As Weber (1958: 270) pointed out, pariah groups tend to develop compensatory cultural mechanisms to protect their self-worth. Among these is an ideology of brotherhood and egalitarianism, coupled with a concern for normative cultural propriety. The latter is most obvious in the strict sexual mores that most Semang groups assert, especially in comparison to what they see as the Temiars' immoral ways (Benjamin 1985a: 258). But it would be surprising if it did not also show in an attachment to their own distinctive languages.

A second feature complements this. I commented earlier on the mesh-like relation between the different varieties of Northern Aslian, which is as much idiolectal as dialectal. The Semang have therefore long been used to linguistic *non*-uniformity. Bishop & Peterson (1993a:1) reported of one Semang village in southern Thailand that 'there were 6 [Northern Aslian] languages and/or dialects represented among the 13 adults present in this particular village of

28 people.' Endicott (1997) reported very similar circumstances among the Bateks of Kelantan. Such variation is unlikely to make Northern-Aslian speakers feel that their languages are too 'small' to maintain. (In this respect, they will probably show an opposite reaction from that of some of the smaller Semai dialect-groups.) In addition, they are used to an overt and unusually high rate of linguistic change: Northern-Aslian speakers have reported to both Endicott and myself that they have consciously changed their way of speaking during their lives, depending on whom they married and where they lived.

Thus, we may reasonably expect some of the Northern Aslian languages both to survive and undergo rapid change in future. This claim gains some support from the findings reported earlier that there may be as many as five such speech-varieties spoken in southern Thailand, most of which no longer retain contact with the Northern Aslian languages of Malaysia, and which have survived virtually in secret until recently.

3.3. Aslian literacy and literature

Until very recently, none of the Aslian languages had ever been the basis of a written literature, and there exist few written sources of any kind. This makes it difficult for these languages to be employed in (primary) education. Some Baha'i and Christian missionaries, as well as the writers of the JHEOA newsletters, have nevertheless produced printed materials in Aslian languages. (For the JHEOA materials, see the entries under *Nong Pai* in the References.) Orang Asli appreciate literacy, but they hardly ever make the effort to sustain it in their own language, relying on literacy in Malay and English. Some individuals, such as radio broadcasters and the occasional letter-writer, have struggled to get their words onto paper; but the transcriptional difficulties are so great that what they write has to be *decoded* rather than read in any regular manner.

As an example, consider sentence (1) from a letter sent to me by a Temiar friend. Its English translation is: 'One more thing, Tataaʔ: my children Kidul, Pipi and Jonney are asking you to send the photos taken with Sarah when we were in your house that time.' A comparison of the letter text with the 'decoded' phonemic version demonstrates that they do not stand in a predictable one-to-one relationship:

(1)

ORIGINAL:	Nei	weal	tatak,	kewas	yek	nak
PHONEMIC:	*Ney*	*wɛl*	*tataaʔ,*	*kəwããs*	*yeeʔ*	*naʔ*
GLOSS:	One	again	old.man,	child	1SG	DEM

Kidul,	Pipi	bok	Jonney	napetak	ham kirif	gambar
Kiduul,	*Pipiih*	*bɔʔ*	*Joniih*	*na-pətaaʔ*	*ham-kirip*	*gambar*
Kidul,	Pipi	and	Jonney	3SG-request	2SG.IRR-send	photo

bil	kanek	an deek	haak	nak	nengning	naak	hatop
beel	*kanɛɛˀ*	*ˀɛn-deek*	*hããˀ*	*naˀ*	*nɛŋnɛŋ*	*naˀ*	*hatɔp*
when	1PL.EXCL	in-house	2SG	DEM	previously	DEM	former

bersamak	Saharah.
bɛrsamaaˀ	*Səharaah.*
together.with	Sarah.

The letter-writer also fails to notate certain reductions that would have been made in ordinary Temiar speech, where *kəwããs ˀi-naˀ* 'my children' and *ˀɛn-deek ha-naˀ* 'in your house' would have appeared in place of the non-reduced forms (*kəwããs yeeˀ naˀ*, *ˀɛn-deek hããˀ naˀ*) used here. This cannot be considered an error. Rather, it indicates that the writer's experience with written Standard Malay has influenced the way he thinks written Temiar should look – employing the fuller 'dictionary-entry' forms rather than the reduced proclitics common in ordinary speech.

Despite these orthographic difficulties, there have been official moves for Temiar and Semai to be introduced as languages of instruction into Orang Asli schools in Perak state. Materials had been prepared by a committee of Semai schoolteachers for use up to Primary Two level, and a special 'orientation' meeting in support of the programme was held in Tapah in March 1999. Given the technical difficulties involved (including the considerable dialectal variation of Semai)[33] and the uncertain level of political will behind it, it is too early to tell whether the plan will succeed, but a start has begun with Semai in several schools. Smith (2003: 58-59) presents a brief history of these efforts. The various Semai dictionaries mentioned earlier are tentative steps in that direction. A necessary procedure might be for Temiar and Semai to be 'gazetted' as officially recognised Malaysian languages, but I can find no information on whether this has yet been done.

As to political will, the question of using Semai in Malaysian primary schools was raised in the Malaysian Senate (Dewan Negara) on 5 June 2001 by the late Senator Long Jidin, himself an Orang Asli but not an Aslian-speaker (http://www.parlimen.gov.my/hindex/pdf/DN-05-06-2001.pdf). He pointed out that, since only 30 percent of the Orang Asli understood Semai,

[33] Diffloth (1968: 65) remarks of Semai variability, 'The people themselves are quite aware of these linguistic variations, and many of them have a remarkable knowledge of dialect distribution, professing that even though sounds and words differ, *nanəˀ ˀirtiˀ*, "there is one understanding". As a practical consequence, it would be extremely difficult to implement a literacy campaign in their own idiom until one of the dialects has been accepted by them as standard.'

the time would be better spent in teaching Chinese, English or Malay. The Minister replied that the purpose of introducing Semai into the schools was 'to protect Semai language, culture and traditions as one of the ethnic languages of Malaysia'. Moreover, the Ministry had received a 'consensus' from the Orang Asli that the majority of them understood Semai. The language was used on Radio Malaysia's Network 7 (*Asyik FM*), and according to the Education Act of 1996, at any school in which some 15 parents of Orang Asli pupils requested classes in their ethnic language these would have to be provided. So there is a paradox: an Orang Asli senator resists the idea of providing Aslian-language lessons in school, while a non-Orang Asli minister expresses great keenness for doing so while failing to recognise the linguistic situation and the great difficulties inherent in such an undertaking.

Regardless of the success or failure of this educational venture (see also Alias Abd Ghani et al. 2003), it is reasonable to hope that some interested scholars who are themselves Orang Asli might develop sufficient literacy in their own languages for the purposes of making cultural records. So far, though, Orang Asli themselves have been among the most resistant to using a regularised and reproducible orthography.[34]

Aslian 'literature' is otherwise almost entirely oral, and very little of it has been published in the original languages. The most accessible example is the volume of Jah Hut stories (Holaday et al. 1985; revised edition, 2003), carefully transcribed in the original language, with an English translation on the facing page and a glossary at the back. The first, commercially produced, edition of this volume won a book-design prize in Singapore. The new edition may have more success in its aim of making their own literature available to Aslian-speakers in printed form in their own language. (Unfortunately, the typographically re-set second edition has failed to mark the phonemically nasal vowels that were carefully indicated in the first edition.) From a linguistic point of view, however, the most useful Aslian material yet published is to be found in the carefully transcribed Semelai texts with interlinear glossing presented by Gianno (1990: 193-216) and Kruspe (2004a: 420-447). Carey's book on Temiar (1961: 21-85) has authentic conversational texts, but they need phonological re-working. Other published literature in Aslian languages consists of a St John's Gospel in Semai (The Bible Societies 1962) and some Baha'i religious materials in Semai, Temiar and Jahai; also in Semai is a *Sengoi (Sakai) First Primer* (Nabitoepoeloe 1950). These too suffer from most of the transcriptional faults mentioned earlier.

---

[34] This has not prevented the occasional Orang Asli writer from becoming successful as a published author in Malay: an outstanding example is the Semai short-story writer Mahat Anak China, who publishes under the pen-name Akiya (2001).

Less useful linguistically but otherwise of great interest are volumes of Aslian oral literature that have appeared in translation, such as Howells's collection (1982) of Chewong (Ceq Wong) stories, solely in English, with no indication of the original linguistic resources. Another example, with some analysis, is Lye (1994) on Batek Iga' mythology. Another source is the large amount of tape-recorded material, mainly autobiographical and cultural, that has been gathered by ethnographers over the years in almost all Aslian languages. This would make an excellent source for the production of published Aslian texts, and would serve as valuable linguistic and historical resources for the people themselves. Little such work has yet been done, but a start has been made by setting up an Aslian multimedia digital archive in Lund, as mentioned earlier.

A recent development is the emergence of text messaging in Aslian languages. Relay masts are currently being built in Orang Asli areas, and mobile phone usage is now widespread. Kruspe (2010: 6) mentions that the Mah Meri send text messages in their own language, in an *ad hoc* Malay-looking orthography that they nevertheless manage to decipher when the context is known. Kelantan Temiar – especially the young women – have been texting each other in Temiar since around 2007. Unfortunately, out of shyness or a wish to preserve their privacy, I have not yet been allowed to see how they manage to type the language. I have myself been texting in Temiar for several years: the available glyphs make it possible to do this near-phonemically. My correspondents have so far had no trouble understanding my messages. For example, if I were to send the text presented in (1) above, I would enter it as follows: *Néy wèl tataa': kwããs yéé' na', Kiduul, Pipiih bò' Joniih na-petaa' ham-kirip gambar béél kanèè' 'èn-déék ha-na' nèngnèng ha-tòp*.

Another development is the publication of pop-music recordings in Aslian languages, mostly Temiar and Semai. These can be heard frequently on Asyik FM, the Orang Asli radio channel. A commercially successful example is the album *Asli*, available on both CD and cassette,[35] by the Orang Asli group Jelmol (i.e., *Jɛlmɔl*, the Temiar for 'mountains'). Although most of the lyrics are in Malay, two of the tracks are in Temiar. Unfortunately, the sleeve-notes present a confused transcription of the Temiar lyrics, and provide no translation. With the help of the singers themselves, I have been able to prepare a fair version of their lyrics, of which (2) is a sample. The first line gives the original sleeve-note transcription, the second is the corrected phonemic transcription, and the third is the word-for-word gloss:

---

[35] On CD: Life Records HSP 01079-2; on cassette: Life Records HSP 01079-4.

(2) BENBEJON … BABOK YEK
Bɛnbəjɔn *baboo²* *yee²*
Unattainable woman 1SG
'Unattainable…my woman'

BERAJON TUI KAL
Bərajɔn *tuuy* *kaal*
Out.of.reach over.there from.now.on
'Far away and out of reach from now on'

BEJBEHEJ HENUM YEK
Bɛjbəhɯj *hənum* *yee²*
pining breath 1SG
'My heart pines for her'

BENBEJON … BABOK YEK
Bɛnbəjɔn *baboo²* *yee²*
Unattainable woman 1SG
'Unattainable…my woman'

BERAJON TUI BILOK
Bərajɔn *tuuy* *bilɔɔ²*
Out.of.reach over.there for.evermore
'Far away and out of reach for evermore'

4. Aslian from a linguistic point of view

What would be lost, or never known, if the Aslian languages disappear? In discussing this issue, we should distinguish between what would be lost to the Aslian-speakers themselves and what would be lost to the wider world, including the world of scholarship – not that the two concerns are incompatible.

Here I present examples of what can be learnt about language and also about Southeast Asia through serious study of the Aslian languages, adding to work that has been carried out by Schmidt (1901), Blagden (1906), and Diffloth (see References); Matisoff (2003) and Bulbeck (2004, 2011) in particular have made important contributions to this question.[36] An example

[36] As examples of what the study of language can do for historical, cultural and geographical research that no other approach can tackle, see John Wolff's extended essay (1976) on (Bornean) Malay loanwords in Tagalog and Adelaar's study (1989) of

of the relevance of Aslian to broader Mon-Khmer, and vice versa, is given at the end of this paper, in the discussion of Temiar animistic terminology.

The following sections should be seen as complementary to the comprehensive survey by Matisoff (2003) of Aslian linguistics as known up to the mid 1980s.

4.1. Historical phonology

Aslian has proved to be of great value in the reconstruction of Mon-Khmer, the basal language family of mainland Southeast Asia. Aslian languages are phonologically very conservative, retaining much that has been lost elsewhere: full pre-syllables (Diffloth 1976d: 232), 'original' long vowels (Central Aslian only), typically iambic sesquisyllabic (one-and-a-half-syllable) accentual patterns, initial voiced stops, and rich consonantal and vocalic systems. On the other hand, they appear to have lost any original register or phonation-type alternations that might have existed in Proto-Mon-Khmer.

As noted earlier, there have been no publications on intra-Aslian historical phonology in the decades following Diffloth's key papers. His study of Aslian as a whole (Diffloth 1975) was written when knowledge of the languages was limited, but it remains the only source to date that can be consulted when dealing with problems of phonological change in these languages. I found it invaluable, for example, when attempting to uncover the complicated etymology of the ethnonym 'Temiar' (Benjamin 2012a). His paper on the reconstruction of Proto-Semai in relation to the rest of Mon-Khmer (Diffloth 1977), based on his own detailed survey of over 40 dialects of Semai, is also the sole published example so far of such a study for any branch of Aslian. As Diffloth (2011: 298) remarked, historical phonology is the most truly scientific of linguistic sub-disciplines and the one most directly relatable to other fields like archaeology and genetics, but 'it consumes enormous amounts of research-hours and takes years, even generations to mature.' Phillips (2012), is the only other researcher to have undertaken the task.

Blagden's 1906 comparative vocabulary has been used extensively in reconstruction work despite its unavoidable inadequacies. The etymological notes contained in Shorto (1962, 1971) are key examples. Recent comparative studies include Diffloth (1984), Shorto (2006), Sidwell (2007, 2008), Starostin

---

the complicated secondary connections between Malay and Malagasy across the Indian Ocean. For an Austroasiatic example, at much greater time-depth and incorporating some Aslian material, see Diffloth 2011.

et al. (1998-2011). With the exception of Kruspe's Mah Meri dictionary and a few other published vocabulary lists, e.g. Lye (2004: 201-209), most of the reliable material remains hidden in researchers' private notes. Online lexical archives at Lund and SOAS should greatly improve the situation.

4.2. Synchronic phonology

Aslian languages, like Mon-Khmer in general, are characterised by some of the 'fullest' phonemic systems in the world. This is especially true of the major (i.e., stressed, word-final) vowel systems, which are mostly of a 3 x 3 type. In some cases four, or even more, heights of vowel have been reported (e.g., Bishop 1996a: 228 and Table 2), though some may be artefacts of the chosen descriptive technique. (Bishop's and Peterson's work, for example, is often phonetic rather than phonemic.)

In most, and probably all, Aslian languages the word-final major vowels are cross-cut by phonemic nasality. In Jahai and Temiar, nasality occurs with only high and low, not mid, vowels (except for Jahai *ə̃*), and there are restrictions in some of the other languages too. In most of the Central Aslian languages (Temiar, Semai, Semnam and Sabüm) the vowels are also cross-cut by a phonemic distinction between short and long, a feature that Diffloth (p.c.) has shown to derive from Proto-Mon-Khmer, and is therefore highly conservative. Southern Aslian, Jah Hut and Northern Aslian have lost this distinction: suggestions about the possible consequences of this loss in Northern Aslian have been made by Bauer (1991: 316) and Matisoff (2003: 16).

Table 2 presents samples of different Aslian vowel-nuclei patterns; phonemically long (Temiar, Semai) and nasal vowels are not indicated.

Table 2: Aslian vowel nuclei

i	ʉ/ɨ	u
e	ə	o
ɛ	a	ɔ

Temiar, Jah Hut, Jahai

i	ɯ/ʉ	u
e	ə	o
ɛ		ɔ
a/æ		ɒ

Kampar Semai, Semelai, Ceq Wong, Semaq Beri

i	ɯ	u
ɪ		
ẹ	ə̯	ọ
e	ə	
ɛ	ʌ	ɔ
	a	

Yala Kensiw

(Bishop 1996a, probably phonetic)

i̱ i̤	ɯ̱ ɯ̤	u̱ ṳ
e̱ e̤	ə̱ ə̤	o̱ o̤
ɛ̱ ɛ̤	a̱ a̤	ɔ̱ ɔ̤

Mah Meri

Registers 1 ('clear') and 2 ('breathy')
(Modified from Stevens, Kruspe & Hajek 2006)

Depending to some extent on the chosen analytical framework, some Aslian languages have been analysed as possessing opening diphthongal vowel phonemes. Diffloth, for example, does so for several Semai dialects (*ie*), as do Burenhult & Wegener (2009: 287) for Semnam (which has *ie*, *ie:*, *uo*, *uo:*, and *ũõ:*). Most proto-language reconstructions at all levels in Mon-Khmer incorporate such diphthongal phonemes. In several Aslian languages, a choice exists between treating such sequences as diphthongs or as CV sequences (Temiar *hual* ~ *həwal* 'to come out', for example). As several authors have noted, the decision will to some extent be arbitrary, depending on the relative emphasis given to the phonology as opposed to morphology: see Diffloth (1976a: 103, 1977: 474), Matisoff (2003: 16-17), Burenhult & Wegener (2009: 290-291).

Prefinal minor-syllables generally display a narrower range of breve epenthetic vowels, in which phonemic length and nasality do not feature. These minor vowels are frequently, though not always, fully predictable, reflecting the frequent occurrence of sesquisyllabic word patterns.[37] (For more on this complicated issue, see section 4.2.4 below.) Non-predictable vowels can occur in presyllables as inflectional elements or as components of fully disyllabic (and occasionally trisyllabic) roots. These sometimes indicate historical features that have been lost in other branches of Mon-Khmer.

The consonantal systems (Table 3) are also rather full, but usually not more so than, for example, the consonantal array of many Austronesian languages. In general they are similar to Malay. As is typical of Southeast Asian languages, word-final consonants are usually unreleased orally, but may sometimes be released velicly (Benjamin 1976a:133, Matisoff 2003:18-22). The sounds *f* (for bilabial [ɸ]) and *z* are very rare, occurring only in a few Northern Aslian languages, such as Jahai *ləflif* 'fontanelle' and Batek *zeˀ* 'new'. Equally rare are the glottalised sonorants (such as *ˀm*, *ˀy* and *ˀl*) and voiceless nasals (*m̥*, *n̥*, *ŋ̊*, but not **ɲ̊*) found as syllable-initial phonemes in a small number of Semelai words (Kruspe 2004a: 35-36). Glottalised sonorants are also found in eastern dialects of Semai (Diffloth 1972b: 93), but as accidents of morphological processes rather than as unit phonemes – 'the ˀ simply indicates a morpheme boundary' (Diffloth 1976b: 35). Voiceless nasal phonemes are found in Maniq as well (Wnuk 2010: 27), where *ɲ̊* occurs but not **ŋ̊*, and in Mah Meri (Kruspe & Hajek 2009: 243). (The ethnonym Mah Meri is pronounced [m̥ãˀ mᵇəˈɾiː] in that language, Kruspe 2010: 59.) They also occur non-phonemically in Temiar (Benjamin 1976b: 134), and presumably also in some of the other languages. Apparently entirely absent (or lost?) from Aslian are the implosive voiced stops ([ɓ], [ɗ]) that occur elsewhere in the Mon-Khmer family, and which have been reconstructed for the Proto-language (Sidwell 2006: xix).

Until recently, distinctive post-aspirated stops were thought to occur only in Southern Aslian languages, in syllable-initial position. But they are now known to occur in the Northern Aslian languages of southern Thailand as well: see Bishop (1996a: 232-233), Hajek (2010) and Wnuk (2010: 6). Hajek (2010: 365-367) has argued that this is an emergent areal phenomenon related to the increasing dominance of Thai; indeed, Wnuk says that they occur only in Thai loanwords. It has therefore not (yet?) affected the Northern Aslian languages of Malaysia. The superscript notation (*tʰ*, etc) serves to distinguish

[37] The term 'sesquisyllabic', derived from the Latin *sesqui* 'one and a half', was invented by Matisoff, who has characterised such words as 'minimally disyllabic', thereby hinting that such a structure could never be truly non-syllabic.

these from consonantal clusters of unaspirated stop + *h*, which also occur, sometimes with an epenthetic vowel between them. In fact, Phaiboon (1984: 130) insists that the latter characterises the Kensiw consonants that other authors have interpreted as single aspirated phonemes. Diffloth (1975: 14-15) clearly distinguishes between aspiration and the more 'normal' occurrences of *h*; he argues that Southern Aslian aspirated stops evolved from originally prenasalised stops. In Semelai, glottalisation and aspiration features are lost when the consonant is copied as part of a reduplicative prefix, but there are suggestions (Hajek 2010: 361) that it is retained in some Northern Aslian varieties of southern Thailand, at least in Thai loans. (But see the discussion of prestopped final consonants in the next section, where a similar question arises.)

Table 3: Aslian consonants

Phonemes in parentheses occur only in a few Aslian languages or are of undecided status.

	Labial	Dental	Palatal	Velar	Post-velar
Voiceless stops	p	t	c	k	
Aspirated stops	(p^{h})	(t^{h})	(c^{h})	(k^{h})	
Voiced stops	b	d	j	g	
Nasals	m	n	ɲ	ŋ	
Voiceless nasals	(m̥)	(n̥)	(ɲ̊)	(ŋ̊)	
Semi-vowels	w		y		
Fricatives, flaps, etc	(f)	l r (z)	s		ʔ h

Not indicated in Table 3 is that the phonetic realisation of the rhotic phoneme (*r*) displays considerable variation in point of articulation (alveolar, retroflex, velar) and articulatory manner (flap, trill, fricative, approximant), both within and between languages (see Diffloth 1975: 4). This is also true of the local dialects of Peninsular Malay. Diffloth (1976a: 112) briefly discusses this issue with regard to the interaction between Malay and Jah Hut. Also not indicated in Table 3 is that 'dentals' in some languages are articulated as alveolars; but

no phonemic alveolar-dental contrast has yet been reported for any Aslian language.

The main distinctive feature is the relatively frequent occurrence of word-final palatals, which Diffloth (1976a: 76-78) proposed as a criterion proving the Mon-Khmer affiliation of Aslian. (Final palatals are exceedingly rare in Austronesian languages.) In most Aslian languages, the distinction between voiced and voiceless stops is neutralised in syllable-final or word-final position; but both voiced and voiceless stops occur word-finally in Temiar. Glottal stop, *ʔ*, usually occurs as a full consonant, in all possible consonantal positions: even phonemically doubled, or even tripled, in a few cases: Temiar *ʔɛʔʔɛʔ* 'a kind of bird', phonemically /*ʔʔʔɛʔ*/ but pronounced [ʔɛʔɛːʔ], and its presumed Maniq cognate (Wnuk 2010: 9) [ʔaʔʔaʔ] /*ʔʔaʔ*/ (with co-articulatory creaky voice). Word-initial and word-final glottal stop is often omitted from ostensibly phonemic transcriptions, especially by Malaysian linguists, with the consequent appearance in their publications of Aslian words written with an initial vowel. This is an error (perhaps based on conventional omission of word-initial glottal stop in standard Malay orthography), for all Aslian syllables have a consonantal onset (Section 4.2.4).

4.2.1. Prestopped (preploded) final consonants

This feature, notated and labelled in different ways by various authors, was first discussed by Blagden (1906: 772-773). It involves an out-of-synchronisation pronunciation of the various features that make up word-final stop and nasal consonants under certain conditions, especially (and sometimes only) phrase-finally or when a word is enunciated in isolation, e.g. in elicitation. The most usual pattern is: [A] -^{b}m, -^{d}n, -$^{j}ɲ$, -$^{g}ŋ$, normally interpreted as nasals (or sometimes, with reversed timing [-b^{m} etc.], as voiced stops). In many of the languages this is the only series reported, and should be labelled 'prestopped nasals' following the normal phonetic terminology for this feature (Kruspe 2004a: 34-35). Another pattern is [B] the postnasalised stops -$^{p}m̥$, -$^{t}n̥$, -$^{c}ɲ̊$, -$^{k}ŋ̊$ (with a voiceless nasal), usually regarded as voiceless stops. Blagden (1906: 773) referred to these split pronunciations as a 'careless, slovenly' habit, but Matisoff (2003: 22) concluded that such articulations indicate that the speakers are 'veritable velic virtuosos'.

Temiar exhibits all three possibilities. The prestopped articulations [A]: -^{b}m, -^{d}n, -$^{j}ɲ$, -$^{g}ŋ$, are interpreted in Temiar as nasals. Series [B] is interpreted as voiceless stops (pronounced -$^{p}m̥$ etc.), and a third series [C] as voiced stops (pronounced -b^{m} etc.), which are not otherwise common word-finally in Aslian (or in Southeast Asian languages generally). Based on a wide range of reports (including Benjamin (1976b: 133) on the 'velic release' of final consonants), Matisoff – following Diffloth in some cases – showed that these disintegrated pronunciations were of importance in explaining some aspects of

Aslian historical phonology: series [A] derived from Proto-Aslian voiceless stops and series [B] from Proto-Aslian nasals.[38]

Burenhult (2005a: 25-27) presented a detailed preliminary discussion of this feature in Jahai (see also Schebesta 1928b: 805), while Bishop (1996a: 235) described it in the Kensiw of southern Thailand. Like most other commentators, Burenhult interprets the Jahai instance as representing a nasal phoneme: phonetically [wɔ^{g}ŋ], phonemically *wɔŋ* 'child', as does Wnuk (2010: 11) for Maniq, an especially clear account.[39] On distributional grounds, however, Bishop analyses prestopped nasals in Kensiw as syllable-final allophones of voiced-stop consonants: phonetically [wɔgŋ], phonemically *wɔg* 'child'. As already remarked, Temiar contrasts voiced and unvoiced stops word-finally, and accordingly displays both possibilities (Benjamin 1976b: 133-134).

These phenomena raise questions of wider linguistic significance. Adelaar (1995) discussed the presence of such modes of articulation among some Austronesian languages of Sarawak. Similar features are also found in some of the Aboriginal Malay dialects in the Peninsula (Noone 1939: 154-155) and in Sumatra. In the Peninsula, it is reasonable to assume that an Aslian linguistic substratum may be present; but Adelaar proposed that Aslian may once have spilt over into Borneo too. Blust (1997) presented an extensive critical discussion of these 'split' nasals, in Aslian and beyond. Burenhult's account of Jahai phonology (2001a) contains the first, and favourable, response to Blust's analysis by an Aslian specialist.

Bishop (1996a: 243) also finds examples of prestopped nasals in Kensiw *non*-final syllables: *sʌdsid* [sʌdn.sidn] '*gums*', *pagpɛg* [pagŋ.pɛgŋ] 'beard'. This raises fundamental questions about the nature of the underlying morphological processes: are these examples of reduplicative inflection (suggested by parallel forms in other Aslian languages, and as Bishop's orthography implies)? Or do they indicate that speakers have reanalysed them as distinct morphological words? Wnuk (2010: 11) states that (as is usual in Aslian languages) such prestopping does not occur in the prefinal syllables of the variety of Maniq she documented; see also Burenhult (2001a: 35, 42) on Jahai.

---

[38] I express my gratitude to John Trim, who taught me phonetics at Cambridge University in the early 1960s. On returning from fieldwork among the Temiar in 1965, he helped me work through this material, and also suggested that it might have relevance to historical phonology, suggestions taken up in Benjamin (1976b: 150-152).

[39] As indicated by the phonetic superscript glyphs, Burenhult (2005a: 25-26), Kruspe (1999), Blust (1997) and I all regard differences of relative prominence or timing as involved in speakers' interpretation of 'split' consonants.

The Temiar materials have attracted interest from researchers of non-concatenative (discontinuous) morphology (see section 4.3, below), so these apparently obscure data from another Aslian language may have a broader impact on current morphological theory, and on understanding the rather aberrant phonology of this dialect of Kensiw situated at the very northern limit of the Aslian speech-area. Could this be partly in response to Thai and (former) Malay influences? These Northern Aslian languages have very few speakers, but they may yet make an important contribution to the broader understanding of linguistic processes.

Diffloth (1976a: 111) reports what he calls 'decomposed' final stops in Jah Hut:

> when final stops are preceded by nasal vowels, or by vowels preceded by Nasals, they are decomposed into two phonetic segments; the first is a Nasal homorganic with the Stop, the second is a glottal stop: /ŋɔk/ → [ŋɔ̃ŋˀ] to sit, /hlaˀŋac/ → [hlaˀŋaɲˀ] to be shy.

It is unclear whether this is a variant of the prestopped final nasals just described or a distinct form of double articulation peculiar to Jah Hut. Diffloth seems to favour the latter view, as he mentions only a possible Munda parallel.

4.2.2. Tone and register

Innovative phonetic processes have been recorded in some Aslian languages that appear to replicate the kind of change hypothesised as having occurred earlier and more widely in the history of the Mon-Khmer family. Further research into Aslian phonetics may prove valuable for the study of linguistic change more generally.

It has occasionally been claimed that lexical tone exists in some of the Northern Aslian languages (Schebesta 1928b: 205; Phaiboon 1984; Bishop 1996a). Phaiboon (1984) treats the Kensiw of Southern Thailand as a tonal language throughout. By attaching a tone-marker to every syllable, he inadvertently provides the only account so far of intonation – rather than lexical tone – in any Aslian language! According to Bishop (1996a: 238) and Hajek (2003: 62), the single word that means both 'head' and 'language' in all other Aslian languages is differentiated by tone into two different words here: *gúj* 'head' (high tone), *gūj* 'language' (mid tone). (The glyph <*j*> corresponds to the <*y*> of most other Aslian transcriptions.) Compare the Temiar cognate *kuy* 'head, language'. Unfortunately – or perhaps significantly! – this appears to be one of the very few minimal pairs supposedly distinguished by lexical tone so far reported for any Aslian language.

The majority of Mon-Khmer languages, including the majority of Aslian languages, lack lexical tone. It occurs in Vietnamese and some other Mon-Khmer languages, where it is known to be a secondary development (Haudricourt 1954). Hajek (2003) carefully analysed the claims for lexical tone in Northern Aslian and concluded that some of the languages may exhibit lexical tone, but that it carries a low functional load and is expressed differently by different speakers. Wnuk (2010: 27), on the other hand, explicitly denies that lexical tone exists in any of the Maniq varieties she studied, as does Burenhult (2005a: 38) regarding Schebesta's claims about Jahai.

Several Mon-Khmer languages exhibit a phonemic register (phonation-type) distinction between the clear (i.e. normal) and breathy articulation of vowels. Until recently, however, no such register distinction was known for any Aslian language, so its recent discovery in Mah Meri (Kruspe & Hajek 2009; Stevens et al. 2006) should be regarded as an innovation peculiar to that language, not as a Mon-Khmer retention. Kruspe (2010) carefully marks the clear/breathy distinction. (It appears to have gone unnoticed by the other researchers on Mah Meri, in Asmah Haji Omar 2006.)

4.2.3. Loanword phonology

Malay loanwords are sometimes reshaped to indicate that they *are* loanwords, even though they would be perfectly well-formed in Aslian phonology if left unaltered. Consider the following Temiar examples (3):

(3)	Malay:	*kəbun*	*raɲcaŋ*	*naŋka*	*ɲamoˀ*
	Temiar:	*kəbut*	*raɲcak*	*daŋkaaˀ*	*jamoˀ*
		'garden'	'plan'	'jackfruit'	'mosquito'

This is a rather interesting example of conscious control of language change – and Temiars *are* conscious of what they are doing in this respect. They joke about these productive alterations, apply them to new loans, and have no problems with the conventional Malay pronunciations. (Compare the conscious spread of uvular *r* from Parisian French into German, Norwegian and eventually even Hebrew, for reasons of prestige.) Bishop (1996a: 228) reports that this change is currently occurring in the Kensiw of Yala, Southern Thailand 'particularly in some Malay borrowings, where a retroflexed approximant [ɻ] is reanalyzed as a velar fricative [ɣ] … inconsistently.' Presumably, this is a carrying-over of local-Malay pronunciation, which normally uses a velar *r* (Asmah Haji Omar, p.c.; Wolff 2010: 484) rather than the trill or flap of Standard Malay. For Temiar, however, the motivation for the reshaping seems to be to keep foreign words marked as such. There are also

examples where an old loanword has been altered phonetically but where the connection with Malay is no longer recognised by the speakers, perhaps because the meanings have changed too: Temiar *gonig* 'spirit-guide' < Malay *gundik* 'consort'; Temiar *misik* 'thunder-causing action' < Malay *bising* 'disturbing noise'.

Burenhult (2001c: 4) explicitly reports that deliberate alterations of this kind do *not* occur for Malay words incorporated into Jahai. Kruspe (2009a:16) mentions a few such changes in Ceq Wong, but reports that these are neither regular nor explicitly regarded by the speakers as indicating that the words are imported.

4.2.4. Syllables, word-structure and phonotactics

There is a concise summary of Aslian word-structure in Dunn et al. (2011: 300):

> The final syllable in most Aslian languages always has the structure /CVC/. The reason for this … is that Aslian languages, like other Mon-Khmer languages, seldom have suffixes, and the end of the word is therefore usually part of the root and not affected by synchronic or diachronic morphophonemic processes. Also, the final syllable is the most informative part of a word in that it always receives stress and contains the greatest phonemic variation.

Two qualifications must be made here. First, although CVC is by far the most common pattern in final syllables, Southern Aslian languages also permit open syllables (CV) in that position, so some words end in a vowel: Mah Meri *kətu* 'pig, *mbəri* 'forest' (Kruspe 2010: 328), Semelai *t^{h}i* 'hand', *ba^{ʔ}i* 'monitor lizard' (Kruspe 2004a: 53). Some other reports of word-final vowels, in Northern Aslian languages especially, might have resulted from failure to hear word-final unexploded consonants or a lenis *h*.

As already noted, prefinal 'minor' syllables can be sesquisyllabic (C.), open (CV.CVC) or closed (CVC.CVC). Vowel-initial syllables (V. or VC.) do not occur; reports of them are due to a failure to hear initial glottal stop, or considering it as non-phonemic. Closed prefinal syllables are regularly produced by the reduplicative processes ('incopyfixation' etc.) discussed in the next section, and contain a predictable epenthetic vowel. Such words are structurally CC.CVC, even though they are prosodically disyllabic. But there are also disyllabic lexical items with unpredictable vowels in the prefinal syllable, CVC.CVC, as well as trisyllabic words of the form CV(C).CV(C).CVC. Disyllabic lexemes are relatively common in most Aslian

languages, but trisyllabic lexemes are far less common; in some languages they are mostly found in loanwords.

As in most Mon-Khmer languages (Thomas 1992), the sesquisyllable pattern C.CVC is especially common. In such cases the vowel of the presyllable (usually *ə*, but sometimes *i*, *u*, *a* etc., or zero) is epenthetic and predictable, and therefore (according to the standard viewpoint) not phonemic in that position. Kruspe (2004a: 45-47) explicitly takes this 'non-phonemic' approach for Semelai; see also Diffloth (1976d) on the Central Aslian languages and Burenhult on Jahai (2005a: 28-35). Aslian linguists basically agree on this interpretation, but differ on whether and in what context to insert the non-phonemic epenthetic vowel into the orthography. Frequently, both forms are presented, with the abbreviated phonemic transcription next to the fuller phonetic one: Jahai (Burenhult 2005a: 21) *pcɛ̃ʔ* [pəcᵉɛ̃ʔ] 'wet'; Semelai (Kruspe 2004a: 45) *tyək* [tiyək] 'banana'; Jah Hut (Diffloth 1976a: 110) *jʾaŋ* [jaʾaŋ] 'bone'. Another solution is to employ a more-than-phonemic orthography, in which the predictable and fully syllabic vowel of closed presyllables is inserted (*CVCCVC* rather than *CCCVC*), while the brief epenthetic vowel in open presyllables is omitted (*CCVC* rather than, say, *CəCVC*). This increases legibility and represents the prosodic structure more faithfully: the rhythm of a language is just as important as its segmental structure. I have taken this approach in many of the examples presented in this paper, both with my own data (as in Table 4) and with data taken from other publications.

The presence of fully predictable – and hence purportedly non-phonemic – vowels in the presyllables of Aslian words has led to some discussion in the wider literature about whether 'syllable-less' languages exist. An example of this 'phonemically correct' approach is Sloan's (1988) citation of the Semai expressive that she writes, following Diffloth (1976c: 259), as [kckmrʔɛ:c] ('short, fat arms'). However, prosodically, rather than the monosyllabic CCCCCCVC structure implied by Sloan, the word is probably actually pronounced as if written *kiʾkumrɛʾɛɛc* (in some Semai dialects, at least) with a CVC.CVC.CV.CVC syllabic structure (or, with sesquisyllabic marking, CVC.CVC.C.CVC). The fact that all of the prefinal vowels in this word are wholly predictable is, of course, irrelevant to the (prosodic) syllabic structure. The syllabic structure and the morphological structure require rather different analytical approaches. Sloan's use of square brackets as if it were a phonetic transcription – unlike Diffloth – is misleading, but it is the only way to support the claim that there really are syllable-less languages. Burquest & Steven (1992: 100), who also cite data from Diffloth, take a more flexible view:

> For many language situations, of course, there is no significant difference between the [phonological and phonemic representations], but explicit recognition of the possible contradictory nature of phonetic vs. phonological structure may be as useful in discussing syllable patterns as it has been in contrasting underlying and phonetic forms of segments.

This seems reasonable and, as indicated earlier, Aslian linguists have explicitly acknowledged this when appropriate.

However, in Temiar (and presumably also in other Aslian languages) it is not always possible to determine whether certain prefinal vowels are unpredictable or not, and therefore whether they are phonemic in that position. For example, in syllabic terms the monomorphemic lexical word *kɛlwɛɛk*, a tree-name, is clearly of the form CɛC.CVC, paralleling the normal syllabic shape of incopyfixed verb forms (as in *ciib* 'go' → *cɛbciib*), where the *ɛ* of the prefinal syllable is wholly predictable. This would mean that *kɛlwɛɛk*, interpreted strictly, should perhaps be of phonemic shape *klwɛɛk*, even though it is neither a verb nor a reduplicated form.[40] Kruspe's solution (2004a: 47) to a similar but not identical problem in Semelai is to propose the word as derived from a 'postulated non-lexemic root', and accordingly to treat the vowel of the presyllable as non-phonemic: *knŋpuŋ* [kənĩŋpuʲŋ] 'child-in-law', from a root **kpuŋ*. This approach will also work with the majority of CɛC.CVC Temiar words, despite the example of *kɛlwɛɛk* just presented. For example, *rɛŋkaaʔ* 'door' can be interpreted (following the processes outlined in the next section) as derived from the non-occurring root **rkaaʔ* by incopyfixation + nominalisation with *-n-* + assimilation of the *-n-* to the following velar consonant: **rkaaʔ* → **rʔkaaʔ* [*rɛʔkaaʔ] → **rnkaaʔ* [*rɛnkaaʔ] → *rŋkaaʔ* [rɛŋkaaʔ].

Obviously, questions of syllabification and phonemicisation are especially salient in the analysis of Aslian morphology, to which I now turn.[41]

---

[40] Data of this sort raise questions about the principle of phonemic biuniqueness ('once a phoneme, always a phoneme'). There has been little explicit discussion of this in Aslian studies where, as noted, 'phonemic' has in practice usually been taken to mean 'non-predictable'. However, matters are not always so straightforward – as some of the secondary commentators also seem to have suspected. Clearly, this is an area where the phonology cannot be divorced from morphological considerations, and cannot be derived from the phonetics alone.

[41] For further discussions of the broader theoretical implications of these and other features of Aslian phonology, especially as they interact with the reduplicative morphology, see Gafos (1998: 233-249), Gafos (1999: 75-107), Hendricks (2001), Miyakoshi (2006), Yap (2009a).

4.3. Morphology

Affixal morphological processes in Aslian are fairly rich, especially for verbal inflection and derivation. Cliticisation, prefixation, infixation and reduplication are all found, along with a portmanteau device to which Matisoff (2003: 28-32) has given the self-iconic label 'incopyfixation'. Very occasionally, suffixes are found too, as in Semelai *-iʔ* 'applicative, iterative' and *b<...>an* 'simultaneous action' (Kruspe 2004a: 123), but these are Malay loans of undetermined date. Many of the affixes are cognate with those of other Mon-Khmer languages. Others (such as *bar-*, *tər-* and *-r-*) seem to have Austronesian origins. Yet others (such as the infix *-n-* and causatives in *pə-*) are found in both language stocks. There has also been some synthesis of new morphological elements within individual Aslian languages.

I illustrate Aslian morphology with what is possibly the most 'saturated', and almost certainly the most regular, verbal paradigm of any Aslian language, namely Temiar (Table 4).[42] According to Diffloth (p.c.), no other Mon-Khmer language has such a regular and productive morphology, although Semai and Semelai come close. Temiar verbal morphology has occasioned a growing secondary literature, especially for its 'non-concatenative' character.[43] For my response to this, involving a naturalistic and meaning-based approach, see Benjamin (2011b, 2012b, 2012d).

[42] This is a re-analysis of my earlier account of Temiar verb morphology (Benjamin 1976b: 168-180). The notion of 'simulfactive mode' has been replaced by a valency- and semantics-based analysis in terms of 'middle voice' (Benjamin 2011b). For simplicity the table omits the 'progressive' with the proclitic *bar-*, which can attach to any of the listed forms.

[43] The most detailed studies are McCarthy (1982: 208-221), Walther (2000), Gafos (1998: 233-249), Gafos (1999: 75-107), Yap (2009a), and Miyakoshi (2006). Others include Broselow & McCarthy (1983: 38-43), Ter Mors (1983: 280-288), and Idsardi & Raimy (2008: 152-166). Burenhult (2005a: 48-49) presents an accessible summary of the approach favoured by these authors. For further discussion (especially of Gafos on Temiar) see Blust (2007: 40-42).

Table 4: The Temiar verb, inflectional and derivational morphology

VOICE	ASPECT		VERBAL NOUN
	Perfective	Imperfective	
MONOSYLLABIC: *gəl* 'to sit', *tɛrgəl* 'to seat'			
Base	*gəl* 'sit' (completed act)	*gɛlgəl* 'sit' (uncompleted act)	*gɛlnəl ~ nɛlgəl* 'a sitting'
Middle	*gagəl* 'sit' (uncontrolledly *or* all at once)		*gnagəl* 'a sitting' (uncontrolledly)
Causative	*tɛrgəl* 'make s.o. sit' (completed)	*trɛlgəl* 'make s.o. sit' (uncompleted)	*trɛngəl* 'a making s.o. sit'
SESQUISYLLABIC: *slɔg* 'to sleep', *sɛrlɔg* 'to put to sleep'			
Base	*slɔg* 'sleep' (completed act)	*sɛglɔg* 'sleep' (uncompleted act)	*sɛnlɔg, snɛglɔg* 'a sleeping'
Middle	*salɔg* 'fall asleep' (uncontrolledly)		*snalɔg* 'a falling asleep'
Causative	*sɛrlɔg* 'put s.o to sleep' (completed act)	*srɛglɔg* 'put s.o. to sleep' (uncompleted)	*srɛnlɔg* 'a putting s.o. to sleep'

The morphological processes in Table 4 include infixation (*-n-* 'nominalisation', *-r-* 'causative', *-a-* 'middle voice'), prefixation (*tɛr-* ~ *tər-* 'causative'), cliticisation (*bar-* ~ *bə-* ~ *ba-* 'progressive', not shown in the table), and incopyfixation (*gəl* → *gɛlgəl*, phonemically *glgəl*; *slɔg* → *sɛglɔg*, phonemically *slɔg* → *sglɔg*; *tɛrgəl* → *trɛlgəl*, phonemically *trgəl* → *trlgəl*). Not all Aslian languages exhibit all of these processes, and some show yet others (such as the more unusual Temiar form *sɛgsəlɔg*, with precopyfixation, a pattern often employed in expressives). The productive use of *-a-* 'middle voice' seems to be unique to Temiar and Lanoh, although non-productive *-a-* is found with slightly different meanings in several other Aslian languages, as discussed later.[44]

Carey (1961) found that there was 'no very clear distinction' in Temiar between the 'simple stem' (my 'perfective') and what he called the 'modified' form (my 'imperfective'), but suggested that the latter refers 'to the general and the unspecified, while the simple stem is used with single events and well-defined states or conditions' (Carey 1961:17-18). More specifically, he suggested that the 'modified' form as opposed to the 'simple stem' is usually used (italics original):

- when 'the action is *habitual* or *repeated*', as opposed to 'a *single* event or action'
- when 'the stress is on an action rather than on the actor or his object', as opposed to 'when one process word is *immediately followed* by another'
- when 'the object or end of an action is *not clearly specified*', as opposed to 'after words indicating a *completed* action'
- when it 'enables a process word to *stand on its own*', as opposed to 'the *imperative* mood'.

Other Aslian languages all have causative derivations, and all – with varying degrees of productivity – employ incopyfixed reduplication to form imperfectives, duratives or distributives. (The causatives too can undergo incopyfixation.) However, as Carey suggested, the change in meaning from the unreduplicated form often seems to be slight. Bishop (1996a: 243) says that the 'relatively productive' reduplication in Kensiw has 'a semantic

[44] Diffloth (1976d: 238) stated that Temiar *-a-*, which he treated as a frozen 'resultative' affix, is non-productive. But it has turned out that the affix is fully productive and in regular use in both monosyllabic and sesquisyllabic verbs (Benjamin 2011b: 14-18).

component of repetition or is onomatopoeic in nature' while Diffloth (1972b: 91) says the reduplicated final consonant in Semai 'adds to intransitive verbs a sense of generalised action, lacking a precise aim, or repeated, and requires an indirect object with transitive verbs' [my translation].

Jahai verbs display two distinct incopyfixation patterns. A regular one, similar to the rest of Aslian, forms the imperfective aspect by inserting the predictable non-phonemic epenthetic vowel (usually *ə*) in the presyllable. For several verbs the incopyfixed presyllable can alternatively take the vowel *i* as its nucleus. Burenhult (2005a: 102-104) reports that these carry a distributive meaning: 'the state or action designated by the verb is distributed over more than one person, object or location' but not occasions (Burenhult 2011: 164). Thus, *ʔɛk* 'to give' has both regular (temporal) imperfective aspect *ʔkʔɛk* [ʔəkʔɛk] and distributive Aktionsart *ʔikʔɛk* 'to give here and there' (but not temporally 'now and then').

In Semelai, such incopyfixation (Kruspe's 'coda copy') is also used to derive nominals from stative verbs: *kdor* 'to be female' → *krdor* [kərdor] 'a female'; and stative verbs from nominals: *suk* 'hair' → *sksuk* [saksuk] 'to have hair'. Kruspe (2004a: 111ff.) suggests that in some cases this morphological pattern has a detransitivising function. Widely found in the Aslian languages is the fossilised appearance of reduplicative patterns in lexical items where no obvious unreduplicated root can be found (as noted in Section 4.2.4). In Jahai (Burenhult 2005a: 71-72) these are concentrated in body-parts and animal names: *krtlɔ̃t* [kərətlɔ̃t] 'kidney' (< **krlɔ̃t*), *ktlɨt* [kətlɨt] 'glow-worm' (< **klɨt*). In Temiar, there are many such forms, with the infixes *-r-* and *-n-* also playing a part, especially in the names of trees: *prɛlŋaal* (< **pŋaal*); *tɛnrɔs* [tɛndrɔs] (< **trɔs*, with epenthetic *d*).

Productive affixation of *(-)n-* is widespread, as elsewhere in Mon-Khmer, but the meaning is not always 'nominalisation'; in some Aslian languages it also occurs with noun roots as a derivational affix. Its range of meanings across the various languages almost defies analysis:

(4) Kintaq: *ʔilay* 'to bathe (perfective)' → *ʔənlay* 'to bathe (imperfective)'
Temiar: *cɛr* 'to pare' → *cnɛr* 'knife'
Temiar: *lɛk* 'to know' → *lanɛŋ* 'knowledge' (*-a-* + *-n-*, with nasal assimilation)
Temiar: *mɔɔʔ* 'parent's elder sister' → *mnɔɔʔ* 'husband's elder sister'[45]
Lanoh: *səmaʔ* 'person', cf. Temiar *sɛnmaʔ* 'person(s) (in counting)'
Semai: *dɤɤk* 'house' → *dnɤɤk* 'household'
Semai: *bah* 'parent's younger brother (in direct address)' → *mnah* 'parent's younger brother (referential)' (with predictable initial nasalisation)
Jah Hut: *ktɔʔ* 'daylight' → *kəntɔʔ* 'day (unit of time)'.

In Jahai, conversion of ordinary nouns into enumeratives is normal and fully productive (Burenhult 2005a: 75-77). Schebesta (1928b: 811-812) gives several examples with *(-)n-* as infix and prefix (with incopyfixation). Burenhult (2005a: 79-80) shows that it can also be used with numerals to indicate 'the state of being that number':

(5) *makɔʔ* 'egg' → *mnakɔʔ* 'eggs'
gənun 'bamboo' → *gnənun* 'bamboos'
mid 'eye' → *nədmid* 'eyes'
cɔg 'basket' → *nəgcɔg* 'baskets'
limɛʔ 'five' → *lnimɛʔ* 'the state of being five'.

Kruspe (p.c., 2004a: 218-219) found that Semelai too exhibits several different meanings for productive affixed *(-)n-*, even in loanwords, as in:

(6) Dysphemism (with *ʔi-* 'Mr/Ms'): *jalu* 'pig' → *ʔi-jnalu* 'damned pig!'
Enumerative: *rɔʔ* 'basket' → *naʔrɔʔ* 'basketful'; *sudu*ʔ '(Malay:) spoon' → *snudu*ʔ 'spoonful'; *kampoŋ* '(Malay:) village' → *knampoŋ* 'villageful'

'Individuated kind of': *cim* 'bird' → *nəmcim* 'kind of bird'; *smaʔ* 'person' → *sənmaʔ* or *snaʔmaʔ* 'kind of person'.

---

[45] See Benjamin (1999: 15) for an explanation of this linkage.

Inflection (by reduplication or infixation) of stative verbs, adjectives and human nouns for plurality is sometimes found:

(7) Batek: *kdah* 'young woman' → *kradah* 'young women' (infixation of *-ra-*)
Lanoh: *litɔw* 'young man' → *latɔw* 'young men' (infixation of *-a-*)
Temiar: *jroˀ* 'long' (singular) → *jɛˀroˀ* 'long (plural)' (incopyfixation)
Temiar: *taaˀ* 'be old (male)' → *tataaˀ* 'old man' (partial reduplication, plus infixation of *-a-*); *tɛˀtaaˀ* 'old men' (partial reduplication plus incopyfixation).

In a related fashion, affixation of *(-)rˀ-* (with phonologically determined epenthetic vowel *a*) forms the comparative degree of a small number of attributive stative verbs of dimension in Semelai (Kruspe 2004a: 146-147) (8):

(8) Semelai: *jlɛ̃ˀ* 'to be short' → *jraˀlɛ̃ˀ* 'to be shorter'
Semelai: *sey* 'to be thin' → *raˀsey* 'to be thinner'

Thus, three pigs could be described as 'two bigger and one smaller' (Kruspe 2004a: 147):

(9)

dwaˀ	*ˀikur*	*raˀ-tʰəy*	*rɔm*	*mə-ˀikur*	*raˀ-kɛ̃t.*
two	CLF	COMPAR-be.big	with	*one*-CLF	COMPAR-be.small.

4.3.1. Iconic morphology: expressives and deponents

Several Aslian languages possess a large number of words known as 'expressives', which exhibit a considerable range of morphological complexity. Expressives form a distinct syntactic and lexical class, sometimes with a semantics and traceable Mon-Khmer etymology of their own. Diffloth (1997) says expressives are

> a basic class of words distinct from verbs, adjectives, and adverbs in that they cannot be subjected to logical negation. They describe noises, colours, light patterns, shapes, movements, sensations, emotions, and aesthetic feelings. Synesthesia is often observable in these words and serves as a guide for individual coinage of new words. The forms of the expressives are thus quite unstable, and the additional effect of wordplay can create subtle and endless structural variations.

Diffloth's studies include iconically-orientated accounts of expressives in Jah Hut (1976a: 84-85), Semai (1976c) and Mon-Khmer more widely (1979b, 1994).

However, taking Aslian as a whole, it is clear that 'expressives' do not always form such a distinctive class. Rather a cline exists between verbs and expressives, exhibiting varying degrees of expressive-like morphology and varying degrees of propositional quality. As Burenhult (p.c.) reports,

> recent descriptive work (mostly still unpublished) indicates that expressives may not be a pan-Aslian feature after all, at least not expressives as defined by Diffloth and encountered in Semai and Semelai, for example. In languages like Ceq Wong, Jahai, Semaq Beri and Semnam expressive-like vocabulary is not formally distinguishable from verbs.

Nevertheless, Burenhult (2006: 160-164) recorded a range of spontaneously uttered phonologically and morphologically elaborate expressives from a speaker who, perhaps significantly, had previously lived among Temiars. Examples included *riŋrũɔ̃ŋ* 'the appearance of a bird turning it head', and *placiɸcɔ̃ɸ* 'the appearance of impressions in skin'. But other members of his community did not use expressives, and 'assumed a bewildered attitude towards them'. I too have found marked differences between Temiar speakers: some made plentiful use of expressives in storytelling, while others never used them.

Where distinctive expressives regularly occur, they are of considerable interest, both in themselves and for linguistic theory more generally. They point to an iconic dimension in language that contrasts with the mainstream view that linguistic representations of reality are in principle arbitrary. Although expressives are hard to translate into English or Malay, they add a rich dimension of subjective and iconic meaning (more often visual than aural); as Tufvesson (2011: 88) puts it, 'they typically package multiple aspects of a sensory event into a single word'. Syntactically, they function as adjuncts, only loosely attached to the remainder of the sentence. Diffloth (1976c: 258-259) gives the following examples of Semai expressives:

(10) *səpsərajããp* 'many tears falling in large and fast flow'
səpsərijããp 'many tears falling, making many slow rivulets'
riɲruhɔɔɲ 'the appearance of teeth attacked by decay'
siɲsulɔɲ 'the odd appearance of a snake's head, sharp yet not pointed, rounded-off yet not round'
cuwcumrəhaaw 'sound of waterfall'.

Among the iconic formative elements here are the alternation *i~u* for 'small~big' (see Diffloth 1994), and use of *m* for 'massive' and *r* for 'plurality' (as in the 'plural' and 'comparative' formations reported in Section 4.3). The overall shapes of these expressives are clearly related to the reduplicative patterns of ordinary verbs, further transformed through much poetic exuberance. (See also Hendricks 2001.) For example, the first of these could be analysed as derived from a base **sjããp* by infixation of *-r-* and *-a-* plus precopyfixation:

(11) **sjããp* → **srajããp* (infixation of *-ra-*) → *spsrajããp* [*səpsərajããp*]

Compare this with the more usual productive verb inflectional reduplication pattern, which would lead to the following form by incopyfixation:

(12) **sjããp* → **spjããp* [*səpjããp*]

Further examples of Semai expressives (Tufvesson 2011: 90) exhibit the iconic importance of vowel gradation within the same consonantal template (here, *blʔ_k* 'dark colour'):

(13) *blʔik* 'gray'
blʔak 'black'
blʔɛk 'rust-brown'
blʔɨk 'darker rust-brown'
blʔuk 'dark purple'
blʔɔk 'darker purple'

Temiar also possesses 'pure' expressives (Benjamin 1976b: 177-178) with no known verbal source. But there are also expressives that clearly derive from verbal sources, without behaving syntactically as verbs, as well as unusual verb forms that share both features:

(14) *sɛdsəlood* 'unable to breathe' (cf. *səlood* 'to drown)
rɛɲrəŋaɲ 'glowing, reddish colour (mahseer fish)' (cf. *rəŋaɲ*, *rɛɲŋaɲ* 'to glow')

Morphological play of this kind is still much employed, as illustrated in the following example taken from a commercial pop-song record sung in Temiar,

with its Shakespeare-like play on the syllable *hɔj*.[46] By itself *hɔj* 'already' is the past-tense marker, but here the singer plays with its derivative *bəhɔj* 'to fail to reach', through both plain (*bɛjbəhɔj*) and nominalised (*bɛnbəhɔj*) precopyfixation. The verse means 'Out of reach, out of reach! I follow when you are far away. Father [=lover] has passed out of reach.'

(15)	*Bəhɔj*	*bɛnbəhɔj!*	*ʔi-saar*	*beel*	*ha-jəroʔ*
	Out.of.reach	out.of.reach.NOM!	1SG-follow.PFV	when	2SG-distant

ma-yeeʔ	*bəəh*	*hɔj*	*bɛjbəhɔj!*
to-1SG	father	PAST	out.of.reach!

Similar to expressives in certain respects, but much closer to the grammatical 'core', are more than one hundred Temiar deponent (inherently middle-voice)[47] verbs (Benjamin 2011b: 28-31) and nearly three hundred middle-voice (unaccusative) nouns (Benjamin 2011b: 31-36).These verbs and nouns exemplify the kind of special features that, as Hale (1992a, 1992b) and Evans (2010) have reminded us, remain to be discovered in the world's understudied languages. The deponent verbs and middle-voice nouns consist of non-inflecting disyllabic stems with the 'objective' marker *-a-* permanently lodged in the first syllable. As a productive affix, *-a-* marks the middle voice of verbs (Table 4), in which the actor or subject is thought of as undergoing its own or a partner's actions. Closer examination shows that the same semantic dimension underlies the non-inflecting deponent verbs and middle-voice nouns, despite their superficially 'active' meaning, as in:

(16) *halab* 'to transport/be transported downriver': the raft is transported by the river's flow
carəəh 'to walk downhill': walking downhill is hard to control, as the walker to some extent *suffers* the movement.
tarah 'to plane flat': an activity, but the hand also undergoes the resistance of the wood.

Cross-linguistically, it is unusual for nouns to carry a middle-voice marking, but the Temiar lexicon contains many examples. The largest category, at least

---

[46] From track 1, 'Cedet seniroi' (*cədʉt sənirɔy* 'remember echos') of the unnumbered audio CD *Bergoyang Bersama Kajol* (Sand Stream Enterprise, Pasir Puteh, Kelantan). A performance of this song by Kajol (real name, ʔaŋɛh Rəlɔg) can be viewed on http://www.youtube.com/watch?v=BiJ3B5oLqDY. The quoted verse appears at 0'50" into the song.

[47] The term 'deponent' is taken from a parallel feature in the grammar of Latin.

135 items, consists of the names of mammals, reptiles, birds and arthropods. These *make themselves move*, and can therefore be thought of as their own source and undergoer: examples are *ʔapɔŋ* 'pig-tailed macaque', *taroʔ* 'house lizard', *tahããr* 'southern pied hornbill', *jaleed* 'firefly'.[48] Similarly, there are around thirty body part words that have *-a-* in the presyllable, presumably because they can be thought of as moving and being moved: *sabook* 'windpipe', *kaŋɔŋ* 'elbow'.[49] Other middle-voice nouns include:

(17) *cacuh* 'woven roof thatch': it holds itself together
jalaaʔ 'thorn': it both snags and gets snagged on people's clothing
ranɔʔ 'back-basket': it holds itself in place.

An explicitly marked middle voice is not unusual in languages generally, although it is certainly not universal. In Aslian, e.g. it has also been reported for Semelai (Kruspe 2004a: 117-118) and Mah Meri (Kruspe 2010: 16), but with a morphology in both cases transparently borrowed from Malay. In Semai (Diffloth 1974b), the middle voice is indicated syntactically by reference to the 'personal spheres' of the appropriately marked agent or patient, rather than by any marking on the verb.[50] However, the Temiar middle-voice forms in *-a-*, both productive and frozen, are distinctive: they are applied to a wide range of lexical items explainable only in cultural terms; they are marked by an 'objective' infix, rather than the cross-linguistically more usual 'subjective' marker; and a substantial number of nouns as well as verbs can be marked for middle voice.

---

[48] Interestingly, fish names exhibit very few forms in *-a-*, perhaps because they are thought of simply as undergoers of the flow of water in which they find themselves rather than of their own activity.

[49] In Jahai there is also a tendency for animal names and body parts to unexpectedly share a morphological feature, except that there it is 'fossilised reduplicative morphology' (Burenhult 2006: 164). Could it be that the same semantic, relating to self-movement, underlies these two different manifestations? Burenhult goes on to remark:

> A connection between fossilised reduplicative morphology and names of body parts and animals has been noted in other Mon-Khmer languages as well, including Semelai, another Aslian language of Malaysia (Kruspe, 1999, pp. 156-157; Kruspe, 2004), and Minor Mlabri, a Kammuic language spoken in northeastern Thailand (Rischel, 1995: 94-95).

[50] Burenhult (2006: 157-159), however, shows explicitly that there is no morphological middle-voice form in Jahai, and Diffloth's account of Jah Hut (1976a) seems also to preclude it in that language.

Why did Temiar verbs and nouns come to be shaped in this way? And why is Temiar verbal morphology (Table 4) so rich and so regular? As regards the remaking of linguistic form, Diffloth (1976a: 82-83, on Jah Hut) says,

> in microsocieties … where individualism is appreciated, language change can be a matter of deliberate personal creation. Such societies, and such values, may have been common-place in the Austroasiatic past.

Elsewhere (Benjamin 2011b, 2012b, 2012c), I have proposed a cultural theory of Temiar grammar, based on the view that something had 'happened' to the language. Comparative analysis suggests that the grammar and parts of the lexicon differ in many respects from what they 'should' be, and that they have been re-made to encode certain fundamental notions about the dialectic of Self and Other that underpins the mode of orientation ethnographically evidenced (Benjamin 1985a, 1994) for the Temiar cultural regime more generally. This seems to have encouraged the synthesis of new morphology. The marking of the middle voice with the vowel *a* derives its iconic power from the directly-felt ('aesthetic') effects of contrasting deictic oral gestures that pervade several areas of the grammar and lexicon of Temiar (Benjamin 2012b):[51]

- Opening the mouth wide (*a*, *ʔ*, *h*, vowel nasality), as if addressing oneself to OTHER, the rest of the world
- Closing the mouth in SELF-contemplation (*i*, *c*, *r*, *m*), as if in temporary retreat from the world.

I know of no systemic parallel in other Aslian languages to what occurs in Temiar, but there are echoes of the 'middle-voice' meaning of *-a-*. Burenhult (2005a: 104; 2011: 166, 174) mentions the rare affixation of *-a-* to a few dynamic verbs in Jahai to mark 'all-together' and 'reciprocal'. These are among the meanings indicated in Temiar by the much more productive use of the same affix (Benjamin 2011b: 15-16), and hence I suspect that the Jahai usage is borrowed from the latter language. Diffloth (1976d: 242-243)

[51] See Diffloth (1976c: 262), discussing Semai expressives, on the importance of oral gesture as a source of linguistic iconicity: 'What is important here is perhaps not so much the acoustic quality, but the sensations produced in the vocal tract by the articulation of the sound. … /i/ "feels" smaller than /a/, /l/ "feels" continuous and homogeneous, while /r/ "feels" interrupted and plural. Even in the regular morphology, reduplications are acoustically and articulatorily symbolic of their meaning "repetition in time".'

presents a long list of Semai verbs, many with *-a-* in the first syllable, which he regards as connected to the non-productive 'resultative' verbal affix *-a-*. Although he makes no such comment, I believe that many of these meanings can be analysed as inherently middle-voice in character. As Diffloth (1976d: 238) recognises, this is clearly the same as the Temiar middle-voice affix, except that in Semai it is non-productive. Likewise, Nicole Kruspe has informed me of the same 'deponent'-like (but non-productive) feature in Semelai:

> While there is no (synchronic) affix *-a-* in Semelai, penultimate syllables with a vowel *a* are rife: *catək* 'to collapse, come apart', *jarla* 'thorn', and from *kmɔn* 'sibling's offspring' is derived *knanmɔn* 'to be related to each other in this manner (aunt to niece etc)'.

4.4. Aslian syntax

To date, there are only three detailed studies of Aslian syntax: Diffloth's (1976a: 83-95) compressed account of Jah Hut, Kruspe's (2004a: 202-395) study of Semelai, and Burenhult's (2005a: 127-159) account of Jahai. These pay close attention to semantic issues, so that the syntax is treated throughout as a manifestation of the speakers' intended meanings, some of which are peculiar to the particular speech communities. Other syntactic studies are more restricted, limited mostly to the constituent structure of the verb phrase or noun phrase, and paying little attention to semantic issues.[52] Fazal Mohamed's account of the transitive verb phrase in Batek (2009a) and Kensiw (2009b, 2011), and Fazal et al. (2010) on Menriq, are the sole studies to engage with Chomsky's minimalist approach to syntax. These relate to the authors' earlier work of the same features in Malay, and are deliberately contrastive.

The basic unmarked word-order in Aslian is SV(O) in declarative sentences, both transitive and intransitive; more usually, however, stative verbs precede their subjects (VS). In general typological terms, Aslian languages (like all other Mon-Khmer languages) are typical nominative-accusative SVO languages:

---

[52] These include Asmah (1964, 1976) on Kintaq Bong (Northern Aslian), Benjamin (1976a: 154-156, 165-167) on Temiar, Nik Safiah & Ton (1979b: 27-30) on Semaq Beri, and Kruspe (2010: 17-18) on Mah Meri.

- They have prepositions, not postpositions: Temiar *num-deek* [from-house] 'from the house', Jahai *ba-hayɛ̃ʔ* [to-house] 'to the house'
- Modifiers follow the head: Mah Meri *dụk nale*ʔ [house old] 'old house', Temiar *ciib gej* [go fast] 'to walk fast'
- Determiners follow the head: Jahai *hayɛ̃ʔ tun* [house that] 'that house', Semelai *knɔn nɔ*ʔ [child this] 'this child'
- Possessors follow the head: Jah Hut *syɛ̃ʔ hãh* [house 1SG], Kensiw *bəluʔ yɛ*ʔ [thigh my] 'my thigh'
- Numerals precede the head: Jah Hut *nar del syɛ̃ʔ* [two CLASS house] 'two houses', Semelai *msɔŋ tən* [five time] 'five times'
- Relative clauses follow the head: Semelai *smaʔ ma-dəmdəm* [person REL-lie.down] 'the person who is lying down', Temiar *deek tɛnʔɛl yeh* [house make.NMLZ 1SG] 'the house that I built'.

But altered ('permuted') constituent orders are frequently employed even in declarative sentences, sometimes with the addition of special case-marking and role-marking particles to keep the syntactic relations clear.[53] In the following example from Jah Hut (Diffloth 1976a: 93) both the complement ('my house') and the agent ('you') are shifted from their normal position, respectively after and before the verb. The sentence means 'Come to my house to sleep, won't you':

(18)	*Syɛ̃ʔ*	*hãh*	*dɔh*	*cʔcyɛk*	*mɛh*	*naʔ*	*ʔimãh.*
	house	my	here	sleep	CLITIC	AGENTIVE	you.
	Complement			Verb		Agent	

Altered word orders can be used pragmatically to emphasise a constituent. But, as Diffloth remarks of such examples,

> it is difficult to find out what nuance of meaning is introduced by having the Agent placed in post-verbal position. Since it is especially common in answers to 'Who?' questions, we may

[53] Case-marking particles ('nominative', 'instrument', 'direct object', 'agentive' etc.) proclitic to nouns and/or pronouns are found in several Aslian languages – Jahai, Temiar, Semai, Jah Hut, Semelai, and possibly others. They are employed with varying degrees of optionality, and sometimes display interesting syncretisms with other morphological elements, such as pronouns or prepositions. See, for example, Schebesta (1928b: 813) on Jahai, Benjamin (1976b: 164; 2012b) on Temiar, and Diffloth (1976a: 92) on Jah Hut.

assume that it is the preferred position when the Agent represents 'new information'.

Burenhult (2006: 130-131) also comes to no firm conclusion on the function of such permutations in Jahai. Kruspe (2004a: 256ff.), in a detailed discussion of what she calls 'fronting' in Semelai, concludes that it is a pragmatic rather than syntactic device. In some instances, however, shifts in the word-order signal semantic subtleties that are hard to recognise. For example, Diffloth goes on to claim that certain verb-first patterns in both Jah Hut and Semai (1974b: 129-130) are instances of 'ergative constructions of a new kind … restricted to actions with a Direction or a Complement located outside the Agent'. I am not aware that this has been followed up elsewhere in the literature, but it involves issues about the interplay of semantics and syntax similar to those that have been raised by the equally puzzling pseudo-'passives' of Acehnese (Durie 1988) – a language that some commentators suspect to have had contact with Aslian in earlier times.

4.5. Aslian semantics

A concern with meaning is taken as basic in most of the published studies of Aslian grammar. As Matisoff (2003: 48-51) pointed out, Aslian lexicons reveal unusual or unique semantic features of wider linguistic interest. In particular, he was impressed by the remarkable 'penchant for encoding semantically complex ideas into unanalysable, monomorphemic lexemes' that the Aslian languages share with their Austroasiatic relatives. This is especially apparent when contrasted with the Tibeto-Burman preference for generating lexical breadth through the compounding of a small number of lexical roots, a process barely employed in Aslian. Matisoff's examples, taken from Diffloth's Semai data, include *slɔɔr* 'to lay flat objects into a round container, as banana leaves in a back-basket', *thãl* 'to make fun of elders sexually', *kəlkənareel* 'stiffly vibrating, as an arrow or knife after embedding itself into a piece of wood, or the walk of a tall skinny old man, or an erect penis'; to these he added Temiar *rɔɔp* 'to tickle fish'.

4.5.1. 'Orientation' in Northern Aslian

Recent studies have shown that this degree of lexical specificity is the norm in Aslian. The work of Burenhult and colleagues from the Language and Cognition Group at the Max Planck Institute for Psycholinguistics (Nijmegen) has paid special attention to the centrality of body imagery, deictic functions and the expression of spatial orientation in the Northern Aslian languages, in

relation to the environment where they are spoken and the people who speak them.[54] This has not been based solely on interpreting the linguistic data, but also results from cognitive category elicitation techniques involving diagrams and videos for speakers to comment or draw on (see, e.g. Terrill & Burenhult 2008, Kruspe 2011).

In his examination of the Jahai body-part lexicon, Burenhult (2006) shows that, as Matisoff pointed out more generally for Aslian, terms are very specific and not ranked into higher-level categorisations.[55] For example, there is no word for 'face' as a whole, but there are distinct simplex lexemes for its parts, including e.g. *kənhɨr* 'root of the nose' or *wɛ̃s* 'ridge at the side of the forehead'. Body-part words form a template onto which are calqued ways of ordering their view of the world around them (Burenhult 2006: 162-163):

> Landforms, houses, camps, trees, tools, fire and the universe as a whole are talked about in terms of 'bodies' exhibiting features like 'heads', 'eyes', 'noses', 'shoulders', 'bums' and 'feet'. An investigation of the Jahais' conceptualisation of 'thing' and 'space' thus requires an understanding of how they structure the human body.

This intense cognitive involvement in their environment by the speakers of Northern Aslian languages has been explored under the general rubric of spatial 'orientation' as a semantic dimension. Close examination of gesture, demonstratives and deictics, environmental terminology, and reference to Levinson's theoretical work (summarised in Levinson & Wilkins 2006), has shown how language both reflects and drives spatial cognition – with the key proviso that language and cognition must first be investigated independently of each other.

---

[54] Burenhult (2003, 2004, 2005b, 2006, 2008a, 2008b, 2012); Burenhult & Majid (2011); Burenhult & Levinson (2008); Majid et al. (2011); Terrill & Burenhult (2008); Tufvesson (2008, 2011); Wnuk & Majid (in press).

[55] Howell (1984: 216), on Ceq Wong, suggests that relative absence of hierarchical ordering of lexemes into higher-level classes relates to the egalitarian, unranked social organisation. Egalitarianism is a criterial feature of the societal traditions that I have labelled 'Semang' and Senoi', as well as of some of the 'mixed' traditions; the 'Malayic' tradition, on the other hand, is characterised by a degree of social ranking (Benjamin 2011a). It would be interesting to investigate whether these differences have affected the degree of hierarchical organisation in the lexicons of the different languages.

Burenhult's (2008a) wide-ranging but empirically well-illustrated study demonstrates that deixis and spatial reference in Jahai interact to produce a newly recognised class of demonstrative distinctions referred to as 'spatial-coordinate demonstratives'. These eight lexemes express the cross-cutting of four spatial dimensions (Burenhult 2008a: 117), anchored on the speaker or addressee rather than on the absolute spatial reference that underpins most other demonstrative systems.

(19)

ʔəh	Accessible to speaker (proximal, perceptible, reachable, approachable, etc.)
ʔon	Accessible to addressee (familiar, attended to)
ʔaniʔ	Inaccessible to speaker (distal, imperceptible, unreachable, inapproachable, etc.)
ʔũn	Inaccessible to addressee (unfamiliar, unattended to)
ʔadeh	Located outside speaker's side of speech perimeter
ʔɲɨʔ	Located outside addressee's side of speech perimeter
ʔitɨh	Located above speech situation (overhead, uphill, or upstream)
ʔuyih	Located below speech situation (underneath, downhill, or downstream)

Some of these features, especially the above/below distinction, are found also in other Aslian languages. Further investigation taking this approach may yet lead to richer accounts of the demonstrative systems that have already been reported.

Levinson & Burenhult (2008) employed comparative data from Jahai and other languages to suggest the existence of a hitherto unrecognised type of lexical configuration, cutting across different word classes, that they label 'semplates'. These relate primarily to geometrical or other forms of spatial imagery that implicitly underpin the semantics of different lexical domains. One such semplate in Jahai is a fractal pattern consisting of a main axis with branches that themselves can become axes with branches of their own. This underlies the Jahai way of talking about rivers, bodies and kinship relations: *kit tɔm* [bottom water] 'river mouth'; *klapəh tɔm* [shoulder water] 'tributary'; *kuy tɔm* [head water] 'source'; *tɔm bɨʔ* [water mother] 'main river'; *tɔm wɔŋ* [water child] 'tributary'. (See also Burenhult 2008b.) By itself, this is hardly an unusual metaphoric pattern. But the same semplate underlies a set of monolexemic verbs of motion or positioning whose distinctive meanings seem to lack any semantic rationale except that they too share the same fractal river-calqued semplate: *rkruk* 'to move along mother-water'; *piris* 'to move across the water-flow, *dɛy* 'to move up along child-water, *hə̃c* 'to move down along child-water'; *kldĩŋ* 'to be positioned lengthwise in relation to the flow of

water'; *hajɨl* 'to be positioned transversally in relation to the flow of water'. (This is yet another example of the Aslian penchant for monomorphemic lexical elaboration.)[56]

4.5.2. Olfaction in Aslian

Aslian-based studies of cognition-and-language have contributed significantly to reclaiming a central place for olfactory perception in human life. The conventional view has been that, since most languages have a limited and indeterminate lexicon for odour, the sense of smell must be vestigial in humans. According to Burenhult & Majid (2011) and Wnuk & Majid (2012), the reverse is just as likely to be true, namely the lack of an appropriate vocabulary in many languages – especially those spoken in what Henrich et al. (2010) characterise as WEIRD (Western, Educated, Industrialized, Rich, Democratic) communities – has led to downplaying olfactory sensibility in favour of the visual and auditory. Earlier ethnographic research in Orang Asli communities (both nomadic and sedentary) had already pointed to the importance of odour-based discourse as a representation of social and spiritual relations.[57] This has now been supported by carefully gathered lexical evidence, e.g. Wnuk & Majid (2012: 2-5) on Maniq.

Jahai has more than a dozen monolexemic odour-identifying stative verbs (Burenhult & Majid 2011: 23). All the terms are associated with specific behavioural expectations connected with animals, foods or religious taboos. The monolexemic status is overridden to some extent by 'expressive'-like linkages, such as *plʔeŋ* 'to have a blood/fish/meat-like smell' and *plʔɛŋ* 'to have a smell of blood which attracts tigers'. In some other Aslian languages, odour words more usually are expressives rather than verbs, as Tufvesson (2011: 89, 91) shows for Semai:

(20)	*gh_p*	Template for 'acrid odour'
	ghuup	'acrid: neutral'
	ghoop	'acrid: intense'
	ghɒɒp	'acrid: very intense'

[56] The 'middle-voice' semantic that underlies deponent verbs and middle-voice nouns in Temiar (Section 4.3.1; Benjamin 2011b) could also be considered to constitute a semplate.

[57] See, e.g., Roseman (1991: 36-40) on the Temiars, Endicott (1979: 74-77) on the Bateks, Kroes (2002 : 29-32) on the Semais.

pŋ_s	Template for 'musty odour'
pŋɛ̃s	'of dirty bird plumage'
pŋũs	'of mould; wet fur'
pŋɔ̃s	'of stale rice'
sʔ_k	Template for 'rank odour'
sʔɛ̃ɛ̃k	'of onion; unwashed hair'
sʔɛ̃ɛ̃k	'of rancid fish or meat'
sʔũũk	'of rotten animal'

Wnuk & Majid (2012: 3-4) add further information. Statistical discrimination of elicited results showed that the terms were fairly obviously arrayed along a pleasant/unpleasant dimension (fitting with research on the physiology of olfaction), cross-cut with what seemed to be interpretable as an edibility/inedibility dimension, related to cultural concerns. As the authors remark, research in communities that have not yet been 'deodorised' and in which smells mediate their environmental, alimentary and religious ideas, has a significant role to play in correcting and refining several themes in cognitive science. The Northern-Aslian-speaking communities have proved very important in this respect, but their current sedentisation may lead to a rapid loss of this knowledge.

Comparative evidence shows that many of the form–meaning linkages for smells have retained remarkable consistency throughout Aslian and Mon-Khmer, attesting to deep-seated shared cultural concerns with smell at a very early stage in mainland Southeast Asia. Burenhult & Majid (2011: 25-26) provide evidence for this by comparing odour terms from Kammu, separated from Aslian by thousands of years, with Jahai:

(21)	Kammu: *pʔúus* 'to smell badly'	Jahai: *pʔus* 'to smell musty'
	Kammu: *hʔú* 'to stink'	Jahai: *haʔɛ̃t* 'to stink'
	Kammu: *hʔɨŋ* 'stink of urine'	Jahai: *sʔĩŋ* 'to have a urine smell'

5. Aslian languages and southeast asian culture-history

Aslian provides a valuable window onto wider Peninsular and Southeast Asian history and prehistory. If there were no Aslian languages to suggest otherwise, scholars would persist in a too simple view of Peninsular history, and fail to recognise that its current relatively uniform Malay character has been brought about by very intensive cultural and demographic engineering.

Blagden (1906: 432-472) devoted a chapter to discussing the significance of the Aslian languages in uncovering the early history of mainland Southeast Asia and India. Some of his hypotheses have not stood the test of time, but he correctly noted that what he called the 'Mon-Annam element' in these

languages both posed historical problems and supplied data towards their solution.[58] Diffloth (1976b) has since shown that Blagden, lacking sufficient understanding of Aslian morphology, was incorrect in some of his argumentation. But there is other evidence that Aslian is not the sole source of Mon-Khmer lexicon in the Peninsula. This, if proven, will probably relate to trade and political links maintained during the heights of the Mon and Angkorian kingdoms. Blagden would indeed then be proven correct in insisting that linguistic research provides the best entrée into that period of Peninsular history.

Diffloth (1979a) has argued that Aslian languages deserve a bigger place in Southeast Asian studies (see also Diffloth 2011, 1977: 490-495). Others too have deployed Aslian data to produce fine-grained reconstructions of pre-historic and proto-historic events in the Peninsula, e.g. Burenhult, Kruspe & Dunn (2011), Dunn et al. (2011) on the contribution of Aslian linguistics to understanding long-term nomadism in the Peninsula. The broader archaeological and biological picture, including the effects of farming and trade in the Peninsula, has been studied by Alan Fix and David Bulbeck, who insist their interpretations must make sense in relation to what is known of Aslian linguistic history. They have employed the data on Aslian linguistic differentiation to generate hypotheses on the population-genetic pattern (Fix 2011) and, in considerable detail (Bulbeck 2011), on the relation of the Peninsula's archaeology and physical anthropology to that of Southeast Asia more widely.

On working through selected areas of vocabulary, such as religious and psychological terms, or kinship terminology (Benjamin 1999), I have been surprised at how thoroughly mainland-Mon-Khmer much Orang Asli ethnography seems to be. Conventionally, Malaysia (Orang Asli included) has been treated ethnologically as belonging with the islands, rather than the mainland. But it is clear that the Aslian-speakers retain a great deal of mainland cultural baggage. This raises an important question: was contact with mainland cultures really broken off at the time of the first emergence of the Aslian group 3,000 or more years ago, or did it continue? I believe that the contact continued, and that the Mon language in particular is crucial to understanding the northern parts of Peninsular Malaysia – later Malay culture included – for Mon was probably the main language in the Isthmian lowlands until perhaps 1200 or 1300 CE.

---

[58] As both Diffloth (1975: 2) and Matisoff (2003: 5) have commented since, Blagden seemed reluctant to accept that Aslian (his 'Sakai' and 'Semang') might actually belong in the same family as Mon-Khmer, despite several times citing with approval Schmidt's (1901) pioneering work establishing that relationship.

The mainland connection can be seen in the etymological backgrounds to two words at the core of Temiar animistic terminology, namely *rəwaay* and *hup*, which label the two kinds of souls that human beings are thought to possess. The relation between these two entities is imaged by Temiar in a complexly dialectical manner, having to do with Self/Other relations. Animate beings are said each to possess two souls, one (the *rəwaay*) associated with the upper part of the body (the head-hair roots of humans and animals, the leaves of trees, and the summits of mountains) and the other (the *hup*) with the lower part (the heart, breath and blood, the roots, and subterranean mass). Dreamers and spirit-mediums report that upper-body souls when met with as spirit-guides (*gonig*, from Malay *gundik* 'consort') look like young men or women, but spirit-guides derived from lower-body souls are like tigers.

In other words, upper-body souls are seen as familiar, domestic and Self-like, while lower-body souls are seen as strange, wild and Other-like. Yet it is a person's heart-tiger soul, the *hup*, that is claimed to be the source of his or her will and agency: it is *hup* that makes one do things or lacks the desire to do something. Tigers, of course, are clearly Other. But it is also possible to perceive as 'other' the usually autonomous beating of one's heart (also *hup*) or one's breathing (*hɛmnum*, an *-n-* infixed form of *hup*), since these can be directly monitored by the individual, without needing to be controlled. The head-soul, *rəwaay*, is clearly 'Self'-like in its association with the incessant but unobserved growth of the hair – the marker of bodily integrity. But the *rəwaay* takes the role of a patient-like, non-controlling *experiencer* of whatever befalls the individual in dreams, trance and sickness (one form of which is *rɛywaay*, 'uncontrolled soul-loss').

At first glance, the Temiar religious framework seems to fall well within the expected range of Malay-World animistic practice. However, while examining the Mon-Khmer connections of Aslian with Gérard Diffloth several years ago, it became apparent that the Temiar had built up these ideas from a very 'Mainland' set of notions. The following etymologies from Diffloth imply a complex conjoining of Mon-Khmer (*hup*, *rəwaay*) and Austronesian (*gonig*) materials into a thoroughly dialectical framework of in::out, human::animal, nature::culture relations:

'Head-soul' (human) *rəwaay* ← Austroasiatic 'tiger' (non-human)
'Tiger-spirit' (non-human) *gonig* ← Austronesian 'consort' (human)

Below is the linguistic data that bears out this analysis.[59] The cognates of *rəwaay* clearly demonstrate that the connection between tiger-beliefs, animistic conceptions and ritual performance ('sing') is not unique to Aslian culture, but was developed anciently in Mon-Khmer-speaking mainland Southeast Asia. The same applies to the connection between breath, heart, agency and subjectivity that is contained in the various cognates of *hup*.

TEMIAR: *Rəwaay* (noun) 'soul (generic)', 'head-soul' (considered as the seat of personal experience and bodily growth, and as the source of humanoid upper-body spirit-guides, whose songs are sung at communal séances). *Rɛywaay* (verb) 'to suffer the lessened consciousness of soul-loss', 'startle response'.

PROTO-MON-KHMER: **rəwaay* 'head-soul'[60]

Northern Mon-Khmer: **rəwaay* 'sing ~ tiger ~ head-soul'

Khasian:	Khasi *rwaay*, Amwi *rwua* 'to sing'
Palaungic:	Ta-ang *rəway*, Rumai *ləvaay*, Na-ang *navaay*, Ka-ang *ʔavaay*, Riang *rəwɒy*, Southern U *ʔavay*^{53}, U *ʔarvay*^{53}, Man-Met *ʔawaay*^{32}, Mok *ləwaay*^{31} 'tiger'
Lamet:	*rəwa̤a̤y* 'tiger'
Waic:	Bulang *kəvày*, Phang *khəvày*, Paraok *sivay*, Phalok *ləvay*, Avïa *vi*, Lawa *ləwia* 'tiger'
Khmuic:	Khmu *hrvaay*, Mal *waay*, Ksing-Mul *həwaay* 'tiger'

Eastern Mon-Khmer: **rəwa̤a̤y* 'soul'

Vietic:	... 'soul'
Katuic:	... 'soul'
Khmer:	*rəvəə-rəveay* 'delirious, incoherent...' (cf. Temiar *rɛywaay*, see above)
Pearic:	... 'tiger'

[59] The superscript numerals on some of the Northern Mon-Khmer words are tone markers; the subscript symbols under some of the vowels are markers of phonemic voice-quality and register.

[60] See also Shorto (2006: 412, no. 1535): '**rwaay* tiger, large felid. A: (Khmuic, Palaungic, Katuic) Kammu-Yuan [ró:y] *rəwà:y* tiger [spirit], Thin *rwaị*, Palaung *rəvay*, Riang-Lang *rəway*, Praok *sivay*, Lawa Bo Luang *ʔawia*, Lawa Umphai *rawia*, Mae Sariang *ɣawia*, *ɣawuai* tiger…, Bru *rawìay* spirit.'

Southern Mon-Khmer: **rəwaay* 'head-soul'

Aslian:	... 'head-soul'
Nicobarese:	... 'ghost'

TEMIAR: *Hup* (noun) 'heart (considered the organ of breath)', 'soul (generically)', 'heart-soul (considered as the seat of personal agency, and as the source of tiger-spirit-guides, whose songs are sung at very special séances). *Hɛmnum* (verb) 'to breathe' [*hup* → **hɛmhup* → **hɛmnup* → *hɛmnum*].

PROTO-MON-KHMER (provisional): **-həəm* 'breathe'[61]

Northern Mon-Khmer:

Khasi:	Khasi *mnseem*, Amwi *hnsua* 'spirit, soul, life'
Palaungic:	Re'ang *phɤɤm* 'heart; air; breathe'
Waic:	Paraok (China) *phɔ̱m* 'air, breath'
Khmuic:	Khmu' *hrɲiəm* 'lung, heart, mind', Ksing Mul *hoñuum* 'lung'

Eastern Mon-Khmer: **phəəm* 'to breathe; heart; mind'

Kuay:	*pəhʊm* 'heart, mind'
Khmer:	Khmer *dɔŋhaəm* 'to breathe', *haəm* 'to swell, be swollen (body part)', *phaəm* 'to be pregnant'
Bahnaric:	Tarieng *cəhɨm* 'to breathe (normally)', Brou *səəm* 'to breathe'
Vietic:	Thaveung *pasʌ̱ʌ̱mʔ* 'to breathe'
Katuic:	Pacoh *palhǫ̱ǫ̱m* 'breath', Bru *pahʌ̱ʌ̱m* 'chest, heart (?anat., emot.)'

[61] See also Shorto (2006: 357-358, no. 1299, and the long article that follows it): 'Pre-Proto-Mon-Khmer **iʔim* > **yəm* by dissimilation, Pre-Palaungic &c. **ʔiim*; **ih[i]m* > **yhə(ə)m*, secondarily > **yhu(ə)m* to breathe, live'.

Southern Mon-Khmer

Monic:	Spoken Mon *həm* ‘to speak, to say’
Aslian:	Temiar *hup* (<**hum*) ‘heart, heart-soul’, *hənum* ~ *hɛmnum* ‘to breathe’, Lanoh *həmnum* ‘soul’, *hənum* ‘heart’, Semai *ləhə̃m* ‘breath’, Jah Hut *pəhɔm* ~ *pəhum* ‘to sigh’, *gərhum* ‘to growl (tiger)’, Semelai *ʔamhɔm* ~ *ʔahɔm* ‘to breathe’ [Temiar *gɛrhɛɛm* ‘to clear one’s throat’?]
Nicobarese:	Shom Peng *hɔɔm* ‘to be alive’, Car *hɔɔm* ‘to breathe in and out’ → *unhɔɔmə* ‘life, breath, spirit’

6. Conclusion

The primary reason for undertaking this review has been the belief – or at least hope – that linguistic research will help raise the status of the Aslian languages, with consequent benefit, directly and indirectly, to the Aslian-speaking populations themselves. Detailed grammars of two or three Aslian languages have now been published, with dictionaries to follow. But this does not mean that we now know all that needs to be known about Aslian, for there is still a great deal to be done. Many of the languages have not been studied at all; enough is known about others to show that they will repay further study; almost all the languages need reliable dictionaries; sociolinguistic research is mostly anecdotal; language-and-culture studies have been fruitful, but are in danger of not keeping up with rapid sociocultural change; and the relevance of Aslian linguistics to other fields of study deserves wider recognition.

In the hope that more linguists (including those who are themselves Orang Asli) will be attracted into Aslian studies, I have surveyed what is already known, with special attention to some of the unusual and theoretically interesting features. Although linguistics is a well-established field of study in Malaysia, the most detailed fieldwork-based research on Aslian has so far been carried out by non-Malaysians. This currently shows signs of changing, but there are still some stumbling blocks. (As remarked earlier, the situation in Thailand is better in this regard.)

I have also sought to inform researchers working in other fields about current knowledge of the Aslian languages. Despite the increasing frequency of references to Aslian in the scholarly literature, these languages are still often regarded as somewhat mysterious. Aslian linguistics matters not just to linguists but also to anthropologists, archaeologists, human biologists, historians, geographers, sociologists, development specialists, and administrators. This is why, in the preceding section, I provided examples of how attention to the linguistic data enables understandings that cannot be obtained through any other approach.

References

Adams, Karen Lee. 1989. *Systems of numeral classification in the Mon-Khmer, Nicobarese and Aslian subfamilies of Austroasiatic.* Canberra: Pacific Linguistics.

Adelaar, K. Alexander. 1989. Malay influences on Malagasy: linguistic and culture-historical implications. *Oceanic Linguistics* 28, 1-46.

Adelaar, K. Alexander. 1995. Borneo as a cross-roads for comparative Austronesian linguistics. In Peter Bellwood, James J. Fox & Darrell Tryon (eds.), *The Austronesians: historical and comparative perspectives*, 75-95. Canberra: Australian National University, Department of Anthropology (RSPAS).

Adi bin Haji Taha. 1985. The re-excavation of the rockshelter of Gua Cha, Ulu Kelantan, West Malaysia. *Federation Museums journal (new series)* 30, 1-134.

Akiya (Mahat Anak China). 2001. *Tuntut: kumpulan cerpen Orang Asli*. Kuala Lumpur: Dewan Bahasa dan Pustaka.

Alias Abd Ghani. 2008. Pengajaran bahasa Semai: model revitalisasi dan pelestarian bahasa Orang Asli di Malaysia. *Prosiding Persidangan Perdana Pribumi Malaysia*. Kuala Terengganu: Penerbit Universiti Darul Iman Malaysia.

Alias Abd Ghani. 2010. The Semang group of Orang Asli: language and cultural endangerment. 8 pages. Paper presented at the *International Conference on Minority and Majority: Language, Culture and Identity*, Kuching, 23-24 November 2010. http://www.mymla.org/icmm2010papers.

Alias Abd Ghani & Salasiah Che Lah. 2010. Language choice and shift of the Kensiu Orang Asli community of Ulu Legong, Baling. In Shakila Abdul Manan, Hajar Abdul Rahim & Sarjit Kaur (eds.), *Millenium realities and innovative practices in Asia*. Proceedings of International Conference on Language, Literacy and Culture. Penang: Penerbit Universiti Sains Malaysia.

Alias Abd Ghani, Norizan Rajak & Bahari Belaton. 2003. Pengajaran bahasa Orang Asli di Malaysia: engrok Semai'. In Ambigapathy Pandian (ed.), *Bahasa dan literasi: penyelidikan dan peninjauan untuk pendidikan*. Serdang: Penerbitan Universiti Putra Malaysia.

Anderson, Gregory D. S. (ed.). 2006. *The Munda languages*. London: RoutledgeCurzon.

Anderson, Gregory D. S. 2009. Mon-Khmer languages. In Keith Brown & Sarah Ogilvie (eds.), *Concise encyclopedia of languages of the world*, 724-730. Oxford: Elsevier.

Asmah Haji Omar. 1964. Bahasa Semang: dialek Kentakbong. Unpublished Honours thesis, University of Malaya.

Asmah Haji Omar. 1976. The verb in Kentakbong. In Philip N. Jenner, Laurence C. Thompson & Stanley Starosta (eds.), *Austroasiatic studies, part II*, 951-970. Honolulu: University Press of Hawaii.

Asmah Haji Omar. (ed.). 2006. *Bahasa Mah Meri*. Kuala Lumpur: Penerbit Universiti Malaya.

Baer, Adela. 2000. The genetic history of the Orang Asli: uniting patchwork data. *Indo-Pacific Prehistory Association Bulletin* 19, 3-10.

Basrim bin Ngah Aching, Adif a/l Bah Yit, Bah Hashim bin Bah Hindan, Wak Itam a/p Bah Long & Rosni a/p Longhit. 2008. *Kamus Engròq Semay – Engròq Bahasa Malaysia, Bahasa Semai – Bahasa Malaysia*. With linguistic advisors James T. Collins & Timothy C. Phillips. Bangi: Institut Alam dan Tamadun Melayu, Universiti Kebangsaan Malaysia.

Bauer, Christian. 1991. Kensiw: a Northern Aslian language of southern Thailand. In Surin Pookajorn et al., *Preliminary report of excavations at Moh-Khiew Cave, Krabi Province, Sakai Cave, Trang Province and ethnoarchaeological research of hunter-gatherer group, socall* [sic] *'Sakai' or 'Semang' at Trang Province*, 310-335. Bangkok: Silpakorn University, Faculty of Archaeology.

Bauer, Christian. 1992a. Mon-Aslian contacts. *Bulletin of the School of Oriental and African Studies* 55, 532-537.

Bauer, Christian. 1992b. Aslian phonetic terminology. *Orang Asli Studies Newsletter* 10, 2. (Department of Anthropology, Dartmouth College, Hanover NH.)

Bauer, Christian. 1993. Review of Parkin 1991. *Bulletin of the School of Oriental and African Studies, University of London* 56, 193-194.

Bellwood, Peter. 1993. Cultural and biological differentiation in Peninsular Malaysia: the last 10,000 years. *Asian Perspectives* 32, 37-60.

Benjamin, Geoffrey. 1973. Introduction. In Paul Schebesta, *Among the forest dwarfs of Malaya,* Second impression, v-xii. Kuala Lumpur: Oxford University Press.

Benjamin, Geoffrey. 1976a. Austroasiatic subgroupings and prehistory in the Malay Peninsula. In Philip N. Jenner, Laurence C. Thompson & Stanley Starosta (eds.), *Austroasiatic studies, part I*, 37-128. Honolulu: University Press of Hawaii.

Benjamin, Geoffrey. 1976b. An outline of Temiar grammar. In Philip N. Jenner, Laurence C. Thompson & Stanley Starosta (eds), *Austroasiatic studies, part I*, 129-187. Honolulu: University Press of Hawaii.

Benjamin, Geoffrey. 1983. Peninsular Malaysia *and part of* Southern Mainland Southeast Asia, with notes. In Stephen A. Wurm & Shiro Hattôri (eds.), *Language atlas of the Pacific area, volume 2*, Maps 36 and 37. Canberra: Australian Academy of the Humanities *and* Tokyo: The Japan Academy. Republished digitally in *ECAI Digital Language Map of the Pacific*, University of California, Berkeley: Electronic Cultural Atlas Initiative. Map 37 and the accompanying text are downloadable from: http://ecai.berkeley.edu/austronesiaweb/maps/pacificatlas/Pacific_leaves/L037.jpg and: http://ecai.berkeley.edu/austronesiaweb/maps/pacificatlas/Pacific_Backs/L037_B.jpg.

Benjamin, Geoffrey. 1985a. In the long term: three themes in Malayan cultural ecology. In Karl L. Hutterer, A. Terry Rambo & George Lovelace (eds.), *Cultural values and human ecology in Southeast Asia*, 219-278. Ann Arbor, MI: Center for South and Southeast Asian Studies, University of Michigan.

Benjamin, Geoffrey. 1985b. On pronouncing and writing Orang Asli languages: a guide for the perplexed. Part 1. *Orang Asli Studies Newsletter* 4, 4-16. [Department of Anthropology, Dartmouth College, Hanover NH, USA.]

Benjamin, Geoffrey. 1986a. On pronouncing and writing Orang Asli languages. Part 2. *Orang Asli Studies Newsletter* 5, 4-29. [Department of Anthropology, Dartmouth College, Hanover NH, USA.]

Benjamin, Geoffrey. 1986b. Between Isthmus and Islands: reflections on Malayan palaeo-sociology. *Department of Sociology working papers* no. 71, Singapore: National University of Singapore.

Benjamin, Geoffrey. 1989. Achievements and gaps in Orang Asli studies. *Akademika* 35, 7-46. [Universiti Kebangsaan Malaysia.]

Benjamin, Geoffrey. 1993. Grammar and polity: the cultural and political background to Standard Malay. In W. A. Foley (ed.), *The role of theory in language description* [= *Trends in linguistics, studies and monographs* 69], 341-392. Berlin: Mouton de Gruyter.

Benjamin, Geoffrey. 1994. Danger and dialectic in Temiar childhood. In Josiane Massard-Vincent & Jeannine Koubi (eds), *Enfants et Sociétés d'Asie du Sud-est*, 37-62. Paris: L'Harmattan.

Benjamin, Geoffrey. 1997. Issues in the ethnohistory of Pahang. In Nik Hassan Shuhaimi bin Nik Abdul Rahman, Mohamed Mokhtar Abu Bakar, Ahmad Hakimi Khairuddin & Jazamuddin Baharuddin (eds.), *Pembangunan arkeologi pelancongan Negeri Pahang*, 82-121. Pekan: Muzium Pahang.

Benjamin, Geoffrey. 1999. Temiar kinship terminology: a linguistic and formal analysis. *Occasional paper* no. 1, *AKASS heritage paper series*. Penang: Academy of Social Sciences.

Benjamin, Geoffrey. 2009a. Non-Malay Austronesian loanwords in Aslian. Paper presented at the *Workshop on Loanwords in the Asian Context*. Division of Linguistics and Multilingual Studies, Nanyang Technological University, Singapore, 9 March 2009.

Benjamin, Geoffrey. 2009b. Affixes, Austronesian and iconicity in Malay. *Bijdragen Tot de Taal-, Land- en Volkenkunde* 65, 291-323.

Benjamin, Geoffrey. 2011a. Egalitarianism and ranking in the Malay World. In Kenneth Sillander & Thomas Gibson (eds.), *Anarchic solidarity: autonomy, equality and fellowship in Southeast Asia*, 170-201. New Haven: Yale University Southeast Asia Studies, Monograph no.60.

Benjamin, Geoffrey. 2011b. Deponent verbs and middle-voice nouns in Temiar. In Sophana Srichampa & Paul Sidwell (eds.), *Austroasiatic Studies: Papers from ICAAL4* (=*Mon-Khmer studies, Special issue* no. 2), 11-37. Dallas: SIL International; Salaya: Mahidol University; Canberra: Pacific Linguistics E-8. http://icaal.org/ICAAL-4.1.pdf.

Benjamin, Geoffrey. 2012a. The peculiar history of the ethnonym 'Temiar'. *SOJOURN: Journal of social issues in Southeast Asia* 27, 205-233.

Benjamin, Geoffrey. 2012b. Aesthetic elements in Temiar grammar. To appear in Jeffrey Williams (ed.), *The aesthetics of grammar: sound and meaning in the languages of Mainland Southeast Asia*. Cambridge: Cambridge University Press.

Benjamin, Geoffrey. 2012c. Irrealis in Temiar. MS.

Benjamin, Geoffrey. 2012d. The Temiar causative (and related features). *Mon-Khmer Studies 41, 32-45*.

Benjamin, Geoffrey. 2012e. Kenaboi: A lost Malayan language? MS.

Benjamin, Geoffrey. 2013. Why have the Peninsular 'Negritos' remained distinct? In Phillip Endicott (ed.), *Human Biology*, April–June 2013 (special double issue on *The first people of Southest Asia*).

Bishop, Nancy M. 1992. A preliminary analysis of Maniq oral monophthongs. Paper presented at the *Sino-Tibetan Conference 1991*. Bangkok; Thammasat University. http://msea-ling.info/digidata/495.992/b07274.pdf.

Bishop, Nancy M. 1996a. A preliminary description of Kensiw (Maniq) phonology. *Mon-Khmer Studies* 25, 227–253.

Bishop, Nancy M. 1996b. Who's who in Kensiw: terms of reference and address in Kensiw. *Mon-Khmer Studies* 26, 245–253.

Bishop, Nancy M. & Mary M. Peterson. 1992. Bibliography, Maniq (Kensiw) language project. 12pp. Bangkok: TUSIL-LRDP, Thammasat University. http://msea-ling.info/digidata/495.992/b07276.pdf.

Bishop, Nancy M. & Mary M. Peterson. 1993a. Maniq language survey report. Bangkok: TUSIL-LRDP, Thammasat University. http://msea-ling.info/digidata/495.992/b06687.pdf.

Bishop, Nancy M. & Mary M. Peterson. 1993b. Kensiw Health Glossary. Bangkok: TUSIL-LRDP, Thammasat University. http://msea-ling.info/digidata/495.992/b06698.pdf.

Bishop, Nancy M. & Mary M. Peterson. 1994a. Kensiw glossary. *Mon-Khmer Studies* 23, 163–195.

Bishop, Nancy M. & Mary M. Peterson. 1994b. Kensiw rhyming list. 63pp. Bangkok: TUSIL-LRDP, Thammasat University. http://msea-ling.info/digidata/495.992/b07275.pdf.

Bishop, Nancy M. & Mary M. Peterson. 1995a. ฆฺยมานิ. 22pp. Bangkok: TU-SIL LRDP. http://msea-ling.info/digidata/495.992/b07449.pdf.

Bishop, Nancy M. & Mary M. Peterson. 1995b. ฆฺยมานิ. 11pp. Bangkok: TU-SIL LRDP. http://msea-ling.info/digidata/495.992/b07450.pdf.

Bishop, Nancy M. & Mary M. Peterson. 1997. Kensiw orthography statement. 26pp. http://msea-ling.info/digidata/495.992/b07565.pdf.

Bishop, Nancy M. & Mary M. Peterson. 1999. Language survey instruments for Northern Aslian language. 64pp. Bangkok: TU-SIL-LRDP, Thammasat University. http://msea-ling.info/digidata/495.992/b08712.pdf.

Bishop, Nancy M. & Mary M. Peterson. 2002. The Kensiw (Maniq) people: Kensiw orthography. In *Minority language orthography in Thailand: five case studies*, 55–68. Bangkok: Thammasat University and SIL International.

Bishop, Nancy M. & Mary M. Peterson. 2003. Northern Aslian Language Survey: Trang, Satul and Phatthalung Provinces, Thailand. 231pp. Bangkok: TU-SIL-LRDP, Thammasat University. http://msealing.info/digidata/495.992/b10811.pdf.

Bishop, Nancy M. & Benjawan Srihanto. 1997. *หนังสือการใช้สมุนไพรของชาวซาไก*. [A booklet on Kensiw medicinal plants, in Thai and Thai-orthographed Kensiw.] 26pp. Bangkok: SIL Thailand. http://msealing.info/digidata/495.992/b08555.pdf.

Blagden, C. O. 1906. Language. In W. W. Skeat & C. O. Blagden, *Pagan races of the Malay Peninsula, volume 2*, 379-775. London: MacMillan.

Blust, Robert. 1994. The Austronesian settlement of Mainland Southeast Asia. In K. L. Adams & T. J. Hudack (eds.), *Papers from the Second Annual Meeting of the Southeast Asian Linguistics Society*, 25-83. Tempe AZ: Program for Southeast Asian Studies, Arizona State University.

Blust, Robert. 1997. Nasals and nasalization in Borneo. *Oceanic Linguistics* 36, 149-179.

Blust, Robert. 2007. Supertemplatic reduplication and beyond. In Henry Y. Chang, Lillian Huang & Dah-an Ho (eds.) *Streams converging into an ocean: in honour of Professor Paul Jen-Kuei Li on his 70th birthday,* 439-460. Taipeh: Institute of Linguistics, Academia Sinica.

Bowrey, Thomas. 1701. *A dictionary / English and Malayo, Malayo and English / To which is added some short grammar rules and directions for the better observation of the propriety and elegance of this language*. London: Sam Bridge, for the Author.

Broselow, Ellen & John J. McCarthy. 1983. A theory of internal reduplication. *The Linguistic Review* 3, 25-88.

Bulbeck, David. 1996. Holocene biological evolution of the Malay Peninsula Aborigines (*Orang Asli*). *Perspectives in Human Biology* 2, 37-61.

Bulbeck, David. 2004. Indigenous traditions and exogenous influences in the early history of Peninsular Malaysia. In Ian Glover & Peter Bellwood (eds.), *Southeast Asia: from prehistory to history*, 314-336. London: RoutledgeCurzon.

Bulbeck, David. 2011. Biological and cultural evolution in the population and culture history of *Homo sapiens* in Malaya. In N. J. Enfield (ed.), *Dynamics of human diversity: the case of Mainland Southeast Asia*, 207-257. Canberra: Pacific Linguistics.

Burenhult, Niclas. 1999. A bibliographical guide to Aslian linguistics. *Mon-Khmer Studies* 29, 133-141.

Burenhult, Niclas. 2001a. Jahai phonology: a preliminary survey. *Mon-Khmer Studies* 31, 29-45.

Burenhult, Niclas. 2001b. Linguistic aspects of the Semang. In Razha Rashid & Wazir Jahan Karim (eds.), *Minority cultures of Peninsular Malaysia: survivals of indigenous heritage*, 75-82. Penang: Academy of Social Sciences (AKASS).

Burenhult, Niclas. 2001c. Loanword phonology in Jahai. *Working Papers* 48, 5-14. Lund University: Department of Linguistics.

Burenhult, Niclas. 2003. Attention, accessibility, and the addressee: the case of the Jahai demonstrative 'ton'. *Pragmatics* 13, 363-379.

Burenhult, Niclas. 2004. Spatial deixis in Jahai. In Somsonge Burusphat (ed.), *Papers from the Eleventh Annual Meeting of the Southeast Asian Linguistics Society 2001*, 87-100. Arizona State University: Program for Southeast Asian Studies.

Burenhult, Niclas. 2005a. *A grammar of Jahai*. Canberra: Pacific Linguistics.

Burenhult, Niclas. 2005b. Landscape terms and toponyms in Jahai: a field report. *Working Papers* 51, 17-29. Lund University: Department of Linguistics.

Burenhult, Niclas. 2006. Body part terms in Jahai. *Language Sciences* 28, 162-180.

Burenhult, Niclas. 2008a. Spatial coordinate systems in demonstrative meaning. *Linguistic Typology* 12, 99-142.

Burenhult, Niclas. 2008b. Streams of words: hydrological lexicon in Jahai. *Language Sciences* 30, 182-199.

Burenhult, Niclas. 2011. The coding of reciprocal events in Jahai. In Nicholas Evans, Alice Gaby, Stephen C. Levinson & Asifa. Majid (eds.), *Reciprocals and semantic typology*, 163-176. Amsterdam: John Benjamins.

Burenhult, Niclas. 2012. The linguistic encoding of placement and removal events in Jahai. In Annetta Kopecka & Bhuvana Narasimhan (eds.), *Events of putting and taking: a crosslinguistic perspective*, 21-36. Amsterdam: John Benjamins.

Burenhult, Niclas. In press. Foraging and the history of languages in the Malay Peninsula. In Tom Güldeman, Patrick McConvell & Richard Rhodes (eds.), *Historical linguistics and hunter-gatherer populations in global perspective*.

Burenhult, Niclas & Asifa Majid. 2011. Olfaction in Aslian ideology and language. *The Senses and Society* 6, 19-29.

Burenhult, Niclas & Claudia Wegener. 2009. Preliminary notes on the phonology, orthography and vocabulary of Semnam (Austroasiatic, Malay Peninsula). *JSEALS: Journal of the Southeast Asian Linguistics Society* 1, 283-312.

Burenhult, Niclas, Nicole Kruspe & Michael Dunn. 2011. Language history and culture groups among Austroasiatic-speaking foragers of the Malay Peninsula. In Nicholas J. Enfield (ed.), *Dynamics of human diversity: the case of Mainland Southeast Asia*, 257-277. Canberra: Pacific Linguistics.

Burenhult, Niclas & Stephen C. Levinson. 2008. Language and landscape: a cross-linguistic perspective. *Language Sciences* 30, 135-150.

Burquest, Donald A. & Lee A. Steven. 1992. Prosodic structure and reduplication: data from some languages of Maluku. In Karen L. Adams & Thomas John Hudak (eds.), *Papers from the Second Annual Meeting of the Southeast Asian Linguistics Society*, 85-101. Arizona State University: Program for Southeast Asian Studies.

Carey, Iskandar. 1961. *Tengleq Kui Seroq: a study of the Temiar language with an ethnographical summary*. Kuala Lumpur: Dewan Bahasa dan Pustaka.

Carey, Iskandar.1970. A Mendrik vocabulary. *Federation Museums Journal (New Series)* 15, 185-191.

Collings, H. D. 1949. A Tĕmoq word list and notes. *Bulletin of the Raffles Museum, Series B* 4, 69-85.

Collins, James T. 1998. *Malay, world language: a short history.* Kuala Lumpur: Dewan Pustaka dan Pustaka.

Craig, Collette. 1992. A constitutional approach to language endangerment: the case of Nicaragua. *Language* 68, 17-24.

Dallos, Csilla. 2011. *Roots of inequality: social change among newly sedentary Lanoh hunter-gatherer traders of Peninsular Malaysia.* Toronto: University of Toronto Press.

Dentan, Robert Knox. 1968. *The Semai: a non-violent people of Malaysia.* New York: Holt, Rinehart, Winston. [Second, 'fieldwork', edition, 1979.]

Dentan, Robert Knox. 2003 [1963]. Preliminary field notes on the Semai language. Edited by Thomas Doyle. 91 pages. Unpublished typescript, www.keene.edu/library/OrangAsli/DentanBaer.pdf.

Diffloth, Gérard. 1968. Proto-Semai phonology. *Federation Museums Journal (New Series)* 13, 65-74.

Diffloth, Gérard. 1972a. Notes on expressive meaning. *Papers from the 8th Regional Meeting, Chicago Linguistic Society*, 440-447.

Diffloth, Gérard. 1972b. Ambiguïté morphologique en semai. In Jacqueline M. C. Thomas & Lucien Bernot (eds.), *Langues et techniques, nature et société: approche linguistique,* 91-93. Paris: Klincksieck.

Diffloth, Gérard. 1974a. 'Austroasiatic languages'. *Encyclopedia Britannica (15th edition)* 2, 480-484.

Diffloth, Gérard. 1974b. Body moves in Semai and in French. *Papers From the 10th Regional Meeting, Chicago Linguistic Society*, 128-138.

Diffloth, Gérard. 1975. Les langues mon-khmer de Malaisie: classification historique et innovations. *Asie du Sud-est et Monde Insulindien* 6 (4), 1-19.

Diffloth, Gérard. 1976a. Jah-Hut: an Austroasiatic language of Malaysia. In Nguyen Dang Liem (ed.), *South-east Asian Linguistic Studies, Volume 2*, 73-118. Canberra: Pacific Linguistics.

Diffloth, Gérard. 1976b. Mon-Khmer numerals in Aslian languages. *Linguistics* 174, 31-38.

Diffloth, Gérard. 1976c. Expressives in Semai. In Philip N. Jenner, Laurence C. Thompson & Stanley Starosta (eds.), *Austroasiatic studies, part I*, 249-264. Honolulu: University Press of Hawaii.

Diffloth, Gérard. 1976d. Minor-syllable vocalism in Senoic languages. In Philip N. Jenner, Laurence C. Thompson & Stanley Starosta (eds), *Austroasiatic studies, part I*, 229-247. Honolulu: University Press of Hawaii.

Diffloth, Gérard. 1977. Towards a history of Mon-Khmer: Proto-Semai vowels. *Tonan Ajia Kenkyu [South East Asian studies]* 14, 463-495.

Diffloth, Gérard. 1979a. Aslian languages and Southeast Asian prehistory. *Federation Museums Journal (New Series)* 24, 2-16.

Diffloth, Gérard. 1979b. Expressive phonology and prosaic phonology in Mon-Khmer. In Therapan L. Thongkum et al. (eds.), *Studies in Tai and Mon-Khmer phonetics and phonology*, 49-59. Bangkok: Chulalongkorn University Press.

Diffloth, Gérard. 1984. *The Dvaravati Old Mon language and Nyah Kur*. Bangkok: Chulalongkorn University Printing House.

Diffloth, Gérard. 1994. /i/: big, /a/: small. In Leanne Hinton, Joanna Nichols & John J. Ohala (eds.), *Sound symbolism*, 107-114. Cambridge: Cambridge University Press.

Diffloth, Gérard. 1997. Austroasiatic languages. *Encyclopædia Britannica Online*, http://www.britannica.com/EBchecked/topic/44541/Austroasiatic-languages.

Diffloth, Gérard. 2005. The contribution of linguistic palaeontology to to the homeland of Austroasiatic. In Laurent Sagart, Roger Blench & Alicia Sanchez-Mazas (eds.), *The peopling of East Asia: putting together archaeology, linguistics and genetics*, 77-80. London: RoutledgeCurzon.

Diffloth, Gérard. 2011. Austroasiatic word histories: boat, husked rice and taro. In Nicholas J. Enfield (ed.), *Dynamics of human diversity: the case of Mainland Southeast Asia*, 295-313. Canberra: Pacific Linguistics.

Diffloth, Gérard & Norman Zide. 2005. Austro-Asiatic languages. In William J. Frawley (ed.), *International Encyclopedia of Linguistics*, Second edition, 177-181. Oxford: Oxford University Press.

Dunn, Michael, Niclas Burenhult, Nicole Kruspe, Neele Becker & Sylvia Tufvesson. 2011. Aslian linguistic prehistory: A case study in computational phylogenetics. *Diachronica* 28, 291-323.

Dunn, Michael, Nicole Kruspe & Niclas Burenhult. 2013. Time and place in the prehistory of the Aslian language family. In Phillip Endicott (ed.), *Human Biology*, April–June 2013, (special double issue on *The first people of Southeast Asia*).

Durie, Mark. 1988. The so-called passive of Acehnese. *Language* 64, 104-113.

Endicott, Kirk M. 1972a. A wordlist of Batek Dèʔ. Unpublished manuscript.

Endicott, Kirk M. 1972b. *An analysis of Malay magic.* Oxford: Oxford University Press.

Endicott, Kirk M. 1979. *Batek Negrito religion*. Oxford: Clarendon Press.

Endicott, Kirk M. 1997. Batek history, inter-ethnic relations and subgroup dynamics. In R. Winzeler (ed.), *Indigenous peoples and the state: politics, land, and ethnicity in the Malayan Peninsula and Borneo*, 30-50. New Haven: Yale Southeast Asia Studies Monograph no. 46.

Evans, I. H. N. 1936. *The Negritos of Malaya*. Cambridge: Cambridge University Press. [Reprinted 1968, London: Frank Cass.]

Evans, Nicholas. 2010. *Dying words: endangered languages and what they have to tell us*. Oxford: Blackwell.

Fauziah Mat. 1993. *Dialek Jahai: satu kajian fonologi*. Unpublished academic exercise. Bangi: Universiti Kebangsaan Malaysia.

Fazal Mohamed Mohamed Sultan. 2009a. Struktur sintaksis frasa nama bahasa Bateq. *GEMA Online Journal of Language Studies* 9, 47-61.

Fazal Mohamed Mohamed Sultan. 2009b. Sintaksis frasa nama bahasa Kensiu. *Jurnal Bahasa* 9, 155-172.

Fazal Mohamed Mohamed Sultan. 2011. The syntactic structure of a noun phrase: Austroasiatic vs. Austronesia. *Pertanika Journal of Social Sciences and Humanities* 19, 263-271. [English version of Fazal 2009b.]

Fazal Mohamed Mohamed Sultan, Nor Hashimah Jalaluddin & Harishon Radzi. 2010. Sintaksis kata soal argumen bahasa Mendriq: Kerangka minimalis. *Jurnal Bahasa* 10, 155-172.

Fazal Mohamed Mohamed Sultan, Zaharani Ahmad, Nor Hashimah Jalaluddin & Harishon Radzi. 2010. Kata kerja transitif bahasa Mendriq: kerangka program minimalis. *Jurnal Melayu* 5, 409-423.

Fix, Alan. 1995. Malayan paleosociology: implications for patterns of genetic variation among the Orang Asli. *American Anthropologist* 97 (2), 313-323.

Fix, Alan. 2011. Origin of genetic diversity among Malaysian Orang Asli: an alternative to the demic diffusion model. In Nicholas. J. Enfield (ed.), *Dynamics of human diversity: the case of Mainland Southeast Asia*, 277-291. Canberra: Pacific Linguistics.

Gafos, Adamantios I. 1998. Eliminating long-distance consonantal spreading. *Natural Language and Linguistic Theory* 16, 223-278.

Gafos, Adamantios I. 1999. *The articulatory basis of locality in phonology*. New York: Garland.

Gianno, Rosemary. 1990. *Semelai culture and resin technology*. New Haven CT: Connecticut Academy of Arts and Sciences.

Grimes, Barbara (ed.). 2000. *Ethnologue: languages of the world.* 14th edition. Dallas: SIL International.

Hajek, John. 1998. Kenaboi: an extinct unclassified language of the Malay Peninsula. *Mon-Khmer Studies* 28, 137-162.

Hajek, John. 2003. Evidence of tonal activity in Aslian. *Mon-Khmer Studies* 33, 59-66.

Hajek, John. 2010. Aspiration in some varieties of Northern Aslian. *Oceanic Linguistics* 49, 359-368.

Hale, Kenneth. 1992a. On endangered languages and the safeguarding of diversity. *Language* 68, 1-3.

Hale, Kenneth. 1992b. Language endangerment and the human value of linguistic diversity. *Language* 68, 35-42.

Hamilton, Annette. 2002. Tribal people on the southern Thai border: internal colonialism, minorities, and the state. In Geoffrey Benjamin & Cynthia Chou (eds.), *Tribal communities in the Malay World: historical, social and cultural perspectives*, 77-96. Singapore: ISEAS *and* Leiden: IIAS.

Harishon Radzi, Fazal Mohamed Mohamed Sultan, Zaharani Ahmad & Nor Hashimah Jalaluddin. 2011. Penggolongan kata bahasa Mendriq Negrito: satu penelitian rintis. *Jurnal Linguistik* 12, 134-154.

Haudricourt, André-Georges. 1954. De l'origine des tons en vietnamien. *Journal Asiatique* 242, 69-82.

Hendricks, Sean. 2001. Bare-consonant reduplication without prosodic templates: expressive reduplication in Semai. *Journal of East Asian Linguistics* 10, 287-306.

Henrich, Joseph, Steven J. Heine & Ara Norenzayan. 2010. The weirdest people in the world. *Behavioral and Brain Sciences* 33, 61-83.

Holaday, Duncan, Chin Woon Ping & Teoh Boon Siong. 1985. *Tales of a shaman: Jah Hut myths as told by Batin Long bin Hok*. Singapore: Times Books International.

Holaday, Duncan, Chin Woon Ping & Teoh Boon Siong. 2003. *Batin Long bin Hok's 'Bés Hyang Dnèy' and other Jah Hut stories*. Subang Jaya: Centre for Orang Asli Concerns.

Howell, Signe. 1982. *Chewong myths and legends*. Kuala Lumpur: Malaysian Branch, Royal Asiatic Society.

Howell, Signe. 1984. *Society and cosmos: Chewong of Peninsular Malaysia*. Singapore: Oxford University Press.

Idsardi, William J. & Eric Raimy. 2008. Reduplicative economy. In Bert Vaux & Andrew Nevins (eds.), *Rules, constraints, and phonological phenomena*, 149-185. New York: Oxford University Press.

Ikram Jamaluddin, Haji. 1997. Kenyataan Ketua Pengarah Jabatan Hal Ehwal Orang Asli Malaysia pada perjumpaan dengan wakil-wakil media massa pada 31hb Oktober, 1997. Typescript, 35 pages. Kuala Lumpur: JHEOA.

Jennings, Sue. 1995. *Theatre, ritual and transformation: the Senoi Temiars*. London: Routledge.

Jenny, Mathias & Paul Sidwell (eds.). In preparation. *The Austroasiatic Mon-Khmer languages*. Leiden: Brill.

Kähler, Hans. 1946-49. Ethnographische und linguistische Studien von den Orang Laut auf der Insel Rangsang an der Ostküste von Sumatra. *Anthropos* 61-64, 1-31, 757-785.

Kähler, Hans. 1960. *Ethnographische und linguistische Studien über die Orang Darat, Orang Akit, Orang Laut und Orang Utan im Riau-Archipel und auf den Inseln an der Ostküste von Sumatra*. Berlin: Dietrich Riemer.

Kroes, Gerco. 2002. *Same hair, different hearts: Semai identity in a Malay context*. Leiden: CNWS Publications.

Kruspe, Nicole. 1999. *Semelai*. PhD thesis, Department of Linguistics and Applied Linguistics, University of Melbourne.

Kruspe, Nicole. 2004a. *A grammar of Semelai*. Cambridge Grammatical Descriptions. Cambridge: Cambridge University Press.

Kruspe, Nicole. 2004b. Adjectives in Semelai. In R. M. W. Dixon & A. Y. Aikhenvald (eds.), *Adjective classes: a cross-linguistic typology*, 283-305. Oxford: Oxford University Press.

Kruspe, Nicole. 2009a. Loanwords in Ceq Wong, an Austroasiatic language of Peninsular Malaysia. In Martin Haspelmath & Uri Tadmor (eds.), *Loanwords in the world's languages: a comparative handbook of lexical borrowing*, 659-685. Berlin: Mouton de Gruyter.

Kruspe, Nicole. 2009b. Ceq Wong vocabulary. *World loanword database*. http://wold.livingsources.org/vocabulary/26.

Kruspe, Nicole. 2010. *A dictionary of Mah Meri, as spoken at Bukit Bangkong*. Illustrated by Azman Zainal. Oceanic Linguistics special publication no. 36. Honolulu: University of Hawai'i Press.

Kruspe, Nicole. 2011. Reciprocal constructions in Mah Meri. In Nicholas Evans, Alice Gaby, Stephen C. Levinson & Asifa Majid (eds.), *Reciprocals and semantic typology*, 149-162. Amsterdam: John Benjamins.

Kruspe, Nicole. In preparation a. *A grammar of Ceq Wong*.

Kruspe, Nicole. In preparation b. *A grammar of Mah Meri*.

Kruspe, Nicole. In preparation c. Semaq Beri. To appear in Matthias Jenny & Paul Sidwell (eds.), *The Mon-Khmer Austroasiatic languages*. Leiden: Brill.

Kruspe, Nicole & John Hajek. 2009. Mah Meri. Illustrations of the IPA. *Journal of the International Phonetic Association* 39, 241-248.

Levinson, Stephen C. 2003. Language and mind: let's get the issues straight! In Dedre Gentner & Susan Goldin-Meadow (eds.), *Language in mind: advances in the study of language and thought*, 25-46. Cambridge MA: Bradford Books.

Levinson, Stephen C. & Niclas Burenhult. 2008. Semplates: a new concept in lexical semantics? *Language* 85, 150-172.

Levinson, Stephen C. & David P. Wilkins. 2006. The background to the study of the language of space. In Stephen C. Levinson & David P. Wilkins (eds.), *Grammars of space: explorations in cognitive diversity*, 1-23. Cambridge: Cambridge University Press.

Lloyd, R. J. 1921-23. On the phonology of the Malay and Negrito dialects spoken in the Malay states of Lower Siam. *Bulletin of the School of Oriental Studies, London Institution* 2, 27-38.

Lye, Tuck-Po. 1994. Batek Hep: culture, nature, and the folklore of a Malaysian forest people. MA thesis, University of Hawai'i at Manoa.

Lye, Tuck-Po. 1997. Knowledge, forest, and hunter-gatherer movement: the Batek of Pahang, Malaysia. PhD thesis, University of Hawaii at Manoa.

Lye, Tuck-Po. (ed.). 2001. *Orang Asli of Peninsular Malaysia: a comprehensive and annotated bibliography*. [Linguistics, 158-168.] Kyoto: Kyoto University Center for Southeast Asian Studies.

Lye, Tuck-Po. 2004. *Changing pathways: forest degradation and the Batek of Pahang, Malaysia*. Lanham MD: Lexington Books.

Lye, Tuck-Po. 2011. A history of Orang Asli studies: landmarks and generations. *Kajian Malaysia* 29 (Supp. 1), 23-52.

Majid, Asifa, Nicholas Evans, Alice Gaby & Stephen C. Levinson. 2011. The grammar of exchange: a comparative study of reciprocal constructions across languages. *Frontiers in Psychology, Cultural Psychology* 2 (article 34), 1-15.

Matisoff, James A. 2003. Aslian: Mon-Khmer of the Malay Peninsula. *Mon-Khmer Studies* 33, 1-58.

McCarthy, John. 1982. Prosodic templates, morphemic templates, and morphemic tiers. In Harry van der Hulst & Norval Smith (eds.), *The structure of phonological representations (part 1)*, 191-223. Dordrecht: Foris.

Means, Nathalie. 1998. *Temiar-English, English-Temiar dictionary*. Edited by Gordon P. Means. (With the assistance of Temiar collaborators: Abuk Alang, Aleg B. Along, Angek Asoi, Ahud Duga, Kamaruddin, Awin Pedik, Halimah Yunus.) St Paul MN: Hamline University Press.

Means, Nathalie & Paul B. Means. 1986. *Sengoi-English, English-Sengoi dictionary*. [Semai.] University of Toronto *and* York University: The Joint Centre on Modern East Asia.

Miklucho-Maclay, N. von. 1878. Dialects of the Melanesian tribes in the Malay Peninsula. *Journal of the Straits Branch, Royal Asiatic Society* 1, 38-44.

Miyakoshi, Koichi. 2006. Aslian reduplication as the emergence of the unmarked. *Gengo Kenkyu* 129, 43-89. http://www3.nacos.com/lsj/modules/documents/LSJpapers/journals/129_miyakoshi.pdf

Mon-Khmer Studies: A Journal of Southeast Asian Languages. [Periodical publication initiated in 1964, and currently published (since volume 20) jointly by SIL International (Dallas, Texas) and the Institute of Language and Culture for Rural Development of Mahidol University (Thailand).] Back issues available from http://www.mksjournal.org/.

Nabitoepoeloe, B. W. F. 1950. *Sengoi (Sakai) first primer*. [Semai.] Trichinopoly: Federation Evangelical Lutheran Churches in India.

Nik Safiah Karim. 1968. Penyelidikan bahasa-bahasa Orang Asli di Tanah Melayu. *Dewan Bahasa* 12.

Nik Safiah Karim & Ton Bte Ibrahim. 1979a. Semaq Beri: some preliminary remarks. *Journal of the Malaysian Branch, the Royal Asiatic Society* 47, 123-129.

Nik Safiah Karim & Ton Bte Ibrahim. 1979b. Semoq Beri: preliminary remarks. *Federation Museums Journal (New Series)* 24, 17-38.

Nong Pai. Newsletter published by the Malaysian Department of Aboriginal Affairs (JHEOA). The following is a partial list of language materials from this publication (by courtesy of Shuichi Nagata):

Anjang Sadang bin Itam. 1982a. Bahasa Semai Perak. *Nong Pai* 1, 83, 8.

Anjang Sadang bin Itam. 1982b. Bahasa Semai Perak: pelajaran sepuluh. *Nong Pai* 2, 82, 24-29.

Anjang Sadang bin Itam. 1982c. Bahasa Semai Perak: pelajaran empat belas. *Nong Pai* 4, 82, 22.

Anjang Sadang bin Itam. 1984. Semai and Temiar. *Nong Pai* 8, 84, 28-29.

Anjang Sadang bin Itam. 1985a. Bahasa Semai Perak. *Nong Pai* 5, 81, 10-11.

Anjang Sadang bin Itam. 1985b. Bahasa Semai: hek bulih hajimat duit tenagaru. *Nong Pai* 2, 85, 12.

Gasur. 1991. Bahasa Temiar. *Nong Pai* 1, 91, 25-26.

Jabatan Hal Ehwal Orang Asli. 1982a. Bahasa Semai Perak: pelajaran sepuluh. *Nong Pai* 1, 82, 16-22.

Jabatan Hal Ehwal Orang Asli. 1982b. Bahasa Semai Perak: pelajaran dua belas. *Nong Pai* 3, 82, 18-19.

Jabatan Hal Ehwal Orang Asli. 1987a. Belajar bahasa Temiar. *Nong Pai* 1, 89, 16.

Jabatan Hal Ehwal Orang Asli. 1987b. Belajar bahasa Temiar. *Nong Pai* 3, 89, 16.

Long Alang Chukas. 1981. Bahasa Semai. *Nong Pai* 81, 14.

Noone, H. D. 1936. Report on the settlements and welfare of the Ple-Temiar Senoi of the Perak-Kelantan watershed. *Journal of the Federated Malay States Museums* 19, 1-85.

Noone, H. D. 1939. Notes on the Benua Jakun language, spoken at Sungai Lenga, Ulu Muar, Johore. *Journal of the Federated Malay States Museums* 15, 139-162.

Parkin, Robert. 1991. *A guide to Austroasiatic speakers and their languages*. Honolulu: University of Hawaii Press.

Pelras, Christian. 1972. Notes sur quelques populations aquatiques de l'archipel nusantarien. *Archipel* 3, 133-168.

Peterson, Mary M. 1993. Lost in space: discovering spatial locatives in Kensiw. Draft paper. Thammasat University. http://msea-ling.info/digidata/495.992/b10212.pdf.

Peterson, Mary M.1997. Kensiw grammar: basic declarative, interrogative and imperative propositions. Thammasat University: TU-SIL-LRDP. http://msea-ling.info/digidata/495.992/b07428.pdf.

Phaiboon Duangchand. 1984. A phonological description of the Kansiw (Sakai) language. Unpublished MA thesis. Mahidol University.

Phaiboon Duangchand. 2002. The dialects of Aslian language in southern Thailand. Paper presented at First International Conference on Southern Thai Studies, Pattani, Thailand, 13-15 June 2002.

Phaiboon Duangchand. 2006. Glossary of Aslian languages: the Northern Aslian languages of southern Thailand. *Mon-Khmer Studies* 36, 207-224.

Phaiboon Duangchand. In preparation. *Aslian language in Southern Thailand dictionary*.

Phillips, Timothy C. 2005a. Linguistic comparison of Semai dialects. Unpublished typescript. Economic Planning Unit, Prime Minister's Department, Malaysia.

Phillips, Timothy C. 2005b. A Survey of nasal preplosion in Aslian languages. 22 pages. Paper presented at the *Conference on Indigenous People*, 4-5 July 2005, Kuala Lumpur.

Phillips, Timothy C. 2009. A brief phonology of Semai Betau dialect. In Chong Shin, Karim Harun, Yabit Abas (eds.), *Reflections in Borneo rivers: essays in honour of Professor James T. Collins (book II)*.1-22. Pontianak: STAIN Pontianak Press.

Phillips, Timothy C. 2011. Retention and reduction in reduplicants of Semai. In Sophana Srichampa & Paul Sidwell (eds), *Austroasiatic studies: papers from ICAAL4* (=*Mon-Khmer Studies, Special issue* no. 3), 178-197. Canberra: Pacific Linguistics E-9. http://icaal.org/ICAAL-4.2.pdf.

Phillips, Timothy C. 2012. Proto-Aslian: towards an understanding of its historical linguistic systems, principles and processes. PhD thesis, Institut Alam dan Tamadun Melayu, Bangi: Universiti Kebangsaan Malaysia.

Pinnow, Heinz-Jürgen. 1959. *Versuch einer historischen Lautlehre der Kharia-Sprache*. Wiesbaden: Harrassowitz.

Pothisarn, Chumphol. 2011. Investigating the lexical concept of *haɲa?* 'residence' in Maniq language. Paper prepared for presentation at the *Fifth International Conference on Austroasiatic Linguistics*, Mahidol University, Salaya, Thailand, November 2011.

Pusat Perkembangan Kurikulum. 2000. *Kamus Semai-Melayu-Inggeris*. 29 pages, duplicated. Kuala Lumpur: Kementerian Pendidikan Malaysia.

Pusat Perkembangan Kurikulum. 2004. *Kamus trilingual, Semai, Melayu, Inggeris*. Draf edisi terhad (Kursus Orientasi). 158 pages. Kuala Lumpur: Kementerian Pendidikan Malaysia.

Rischel, J. 1995. *Minor Mlabri: a hunter-gatherer language of northern Indochina*. Copenhagen: Museum Tusculanum Press.

Roseman, Marina. 1991. *Healing sounds from the Malaysian rainforest: Temiar music and medicine*. Berkeley: California University Press.

Sandbukt, Øyvind. 1983. The sea nomads of Southeast Asia: new perspectives on ancient traditions. *Annual Newsletter of the Scandinavian Institute of Asian Studies* 17, 3-13.

Schebesta, Paul. 1928a. The jungle tribes of the Malay Peninsula. Translated by C. O. Blagden. *Bulletin of the School of Oriental Studies, London Institution* 4, 269-278.

Schebesta, Paul. 1928b. Grammatical sketch of the Jahai dialect as spoken by a Negrito tribe of Ulu Perak and Ulu Kelantan. *Bulletin of the School of Oriental Studies, London Institution* 4, 803-826.

Schebesta, Paul. 1931. Grammatical sketch of the Ple-Tĕmer language. *Journal of the Royal Asiatic Society 1931,* 641-652.

Schmidt, Wilhelm. 1901. Die Sprachen der Sakei und Semang auf Malacca und ihr Verhältnis zu den Mon-Khmēr-Sprachen. *Bijdragen Tot de Taal-, Land- en Volkenkunde* 52, 399-583.

Shorto, H. L. 1962. *A dictionary of modern Spoken Mon*. London: Oxford University Press.

Shorto, H. L. 1971. *Dictionary of the Mon inscriptions from the sixth to the sixteenth centuries*. London: Oxford University Press.

Shorto, H. L. 2006. *A Mon-Khmer comparative dictionary*. Edited by Paul Sidwell, Doug Cooper & Christian Bauer. Canberra: Pacific Linguistics.

Sidwell, Paul. 2006. Preface. In H. L. Shorto, 2006. *A Mon-Khmer comparative dictionary*, vii-xxv. Edited by Paul Sidwell, Doug Cooper & Christian Bauer. Canberra: Pacific Linguistics

Sidwell, Paul. 2007. Lecture notes on Aslian. Available at http://people.anu.edu.au/~u9907217/languages/aslian_langs.html.

Sidwell, Paul. 2008. Issues in the morphological reconstruction of Proto-Mon-Khmer. In Claire Bowern, Bethwyn Evans & Luisa Miceli (eds.), *Morphology and language history: in honour of Harold Koch*, 251-265. Amsterdam: John Benjamins.

Sidwell, Paul. 2010. The Austroasiatic central riverine hypothesis. *Journal of Language Relationship* / *Вопросы Языкового Родства* 4, 117-134.

Sidwell, Paul & Roger Blench. 2011. The Austroasiatic Urheimat: the southeastern riverine hypothesis. In N. J. Enfield (ed.), *Dynamics of human diversity: the case of Mainland Southeast Asia*, 315-343. Canberra: Pacific Linguistics.

SIL International, Malaysia Branch. 2002. *Kamus Semai-Melayu-Inggeris*. Unpublished typescript.

Skeat W. W. & C. O. Blagden. 1906. *Pagan races of the Malay Peninsula*. Two volumes. London: MacMillan.

Sloan, Kelly. 1988. Bare-consonant reduplication: implications for a prosodic theory of reduplication. In H. Borer (ed.), *Proceedings of the Seventh West Coast Conference on Formal Linguistics*, 319-330. Stanford, CA: Stanford Linguistic Association.

Smith, Karla J. 2003. Minority language education in Malaysia: four ethnic communities' experiences. *International Journal of Bilingual Education and Bilingualism* 6, 52-65.

Solheim, Wilhelm G., II. 1980. Searching for the origin of the Orang Asli. *Federation Museums Journal (New Series)* 25, 61-75.

Starostin, S., G. Bronnikov & Phil Krylov. 1998-2011. Aslian etymology. http://starling.rinet.ru/cgi-bin/response.cgi?root=config&morpho=0&basename=\data\aas\asliet&first=1.

Stevens, M., N. Kruspe & J. Hajek. 2006. Register in Mah Meri: a preliminary phonetic analysis. In R. Hoffmann & H. Mixdorff (eds.), *Speech Prosody. 3rd International Conference Dresden, May 2-5, 2006,* 269-272. Dresden: TUD Press. http://aune.lpl.univ-aix.fr/~sprosig/sp2006/contents/papers/PS3-01_0256.pdf.

Tengku Intan Suzila Tengku Sharif, Mohd Yusri Mohamad Noor & Badli Esham Ahmad. 2012. Ethical concerns in language documentation: a case study of the Tembeling River Batek Aslian. *Second International Conference on Social Science and Humanity IPEDR* 31, 154-158. Singapore: IACSIT Press.

Ter Mors, Christine. 1983. Affix to X. Broselow & McCarthy versus Marantz: a reinterpretation. *The Linguistic Review* 3, 275-297.

Terrill, Angela & Niclas Burenhult. 2008. Orientation as a strategy of spatial reference. *Studies in Language* 32, 93-136.

The Bible Societies in Malaya. 1962. *Gah selamat du ki-tulis ya Bah Markus*. [St Mark's Gospel in Semai.]

Theraphan L. Thongkum. N.d. Ten'en wordlist. Unpublished typescript.

Thomas, David. 1992. On sesquisyllabic structure. *Mon-Khmer Studies* 21, 206-210.

Thurgood, Graham. 1999. *From ancient Cham to modern dialects: two thousand years of language contact and change*. Honolulu: University of Hawai'i Press.

Tufvesson, Sylvia. 2008. Abstract: Semantics and sensory encoding in Semai expressives. (Retrieved 21 September 2011) http://lingweb.eva.mpg.de/jakarta/seals_xviii/abstracts/sylvia.pdf.

Tufvesson, Sylvia. 2011. Analogy-making in the Semai sensory world. *The Senses and Society* 6, 86-95.

van Reijn, E. O. 1975. Les emprunts du vieux-malais dans les dialectes autochthones de la péninsule malaise et dans l'atchinois. *Sari, Occasional Papers* 1, 152-156. Kuala Lumpur: Institute of Malay Language, Literature and Culture, National University of Malaysia.

Walther, Markus. 2000. Temiar reduplication in one-level prosodic morphology. *Proceedings of the SIGPHON 2000 Workshop on Finite-State Phonology*, 13-21. Luxembourg, 6 August 2000. www.arxiv.org/pdf/cs.CL/0008015.

Watahomigie, Lucille J. & Akira Y. Yamamoto. 1992. Local reactions to perceived language decline. *Language* 68, 10-17.

Weber, Max. 1958 [1922-23]. The social psychology of the world religions. In H. H. Gerth & C. Wright Mills (eds), 1958, *From Max Weber: essays in sociology*, 267-301. New York: Oxford University Press.

Wilkinson, R. J. 1910. *The Aboriginal tribes*. Papers on Malay Subjects, Supplement. Reprinted 1926. Kuala Lumpur: FMS Government Press.

Wilkinson, R. J. 1915. *A vocabulary of Central Sakai*. Papers on Malay Subjects, Second series no. 3. Kuala Lumpur: Government Press.

Williams-Hunt, P. D. R. 1952. *An introduction to the Malayan Aborigines*. Kuala Lumpur: Government Press.

Wnuk, Ewelina. 2010. The phonology of Maniq (southern Thailand): a preliminary account. Unpublished Master's Thesis. University of Amsterdam.

Wnuk, Ewelina. Webpage on Maniq. http://www.mpi.nl/institute/research-departments-groups/language-and-cognition-ep/fieldsites/maniq.

Wnuk, Ewelina & Asifa Majid. 2012. Olfaction in a hunter-gatherer society: insights from language and culture.' [On Maniq.] In N. Miyake, D. Peebles & R. P. Cooper (eds.), *CogSci 2012: Proceedings of the 34th Annual Conference of the Cognitive Science Society*, Austin, TX: Cognitive Science Society. http://mindmodeling.org/cogsci2012/papers/0207/index.html.

Wolff, John U. 1976. Malay borrowings in Tagalog. In C. D. Cowan & O. W. Wolters (eds.), *Southeast Asian history and historiography: essays presented to D. G. E. Hall*, 345-367. Ithaca: Cornell University Press.

Wolff, John U. 2010. *Proto-Austronesian phonology with glossary*. Two volumes. Cornell University: Southeast Asia Program Publications.

Yager, Joanne. In preparation (2013). Northern Aslian prehistory: tracing the effects of contact. Master's thesis in General Linguistics. Lund University: Centre for Language and Literature.

Yap Ngee Thai. 2009a. Non-exhaustive syllabification in Temiar. *JSEALS: Journal of the Southeast Asian Linguistics Society* 1, 267-282.

Yap Ngee Thai. 2009b. The status of surface vowels in Temiar: evidence from a case study. Paper presented at *Seminar Pembentangan Hasil Penyelidikan Pensyarah Lantikan Baharu 2007*, Faculty of Modern Languages and Communication, Universiti Putera Malaysia.

Yap Ngee Thai. *Online basic Temiar course*. Universiti Putra Malaysia, Faculty of Modern Languages and Communication. www.fbmk.upm.edu.my/ccms/courses/CLad18/ (Accessed in 2009, but currently not available).